

中图法分类号: TP391.41; TN448.2 文献标识码: A 文章编号: 1006-8961(2026)07-2471-34

论文引用格式: Jiang J X and Wu Y Q. 2026. Progress of research on noncooperative space object recognition and pose estimation on the basis of vision and deep learning. Journal of Image and Graphics, 31(7):2471-2504(姜建湘, 吴一全. 2026. 基于视觉与深度学习的非合作空间目标识别与位姿估计方法研究进展. 中国图象图形学报, 31(7):2471-2504)[DOI:10.11834/jig.250435]

基于视觉与深度学习的非合作空间目标识别与位姿估计方法研究进展

姜建湘, 吴一全*

南京航空航天大学电子信息工程学院, 南京 211106

摘要: 为提升非合作空间目标在轨服务与态势感知中的感知能力, 本文系统综述了基于视觉与深度学习的非合作空间目标识别与位姿估计方法研究进展。首先阐明非合作空间目标识别与位姿估计的任务特点, 简要回顾了各自的发展历程。其次围绕非合作空间目标识别与位姿估计等关键任务, 梳理了当前主流方法。在目标识别方面, 重点分析了基于多模态视觉融合、目标检测与分割、迁移学习与少样本学习的识别模型及其在复杂空间环境中的适应性与鲁棒性; 在位姿估计方面, 主要包括直接回归、关键点检测、无监督与域自适应以及多模态视觉融合等技术。随后梳理了空间目标公开数据集及自建数据集, 评估了目标识别与位姿估计方法的性能, 并分别对其优缺点进行比较。结果表明, 深度学习方法相较于传统方法在弱纹理、遮挡和姿态变化等复杂空间环境下具有更强的鲁棒性, 多模态融合与无监督策略进一步提升了跨域泛化能力。最后总结了目标识别与位姿估计方法面临的主要挑战, 并展望了未来的发展方向, 旨在为该领域的技术研究与工程应用提供参考。本文提及的算法、数据集和评价指标已汇总至 <https://github.com/viskyl/openResource/tree/main>。

关键词: 非合作空间目标; 目标识别; 位姿估计; 视觉; 深度学习; 多模态视觉融合; 直接回归; 关键点检测

Progress of research on noncooperative space object recognition and pose estimation on the basis of vision and deep learning

Jiang Jianxiang, Wu Yiquan*

College of Electronic and Information Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China

Abstract: With the recent increase in space missions and the proliferation of space debris, the perception of noncooperative objects has become a critical issue for on-orbit servicing and space situational awareness. Noncooperative objects are objects in space that lack any cooperative markers, communication channels, or navigation aids, such as defunct satellites, orbital debris, small debris, or unknown spacecraft. Identifying such objects and estimating their six-degree-of-freedom pose is crucial for various space activities, including debris removal, rendezvous and docking, satellite servicing, and space security. Reliable identification provides prior semantic and structural information for subsequent pose estimation, and accurate pose estimation supports autonomous navigation, capture, and manipulation. Therefore, progress in this field is not only of scientific value but also of practical importance for ensuring the safety and sustainability of space

收稿日期: 2025-09-17; 修回日期: 2025-11-17; 预印本日期: 2025-11-23

* 通信作者: 吴一全 nuaaimage@163.com

基金项目: 国家自然科学基金项目(61573183)

Supported by: National Natural Science Foundation of China (61573183)

