

中图法分类号: 文献标识码: A 文章编号: 1006-8961(XXXX)XX-0001-26

论文引用格式: Ying Xinyi, Zhang Xin, Zhou Xiang, Cao Xu, Lin Zaipin, Li Miao, Zhou Shilin, An Wei. A Simulation Dataset for Infrared Dim and Small Target Detection and Tracking in Space-Based Surveillance IRSatVideo[J/OL]. Journal of Image and Graphics, XXXX: 1-26. DOI: 10.11834/jig.250617. (应昕怡, 张鑫, 周翔, 曹旭, 林再平, 李淼, 周石琳, 安玮. 面向天基监视的红外弱小目标检测与跟踪仿真数据集 IRSatVideo[J/OL]. 中国图象图形学报, XXXX: 1-26. DOI: 10.11834/jig.250617. ) [DOI: 10.11834/jig.250617]

# 面向天基监视的红外弱小目标检测与跟踪仿真数据集 IRSatVideo

应昕怡, 张鑫, 周翔, 曹旭, 林再平, 李淼, 周石琳, 安玮

国防科技大学 电子科学学院, 长沙 410073

**摘要:** 天基红外序列图像中的弱小目标检测与跟踪技术在空间态势感知等领域具有重要应用价值。然而, 该领域的发展长期受限于真实数据获取困难、现有数据集规模小、场景单一且缺乏物理真实性的挑战。本文旨在构建一个大规模、高保真、涵盖多样轨道场景的红外弱小目标序列数据集, 以促进相关算法的研究与发展。具体而言, 本文提出一种半仿真的数据构建方法, 利用 Landsat 和高分四号卫星的真实红外影像作为背景, 通过精确建模卫星平台运动(姿态、轨道)、目标特性(形态、轨迹、辐射强度)以及复杂的背景杂波, 生成了包含低地球轨道(LEO)和地球同步轨道(GEO)两种场景的大规模序列数据集 IRSatVideo-LEO 与 IRSatVideo-GEO。IRSatVideo-LEO 与 IRSatVideo-GEO 数据集分别包含 200/200 段序列、220126/344285 个目标实例, 并提供实例级像素掩码和矩形框标注。实验显示, 在 IRSatVideo-LEO 与 IRSatVideo-GEO 数据集上, 典型基线算法可实现 20.38/33.70 的 IoU、41.05/73.39 的检测率与 8.56E-8/6.31E-6 的虚警率。IRSatVideo 是首个专为天基红外弱小目标检测与跟踪设计的大规模仿真数据集, 其创新性体现在三个方面: 动态模拟卫星平台运动引起的背景变化; 涵盖从机动目标到信杂比低至 1 的极弱目标; 包含多红外波段及全球多样背景。通过对目标尺寸、信杂比、运动特性及背景复杂度等关键属性的统计分析, 并与现有数据集对比, 表明 IRSatVideo 在规模、多样性和物理真实性方面具有显著优势。本数据集为天基红外弱小目标检测研究提供了一个接近真实、挑战性强的基准平台, 有望显著加速相关算法的研发与评测进程, 对推动空间态势感知技术进步具有重要意义。数据集下载链接: <https://github.com/XinyiYing/IRSatVideo>

**关键词:** 天基红外目标检测; 半仿真数据集; 弱小目标; 地球同步轨道; 低地球轨道

## A Simulation Dataset for Infrared Dim and Small Target Detection and Tracking in Space-Based Surveillance IRSatVideo

Ying Xinyi, Zhang Xin, Zhou Xiang, Cao Xu, Lin Zaipin, Li Miao, Zhou Shilin, An Wei

College of Electronic Science and Technology, National University of Defense Technology, Changsha 410073

**Abstract:** The detection and tracking of infrared dim and small targets within sequential imagery acquired by space-based platforms constitutes a cornerstone technology for critical applications such as space situational awareness (SSA), satellite collision avoidance, and the monitoring of space debris. However, the rapid development and rigorous benchmarking of advanced algorithms, particularly data-driven deep learning models, in this domain have been fundamentally hampered by a severe scarcity of high-quality, large-scale, and physically realistic datasets. The acquisition of real, annotated infrared satellite video data containing genuine dim targets is prohibitively difficult, expensive, and often restricted. Consequently, the field has relied on limited-scale datasets, which frequently suffer from a lack of scenario diversity, simplified or static

收稿日期: 2025-12-05; 修回日期: 2026-02-20

© 中国图象图形学报版权所有

backgrounds, and an absence of authentic sensor and platform motion dynamics, leading to a significant "reality gap" between simulation and operational conditions. To bridge this critical gap and propel the field forward, this paper introduces IRSatVideo, the first publicly available, large-scale, high-fidelity simulation dataset specifically designed for infrared dim and small target detection and tracking in space-based surveillance contexts. The core contribution lies in a novel semi-physical simulation methodology that ingeniously blends real-world elements with precise physical modeling to achieve unprecedented authenticity. Instead of relying on purely synthetic or simplistic backgrounds, our approach utilizes real infrared background imagery acquired by established Earth observation satellites: Landsat series satellites (for Low Earth Orbit scenarios) and the Gaofen-4 (GF-4) satellite (for Geostationary Earth Orbit scenarios). This ensures that the complex, heterogeneous clutter stemming from various Earth surfaces (land, water, urban areas), atmospheric phenomena (clouds of varying types and coverage), and deep space is accurately represented. Upon these authentic backgrounds, we generate synthetic target sequences through a comprehensive and physically-grounded modeling pipeline. This pipeline meticulously simulates: (1) Satellite Platform Dynamics: We model the precise motion characteristics of the sensing satellite itself. For the Geostationary Earth Orbit (GEO) scenario, this involves simulating subtle camera line-of-sight jitter induced by onboard mechanisms such as cryocoolers and momentum wheels, derived from real sensor data analysis. For the Low Earth Orbit (LEO) scenario, we model the continuous, predictable background translation caused by the satellite's orbital motion, calculated using orbital mechanics. (2) Target Characteristics: A diverse spectrum of target properties is simulated, including realistic trajectories (encompassing both linear and maneuvering motions), morphology (approximated via Gaussian kernels to model the point-spread function), and crucially, time-varying radiation intensity. Special emphasis is placed on generating extremely dim targets with Signal-to-Clutter Ratios (SCR) as low as 1, representing the most challenging detection cases critical for early warning systems. (3) Sensor and Imaging Effects: The imaging process for specific infrared wavebands (short-wave infrared for LEO, mid-wave infrared for GEO) is simulated, and sensor-specific anomalies like dead pixels in GF-4 imagery are detected and compensated for using advanced algorithms to enhance data quality. Applying this methodology, we construct two distinct, large-scale datasets tailored to the primary orbital regimes for space-based observation: IRSatVideo-LEO: Comprising 200 video sequences (91,021 total frames) synthesized from Landsat backgrounds. It contains 220,126 annotated target instances and simulates the dynamic background motion and higher-fidelity targets typical of LEO surveillance. IRSatVideo-GEO: Comprising 200 video sequences (94,136 total frames) synthesized from GF-4 backgrounds. It contains 344,285 annotated target instances and focuses on the extreme challenge of detecting ultra-dim, slow-moving targets against a relatively static but cluttered background, characteristic of GEO staring observation. Both datasets are meticulously partitioned into training and testing sets (160/40 sequences each) following strict principles to ensure diversity and prevent bias. Each dataset provides rich, instance-level annotations, including precise pixel-level masks and bounding boxes for every target in every frame, supporting both detection and tracking tasks, as well as pixel-wise segmentation research. **The innovations of the IRSatVideo dataset are manifold** 1) It dynamically simulates background changes caused by actual satellite platform motion (both jitter and translation), a critical factor absent in prior static-background datasets. 2) It encompasses a vast continuum of target challenges, from maneuvering objects to extremely dim targets with SCR values as low as 1. 3) It incorporates multiple infrared spectral bands and leverages globally diverse, real satellite backgrounds, ensuring broad generalization capability for trained algorithms. Statistical analyses confirm the dataset's extensive coverage in terms of target size (mostly sub-10x10 pixels), SCR distribution, motion attributes, and background complexity levels.

**Key words:** Space-based infrared target detection; Synthetic dataset; Dim and small target; Geostationary Orbit (GEO); Low Earth Orbit (LEO)

## 0 引言

高质量且多样化的公开数据集是深度学习算法

发展的基石,然而,在红外弱小目标检测与跟踪领域,虽然有Wang等人(2019)的公开可用的数据集,Wang等人(2024)为算法学习精确的特征表示提供了必要条件,但仍然十分有限,这一现状对该方向的

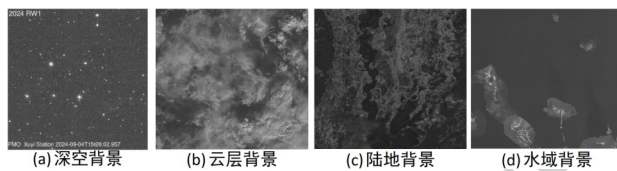
研究构成了显著挑战。根据 Ying 等人(2025)的研究,在天基红外监视场景下,由于真实卫星视频数据的获取受到条件限制,公开数据集几乎空白,极大地阻碍了相关算法的研究和验证。因此,根据 Li 等人(2025)对红外弱小目标检测方法 with 数据研究的结果,构建高保真、规模化的天基红外弱小目标数据集对于推动该领域的发展具有重要意义, Ying 等人(2023)实现了在复杂多变的场景中稳定且可靠的预测与决策, Ying 等人(2025)给出了前沿视觉目标检测相关数据集。

近年来,得益于红外成像技术与数据处理能力的不断提升, Li 等人(2022)开始尝试采集、合成并公开红外弱小目标检测数据集,这些工作在一定程度上推动了该领域的发展。现有数据集大致可分为单帧图像数据集,如 Wang 等人(2019)所使用,与序列图像数据集,如 Hui 等人(2019)所使用,其中单帧数据集由独立时间点的图像组成,标注成本较低,但通常规模有限且目标尺寸偏大,难以满足极弱小目标的检测需求;序列数据集由连续帧构成,能够利用时间信息刻画目标运动特征,更适用于动态检测与跟踪任务,且通常包含更多小尺寸、低信杂比(Signal-Clutter Ratio, SCR)的目标。然而,当前大多数序列数据集仅提供矩形框或点标签,缺乏实例级像素掩码标注,限制了方法间的全面比较,尤其不利于将红外弱小目标检测视为逐像素任务的研究进行有效评估。

在已有数据集中, Zhang 等人(2024)所提出的 IRDST 数据集提供了掩码标签,但其采集方式主要基于近景成像系统和云台摄像机,导致其存在视野局限、视差变化及非规律性背景运动等特征,与天基红外卫星视频的实际应用场景存在显著差异。卫星遥感系统往往具备更大视场、更稳定的成像几何关系,以及由轨道与姿态控制决定的慢速、规律性背景运动,因此现有数据集无法满足天基红外监视场景下弱小目标检测与跟踪的需求。

为攻克上述数据瓶颈,受 Ying 等人(2025)

的其他数据稀缺领域的启发,本文提出并构建了 IRSatVideo 数据集。该数据集是首个面向天基监视的大规模高保真红外弱小目标仿真数据集。IRSatVideo 采用半仿真构建方法,以真实的卫星红外影像作为背景,结合精确建模的卫星轨道与姿态、目标外观、运动轨迹和辐射强度,生成符合低地球轨道



((a) deep space background; (b) cloud background; (c) land background; (d) water background)

图1 典型的的天基红外背景图像

Fig. 1 Typical space-based infrared background image

(Low Earth Orbit, LEO)与地球同步轨道(Geostationary Earth Orbit, GEO)观测特性的红外序列数据。该数据集不仅包含大量尺寸微小、SCR极低的目标,还提供了实例级像素掩码与矩形框双重标注,能够更好地支持检测与跟踪方法的研究与评测。IRSatVideo 的发布为天基红外弱小目标检测与跟踪算法提供了一个接近真实、挑战性强的基准平台,有望推动空间态势感知和天基监视等领域的技术发展。

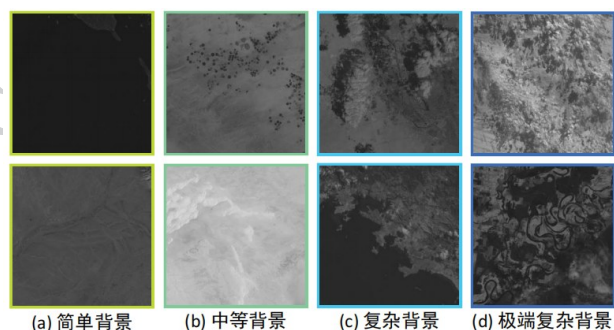
## 1 数据集仿真方法

为解决天基红外序列图像弱小目标数据稀缺的问题,本文借鉴了 Li 等人(2022)的其他数据受限领域的策略,提出了一种半仿真数据生成方法。该方法以真实卫星红外图像为背景,结合卫星轨道与姿态建模、目标时空域强度分布模型及天基成像噪声特性,生成高保真度的红外弱小目标仿真序列。整体流程包括五个关键步骤:背景图像采集、背景序列仿真、目标形态序列仿真、目标轨迹仿真以及目标背景混叠。其核心思想是在真实背景的基础上,精确刻画卫星成像几何、目标辐射特性及成像噪声特性,从而生成既具物理真实性又能支持算法评测的卫星序列数据。

### 1.1 背景图像采集

在天基红外成像中,根据背景特征可将图像背景分为深空背景(宇宙辐射、恒星点源)、大气背景(大气辉光、气辉、云层)和地球表面背景(陆地、水域、城市等)。深空背景包括均匀低辐射的宇宙背景和呈点状分布的恒星背景,后者可能因大

气扰动或传感器噪声出现轻微闪烁或形态变化。大气背景涵盖大气辉光、气辉等现象以及具有不同纹理与灰度特征的云层,云边缘通常因透光性更好而显得更亮。地球表面背景包含陆地区域(如



(a) simple background; (b) moderate background; (c) complex background; (d) extremely complex background

图2 不同复杂度下的背景图像

Fig. 2 Background images under different complexity levels

山脉、森林、城市等)和水域,其中陆地因地表材质不同呈现多样辐射特性,水体则多表现为均匀的低温区域。地球表面背景的纹理复杂性与灰度多样性对红外目标探测构成重要挑战。图1展示了典型的深空、大气和地球表面背景示例。为量化不同背景的复杂程度,本文采用图像的平均信息熵和方差来评估:

$$C = - \sum_{s=0}^{255} [s - \bar{s}(I)] \cdot p_s(I) \cdot \log[p_s(I)] \quad (1)$$

式中, $I$ 表示图像, $s$ 表示图像直方图中的灰度级。 $p_s(I)$ 表示灰度 $s$ 在图像中出现的概率。背景复杂度可以粗略分为四个级别:简单背景 $C \in [0, 200)$ 、中等复杂背景 $C \in [200, 1000)$ 、复杂背景 $C \in [1000, 2000)$ 和极端复杂背景 $C \in [2000, \infty)$ 。图2展示了不同复杂度下的天基红外背景图像。这些背景类别和复杂度信息直接用于后续仿真场景的选取与难度分级。

背景数据的真实性直接决定了数据集的物理可信度。本文选取Landsat系列卫星(LEO,低轨)和高分四号卫星(GEO,地球同步轨道)获取的真实红外影像作为背景。

1)LEO背景:Landsat 8与Landsat 9提供了覆盖全球的短波红外影像,其轨道高度约705km,倾角 $98.2^\circ$ ,重访周期16天。本文通过随机抽样从全球不同地理区域(涵盖多种地形和云量条件,云覆盖率0-61.25%)获取背景影像,确保了背景场

## 1.2 背景序列仿真

表1 GEO卫星相机参数

Table 1 Parameters of GEO satellite camera

| 参数     | 数值               | 备注                   |
|--------|------------------|----------------------|
| 波长     | 2.85微米           | 短波红外                 |
| 量子效率   | 0.6              | 探测器将吸收的光子转换为电光子的概率   |
| 普朗克常数  | $6.626E-34$ 焦耳·秒 | 粒子的能量与其频率之间的关系       |
| 光速     | $3E8$ 米/秒        | 光在真空中的传播速度           |
| 系统透过率  | 0.6              | 光学系统允许通过的光的比例        |
| 地面分辨率  | 600米             | 相机能够分辨的两个地面目标之间的最小距离 |
| 探测器尺寸  | 15微米×15微米        | 单个探测器元素(像素)的物理尺寸     |
| 光学孔径   | 0.35米            | 相机镜头的直径              |
| 卫星轨道高度 | 35786千米          | 卫星传感器到地球表面的距离        |
| 地球半径   | 6371千米           | 地球半径的平均值             |
| 积分时间   | 0.4秒             | 探测器收集光信号的时间长度        |
| 成像帧率   | 1赫兹              | 相机每秒能够采集的图像帧数        |

景的多样性和复杂性。

2)GEO背景:高分四号卫星运行于地球同步轨道,具备400km幅宽和20秒重访周期。本文选取其中波红外影像,随机抽样自中国典型地理位置,以反

映不同地表特征和大气条件。

本小节将系统阐述天基红外序列图像的背景仿真方法。首先,为实现对卫星观测几何关系的精确模拟,本文分别对GEO与LEO的轨道动力学模型进

行建模。GEO轨道模型侧重于描述卫星相对于地球表面的微小抖动,而LEO轨道模型则需精确计算卫星运动及其引起的像面连续移动。在此基础上,本文将详细介绍背景序列的仿真过程:利用真实卫星红外影像模拟生成由卫星平台运动引起的背景场景在时序上的连续变化,为后续注入合成目标提供了物理可信的动态背景环境。

#### 地球同步轨道卫星平台运动特性分析

由于地球同步轨道上的凝视型红外相机相对地球静止,因此本小节主要分析了相机的视线抖动特性:首先利用角位移传感器数据和相关相机参数(见表1),分析了相机的视线抖动频谱,随后计算推导了像平面抖动量。

##### 1) 视线抖动频谱分析

地球同步轨道凝视型红外相机因较长的积分时间对抖动更敏感。卫星的活动部件如数传天线、太阳能电池阵驱动机构、动量轮和脉冲管制冷机都可能引起相机抖动。其中,数传天线在成像时静止,而太阳能电池阵驱动机构的扰动主要在低频段(0.2~20Hz)且影响较小。因此,本小节主要聚焦于制冷机和动量轮作为主要因素,通过主镜角位移传感器

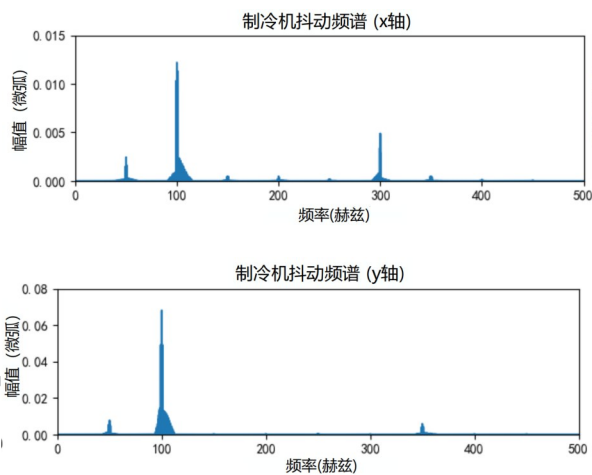


图3 制冷机导致的相机抖动频谱示意图

Fig. 3 Schematic diagram of camera vibration spectrum caused by refrigerator

数据,分析了这两者的视线抖动频谱。

卫星在轨期间,制冷系统负责保持红外焦平面的低温运作,以优化探测器的成像效能。制冷机工作期间,高压气体的压力波动以及活塞等机械部件的动量不均衡,引发了相机视线的抖动现象。图3

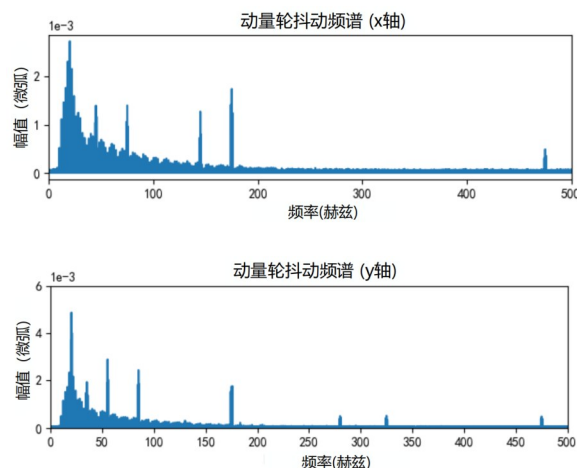


图4 动量轮导致的相机抖动频谱示意图

Fig. 4 Schematic diagram of camera vibration spectrum caused by momentum wheel

展示了由制冷机因素引起的相机在  $x$  轴和  $y$  轴方向上的视线抖动频谱。从图中可以得知,制冷机引起的相机视线抖动呈现出典型的谐波特性,其频率与制冷系统的操作频率及其倍频相对应。

在动量轮中,高速旋转的惯性元件由于制造工艺的限制,其质心与旋转轴之间存在偏差,引发了旋转不平衡,进而导致了相机视线的抖动现象。图4展示了动量轮因素导致相机在  $x$  轴和  $y$  轴方向上的视线抖动频谱。从图中可以得知,动量轮引起的相机视线抖动呈现出典型的谐波特性和宽频噪声特性。

##### 2) 像平面抖动量分析

根据制冷机和动量轮的相机视线抖动频谱和序列图像长度,得到相机视线抖动的时域特性,图5展示了100s时间范围内的相机抖动时域特性。

对上述制冷机和动量轮导致的相机视线抖动时域特性曲线,按相机帧频(1Hz)进行采样并叠加得到每帧的抖动角:

$$\sigma'_x = \text{Sample}_t(\tilde{\sigma}_{J_x}) + \text{Sample}_t(\tilde{\sigma}_{C_x}) \quad (2)$$

$$\sigma'_y = \text{Sample}_t(\tilde{\sigma}_{J_y}) + \text{Sample}_t(\tilde{\sigma}_{C_y}) \quad (3)$$

式中,  $\text{Sample}_t(\cdot)$  为按时间  $t$  对相机视线抖动时域特性曲线进行采样。  $(\sigma'_x, \sigma'_y)$  为  $x$  轴和  $y$  轴的相机视线抖动时域特性曲线,  $(\tilde{\sigma}_{C_x}, \tilde{\sigma}_{C_y})$  为制冷机导致的  $x$  轴和  $y$  轴的相机视线抖动时域特性曲线,  $(\tilde{\sigma}_{J_x}, \tilde{\sigma}_{J_y})$  为动量轮导致的  $x$  轴和  $y$  轴的相机视线抖动时域特性曲线。

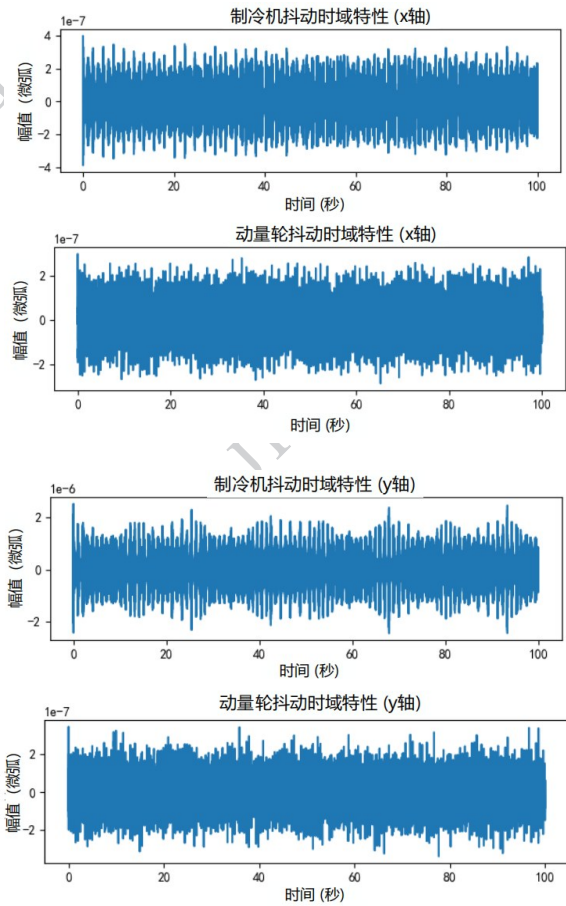


图5 相机抖动时域特性示意图

Fig. 5 Schematic diagram of camera vibration time-domain characteristics

结合相机视线抖动量 $(\sigma'_x, \sigma'_y)$ 、卫星轨道高度 $h$ 、地球半径 $R$ 和相机对地分辨率 $K$ ,可以计算得到每一时刻像平面抖动量。随后对其按时间进行

积分即可得到序列图像中每一帧的像平面抖动量:

$$\sigma'_{lx} = (h + R) \cdot \sin\left(\sum_{i=1}^t \sigma'_x\right) / K \quad (4)$$

$$\sigma'_{ly} = (h + R) \cdot \sin\left(\sum_{i=1}^t \sigma'_y\right) / K \quad (5)$$

式中, $\sigma'_{lx}, \sigma'_{ly}$ 为首帧参考系下序列图像中每一帧的像平面抖动量。通过计算,可得相机视线抖动在 $x$ 轴的最大值为 $3.087\text{E}-06\mu\text{rad}$ ,平均值为 $4.761\text{E}-08\mu\text{rad}$ , $y$ 轴的最大值为 $1.377\text{E}-5\mu\text{rad}$ ,平均值为 $1.264\text{E}-7\mu\text{rad}$ 。像平面偏移量在 $x$ 轴的最大值为 $0.087$ 像素,平均值为 $0.061$ 像素,像平面偏移量在 $y$ 轴的最大值为 $0.214$ 像素,平均值为 $0.145$ 像素,模值的最大值 $0.231$ 像素,均值为 $0.157$ 像素,符合实

际情况。

### 1.1.1 低地球轨道卫星平台运动特性分析

低地球轨道卫星绕地球的轨道可近似为圆轨道,即偏心率为0的椭圆轨道。卫星在圆轨道上任意一点的线速度均相同:

$$V = \sqrt{\frac{GM_e}{h + R}} \quad (6)$$

低地球轨道周期为:

$$T = 2\pi \sqrt{\frac{(h + R)^3}{GM_e}} \quad (7)$$

常见的地球圆轨道的轨道高度、速度和周期见表2。

卫星与地球之间均存在相对运动,然而由于观测距离较远(200-36000km)且观测周期较长(88.3-718.3min),因此像平面的背景运动不存在近景采集平台固有的视差问题,近似为全局一致的匀加速/匀速直线运动,线速度数值见公式(5)和(6)。结合卫星线速度 $V$ 、卫星轨道高度 $h$ 、地球半径 $R$ 和相机对地分辨率 $K$ ,相机帧率 $f$ ,可以得到每帧的像平面速度 $\sigma$ 。

$$\sigma = (h + R) \cdot \sin\left[\frac{V}{(h + R) \cdot f}\right] / K \quad (8)$$

### 1.1.2 背景序列仿真过程

背景序列具体步骤见图6,仿真参数设置如表3所示。考虑到卫星运动的复杂性,本文将建模为包含六个自由度的动态系统:三维姿态(俯仰角 $\alpha$ 、偏航角 $\beta$ 、滚转角 $\gamma$ )和三维位置(沿 $x$ 轴、 $y$ 轴和 $z$ 轴的平移)。首先对全局背景图像应用基于三维姿态序列的单应性变换以生成全局背景序列。

$$I_t^{CB} = I^{CB} \otimes H(\alpha_t, \beta_t, \gamma_t) \quad (9)$$

式中, $I^{CB} \in R^{h_0 \times w_0}$ 为全局背景图像, $I_t^{CB}, t \in [1, T]$ 为全局背景序列, $T$ 为序列长度。

$\alpha_t, \beta_t, \gamma_t$ 代表 $t$ 时刻的俯仰角、偏航角和滚转角, $H(\cdot)$ 代表单应性变换, $\otimes$ 表示矩阵乘法。

在研究卫星姿态控制问题中,本文考虑了两种不同的卫星姿态模式:无调度(固定姿态)和有调度(姿态缓慢变化)。对于无调度的卫星,本文随机设置固定的三维姿态,记为 $(\alpha, \beta, \gamma) = (\alpha_0, \beta_0, \gamma_0)$ ,其中 $\alpha_0 \in [-10, 10]$ , $\beta_0 \in [-10, 10]$ , $\gamma_0 \in [-10, 10]$ 。对于有调度的卫星,本文设定了初始和最终的三维姿态,分别为 $(\alpha_1, \beta_1, \gamma_1)$ 和

表2 常见的地球圆轨道的轨道参数

Table 2 Orbital parameters of common circular earth orbits

| 轨道类型   | 轨道高度 (km) | 卫星速度 (km/s) | 轨道周期 (min)   |
|--------|-----------|-------------|--------------|
| 低地球轨道  | 200       | 7.8         | 88.3         |
|        | 500       | 7.6         | 94.4         |
|        | 1000      | 7.4         | 104.9        |
| 中轨道    | 5000      | 5.9         | 201.1        |
|        | 10000     | 4.9         | 347.4        |
|        | 20200     | 3.9         | 718.3(12小时)  |
| 地球同步轨道 | 35800     | 3.1         | 1436.2(24小时) |

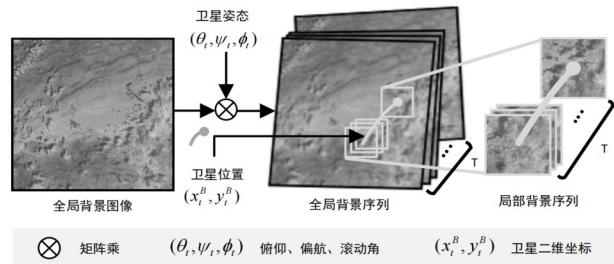


图6 天基红外运动小目标数据仿真方法中背景序列仿真的详细步骤示意图

Fig.6 Schematic diagram of detailed steps for background sequence simulation in the space-based infrared moving small target data simulation method

$(\alpha_T, \beta_T, \gamma_T)$ , 并采用线性插值方法生成中间姿态。其中, 姿态调度的范围  $(\Delta\alpha, \Delta\beta, \Delta\gamma) = (\alpha_T - \alpha_1, \beta_T - \beta_1, \gamma_T - \gamma_1)$  中的每个分量取值范围为  $[-5, 5]$ 。这两种姿态模式的设置概率均等, 为 0.5。

在研究卫星的三维坐标问题中, 由于天基成像系统的成像距离较远, 因此沿  $z$  轴的平移被忽略, 仅关注二维平面上的运动。针对地球同步轨道卫星, 二维卫星位置序列  $(x_i^b, y_i^b)$  由公式(4)和(5)求得。针对大椭圆轨道,  $(x_i^b, y_i^b)$  由公式(6)和(9)求得像平面移动距离, 从而确定序列

的起始点  $(x_1^b, y_1^b)$  和终点  $(x_T^b, y_T^b)$ , 随后通过二次插值方法构造二维卫星位置序列。针对中低圆轨道,  $(x_i^b, y_i^b)$  由公式(8)和(9)在  $x$  轴和  $y$  轴分

解求得。计算求得像平面移动距离, 从而确定序列的起始点  $(x_1^b, y_1^b)$  和终点  $(x_T^b, y_T^b)$ , 随后通过二次插值方法构造二维卫星位置序列。根据二维卫星位置序列, 基于预定义的视场范围  $(H_0, W_0)$  进行图像裁剪, 得到局部背景序列:

$$I_{LB}^i = \text{Crop}(I_{CB}^i, x_i^b, y_i^b, H_0, W_0) \quad (10)$$

式中,  $I_{LB}^i$  为局部背景序列,  $\text{Crop}$  表示从全局背景序列  $I_{CB}^i$  中按卫星二维位置  $x_i^b, y_i^b$  裁剪局部背景序列。

### 1.3 目标形态序列仿真

由于天基成像系统和光学衍射效应的影响, 红外小目标在成像中通常呈现为点状或艾里斑的形式, 目标的形态可以通过高斯核函数  $G(a, b, \sigma)$  进行近似表示, 其中  $a, b$  分别表示高斯核的长轴和短轴,  $\sigma$  表示标准差, 决定了核函数的平滑程度及其影响范围。目标形态序列仿真具体步骤见图7, 参数如表4所示。

在构建目标形态序列的过程中, 本文在每个图像序列中随机生成  $N$  个高斯目标,  $N$  取值范围是

$[1, 6]$ 。以第  $n$  个高斯目标形态序列

$G_{nt}(a_i, b_i, \sigma_i)$  为例, 首先随机选取  $K$  个关键形态表征  $\{G_{nt_1}, G_{nt_2}, \dots, G_{nt_K}\}$ , 分别对应于序列的

起始点  $t_1 = 1, t_K = T$ , 以及中间的时间点  $\{t_2, \dots, t_{K-1} \in (1, T)\}$ , 其中  $K$  的取值范围是  $[2, 5]$ 。

表3 背景序列仿真的参数设置

Table 3 Parameter settings for background sequence simulation

| 符号         | 描述                | 分布                         | 值  |
|------------|-------------------|----------------------------|--|
| $t$        | 时间                | 线性                         | $t = \{1, 2, \dots, T-1, T\}, T \in [200, 1200]$               |
| $\alpha_t$ | 时刻 $t$ 的俯仰角       | 常数                         | $\alpha_t = \alpha_0, \alpha_0 \in [-10, 10]$                  |
|            |                   | 线性                         | $\Delta\alpha = \alpha_T - \alpha_1, \Delta\alpha \in [-5, 5]$ |
| $\beta_t$  | 时刻 $t$ 的偏航角       | 常数                         | $\beta_t = \beta_0, \beta_0 \in [-10, 10]$                     |
|            |                   | 线性                         | $\Delta\beta = \beta_T - \beta_1, \Delta\beta \in [-5, 5]$     |
| $\gamma_t$ | 时刻 $t$ 的滚动角       | 常数                         | $\gamma_t = \gamma_0, \gamma_0 \in [-10, 10]$                  |
|            |                   | 线性                         | $\Delta\gamma = \gamma_T - \gamma_1, \Delta\gamma \in [-5, 5]$ |
| $x_t^B$    | 时刻 $t$ 沿 $x$ 轴的平移 | GEO 抖动<br>HEO 位移<br>LEO 位移 | 公式(4), 公式(5)<br>公式(6), 公式(9)<br>公式(8), 公式(9)                   |
| $y_t^B$    | 时刻 $t$ 沿 $y$ 轴的平移 | GEO 抖动<br>HEO 位移<br>LEO 位移 | 公式(4), 公式(5)<br>公式(6), 公式(9)<br>公式(8), 公式(9)                   |
| $z_t^B$    | 时刻 $t$ 沿 $z$ 轴的平移 | 常数                         | $\approx 0$  |
| $H_0$      | 视场的高              | 常数                         | 处理系统的尺寸, 常见为 1024, 2048, 4096                                  |
| $W_0$      | 视场的宽              | 常数                         | 处理系统的尺寸, 常见为 1024, 2048, 4096                                  |

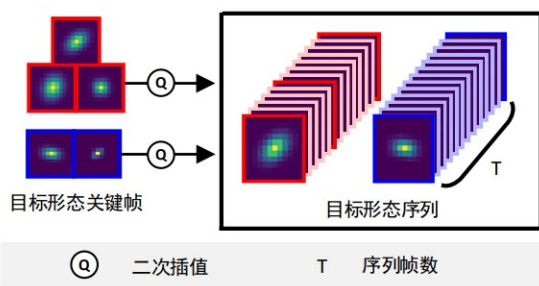


图7 天基红外运动小目标数据仿真方法中目标形态序列仿真的详细步骤示意图

Fig. 7 Schematic diagram of detailed steps for target morphology sequence simulation in the space-based infrared moving small target data simulation method

随后, 本文对参数  $(a_i, b_i, \sigma_i)$  进行二次插值, 计算出序列中各个关键形态表征点之间的中间形态, 从而生成一个连续的目标形态序列。

#### 1.4 目标轨迹仿真

图8展示了卫星视频序列中机动目标的轨迹示例, 目标的真实轨迹在短时间内表现为连续且相对平滑的曲线, 可以通过低阶曲线进行近似, 阶数  $p \in (0, 3)$ 。

目标轨迹仿真具体步骤见图9, 参数如表5所示。考虑到机动目标可能会改变其移动方向和速度, 为了模拟具有  $S$  次机动的第  $n$  个目标轨迹, 首先通过低阶曲线合成  $S$  个局部目标轨迹, 并将其顺序连接。然后, 通过一维插值方法平滑两个相邻轨迹的交点, 从而得到一个全局平滑的轨迹  $(x_n^T, y_n^T)$ 。图10展示了首帧参考系下三个机动目标和一个无机动目标的轨迹。从图中可以看出, 目标轨迹在局部区域内是连续的, 而在整体上是平滑的, 这与实际目标运动的特性相符合。

值得注意的是, 在图

表4 目标形态序列仿真的参数设置

Table 4 Parameter settings for target morphology sequence simulation

| 符号         | 描述               | 分布         | 值                            |
|------------|------------------|------------|------------------------------|
| $a_t$      | 时间 $t$ 下高斯目标的长轴  | 连续多个二次状态级联 | $a_t \in [1, 9]$             |
| $b_t$      | 时间 $t$ 下高斯目标的短轴  | 连续多个二次状态级联 | $b_t \in [1, 9]$             |
| $\sigma_t$ | 时间 $t$ 下高斯目标的标准差 | 连续多个二次状态级联 | $\sigma_t \in [0.1, 1]$      |
| $N$        | 序列中目标的数量         | 常数         | $N \in \{1, 2, 3, 4, 5, 6\}$ |
| $K$        | 序列中某一目标的关键形态表征数量 | 常数         | $K \in [2, 3, 4, 5]$         |



((a) target trajectory (no maneuvering); (b) target trajectory (one maneuver); (c) target trajectory (two maneuver))

图8 卫星视频中的目标轨迹示例图

Fig. 8 Example diagram of target trajectories in satellite videos

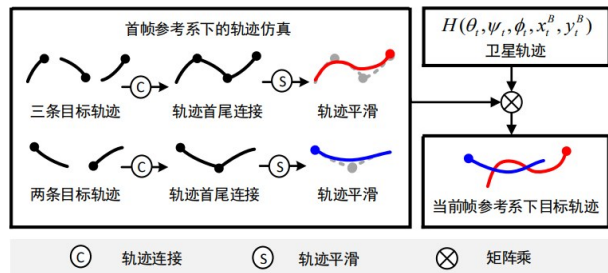


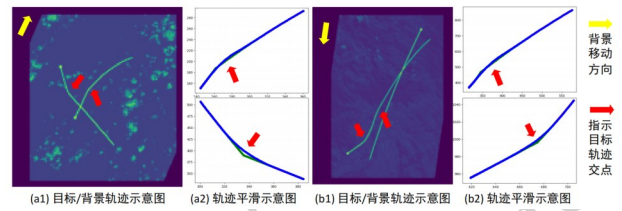
图9 天基红外运动小目标数据仿真方法中目标轨迹仿真的详细步骤示意图

Fig. 9 Schematic diagram of detailed steps for target trajectory simulation in the space-based infrared moving small target data simulation method

10 (b1)中, 一个目标的移动方向与背景的移动方向相反, 这表明在仿真序列的某些帧中, 目标可能会暂时消失。

在考虑卫星运动的情况下, 本文采用单应性变换来计算图像平面中的目标位置。具体来说, 首帧参考系中的目标轨迹通过卫星运动的单应性矩阵进行变换, 以生成当前帧参考系中的轨迹。该变换过程可以表示为:

$$(\hat{x}_T^n, \hat{y}_T^n) = (x_T^n, y_T^n) \otimes H(\alpha_t, \beta_t, \gamma_t, x_B^t, y_B^t) \quad (11)$$



((a1) schematic diagram of target/background trajectory; (a2) schematic diagram of trajectory smoothing; (b1) schematic diagram of target/background trajectory; (b2) schematic diagram of trajectory smoothing)

图10 首帧参考系下仿真机动目标和背景运动轨迹示意图

Fig. 10 Schematic diagram of simulated maneuvering target and background trajectories under the first-frame reference system

式中  $(x_{nt}^T, y_{nt}^T)$  和  $(\hat{x}_{nt}^T, \hat{y}_{nt}^T)$  分别表示当首帧参考系和当前帧参考系下的目标轨迹,  $(\alpha_t, \beta_t, \gamma_t, x_B^t, y_B^t)$  是卫星运动的参数。

### 1.5 目标背景混叠

目标背景混叠的具体步骤如图 11 所示, 参数设置如表 6 所示。首先, 将第  $n$  目标的外观序列  $G_n$  和强度序列  $E_n$  进行逐元素相乘, 以生成目标模板序列  $I_n$ :

$$I_n = G_n \odot E_n \quad (12)$$

$$E_n = [scr \times \sigma(T_1^{LB}) + \mu(T_1^{LB})] \times (1 + a_n) \quad (13)$$

式中  $\mu(T_1^{LB})$  和  $\sigma(T_1^{LB})$  分别代表起始时刻  $t = 1$  目标局部背景的均值和标准差,  $scr$  为目标信杂比, 其值从  $[1, 20]$  区间随机采样。  $a_n$  为目标加速度序列,  $\odot$  代表逐元素乘法。最后, 将目标模板序列和局部背景序列进行自适应加权求和, 并应用高斯模糊函数以实现图像的平滑处理。该过程可表述为:

表5 目标轨迹仿真的参数设置

Table 5 Parameter settings for target trajectory simulation

| 符号               | 描述                  | 分布         | 值                 |
|------------------|---------------------|------------|-------------------|
| $x_{nt}^T$       | 首帧参考系下 t 时刻的 x 轴坐标  | 多个平滑低阶曲线级联 | -                 |
| $y_{nt}^T$       | 首帧参考系下 t 时刻的 y 轴坐标  | 多个平滑低阶曲线级联 | -                 |
| $\hat{x}_{nt}^T$ | 当前帧参考系下 t 时刻的 x 轴坐标 | 多个平滑低阶曲线级联 | -                 |
| $\hat{y}_{nt}^T$ | 当前帧参考系下 t 时刻的 y 轴坐标 | 多个平滑低阶曲线级联 | -                 |
| $p$              | 曲线阶数                | 常数         | $p \in [0,3]$     |
| $S$              | 目标机动次数              | 常数         | $S \in \{0,1,2\}$ |

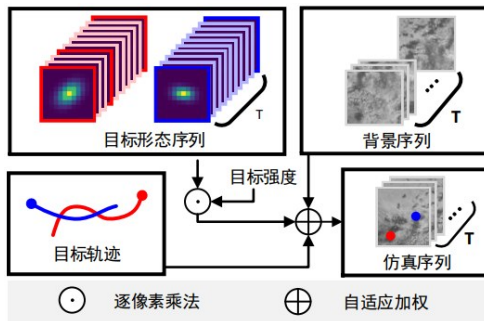


图11 天基红外运动小目标数据仿真方法中目标背景混叠的详细步骤示意图

Fig. 11 Schematic diagram of detailed steps for target-background aliasing in the space-based infrared moving small target data simulation method

式中,  $\hat{I}_{nt}^{sim}$  和  $I_{nt}^{LB}$  分别代表仿真生成的图像序列和局部背景序列在坐标  $(\hat{x}_{nt}^T, \hat{y}_{nt}^T)$  处的图像块。由于坐标  $(\hat{x}_{nt}^T, \hat{y}_{nt}^T)$  的值通常是分数, 因此需要采用双线性插值方法以获取准确的像素值。  $Norm(I) = I/\max(I)$  表示图像归一化,  $\max(I)$  表示图像的最大

像素值。  $GaussianBlur$  表示高斯模糊函数以实现图像平滑, 其尺寸  $k_{gb}$  从  $\{3, 5, 7\}$  集合随机采样, 标准差  $\sigma_{gb}$  从  $[0.2, 0.6]$  区间随机采样。最后, 通过随机采样并混叠高斯白噪声和低频噪声, 来模拟图像中的噪声分布, 从而生成含有噪声的序列图像。

## 2 数据集属性

根据第二章所述的天基红外序列图像小目标仿真方法, 本章构建了两个面向天基监视任务的红外弱小目标检测与跟踪仿真数据集: IRSatVideo-LEO 和 IRSatVideo-GEO。本章将系统介绍这两个数据集的构成、属性及其在目标检测与跟踪任务中的特点, 并与现有数据集进行对比分析, 以凸显其在天基红外监视研究中的独特价值。

### 2.1 IRSatVideo-LEO 数据集

IRSatVideo-LEO 仿真数据集基于 200 张由 Landsat7-8 采集的短波红外波段 ( $2.1 - 2.3\mu m$ ) 背景图

表6 目标背景混叠的参数设置

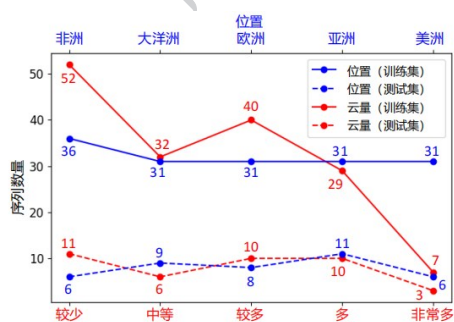
Table 6 Parameter settings for target-background aliasing

| 符号            | 描述                              | 分布       | 值                           |
|---------------|---------------------------------|----------|-----------------------------|
| $scr$         | SCR value of a target at time 1 | Constant | $scr \in [1,20]$            |
| $k_{gb}$      | kernel size of Gaussian blur    | Constant | $k_{gb} \in \{3,5,7\}$      |
| $\sigma_{gb}$ | Sigma of Gaussian blur          | Constant | $\sigma_{gb} \in [0.2,0.6]$ |

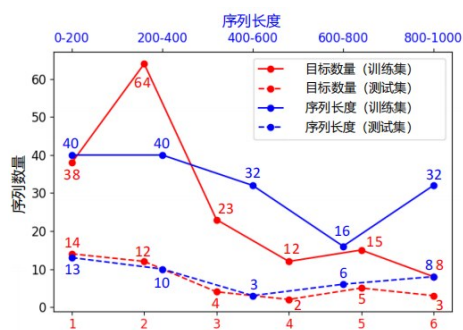
$$\hat{I}_{nt}^{sim} = GaussianBlur(I_{nt}^{sim}, k_{gb}, \sigma_{gb}) \quad (14)$$

$$I_{nt}^{sim} = Norm(I_{nt}^T) \odot I_{nt}^T + (1 - Norm(I_{nt}^T)) \odot I_{nt}^{LB} \quad (15)$$

像,合成了200个视频序列,共计91021帧图像。为了减少数据偏差和防止过拟合,IRSatVideo-LEO数据集被划分为160个视频的训练集和40个视频的测试集,并遵循以下划分原则:1)各子集均包含目标的全面属性,包括目标数量、信杂比、移动速度和机动次数;2)各子集亦涵盖了背景的完整属性,如背景复杂性和背景运动速度;3)训练集与测试集序列不存在交集。IRSatVideo-LEO数据集在训练集与测试集中均展现出良好的多样性与覆盖广度。如图12(a)所示,背景图像涵盖全球多



(a) 云量和位置分布



(b) 目标个数和序列长度分布

((a) cloud cover and position distribution; (b) the number of targets and the distribution of sequence length)

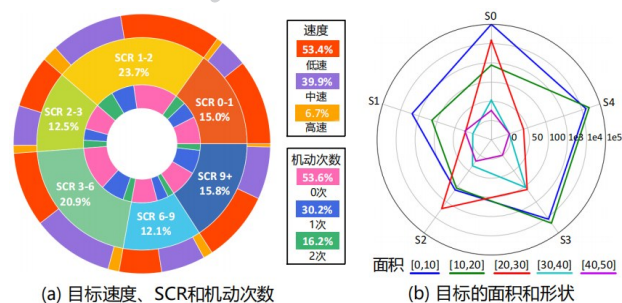
图12 IRSatVideo-LEO仿真数据集序列属性示意图

Fig. 12 Schematic diagram of sequence attributes of the IRSatVideo-LEO simulation dataset

个大陆,云量覆盖率跨度从0%至61.25%,体现了数据集在地理和气象条件上的丰富性。从目标分布来看(图12(b)),序列中目标数量范围为1到6个,其中包含2个目标的序列占比最高,达38%,且目标数不超过3的序列占总数的77.5%,显著多于目标数超过3的序列占总数的22.5%。在序列长度方面,帧数分布在104至993之间,其中短序列(0-200帧)占26.5%,中长序列(200-600帧)占

42.5%,长序列(600-1000帧)占31.0%,整体分布均衡,适用于不同长度的时序分析任务。

IRSatVideo-LEO数据集提供了实例级别的矩形框标签和掩码标签两种标注形式,分别适用于红外小目标的检测与跟踪任务,以及更精细的逐像素处理需求。该数据集具有广泛的覆盖范围和多样的背景特性,结合详尽的标注信息,为低地球轨道红外小目标检测算法的训练与评估提供了有效支持。以下将详细阐述IRSatVideo-LEO数据集中目标和背景属性的统计分析,并与现有数据集进行量化比较,以凸显其独特性和价值。



(a) 目标速度、SCR和机动次数

(b) 目标的面积和形状

1) 丰富的目标种类

((a) target speed, SCR and number of maneuvers; (b) the area and shape of the target)

图13 IRSatVideo-LEO仿真数据集目标属性示意图

Fig. 13 Schematic diagram of target attributes of the IRSatVideo-LEO simulation dataset

图13(a)展示了在IRSatVideo-LEO数据集中,目标的信杂比、速度和机动次数的分布情况。目标的移动速度被划分为四个级别:低速(每帧平均移位像素在0到3之间)、中速(3到6之间)、高

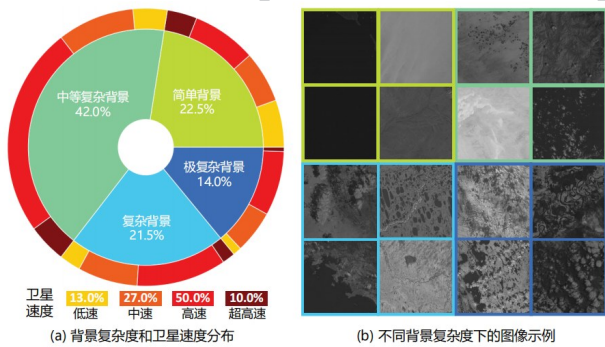
速(6以上)。分析结果表明,SCR小于3的目标占比达到34.8%,这表明数据集不仅面临小尺寸目标(目标面积小于 $9 \times 9$ 像素)的检测挑战,还面临弱信号目标的挑战。此外,超过半数的目标在运动过程中未表现出机动行为,且速度较低,而有一定比例的目标出现1-2次的转弯动作(46.4%)以及中到高速的运动(38.5%)。这种多样性的目标行为增加了数据集的复杂性和检测难度。总体而言,IRSatVideo-LEO数据集涵盖了在数量、SCR、速度和机动次数上具有多样性的目标,其属性分布合理,贴近实际应用场景。

图13(b)进一步展示了目标面积和形状的情况。鉴于数据集中的目标是通过高斯核模拟的,

因此采用椭圆的偏心率  $e$  来描述目标形状与圆形的偏差程度,其计算公式如下所示:

$$e = \sqrt{1 - \frac{b^2}{a^2}} \quad (16)$$

式中,  $a$  是椭圆的半长轴,而  $b$  是椭圆的半短轴。偏心率  $e$  的值介于 0 到 1 之间,当  $e=0$  时,目标呈圆形,随着  $e$  值的增大,椭圆的扁平程度增加。将目标的形状分为五个等级:  $S_0$ : ( $0 \leq e < 0.2$ )、 $S_1$ : ( $0.2 \leq e < 0.4$ )、 $S_2$ : ( $0.4 \leq e < 0.6$ )、 $S_3$ : ( $0.6 \leq e < 0.8$ )、 $S_4$ : ( $0.8 \leq e < 1$ )。从图中可以看出,面积小于 10



(a) background complexity and satellite velocity distribution;  
(b) image examples under different background complexity levels

图 14 IRSatVideo-LEO 仿真数据集背景属性示意图

Fig. 14 Schematic diagram of background attributes of the IRSatVideo-LEO simulation dataset

像素的目标占据了较大比例,并且大多数目标呈现出较为圆形的形状。

## 2) 复杂的背景杂波

对 IRSatVideo-LEO 数据集的背景复杂度特性进行详细分析,分析结果如图 14 所示,表明了

IRSatVideo-LEO 数据集覆盖了广泛的背景复杂度范围 ( $C \in [9, 5596]$ )。通过可视化不同复杂度级别的示例背景图像,可以发现简单背景序列波动较小,而随着复杂度级别的提升,背景波动显著增加。

此外,本节通过将背景移动速度分为四个级别:低速(每帧平均移位像素在 0 到 1/10 之间、中速(1/10 到 1/3 之间)、高速(1/3 到 1 之间)和超高速(1 到 2 之间)。从图中可以看出,每个背景复杂度级别都包含了不同背景移动速度的序列,且超过 60% 的序列背景移动速度处于高速或更高级别。综上所述,IRSatVideo-LEO 数据集包含了多样化复杂度和移动

速度的背景,为红外小目标检测提供了丰富的训练和测试场景。

## 2.2 IRSatVideo-GEO 数据集

IRSatVideo-GEO 仿真数据集是由 200 张高分四号卫星在中波红外波段 ( $3.5 - 4.1 \mu\text{m}$ ) 采集的背景图像合成的,包含了 200 个视频序列,共计 94136 帧图像。为了减少数据偏差和防止过拟合,该数据集被划分为 160 个视频的训练集和 40 个视频的测试集。划分过程遵循了以下原则:1) 确保每个子集均包含目标的完整属性,如目标数量、信杂比、移动速度;2) 确保背景的复杂度属性在各个子集中得到了均衡的覆盖;3) 确保训练集与测试集之

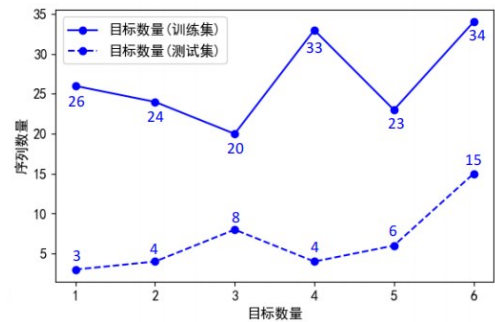
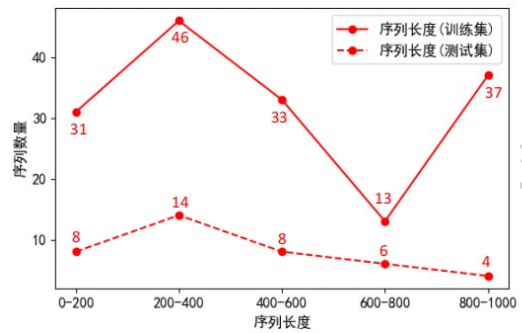


图 15 IRSatVideo-GEO 仿真数据集序列属性示意图

Fig. 15 Schematic diagram of sequence attributes of the IRSatVideo-GEO simulation dataset

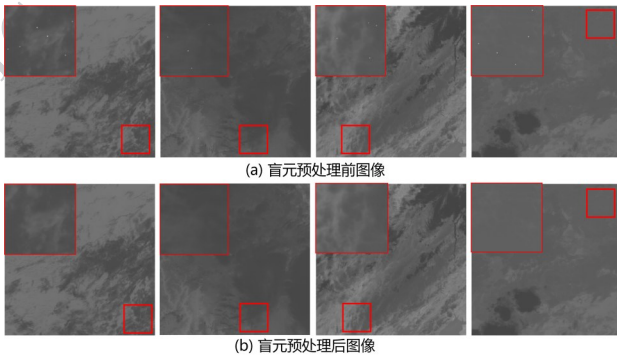
间的序列索引不存在交集,以保持数据的有效性和可靠性。

图 15 展示了 IRSatVideo-GEO 数据集训练集和测试集内序列长度和目标数量分布情况。图中数据显示,序列长度介于 105 到 998 帧之间,其中短序

列(0 到 200 帧)、中等长度序列(200 到 600 帧)和长序列(600 到 1000 帧)的比例分别为 19.5%、50.5% 和 30.0%。在目标数量方面,其范围在 1 到 6 之间,其中目标数量为 6 的序列占比最高,达到

24.5%。目标数量不超过3的序列数量是超过3的序列数量的1.4倍(57.5% vs. 42.5%)。

IRSatVideo-GEO数据集涵盖了多种不同的观测场景和背景条件,确保了在实际应用中的普遍适用性和鲁棒性;提供了实例级别的矩形框标签和像素级掩码标签,为红外小目标的检测、跟踪以及逐像素处理等任务的算法研究提供了强有力的数据支持。然而,值得注意的是,高分四号卫星采集的背景图像中存在大量的盲元(Dead Pixel, DP),这些盲元是由于制造缺陷或老化导致的红外焦平面阵列中的失效像素。如图16第一行所示,盲元表现为图像中的异常像素(常亮、常暗或随机噪声),这会严重影响图像质量及后续的目标检测、跟踪等处理。因此,对盲元进行预处理是必要的,这主要包括盲元检测、盲元补偿和图像优化三个步骤。



((a) image before dead pixel preprocessing; (b) image after dead pixel preprocessing)

图16 IRSatVideo-GEO仿真数据集序列属性示意图  
Fig. 16 Schematic diagram of sequence attributes of the IRSatVideo-GEO simulation dataset

### 1) 盲元检测

盲元检测通过动态阈值法实现了对盲元位置的精确定位,从而有效减少了误检和漏检的发生。

该过程主要细分为两个阶段:首先是像素响应的统计分析,其次是基于动态阈值的判定步骤。

#### (1) 像素响应统计

对每个像素 $(i, j)$ 计算其 $5 \times 5$ 局部邻域内的均值 $\mu_{ij}$ 和标准差 $\sigma_{ij}$ :

$$\mu_{ij} = \frac{1}{N} \sum_{k=1}^N I_k(i, j) \quad (17)$$

$$\sigma_{ij} = \sqrt{\frac{1}{N-1} \sum_{k=1}^N (I_k(i, j) - \mu_{ij})^2} \quad (18)$$

正常像素的 $\sigma_{ij}$ 反映背景杂波水平,而盲元的 $\sigma_{ij}$

趋近于零(常亮/常暗)或异常高(随机噪声)。

### (2) 动态阈值判定

首先计算全图标准差均值 $\bar{\sigma}$ :

$$\bar{\sigma} = \frac{1}{M \times M} \sum_{ij} \sigma_{ij}, \quad (19)$$

随后设定阈值 $T_{low} = 0.1\bar{\sigma}$ 检测常亮/常暗盲元,设定阈值 $T_{high} = 3\bar{\sigma}$ 检测随机噪声盲元。最后,标记检测出的盲元 $DP(i, j)$ :

$$DP(i, j) = \begin{cases} 1, & \text{if } \sigma_{ij} < T_{low} \text{ 或 } \sigma_{ij} > T_{high} \\ 0, & \text{otherwise} \end{cases} \quad (20)$$

### 2) 盲元补偿

盲元补偿策略采用了自适应双线性插值技术,利用邻近的有效像素信息来重建盲元像素的值,同时确保图像细节的保留。该过程主要细分为三个阶段:首先是邻域有效性检测,其次是梯度权重计算,最后实施双线性插值以完成盲元补偿。

#### (1) 邻域有效性检测

以盲元 $(i, j)$ 为中心,定义 $3 \times 3$ 邻域窗口。检查邻域内像素是否为盲元:若邻域中有效像素数 $\geq 4$ 则进行插值;否则扩大窗口至 $5 \times 5$ 。

#### (2) 梯度权重计算

首先计算四个方向梯度(水平、垂直、对角线):

$$\begin{aligned} G_x &= |I(i+1, j) - I(i-1, j)|, \\ G_y &= |I(i, j+1) - I(i, j-1)|, \\ G_{d1} &= |I(i+1, j+1) - I(i-1, j-1)|, \\ G_{d2} &= |I(i+1, j-1) - I(i-1, j+1)| \end{aligned} \quad (21)$$

随后进行权重分配,梯度越小,则方向越平滑,权重越高,其公式表达如下:

$$\begin{aligned} w_x &= 1/(G_x + \varepsilon), \\ w_y &= 1/(G_y + \varepsilon), \\ w_{d1} &= 1/(G_{d1} + \varepsilon), \\ w_{d2} &= 1/(G_{d2} + \varepsilon), \end{aligned} \quad (22)$$

式中, $\varepsilon = 1e-6$ 防止除零。

#### (3) 双线性插值补偿

对有效像素值进行加权平均:

$$\hat{I}(i, j) = \frac{\sum_{m, n \in \Omega} w_{m, n} I(m, n)}{\sum_{m, n \in \Omega} w_{m, n}} \quad (23)$$

式中, $\Omega$ 为邻域内有效像素集合, $w_{m, n}$ 为对应方向权重。

### 3) 图像优化

在盲元预处理的基础上,为了进一步提升图像

表7 红外小目标检测数据集

Table 7 Infrared small target detection dataset

|                | 数据集名称          | 类型 | 场景 | 标签 | 波段                   | 分辨率                  | 序列数   | 帧数     | T-Num  | T-Size | SCR   |
|----------------|----------------|----|----|----|----------------------|----------------------|-------|--------|--------|--------|-------|
| 单帧数据集          | MFIRST         | 仿真 | 地基 | 掩膜 | 长波                   | 129×129 <sup>1</sup> | -     | 10100  | 10337  | 40     | 5.40  |
|                | SIRST-v2       | 实测 | 空基 | 掩膜 | 长波                   | 278×366 <sup>2</sup> | -     | 514    | 648    | 37     | 12.07 |
|                | NUDT-SIRST     | 仿真 | 地基 | 掩膜 | 长波                   | 256×256              | -     | 1327   | 1863   | 34     | 5.60  |
|                | IRSTD-1K       | 实测 | 地基 | 掩膜 | 长波                   | 512×512              | -     | 1001   | 1492   | 53     | 8.92  |
|                | NUDT-SIRST-Sea | 实测 | 天基 | 掩膜 | 近红外                  | 998×998 <sup>3</sup> | -     | 5808   | 16929  | 36     | 15.32 |
| 多帧数据集          | Hui            | 实测 | 地基 | 点  | 长波                   | 256×256              | 22    | 16177  | 16944  | -      | 5.33* |
|                | Fu             | 实测 | 空基 | 框  | 长波                   | 640×480              | 87    | 21750  | 89174  | 185    | 4.66* |
|                | Anti-UAV v2    | 实测 | 地基 | 框  | 长波                   | 640×512              | 140   | 152561 | 152561 | 3219   | 1.32* |
|                | IRDST-Real     | 实测 | 地基 | 掩膜 | 长波                   | 992×742              | 85    | 40656  | 41801  | 10     | 6.89  |
|                | IRDST-Simulate | 仿真 | 地基 | 掩膜 | 长波                   | 720×480              | 316   | 106254 | 102077 | 6      | 4.68  |
|                | SIATD          | 仿真 | 空基 | 点  | 长波                   | 640×512              | 350   | 150185 | 247080 | <7×7   | 6.95* |
|                | NUDT-MIRSDT    | 仿真 | 空基 | 掩膜 | 长波                   | 217×302 <sup>4</sup> | 120   | 12000  | 11464  | 34     | 1.87  |
|                | SIRSTD         | 实测 | 地基 | 框  | 长波                   | 640×512              | 48    | 50388  | 48565  | 60     | 10.40 |
|                | IRSatVideo-LEO | 仿真 | 天基 | 掩膜 | 短波                   | 1024×1024            | 200   | 91021  | 220126 | 4      | 6.69  |
| IRSatVideo-GEO | 仿真             | 天基 | 掩膜 | 中波 | 742×733 <sup>5</sup> | 200                  | 94136 | 344285 | 7      | 1.10   |       |

注:表中T-Num、T-Size分别代表目标数量和尺寸。<sup>1234</sup>由于数据集内图像分辨率非统一,因此表格中展示的为数据集中所有图像的平均分辨率。其中,NUST-SIRST数据集 $h \in [96, 327]$ ,  $w \in [101, 442]$ , SIRST-v2数据集 $h \in [96, 1024]$ ,  $w \in [135, 1280]$ , NUDT-SIRST-Sea数据集 $h \in [740, 1024]$ ,  $w \in [740, 1024]$ , NUDT-MIRSDT数据集 $h \in [154, 324]$ ,  $w \in [209, 407]$ , IRSatVideo-GEO数据集 $h \in [154, 324]$ ,  $w \in [209, 407]$ 。

质量,以下优化措施被采纳:

#### (1) 边缘保持策略

针对位于图像边缘的盲元(例如,图像的第一行或第一列),采用镜像边界扩展技术,以防止在插值过程中引入边缘伪影。

#### (2) 密集盲元区域处理

针对连续出现的盲元区域(例如,3×3的完全盲元区块),实施非局部均值滤波算法,通过参考图像中全局范围内的相似块来进行修复,以恢复图像的完整性。

图16详细展示了盲元预处理前后的对比效果。通过对比分析,可以看出基于动态阈值的盲元检测能够精确地识别出图像中的异常盲元点,基于自适应双线性插值的盲元补偿可以有效地对这些盲元点进行值重建。这一系列的盲元检测与补偿流程共同作用,成功地移除了图像中的盲元缺陷,从而在很大程度上改善了图像的整体质量。此外,该处为后续的检测跟踪等应用处理打下了良好的基础。

在对比IRSatVideo-GEO仿真数据集所涵盖的机动目标与平稳运动的背景时,地球同步轨道卫星由于其位于大约36,000公里远的地球表面之上,导致大多数可观测目标展现出弱机动特性。这类目标包括人造卫星、空间碎片、大型航天器、空间站、运载火箭残骸以及深空探测器等。鉴于此,在本研

究环节中,IRSatVideo-GEO仿真数据集将不考虑目标的机动性,即单个目标的机动次数设置为0。

此外,由于地球同步轨道卫星相对于地球表面保持相对静止,根据第二章分析,本节将不进行背景运动的统计分析。IRSatVideo-GEO数据集的核心研究是极弱目标的检测问题,专注于信杂比低于3的目标检测问题,旨在推动在极低信杂比条件下红外小目标检测技术的研究与发展。

以下内容将详尽阐述IRSatVideo-GEO数据集中目标与背景特性的统计分析结果,并与现有数据集进行量化对比,旨在凸显本数据集的独特性及其在学术研究中的价值。

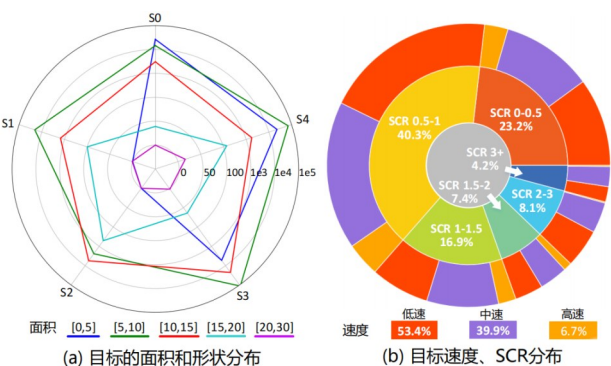


图 17 IRSatVideo-GEO 仿真数据集目标属性示意图  
Fig. 17 Schematic diagram of target attributes of the IRSatVideo-GEO simulation dataset

图 17 IRSatVideo-GEO 仿真数据集目标属性示意图  
Fig. 17 Schematic diagram of target attributes of the IRSatVideo-GEO simulation dataset

### 1) 小尺寸的极弱目标

图 17 (a)展示了 IRSatVideo-GEO 数据集中目标面积与形状分布的特征。目标形状通过椭圆的偏心率来量化,并分为五个类别:  $S0: 0 \leq e < 0.2$ ,  $S1: 0.2 \leq e < 0.4$ ,  $S2: 0.4 \leq e < 0.6$ ,  $S3: 0.6 \leq e < 0.8$ ,  $S4: 0.8 \leq e < 1$ 。从图中可以看出,面积小于 15 像素的目标占据了极大比例,并且大多数目标呈现出较为圆形的形状。

图 17 (b)详细展示了目标的信杂比和移动速度的分布。目标移动速度分为三个等级:低速(每帧平均移位像素在 0 到 1 之间)、中速(1 到 3 之间)和高速(3 像素以上)。分析结果表明,SCR 小于 3 的目标占比达到 95.8%,SCR 小于 1 的目标占一半以上(63.5%)。所有目标的平均 SCR 为 1.1,说明数据集不仅面临极小尺寸目标(绝大多数目标的面积小于 15 像素)的挑战,还需应对极弱信号目标的挑战。综上所述,IRSatVideo-GEO

数据集包含了在数量、SCR 和速度方面具有广泛多样性的目标,其属性分布均匀,与实际应用场景相契合。

### 2) 复杂的背景杂波

对 IRSatVideo-GEO 数据集的背景复杂度特征进行深入分析,分析结果如图 18 所示,可以得出 IRSatVideo-LEO 数据集涵盖了广泛的背景复杂度区间。鉴于高分四号卫星的轨道高度是 Landsat 卫星的 4 倍,高分四号卫星图像的复杂度整体低于 Landsat,集中于简单背景( $C \in [10, 5262]$ )和中等复杂背景( $C \in [200, 1000]$ )。

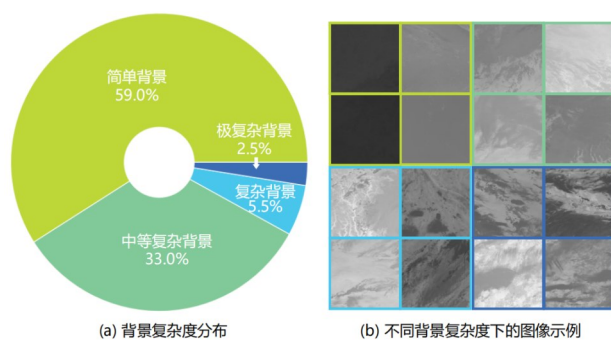


图 18 IRSatVideo-GEO 仿真数据集背景属性示意图  
Fig. 18 Schematic diagram of background attributes of the IRSatVideo-GEO simulation dataset

图 18 IRSatVideo-GEO 仿真数据集背景属性示意图  
Fig. 18 Schematic diagram of background attributes of the IRSatVideo-GEO simulation dataset

## 2.3 与现有数据集对比

本节将 IRSatVideo-LEO 数据集和 IRSatVideo-GEO 数据集与现有的单帧和多帧数据集进行了全面比较。对比结果如表 7 所示,与单帧数据集相比,多帧数据集的尺度更大(例如,帧数和目标号),并专注于具有较低 SCR 的目标,表明了时间信息利用是多帧检测的关键步骤。表中许多多帧数据集仅提供矩形框框或点标签,无法与现有的红外小目标检测方法(即前景和背景细分任务)之间进行全面的性能评估。尽管 IRDST 和 NUDT-MIRSdT 数据集提供了掩膜标签,但陆基成像系统(较小的视场和近景平台导致的视差变化)和云台(较小的视场和近景平台导致的视差变化)和云台摄像机(高自由度平台导致的不规则背景运动)或静态背景动作无法满足卫星视频的实际条件(遥感成像系统视场较大且不存在视差问题,卫星平台运动缓慢且规则)。

此外,本节还将 IRSatVideo-LEO 数据集和

IRSatVideo-GEO 数据集,与 NUDT-MIRSdT 数据集进行详细属性对比,并将结果在表 8 中进行了汇总。从表格表 7 和表 8 中的对比结果可以得出如下结论:

### 1) 数据规模与多样性显著提升

IRSatVideo-LEO/GEO 数据集在序列数量、图像数量和数量上实现了数量级的扩展,为深度学习模型提供了更丰富的训练样本。这一大规模数据集尤其适用于天基场景下的长时序目标跟踪任务,能够有效支持复杂场景下的模型泛化能力。

### 2) 成像条件更加复杂

分辨率提升:图像分辨率大幅提升,增强了对弱

表8 NUDT-MIRS DT 和 IRSatVideo-LEO 数据集之间的关键差异

Table 8 Key differences between the NUDT-MIRS DT and IRSatVideo-LEO datasets

| 属性对比       | NUDT-MIRS DT   | IRSatVideo-LEO   | IRSatVideo-GEO   |
|------------|----------------|------------------|------------------|
| 序列/图像/目标数量 | 48/12000/11464 | 200/91021/220126 | 200/91021/220126 |
| 图像分辨率      | 217×302        | 1024×1024        | 1024×1024        |
| 背景类型       | 地基/空基          | 天基               | 天基               |
| 平台运动       | 静止平台           | 运动平台             | 微动平台             |
| 波段         | 长波             | 短波               | 中波               |
| 目标运动       | 非机动目标          | 机动目标             | 非机动目标            |
| 目标形态       | 非规则            | 高斯核              | 高斯核              |
| 时序上目标形态变化  | 快速不连续变化        | 缓慢连续变化           | 缓慢连续变化           |

小目标的表征能力,但同步增加了计算复杂度;

动态背景干扰:天基背景与运动/微动平台的引入,要求算法具备更强的背景抑制与运动补偿能力,以应对复杂的动态干扰;

波段差异:短波/中波红外波段与长波的辐射特性差异显著,需针对不同波段下的目标辐射特性进行精确建模。

### 3) 目标特性挑战升级

IRSatVideo-LEO 场景中的机动目标:目标的加速度变化对跟踪算法的预测鲁棒性提出了更高要求,需结合轨道动力学约束提升轨迹预测精度;

IRSatVideo-GEO 场景中的极弱目标:极低信杂比目标(SCR<3dB)对检测算法的灵敏度与多帧时空关联能力提出了严苛要求;

目标形态与变化特性:高斯核形态与缓慢连续变化特性更贴近天基观测中远距离目标的成像规律,但削弱了形状特征的可用性,需依赖运动特征与辐射强度分析实现目标表征。

总体而言,IRSatVideo-LEO 和 IRSatVideo-GEO 数据集相较于现有的单帧与多帧数据集,展现出显著的优越性。这些数据集不仅拥有更大的样本规模和更高的图像分辨率,而且还模拟了地球轨道卫星移动平台的动态特性,专注于短波与中波红外波段的成像,并包含了机动目标和极低信杂比目标。这些特点凸显了监测范式从传统的“静态背景强特征目标检测”向“动态背景弱特征机动目标持续跟踪”的转变。此外,IRSatVideo 数据集为天基红外监视任务提供了一个接近实际操作的基准平台,在卫星视频监控和遥感应用的研究与开发领域显得尤为宝

贵,构成了一个具有研究价值的资源库。

## 3 数据集验证与评估

为了对 IRSatVideo 数据集的有效性与挑战性进行系统验证,在 IRSatVideo-LEO 和 IRSatVideo-GEO 数据集下分别开展实验,选取多种典型的单帧与多帧红外小目标检测方法作为基线,对其在不同信杂比条件、不同背景复杂度下的检测性能进行系统评估。实验结果不仅展现了 IRSatVideo 的多样性和物理真实性,也突出了在复杂卫星观测环境下弱小目标检测的困难性与研究价值。

### 3.1 实验实施细节

本节对当前最先进的单帧和多帧红外小目标检测方法进行性能评测,以深入理解不同方法在红外小目标检测任务中的表现和差异。其中,单帧红外小目标检测方法采用十种最新的数据驱动方法:Dai 等人(2021)提出的 ACM, ALCNet, Li 等人(2022)提出的 DNANet, Zhang 等人(2022)提出的 ISNet, Wu 等人(2022)提出的 UIU-Net, Sun 等人(2023)提出的 RDIAN, Hou 等人(2022)提出的 ISTDU-Net, Xiao 等人(2018)提出的 ResUNet, Wu 等人(2024)提出的 RPCANet, Zhang 等人(2023)提出的 AGPCNet 进行性能对比。多帧红外小目标检测方法采用四种先进的数据驱动方法:Yan 等人(2023)提出的 STDMA-Net, Li 等人(2025)提出的 ResUNet\_DTUM 和 ISTDU-Net\_DTUM。

为确保比较的公正性,本节采用了相同的损失函数和训练实施细节,在 IRSatVideo-GEO 数据集上

对所有数据驱动的多帧检测方法进行了从头开始的训练。单帧检测方法训练参数是按照 BasicIRSTD 工具箱的标准默认参数进行设置的,

表9 不同检测方法在 IRSatVideo-GEO 数据集上的性能对比

Table 9 Performance comparison of different detection methods on the IRSatVideo-GEO dataset

|         | 方法名称         | 参数量 ↓  | FLOPs ↓  | 运行时间 ↓ | IoU ↑ | Pd ↑  | Fa ↓    |
|---------|--------------|--------|----------|--------|-------|-------|---------|
| 单帧小目标检测 | ACM          | 0.40M  | 7.40G    | 1.41   | 3.00  | 8.97  | 2.02E-5 |
|         | ALCNet       | 0.43M  | 7.50G    | 0.26   | 4.15  | 9.60  | 3.83E-5 |
|         | DNANet       | 4.69M  | 283.59G  | 0.99   | 3.89  | 8.53  | 3.46E-5 |
|         | ISNet        | 0.97M  | 610.08G  | 0.49   | 2.08  | 8.59  | 1.34E-4 |
|         | UIU-Net      | 50.54M | 1087.52G | 1.28   | 0.90  | 6.55  | 2.47E-4 |
|         | RDIAN        | 0.22M  | 74.15G   | 0.16   | 3.25  | 10.85 | 7.82E-5 |
|         | ISTDU-Net    | 2.75M  | 156.75G  | 0.60   | 2.00  | 7.12  | 7.88E-5 |
|         | ResUNet      | 0.91M  | 76.33G   | 0.16   | 3.24  | 9.87  | 7.38E-5 |
|         | RPCANet      | 0.68M  | 891.15G  | 1.37   | 1.84  | 13.68 | 2.92E-4 |
|         | AGPCNet      | 12.43M | 863.32G  | 2.59   | 4.08  | 10.81 | 5.72E-5 |
| 序列小目标检测 | STDMA Net    | 11.88M | 62.98G   | 1.05   | 20.38 | 41.05 | 8.56E-8 |
|         | DNANet_DTUM  | 1.21M  | 176.66G  | 3.68   | 10.63 | 24.68 | 3.25E-5 |
|         | ResUNet_DTUM | 0.30M  | 40.88G   | 2.18   | 9.37  | 26.74 | 4.58E-5 |
|         | UIU-Net_DTUM | 51.04M | 1151.19G | 6.46   | 4.08  | 6.16  | 3.43E-7 |

表10 不同算法在 IRSatVideo-GEO 数据集不同信杂比场景上对比的定量分析结果

Table 10 Quantitative analysis results of comparisons between different algorithms under different signal-to-clutter ratio scenarios on the IRSatVideo-GEO dataset

|           | 方法名称         | 微弱信杂比 |       |         | 极低信杂比 |       |         | 低信杂比  |       |         |
|-----------|--------------|-------|-------|---------|-------|-------|---------|-------|-------|---------|
|           |              | IoU ↑ | Pd ↑  | Fa ↓    | IoU ↑ | Pd ↑  | Fa ↓    | IoU ↑ | Pd ↑  | Fa ↓    |
| 单帧红外小目标检测 | ACM          | 0.25  | 1.11  | 2.96E-5 | 5.60  | 14.25 | 9.03E-6 | 25.78 | 63.39 | 3.46E-6 |
|           | ALCNet       | 0.29  | 1.02  | 5.14E-5 | 8.47  | 15.54 | 2.11E-5 | 28.32 | 67.56 | 2.43E-5 |
|           | DNANet       | 0.32  | 0.89  | 2.56E-5 | 5.09  | 12.61 | 4.26E-5 | 19.62 | 69.19 | 6.53E-5 |
|           | ISNet        | 0.06  | 0.35  | 9.6E-5  | 2.36  | 13.74 | 1.95E-4 | 16.02 | 68.47 | 1.05E-4 |
|           | UIU-Net      | 0.36  | 0.72  | 3.34E-5 | 0.53  | 8.85  | 5.88E-4 | 10.28 | 58.77 | 1.42E-4 |
|           | RDIAN        | 0.40  | 1.77  | 1.15E-4 | 7.46  | 16.76 | 3.03E-5 | 29.78 | 75.17 | 3.71E-5 |
|           | ISTDU-Net    | 0.05  | 0.30  | 4.75E-5 | 2.53  | 12.24 | 1.23E-4 | 10.06 | 50.47 | 9.46E-5 |
|           | ResUNet      | 0.39  | 1.10  | 4.84E-5 | 3.88  | 15.85 | 1.11E-4 | 18.67 | 69.97 | 7.59E-5 |
|           | RPCANet      | 0.38  | 3.09  | 2.66E-4 | 2.64  | 21.83 | 3.21E-4 | 7.48  | 79.56 | 3.41E-4 |
|           | AGPCNet      | 0.55  | 1.39  | 4.27E-5 | 5.18  | 17.22 | 7.73E-5 | 22.31 | 75.61 | 6.69E-5 |
| 序列红外小目标检测 | STDMA Net    | 14.25 | 30.23 | 6.79E-8 | 26.31 | 52.32 | 1.27E-7 | 46.32 | 87.00 | 3.08E-9 |
|           | DNANet_DTUM  | 0.10  | 0.42  | 2.00E-8 | 6.37  | 13.38 | 1.27E-8 | 39.20 | 64.47 | 1.23E-8 |
|           | ResUNet_DTUM | 0.91  | 1.78  | 9.45E-8 | 12.21 | 20.30 | 1.32E-8 | 49.17 | 74.92 | 1.54E-9 |
|           | UIU-Net_DTUM | 0.67  | 1.08  | 5.74E-7 | 5.44  | 7.62  | 5.75E-8 | 31.39 | 55.44 | 1.54E-9 |

数据驱动的多帧检测方法的训练细节如下:

在 IRSatVideo-GEO 数据集实验中,本文重点关注 GEO 场景下极低信杂比目标的检测任务。在训练过程中,从训练集中随机选取  $K$  个连续帧,并对这些帧进行随机裁剪,形成  $K \times 128 \times 128$  的小块作为输入。

借鉴 Liu 等人(2016)和 Tong 等人(2024)的设置,本章通过随机翻转和旋转操作对训练数据进行增强。为了加速训练,本章维持正负样本的比例为 0.8:0.2。本研究采用时域滑动窗口方法进行数据遍历,输入数据为  $K$  帧,输出数据为滑动窗口末尾帧。网络训练采用焦点损失函数(Focal Loss),并通过 Adam 优化器进行优化,其参数设置为  $\lambda_1 = 0.9, \lambda_2 = 0.999$ 。训练批次大小设为 3,初始学习率为  $5e-4$ ,共训练 20 个轮次,并在每 5 个训练周期后减半。所有实验均在配备 Nvidia GeForce RTX 3090 GPU 的 PC 机上,使用 PyTorch 框架完成。由于数据量较大(共计 344285 帧图像),因此训练样本每轮进行随机采样,采样率为 5。

在 IRSatVideo-LEO 数据集实验中,考虑到 LEO 场景下目标信杂比更高,且背景动态特性更复杂,训练采用更长的输入序列( $K = \{10, 20, 40\}$ )以捕捉长时间依赖关系。其他训练参数与 GEO 实验保持一致,但正负样本比例调整为 0.9:0.1 以加速训练收敛。测试阶段,将整段序列输入检测网络以适应递归算法从而获得更全面的性能表现。

在评估指标方面,本文使用检测率(Pd)、虚警率(Fa)、交并比(IoU)值作为主要性能度量,以便更准确地分析模型在不同数据条件下的表现。此外,本章进一步将 IRSatVideo-GEO 测试集细分为以下三种场景:1)微弱信杂比场景:目标 SCR 小于 1,共 20 段序列;2)极低信杂比场景:目标 SCR 小于 2,共 16 段序列;3)低信杂比场景:目标 SCR 小于介于 2 到 4 之间,共 4 段序列。

通过对数据集进行细分更准确地评估算法在不同信杂比场景下的综合性能。IRSatVideo-LEO 测试集细分为以下三种场景:1)简单场景:目标 SCR

大于 6,背景复杂度小于 1000,共 18 个场景;2)中等场景:目标 SCR 大于 6,背景复杂度小于 1000,共 13 个场景;3)复杂场景:目标 SCR 小于 6,背景复杂度大于 1000,共 9 个场景。通过对数据集进行细分,能够更准确地评估算法在不同难度级别场景下

的性能。

### 3.2 IRSatVideo-GEO 实验结果与分析

表 9 展示了不同算法在 IRSatVideo-GEO 数据集上的定量分析结果。通过对比分析可以发现,单帧方法平均检测率仅为  $9.02 \pm 3.85\%$ ,虚警率为  $1.12 \times 10^{-4} \pm 8.6 \times 10^{-5}$ ,表明依赖于目标表征信息的单帧检测算法在低信杂比目标检测中可靠性较低。极端案例如 UIU-Net 虽参数量高达 50.54M,但其 IoU 仅为 0.90%,揭示单纯增加模型容量无法解决低信杂比目标特征模糊问题。序列检测方法中,STDMA Net 实现显著性能提升,以最低虚警率取得

最高的检测率,验证时间维度信息融合的有效性。

表 10 进一步细化了不同信杂比场景下的性能表现。微弱信杂比(SCR<1)场景下,单帧方法检测率普遍在 1% 左右,说明单帧模型在微弱信杂比下难以有效提取目标特征。STDMA Net 通过时域累积将 Pd 提升至 41.05%,其设计的深度时空频域显著性滤波算法可对 5 帧弱信号进行非相干积累,显著提升目标信杂比。

可视化分析结果如图 19 所示。STDMA Net 算法在所有比较方法中表现最为出色,可以实现精确的目标定位和形状分割,同时最小化错误警报。单帧方法在处理低信杂比场景中性能受限,往往产生较多漏检。STDMA Net 采用时空频联合的目标检测设计,对低信杂比目标具有更强鲁棒性。

表 9 的计算效率评估表明,与模型驱动方法相比,数据驱动方法在检测性能和运行时间上都实现显著提升。多帧红外小目标检测方法不仅展现卓越检测性能,同时保持合理计算成本和运行时间。

整体来看,现有单帧方法在 IRSatVideo-GEO 上普遍表现不佳,难以应对目标极小、背景复杂、运动缓慢的观测条件,而多帧方法凭借对时序一致性的建模能力展现出更强的鲁棒性。这种明显的性能分层不仅验证了时间信息在红外小目标检测中的关键作用,也充分体现了 IRSatVideo-GEO 数据集的判别能力。它能够清晰区分不同检

测范式之间的性能差距,并对各类模型的优劣提供真实、可信的评估反馈,具备作为高挑战性基准数据集的价值。

### 3.3 IRSatVideo-LEO 实验结果与分析

表 11 展示了 LEO 数据集上的定量分析结果。

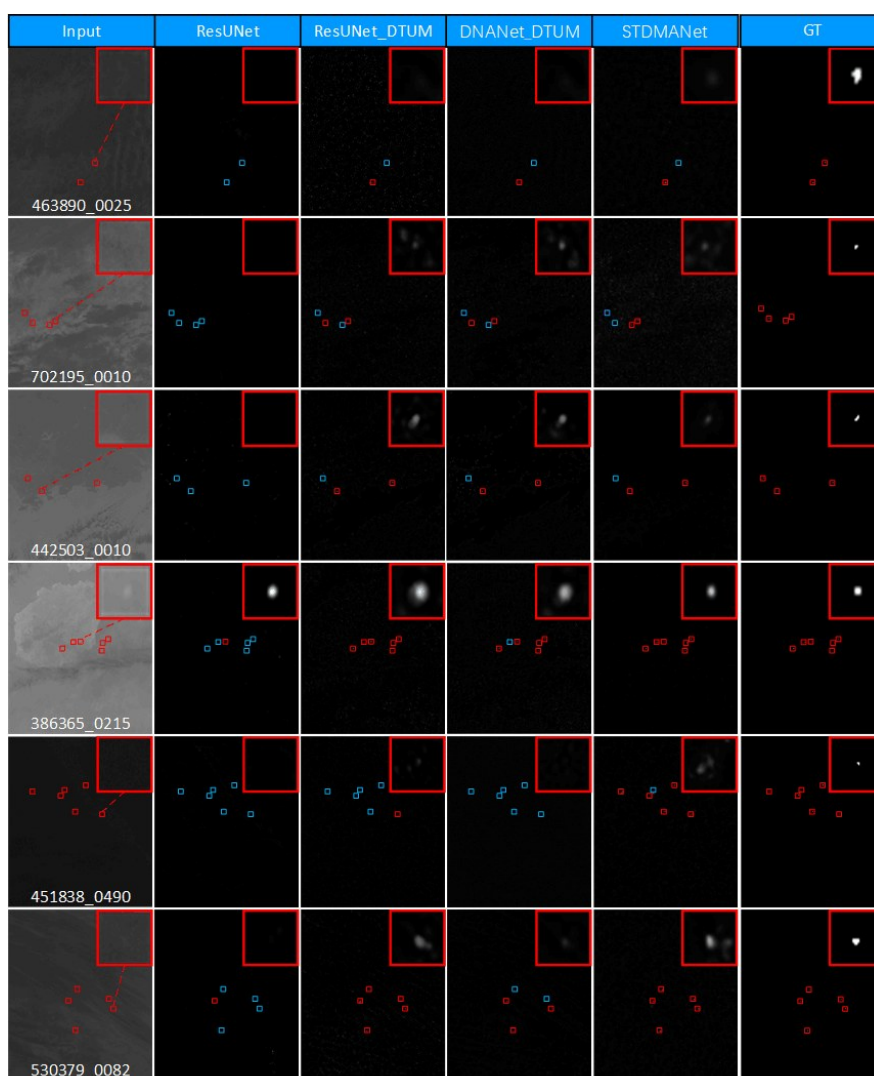


图 19 不同算法在 IRSatVideo-GEO 数据集上对比的可视化结果

Fig. 19 Visualization results of comparison between different algorithms on the IRSatVideo-GEO dataset

与 GEO 数据集相比,LEO 场景中单帧检测方法展现出截然不同的特性。

最优单帧模型 ISTDU-Net 的 IoU (31.82%) 和 Pd (52.88%) 分别达到 GEO 数据集最优值的 12.5 倍和 4.9 倍,揭示轨道高度降低带来的信杂比改善。序列检测方法中,ResUNet\_RFR 通过轻量级残差递归架构,在保持轻量化设计 (1.01M 参数) 的同时实现性能提升, IoU (33.70%) 较单帧最优提升 6.1%, 运行时间 (0.38s/20 帧) 满足近实时处理需求。与传统方法相比,SRSTT 等非深度学习方法虽虚警率较低,但其检测率较 ResUNet\_RFR 低 29.1%, 且运行时间超出 4 个数量级,凸显深度学习方法工程适用性。

表 12 的场景细分分析表明,数据驱动方法较传统模型在中等场景实现 Pd 值 20.7 个百分点的绝对

提升 (60.55% vs. 37.85%), 且虚警率降低 2 个数量级。ResUNet\_RFR 在 LEO 数据集上达成跨场景最优平衡,在简单场景下 IoU (48.12%) 和 Pd (96.13%), 复杂场景 IoU (21.79%) 和 Pd (63.37%)。复杂场景下,单帧方法 Pd 衰减率达  $32.6 \pm 8.4\%$ , 而序列方法通过时域非相干积累将衰减率压制至  $18.2 \pm 5.7\%$ , 证明时空维度信息融合是突破红外序列图像小目标检测理论极限的关键路径。

可视化结果如图 20 所示。ResUNet\_RFR 算法表现最为出色,可实现精确目标定位和形状分割。传统方法在处理复杂背景中性能受限,产生较多虚警。相较于 ResUNet\_DTUM, ResUNet\_RFR 采用联合特征对齐和目标检测设计,对卫星运动具有更强鲁棒性。递归特征精炼算法能够与不同单帧检测技

表 11 不同检测方法在 IRSatVideo-LEO 数据集上的性能对比

Table 11 Performance comparison of different detection methods on the IRSatVideo-LEO dataset

|         | 方法名称          | 参数量 ↓  | FLOPs ↓  | 运行时间 ↓  | IoU ↑ | Pd ↑  | Fa ↓    |
|---------|---------------|--------|----------|---------|-------|-------|---------|
| 单帧小目标检测 | ACM           | 0.40M  | 7.40G    | 0.10    | 21.49 | 61.88 | 1.90E-5 |
|         | ALCNet        | 0.43M  | 7.50G    | 0.09    | 26.52 | 59.35 | 3.69E-6 |
|         | DNANet        | 4.69M  | 283.59G  | 6.54    | 28.95 | 44.33 | 2.36E-6 |
|         | ISNet         | 0.97M  | 610.08G  | 1.81    | 11.32 | 73.18 | 6.87E-5 |
|         | UIU-Net       | 50.54M | 1087.52G | 1.98    | 30.04 | 63.99 | 9.13E-6 |
|         | RDIAN         | 0.22M  | 74.15G   | 1.23    | 15.01 | 67.06 | 4.29E-5 |
|         | ISTDU-Net     | 2.75M  | 156.75G  | 1.41    | 31.82 | 52.88 | 2.67E-5 |
|         | ResUNet       | 0.91M  | 76.33G   | 0.37    | 23.44 | 70.78 | 2.08E-5 |
|         | RPCANet       | 0.68M  | 891.15G  | 3.46    | 17.60 | 71.86 | 3.42E-5 |
|         | AGPCNet       | 12.43M | 863.32G  | 1.73    | 29.19 | 72.74 | 1.19E-5 |
| 序列小目标检测 | MSLSTIPT      | -      | -        | 593.30  | 0.64  | 0.80  | 6.21E-8 |
|         | NFTDGSTV      | -      | -        | 1019.17 | 0.42  | 20.43 | 1.16E-4 |
|         | RCTVW         | -      | -        | 6.73    | 0.04  | 23.15 | 1.73E-3 |
|         | IMNN-LWEC     | -      | -        | 3203.45 | 0.65  | 51.74 | 5.00E-4 |
|         | SRSTT         | -      | -        | 4693.61 | 0.00  | 52.03 | 9.27E-2 |
|         | ASTTV-NTLA    | -      | -        | 1148.70 | 2.72  | 71.96 | 6.46E-5 |
|         | 4DISTD        | -      | -        | 3536.69 | 8.75  | 66.49 | 1.13E-5 |
|         | STDMA-Net     | 11.88M | 62.98G   | 19.72   | 31.44 | 68.83 | 7.99E-6 |
|         | DNANet_DTUM   | 1.21M  | 176.66G  | 23.45   | 17.60 | 72.54 | 2.74E-5 |
|         | ResUNet_DTUM  | 0.30M  | 40.88G   | 3.09    | 21.04 | 67.03 | 1.58E-5 |
|         | UIU-Net_DTUM  | 51.04M | 1151.19G | 11.58   | 22.82 | 70.71 | 1.56E-5 |
|         | ACM_RFR       | 0.50M  | 79.88G   | 0.20    | 26.15 | 63.90 | 7.74E-6 |
|         | ALCNet_RFR    | 0.53M  | 79.39G   | 0.20    | 21.84 | 64.29 | 1.65E-5 |
|         | DNANet_RFR    | 4.80M  | 358.16G  | 6.52    | 29.40 | 71.73 | 1.01E-5 |
|         | ISTUD-Net_RFR | 2.86M  | 231.41G  | 1.54    | 26.62 | 55.74 | 4.12E-6 |
|         | ResUNet_RFR   | 1.01M  | 150.90G  | 0.38    | 33.70 | 73.39 | 6.31E-6 |

术相结合,进一步提升检测性能。

计算效率评估表明,数据驱动方法在保持高性能的同时具有显著效率优势。ResUNet\_RFR 在复杂场景下的 Pd 衰减率较单帧最优模型降低 48.3%,其特征精炼模块使时空特征融合效率达到 STDMA-Net 的 3.1 倍,满足天基监视实时处理需求。

整体来看,IRSatVideo-LEO 数据集能够清晰区分不同技术路线的性能边界,既体现了低轨观测的优势,又保留了动态背景、弱小目标、复杂干扰等核心挑战。现有先进方法在其上的表现差异显著,说

明该数据集具备良好的判别性与评估价值,可作为检验红外小目标检测算法真实鲁棒性的可靠基准。

通过上述实验验证,IRSatVideo 数据集为天基红外弱小目标检测算法研究提供了全面、可靠的评估平台,有效推动了该领域的技术发展。

#### 4 结论及讨论

本文针对天基红外弱小目标检测领域长期存在的真实标注数据稀缺这一核心瓶颈,首次构建并发

表 12 不同算法在 IRSatVideo-LEO 数据集不同复杂度场景上对比的定量分析结果

Table 12 Quantitative analysis results of comparisons between different algorithms under different complexity scenarios on the IRSatVideo-LEO dataset

|             |                      | 简单场景         |         |                | 中等场景  |              |                | 复杂场景         |         |                |
|-------------|----------------------|--------------|---------|----------------|-------|--------------|----------------|--------------|---------|----------------|
|             |                      | IoU ↑        | Pd ↑    | Fa ↓           | IoU ↑ | Pd ↑         | Fa ↓           | IoU ↑        | Pd ↑    | Fa ↓           |
| 单帧红外小目标检测   | ACM                  | 33.94        | 93.83   | 1.10E-5        | 15.65 | 49.15        | 2.13E-5        | 14.73        | 60.34   | 2.42E-5        |
|             | ALCNet               | 36.18        | 86.13   | 3.07E-6        | 20.90 | 50.91        | 6.77E-6        | 20.55        | 64.25   | 1.24E-5        |
|             | DNANet               | 42.73        | 85.86   | 5.03E-6        | 19.14 | 40.07        | 4.27E-6        | 15.99        | 48.89   | 1.41E-5        |
|             | ISNet                | 16.14        | 95.30   | 5.48E-5        | 6.80  | 64.57        | 9.83E-5        | 15.69        | 56.91   | 1.47E-5        |
|             | UIU-Net              | 44.74        | 93.30   | 4.31E-6        | 20.06 | 42.30        | 9.84E-6        | 23.91        | 66.80   | 1.14E-5        |
|             | RDIAN                | 20.49        | 95.29   | 3.89E-5        | 9.36  | 48.98        | 4.82E-5        | 13.44        | 71.10   | 3.34E-5        |
|             | ISTDU-Net            | 45.71        | 89.04   | 5.42E-6        | 23.15 | 51.98        | 6.34E-6        | 20.71        | 61.90   | 1.21E-5        |
|             | ResUNet              | 28.57        | 96.09   | 2.14E-5        | 18.94 | 52.94        | 1.61E-5        | 20.51        | 70.42   | 1.59E-5        |
|             | RPCANet              | 20.43        | 90.64   | 3.90E-5        | 16.10 | 60.46        | 2.23E-5        | 14.39        | 65.01   | 2.75E-5        |
|             | AGPCNet              | 40.27        | 95.45   | 9.32E-6        | 24.56 | 59.80        | 1.14E-5        | 19.24        | 66.68   | 1.59E-5        |
| 序列红外小目标检测   | MSLSTIPT             | 1.18         | 13.34   | <b>1.67E-7</b> | 0.51  | 0.81         | <b>9.93E-9</b> | 0.00         | 0.00    | 7.54E-9        |
|             | NFTDGSTV             | 5.61         | 19.20   | 1.02E-6        | 0.22  | 17.80        | 2.03E-4        | 0.41         | 27.09   | 1.31E-4        |
|             | RCTVW                | 6.04         | 20.13   | 1.37E-6        | 0.05  | 22.86        | 1.20E-3        | 0.01         | 27.92   | 4.49E-3        |
|             | IMNN-LWEC            | 6.24         | 74.30   | 6.36E-5        | 0.42  | 37.85        | 5.97E-4        | 0.28         | 46.32   | 9.17E-4        |
|             | SRSTT                | 0.61         | 70.26   | 6.17E-4        | 0.00  | 45.78        | 7.07E-2        | 0.00         | 38.29   | 2.40E-1        |
|             | ASTTV-NTLA           | 15.15        | 86.57   | 1.37E-6        | 4.09  | 59.07        | 2.59E-6        | 0.94         | 72.22   | 1.90E-4        |
|             | 4DISTD               | 14.59        | 82.83   | 1.27E-5        | 10.12 | 59.38        | 6.10E-6        | 3.61         | 57.03   | 1.72E-5        |
|             | STDMA <sub>Net</sub> | 43.49        | 93.95   | 5.77E-6        | 26.57 | 59.09        | 7.12E-6        | <b>21.98</b> | 61.54   | 1.19E-5        |
|             | DNANet_DTUM          | 24.51        | 95.75   | 2.08E-5        | 16.74 | 59.43        | 2.22E-5        | 10.21        | 62.48   | 4.23E-5        |
|             | ResUNet_DTUM         | 25.18        | 83.03   | 1.42E-5        | 25.12 | 58.36        | 8.81E-6        | 13.14        | 59.33   | 2.70E-5        |
|             | UIU-Net_DTUM         | 30.78        | 77.13   | 1.10E-5        | 22.51 | <b>76.35</b> | 1.77E-5        | 16.45        | 58.90   | 1.88E-5        |
|             | ACM_RFR              | 38.56        | 84.46   | 3.56E-6        | 22.43 | 54.57        | 7.64E-6        | 16.00        | 50.83   | 1.31E-5        |
|             | ALCNet_RFR           | 41.27        | 89.00   | 3.20E-6        | 18.43 | 48.26        | 1.45E-5        | 10.94        | 57.00   | 3.58E-5        |
|             | DNANet_RFR           | 44.85        | 94.71   | 4.74E-6        | 23.78 | 58.78        | 9.82E-6        | 18.59        | 61.56   | 1.72E-5        |
|             | ISTUD-Net_RFR        | 35.87        | 77.28   | 2.42E-6        | 20.67 | 40.45        | 5.27E-6        | 21.24        | 51.69   | <b>5.78E-6</b> |
| ResUNet_RFR | <b>48.12</b>         | <b>96.13</b> | 2.87E-6 | <b>28.19</b>   | 60.55 | 5.88E-6      | 21.79          | <b>63.37</b> | 1.11E-5 |                |

布了大规模高保真仿真数据集 IRSatVideo。本工作提出并采用了创新的“半仿真”技术路径,通过将真实的卫星红外背景影像与基于轨道动力学、目标辐射物理的精确建模相融合,从源头上确保了数据的物理真实性与场景多样性。所构建的数据集包含 LEO 与 GEO 两大子集,规模达到近 19 万帧,是目前该领域规模最大、标注最完整的基准数据集。其核心价值在于,以可控、可扩展的方式,提供了兼具规

模优势(海量多样场景)与物理真实性优势(背景真实、目标辐射与运动特性高保真)的解决方案,系统地缓解了真实数据难以获取与标注的困境,为算法研发与评测提供了坚实可靠的数据基础。

本文系统的实验评估充分验证了本数据集的实用价值与判别能力。在目标信杂比极低( $SCR < 2$ )的 GEO 场景下,实验表明单帧检测方法普遍性能受限(平均检测率低于 10%),这凸显了利用时序信息进

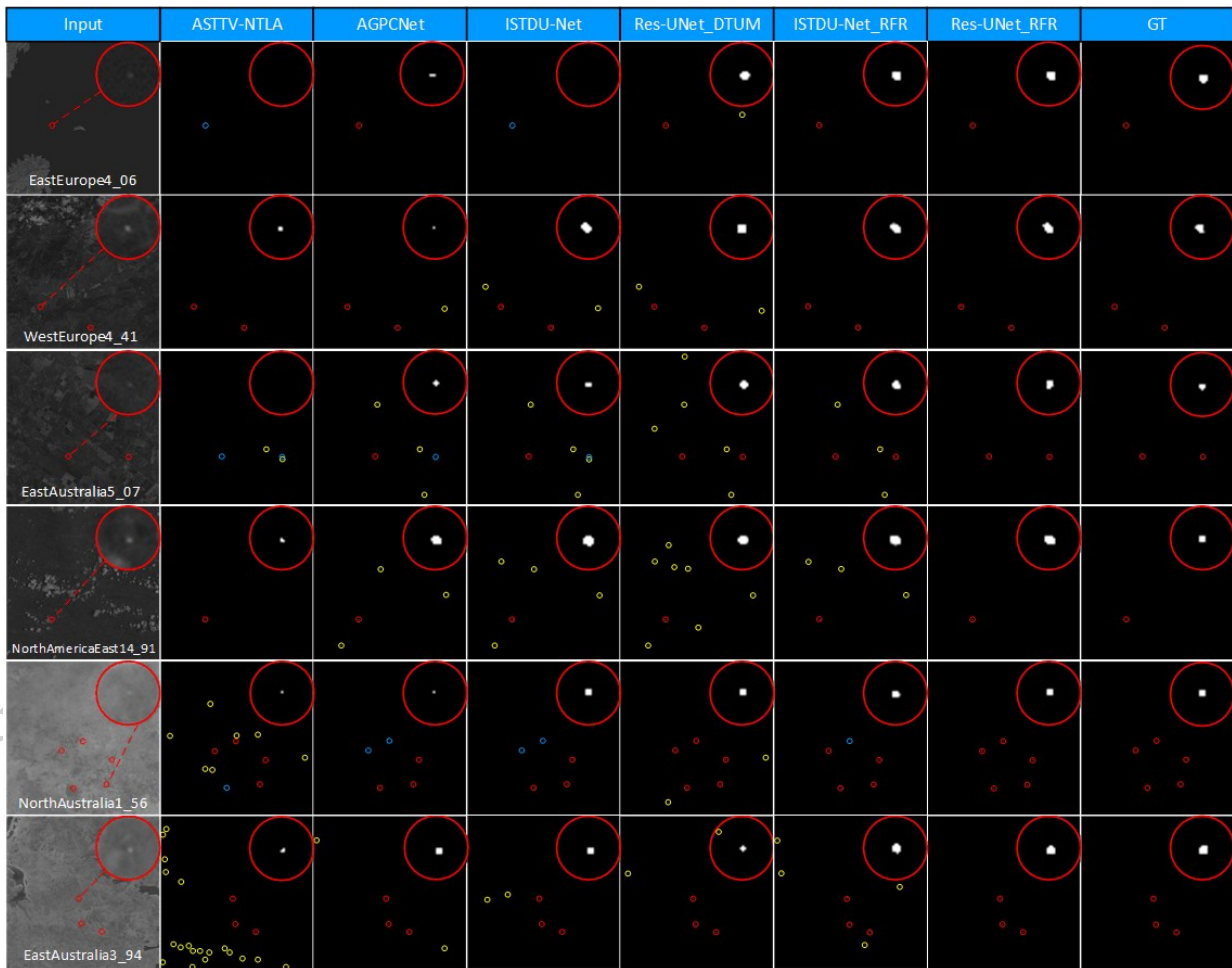


图20 不同算法在IRSatVideo-LEO数据集上对比的可视化结果

Fig. 20 Visualization results of comparison between different algorithms on the IRSatVideo-LEO dataset

行多帧累积检测的必要性;而在背景动态更复杂的LEO场景下,算法需同时应对目标机动性与复杂背景杂波的挑战。评估结果清晰显示,能够有效建模时空上下文关系的先进方法(如STDMA

-Net, ResUNet\_RFR)在本数据集上展现出显著优势,而传统方法或纯单帧方法则表现不佳,这种显著的性能分层效应有力证明了IRSatVideo具备卓越的判别力,能够可靠区分不同算法的适应性与鲁棒性。此外,图21对现有典型算法的失败案例进行了可视化分析,揭示了当前检测技术在面对两类典型挑战时的共性局限:暗弱目标漏检与类目标杂波虚警。在复杂背景或远距离观测条件下,目标辐射信号极其微弱,信噪比接近或低于背景起伏水平(如图中“目标1”与“目标2”所示)。现有算法往往依赖局部对比度或显著性特征,此类目标因强度与背景高度融合、形态模糊而难以被可靠提取,导致漏检。图像中存在的结构化背景干扰,在局部纹理、强度分布

上与真实目标高度相似(如图中“虚警1”“虚警2”“虚警3”)。这类目标杂波易被算法误判为目标,尤其在缺乏时序上下文或光谱辅助信息时,虚警率显著上升。

该案例进一步说明,仅依赖单帧外观特征或简单阈值分割的方法在应对极低信噪比目标与高相似性背景干扰时仍面临根本性挑战,凸显了引入时序建模、多特征融合与物理约束的必要性,也为后续研究指明了改进方向。

与现有公开数据集相比,IRSatVideo在多个维度实现了突破。在规模上,其近19万帧的总量远超SIRST、IRSTD-1K等同类数据集;在物理真实性上,通过引入真实背景影像并模拟卫星轨道与姿态运动,有效还原了天基观测中背景的规律性动态变化,避免了近景成像数据(如IRDST)中存在的视差与非规律运动偏差;在目标特性上,覆盖了从高信杂比到低至1的极弱目标,且包含机动行为,更贴近真实空

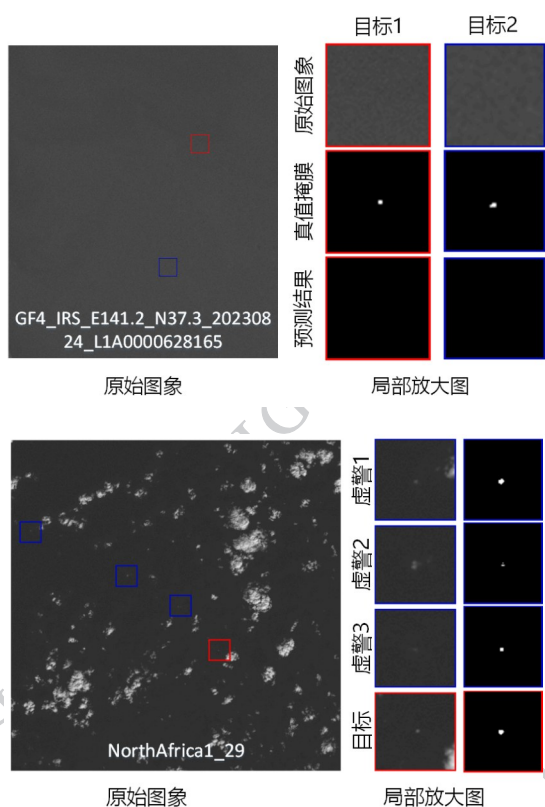


图 21 失败案例  
Fig. 21 Failure Case

间目标的复杂运动模式;在标注精度上,同时提供掩码与框级标注,支持从分割到跟踪的多任务统一评估,弥补了多数现有数据集仅提供粗粒度标签的不足。

然而,该数据集仍存在可改进之处。首先,当前仿真未考虑大气湍流对长距离成像的动态模糊效应,未来可通过引入时变点扩散函数进一步提升成像模型的物理逼真度。其次,目标建模以静态高斯斑点为主,模型类型较为单一,无法准确表征卫星本体、太阳能帆板等组件在红外波段可能呈现的非对称、时变辐射特征;同时,当前方法尚未涵盖目标自旋、翻滚等复杂姿态变化引起的闪烁与形态调制,限制了对动态特征建模算法的评估能力。此外,传感器退化效应仅模拟了盲元,未包括响应非均匀性、串扰、渐晕等更为复杂的器件级失真,与实际成像系统的偏差尚需弥补。在场景建模方面,背景中虽包含云层,但对极端天气条件(如强对流云团)的模拟仍显不足;同时,多目标之间的交互关系(如遮挡、近距离分辨困难等)也未被纳入仿真流程,限制了其在密集、交会场景下的适用性。

未来工作将围绕三方面展开:一是扩展多波段支持,引入长波红外与中波红外的协同观测数据,支持跨波段融合研究;二是构建更大规模的测试集,涵盖更多轨道类型(如MEO、HEO)与复杂交会场景,推动空间态势感知的深入研究;三是建立在线评测平台,支持检测、跟踪与轨迹预测等任务的自动化性能评估,促进算法社区的持续发展。IRSatVideo的发布有望为天基红外监视领域的算法创新提供坚实的数据基础,推动相关技术向更高精度、更强鲁棒性的方向演进。

### 参考文献 (References)

- BIAN Boyuan, ZHOU Feng, LI Xiaoman, JIN Libing, GONG Hui, LIAN Minlong. 2023. Research on Clutter Caused by Jitter of Geostationary Orbit Staring Infrared Camera[J]. *Infrared Technology*, 45(7): 775-783 [DOI: 10.3390/s23115278]
- Chen, Nuo, Li, Boyang, Wang, Yingqian, Ying, Xinyi, Wang, Longguang, Zhang, Chushu, et al. 2025. Motion and Appearance Decoupling Representation for Event Cameras. *IEEE Transactions on Image Processing*, 34: 5964-5977 [DOI: 10.1109/TIP.2025.3607632]
- Chen, Nuo, Xiao, Chao, Dai, Yimian, He, Shiman, Li, Miao and An, Wei. 2025. Event-based Tiny Object Detection: A Benchmark Dataset and Baseline//Proceedings of the IEEE/CVF International Conference on Computer Vision. Xi'an, China: IEEE/CVF: 7209-7218 [DOI: 10.48550/arXiv.2506.23575]
- Chen, Nuo, Zhang, Chushu, An, Wei, Wang, Longguang, Li, Miao and Ling, Qiang. 2025. Event-based Motion Deblurring with Blur-aware Reconstruction Filter. *IEEE Transactions on Circuits and Systems for Video Technology*, 35(9): 8508-8519 [DOI: 10.1109/TCSVT.2025.3551516]
- Dai, Yimian, Li, Xiang, Zhou, Fei, Qian, Yulei, Chen, Yaohong and Yang, Jian. 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, Yimian, Wu, Yiquan, Zhou, Fei and Barnard, Kobus. 2021. Asymmetric contextual modulation for infrared small target detection. *Proceedings of the IEEE Winter Conference on Applications of Computer Vision*, Virtual: IEEE: 949-958 [DOI: 10.1109/WACV48630.2021.00099]
- Dai, Yimian, Wu, Yiquan, Zhou, Fei and Barnard, Kobus. 2021. 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]
- Fu, Ruigang, Fan, Hongqi, Zhu, Yongfeng, Hui, Bingwei, Zhang, Zhilong, Zhong, P, et al. 2022. A dataset for infrared

- time-sensitive target detection and tracking for air-ground application. *Science Citation Database*, 7(2): 206-221 [DOI: 10.11922/sciencedb.j00001.00331]
- Guo, Yulan, Choi, Michael, Li, Kunhong, Boussaid, Farid and Benamoun, Mohammed. 2022. Soft exemplar highlighting for cross-view image-based geo-localization. *IEEE transactions on image processing*, 31: 2094-2105 [DOI: 10.1109/TIP.2022.3152046]
- Guo, Yulan, Wang, Longguang, Mao, Wendong, Dong, Xiaoyu, Wang, Yingqian, Liu, Li et al. 2025. Deep Lookup Network. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1-18 [DOI: 10.1109/TPAMI.2025.3605660]
- Hong Zhang, Lei Zhang, Ding Yuan and Hao Chen. 2018. Infrared small target detection based on local intensity and gradient properties. *Infrared Physics and Technology*, 89: 88-96 [DOI: 10.1016/J.INFRARED.2017.12.018]
- Hou Wang, Mei Feng-Hua, Cheng Guo-Jun and Deng Xi-Wen. 2015. An evaluation criterion of infrared image complexity based on background optimal filter scale. *Acta Physica Sinica*, 64(23): 234202 [DOI: 10.7498/aps.64.234202]
- Hou, Qingyu, Zhang, Liuwei, Tan, Fanjiao, Xi, Yuyang, Zheng, Haoliang and Li, Na. 2022. ISTDU-Net: Infrared Small-Target Detection U-Net. *IEEE Geoscience and Remote Sensing Letters*, 19: 1-5 [DOI: 10.1109/LGRS.2022.3141584]
- Hui Bingwei, Song, Zhiyong Song Zhiyong, Fan, Hongqi Fan Hongqi, Zhong, Ping et al. 2019. A dataset for infrared image dim-small aircraft target detection and tracking under ground / air background. *Science Data Bank Datasets*, 5(3) [DOI: 10.11922/csdata.2019.0074.zh]
- Jiang, Nan, Wang, Kuiran, Peng, Xiaoke, Yu, Xuehui, Wang, Qiang, Xing, Junliang, et al. 2022. Anti-UAV: a large-scale benchmark for visionbased UAV tracking. *IEEE Transactions on Multimedia*, 25: 486-500 [DOI: 10.1109/TMM.2021.3128047]
- Kim, Sungho and Lee, Joohyoung. 2012. Scale invariant small target detection by optimizing signaltoclutter ratio in heterogeneous background for infrared search and track. *Pattern Recognition*, 45 (1): 393-406 [DOI: 10.1016/j.patcog.2011.06.009]
- Li, Boyang, Xiao, Chao, Wang, Longguang, Wang, Yingqian, Lin, Zaiping, Li, Miao, et al. 2022. Dense nested attention network for infrared small target detection. *IEEE Transactions on Image Processing*, 32: 1745-1758 [DOI: 10.1109/TIP.2022.3199107]
- Li, Jie, Zhang, Ping, Zhang, Lingyi and Zhang, Zhiyuan. 2023. Sparse regularization-based spatial - temporal twist tensor model for infrared small target detection. *IEEE Transactions on Geoscience and Remote Sensing*, 61: 1-17 [DOI: 10.1109/TGRS.2023.3234608]
- Li, Ruojing, An, Wei, Xiao, Chao, Li, Boyang, Wang, Yingqian, Li, Miao 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, Weijie, Yang, Wei, Liu, Tianpeng, Hou, Yuenan, Li, Yuxuan, Liu, Zhen, et al. 2024. Predicting gradient is better: Exploring self-supervised learning for SAR ATR with a joint-embedding predictive architecture. *ISPRS Journal of Photogrammetry and Remote Sensing*, 218: 326-338 [DOI: 10.1016/j.isprs.2024.09.013]
- Liu, Li and Fieguth, Paul. 2012. Texture classification from random features. *IEEE transactions on pattern analysis and machine intelligence*, 34(3): 574-586 [DOI: 10.1109/TPAMI.2011.145]
- Liu, Li, Lao, Songyang, Fieguth, Paul W., Guo, Yulan, Wang, Xiaogang and Pietikäinen, Matti. 2016. Median robust extended local binary pattern for texture classification. *IEEE Transactions on Image Processing*, 25(3): 1368-1381 [DOI: 10.1109/TIP.2016.2522378]
- Liu, Ting, Yang, Jungang, Li, Boyang, Wang, Yingqian and An, Wei. 2023. Infrared small target detection via nonconvex tensor tucker decomposition with factor prior. *IEEE Transactions on Geoscience and Remote Sensing*, 61: 1-17 [DOI: 10.1109/TGRS.2023.3298192]
- Liu, Ting, Yang, Jungang, Li, Boyang, Wang, Yingqian and An, Wei. 2023. Representative coefficient total variation for efficient infrared small target detection. *IEEE Transactions on Geoscience and Remote Sensing*, 61:1-18 [10.1109/TGRS.2023.3324821]
- Liu, Ting, Yang, Jungang, Li, Boyang, Xiao, Chao, Sun, Yang, Wang, Yingqian, et al. 2021. Nonconvex tensor low-rank approximation for infrared small target detection. *IEEE Transactions on Geoscience and Remote Sensing*, 60: 1-18 [DOI: 10.1109/TGRS.2021.3130310]
- Wang Longguang, Guo Yulan, Dong Xiaoyu, Wang Yingqian, Ying Xinyi, Lin Zaiping, et al. 2022. Exploring fine-grained sparsity in convolutional neural networks for efficient inference. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(4): 4474-4493 [DOI: 10.1109/TPAMI.2022.3193925]
- Luo, Yuan, Li, Xiaorun, Chen, Shuhan, Xia, Chaoqun and Zhao, Liaoying. 2022. IMNN-LWEC: A novel infrared small target detection based on spatial-temporal tensor model. *IEEE Transactions on Geoscience and Remote Sensing*, 60: 1-22 [DOI: 10.1109/TGRS.2022.3230051]
- Mingjin Zhang, Yuchun Wang, Jie Guo, Yunsong Li, Xinbo Gao and Jing Zhang. 2024. IRSAM: Advancing segment anything model for infrared small target detection//European Conference on Computer Vision. Milan, Italy: Springer: 233-249 [DOI: 10.1007/978-3-031-72855-6\_14]
- Shermeyer, Jacob, Hossler, Thomas, Van Etten, Adam, Hogan, Daniel, Lewis, Ryan and Kim, Daeil. 2021. Rareplanes: Synthetic data takes flight. *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, Virtual: IEEE/CVF: 207-217* [DOI: 10.48550/arXiv.2006.02963]
- Tong, Xiaozhong, Zuo, Zhen, Su, Shaojing, Wei, Junyu, Sun, Xiaoy-

- ong, Wu, Peng et al. 2024. ST-Trans: Spatial-temporal transformer for infrared small target detection in sequential images. *IEEE Transactions on Geoscience and Remote Sensing*, 62: 1-19 [DOI: 10.1109/TGRS.2024.3355947]
- Su, Zhuo, Zhang, Jiehua, Wang, Longguang, Zhang, Hua, Liu, Zhen, Pietikäinen, Mattiet al. 2023. Lightweight pixel difference networks for efficient visual representation learning. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45 (12) : 14956-14974 [DOI: 10.1109/TPAMI.2023.3300513]
- Sun Y, Yang J and An W. 2020. Infrared dim and small target detection via multiple subspace learning and spatial-temporal patch-tensor model. *IEEE Transactions on Geoscience and Remote Sensing*, 59 (5): 3737-3752 [DOI: 10.1109/TGRS.2020.3022069]
- Sun, Heng, Bai, Junxiang, Yang, Fan and Bai, Xiangzhi. 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: 1-13 [DOI: 10.1109/TGRS.2023.3235150]
- Sun, Heng, Bai, Junxiang, Yang, Fan and Bai, Xiangzhi. 2023. Receptive-field and direction induced attention network for infrared dim small target detection with a large-scale dataset IRDST [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 61: 1-13 [DOI: 10.1109/TGRS.2023.3235150]
- Sun, X, Guo, L, Zhang, W, Wang, Z, Hou, Y, Li, Z and Teng, X. 2021. A dataset for small infrared moving target detection under clutter background. *Science Citation Database*, 5(6) : 8 [DOI: 10.11922/sciedb.j00001.00231]
- Wang, Huan, Zhou, Luping and Wang, Lei. 2019. Miss detection vs. false alarm: Adversarial learning for small object segmentation in infrared images. *IEEE/CVF International Conference on Computer Vision*, Seoul, South Korea: IEEE/CVF: 8509-8518 [DOI: 10.1109/ICCV.2019.00860]
- Wang, Longguang, Guo, Yulan, Wang, Yingqian, Dong, Xiaoyu, Xu, Qingyu, Yang, Junganget al. 2024. Unsupervised degradation representation learning for unpaired restoration of images and point clouds. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 47(1): 1-18 [DOI: 10.1109/TPAMI.2024.3471571]
- Li Weijie, Yang Wei, Hou Yuenan, Liu Li, Liu Yongxiang and Li Xiang. 2025. SARATR-X: Towards building a foundation model for SAR target recognition. *IEEE Transactions on Image Processing*, 34: 869 - 884 [DOI: 10.1109/TIP.2025.3531988]
- Wu, Fengyi, Yu, Hang, Liu, Anran, Luo, Junhai and Peng, Zhenming. 2023. Infrared small target detection using spatiotemporal 4-D tensor train and ring unfolding. *IEEE Transactions on Geoscience and Remote Sensing*, 61: 1-22 [DOI: 10.1109/TGRS.2023.3288024]
- Wu, Fengyi, Zhang, Tianfang, Li, Lei, Huang, Yian and Peng, Zhenming. 2024. RPCANet: Deep unfolding RPCA based infrared small target detection. *Proceedings of the IEEE Winter Conference on Applications of Computer Vision*, WAIKOLOA, HAWAII: IEEE: 4809-4818 [DOI: 10.1109/WACV57701.2024.00474]
- Wu, Tianhao, Li, Boyang, Luo, Yihang, Wang, Yingqian, Xiao, Chao, Liu, Ting, et al. 2023. MTU-Net: Multi-level TransUNet for space-based infrared tiny ship detection. *IEEE Transactions on Geoscience and Remote Sensing*, 61: 1-15 [DOI: 10.1109/TGRS.2023.3235002]
- Wu, Xin, Hong, Danfeng and Chanussot, Jocelyn. 2022. UIU-Net: U-Net in U-Net for infrared small object detection. *IEEE Transactions on Image Processing*, 32: 364-376 [DOI: 10.1109/TIP.2022.3228497]
- Xiao, Chao, An, Wei, Zhang, Yifan, Su, Zhuo, Li, Miao, Sheng, et al. 2024. Highly efficient and unsupervised framework for moving object detection in satellite videos. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 46(12): 11532-11539 [DOI: 10.1109/TPAMI.2024.3409824]
- Xiao, Chao, Yin, Qian, Ying, Xinyi, Li, Ruoqing, Wu, Shuanglin, Li, Miao, et al. 2021. DSFNNet: Dynamic and static fusion network for moving object detection in satellite videos. *IEEE Geoscience and Remote Sensing Letters*, 19: 1~5 [DOI: 10.1109/LGRS.2021.3124222]
- Xiao, Xiao, Lian, Shen, Luo, Zhiming and Li, Shaozi. 2018. Weighted Res-UNet for high-quality retina vessel segmentation. *Proceedings of the International Conference on Information Technology in Medicine and Education*, Hangzhou, China: IEEE: 327-331 [DOI: 10.1109/ITME.2018.00080]
- Ying Xinyi, Liu Li, Wang Yingqian, Li Ruoqing, Chen Nuo, Lin Zaiping, et al. 2023. Mapping degeneration meets label evolution: Learning infrared small target detection with single point supervision//*Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. Vancouver, Canada: IEEE/CVF: 15528-15538 [DOI: 10.1109/cvpr52729.2023.01490]
- Yan, Puti, Hou, Runze, Duan, Xuguang, Yue, Chengfei, Wang, Xin and Cao, Xibin. 2023. STDMA Net: Spatio-temporal differential multiscale attention network for small moving infrared target detection. *IEEE Transactions on Geoscience and Remote Sensing*, 61: 1-16 [DOI: 10.1109/TGRS.2023.3241311]
- Yin, Qian, Hu, Qingyong, Liu, Hao, Zhang, Feng, Wang, Yingqian, Lin, Zaiping, et al. 2021. Detecting and tracking small and dense moving objects in satellite videos: A benchmark. *IEEE Transactions on Geoscience and Remote Sensing*, 60: 1~18 [DOI: 10.1109/TGRS.2021.3130436]
- Ying, Xinyi, Liu, Li, Lin, Zaiping, Shi, Yangsi, Wang, Yingqian, Li, Ruoqing, et al. 2025. 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: 1-18 [DOI: 10.1109/TGRS.2025.3542368]
- Ying, Xinyi, Xiao, Chao, An, Wei, Li, Ruoqing, He, Xu, Li, Boyang, et al. 2025. Visible-thermal tiny object detection: A bench-

- mark dataset and baselines. IEEE Transactions on Pattern Analysis and Machine Intelligence, 47 (7) : 6088-6096 [DOI: 10.1109/TPAMI.2025.3544621]
- Ying, Xinyi, Xiao, Chao, An, Wei, Li, Ruoqing, He, Xu, Li, Boyang, et al. 2025. 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]
- Liu Yongxiang, Li Weijie, Liu Li, Zhou Jie, Peng Bowen, Song Yafei, et al. 2025. ATRNet-STAR: A Large Dataset and Benchmark Towards Remote Sensing Object Recognition in the Wild [OL]. [2025-11-24]. <https://arxiv.org/pdf/2501.13354>
- Zhang, Feng, Wang, Xueying, Zhou, Shilin, Wang, Yingqian and Hou, Yi. 2021. Arbitrary-oriented ship detection through center-head point extraction. IEEE Transactions on Geoscience and Remote Sensing, 60: 1~14 [DOI: 10.1109/TGRS.2021.3120411]
- Zhang, Mingjin, Feng, Jin, Yang, Handi, Guo, Jie, Li, Yunsong and Gao, Xinbo. 2025. IRPruneDeXt: Efficient Infrared Small Target Detection via Musical Wavelet-Regularized Channel Pruning. IEEE Transactions on Neural Networks and Learning Systems, 1-14 [DOI: 10.1109/TNNLS.2025.3594958]
- Zhang, Mingjin, Yue, Ke, Guo, Jie, Zhang, Qiming, Zhang, Jing and Gao, Xinbo. 2025. Computational fluid dynamic network for infrared small target detection. IEEE Transactions on Neural Networks and Learning Systems, 36 (8) : 14777-14789 [DOI: 10.1109/TNNLS.2025.3548984]
- Zhang, Mingjin, Zhang, Chi, Zhang, Qiming, Li, Yunsong, Gao, Xinbo and Zhang, Jing. 2024. Unleashing the power of generic segmentation model: A simple baseline for infrared small target detection//Proceedings of the 32nd ACM International Conference on Multimedia. Melbourne, Australia; ACM: 10392-10401 [DOI: 10.1145/3664647.3680609]
- Zhang, Mingjin, Zhang, Rui, Yang, Yuxiang, Bai, Haichen, Zhang, Jing and Guo, Jie. 2022. ISNet: Shape matters for infrared small target detection. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, New Orleans, LA, USA: IEEE: 877-886 [DOI: 10.1109/CVPR52688.2022.00095]
- Zhang, Tianfang, Li, Lei, Cao, Siying, Pu, Tian and Peng, Zhenming. 2023. Attention-guided pyramid context networks for detecting infrared small target under complex background [J]. IEEE Transactions on Aerospace and Electronic Systems, 59(4): 4250-4261 [DOI: 10.1109/TAES.2023.3238703]
- Zhou, Jie, Liu, Yongxiang, Peng, Bowen, Liu, Li and Li, Xiang. 2025. MaDiNet: Mamba Diffusion Network for SAR Target Detection[J]. IEEE Transactions on Circuits and Systems for Video Technology, 35: 10787-10800 [DOI: 10.1109/TCSVT.2025.3574657]

### 作者简介

应昕怡:女,讲师,研究方向为天基光学监视、小目标检测、图像超分辨。Email: yingxinyi18@nudt.edu.cn

张鑫:男,博士研究生,主要研究方向为小目标检测、密集目标计数、密集目标跟踪。Email: zhangxin21@nudt.edu.cn

周翔:男,研究生,主要研究方向为小目标检测、小目标跟踪。Email: zx\_25@nudt.edu.cn

曹旭:男,博士研究生,主要研究方向为多源融合解译、目标检测。Email: cx2020@nudt.edu.cn

林再平:男,副研究员,主要研究方向为天基动目标信息处理、红外图像处理。Email: linzaiping@nudt.edu.cn

李淼:男,副研究员,主要研究方向为红外弱小目标检测、事件相机表征。Email: lm8866@nudt.edu.cn

周石琳:男,教授,主要研究方向为遥感图象处理。Email: zhoushilin@nudt.edu.cn

安玮:女,教授,主要研究方向为图像处理、信号处理。Email: anwei@nudt.edu.cn