operations. This study comprehensively reviews the progress of the latest research on noncooperative space object recognition and pose estimation on the basis of vision and deep learning. Unlike traditional computer vision methods that rely heavily on handcrafted features and geometric modeling, deep learning methods demonstrate high robustness, adaptability, and generalization advantages in complex space environments, such as those with weak textures, drastic illumination variations, rapid pose changes, and partial occlusion. We categorize existing research into two major areas, namely, object recognition and pose estimation, and systematically analyze representative methods along with their strengths, limitations, and applicability. For object recognition, three main research directions are identified. The first is multimodal visual fusion, which integrates data from multiple sensors, such as RGB cameras, infrared imaging, depth maps, or radar. By leveraging complementary features across modalities, these methods can overcome the limitations of single-sensor systems, particularly in low visibility or with incomplete imagery. However, multimodal fusion often incurs additional computational overhead and sensor dependency. The second research direction is object detection and segmentation methods. Powered by deep models, such as you only look once (YOLO), single-shot multibox detector, faster region-based convolutional neural network, and mask region-based convolutional neural network, these methods have substantially improved spatial object localization and fine-grained analysis capabilities. However, these methods typically require large-scale annotated datasets and often struggle to identify very small or distant objects. The third direction is transfer learning and few-shot learning methods, which mitigate data scarcity by adopting pretrained models or meta-learning frameworks. These methods have produced promising results in identifying satellites or debris with limited sample sizes, although domain transfer between synthetic and real images remains a challenge. For pose estimation, this study reviews four major categories of methods. First, direct regression methods regard pose estimation as an end-to-end mapping problem and are highly effective when dealing with symmetric or highly variable objects. Second, keypoint detection methods combine landmark localization with geometric constraints to achieve robust performance. Third, unsupervised and domain adaptation methods reduce the reliance on expensive labeled data and address the gap between synthetic and real domains through reconstruction losses, contrastive learning, or adversarial adaptation. Finally, multimodal fusion methods combine visual data with LiDAR, radar, or infrared sensing to improve robustness in extreme environments, but the complexity of the integration limits their widespread application. In addition to methodological analysis, this study summarizes public datasets, such as SPEED, SPEED+, URSO, and BUAA-SID, and self-constructed datasets. Evaluation metrics, such as position error, pose error, mean average precision, intersection over union, and inference speed, are summarized to provide strong data support for researchers. For spatial object recognition, the traditional principal component analysis feature extraction method combined with random forest classification achieves an accuracy of approximately 89.13%. The deep convolutional neural network (DCNN) approach substantially improves recognition performance, with its accuracy increments ranging from 4.31% to 15.03%. By further employing data augmentation strategies, DCNN's recognition accuracy is as high as 99.90%, demonstrating exceptional generalization and robustness. In the pose estimation task, with the SPEED dataset as a benchmark, the traditional perspective-n-point algorithm achieves a pose angle error of 8.74° , with its performance degrading by over 40% in low-resolution imagery. By contrast, the deep learning-based DSOAE-Net model reduces the pose angle error to 7.53° and maintains stable accuracy in low-resolution scenarios, demonstrating high environmental adaptability and robustness. Comparative analysis shows that deep learning-based methods outperform traditional methods in terms of robustness to illumination variations, occlusions, and structural ambiguity. Multimodal fusion and domain adaptation strategies are particularly effective in improving cross-domain generalization capabilities. In sum, this study summarizes the main challenges that face object recognition and pose estimation methods and proposes future development directions to provide a reference for technical research and engineering applications in this field. The methods, datasets, and evaluation metrics mentioned are linked at <https://github.com/viskyl/openResource/tree/main>.

Key words: non-cooperative space target; target recognition; pose estimation; vision; deep learning; multimodal visual fusion; direct regression; key point detection

0 引言

随着航天活动的日益频繁,非合作空间目标的识别与位姿估计已成为在轨服务与态势感知的关键技术。非合作目标通常指形态未知、无法传递相对状态信息、缺乏合作标识或抓持装置的目标,如太空垃圾、失效卫星、微小碎片及未知航天器等(Bonnal等,2013)。目标识别判别目标类别与结构特征,位姿估计则在此基础上确定目标的空间位置与姿态,二者密切相关,识别结果可为位姿估计提供先验模型,准确的位姿估计亦有助于提升识别的精度与稳定性。

为保障在轨服务质量,国际上先后提出了多项轨道服务与空间碎片清除任务,其中包括英国萨里航天中心的 RemoveDEBRIS 任务(Aglietti等,2020)和美国国防部高级研究计划局(Defense Advanced Research Projects Agency, DARPA)(Dang等,2020)以及美国航空航天局(National Aeronautics and Space Administration, NASA)的 Restore-L 任务等(Lopatka,2018)。这些任务对非合作目标的准确识别与位姿估计提出了迫切需求。识别为位姿估计提供结构先验与语义信息,位姿估计则为自动对接、捕获和操作控制提供核心支撑。在轨碎片与失效航天器的识别与清除可保障航天器安全,维护太空环境;在空间态势感知与国家安全领域,识别与位姿估计对于检测和跟踪潜在威胁目标同样具有重要意义。因此,提升非合作空间目标的感知能力对于增强空间任务智能化与保障空间资产安全具有重要价值。

现有的非合作空间目标识别与位姿估计方法主要分为基于传统计算机视觉的方法和基于深度学习的方法。传统方法依赖图像处理与特征提取,虽然可实现目标识别与位姿估计,但在光照变化、目标遮挡和姿态变化下,精度与鲁棒性易受影响。相比之下,深度学习近年来展现出强大的检测与识别能力,通过构建多层神经网络,显著增强了特征表达能力。典型模型包括卷积神经网络(convolutional neural network, CNN)(Sharma和D'Amico,2019)、循环神经网络(recurrent neural network, RNN)(Liu等,2022a)、生成对抗网络(generative adversarial network, GAN)(Al Najar等,2023)及基于注意力机制的模型。卷积神经网络擅长提取多层次特征,对光照变化具备鲁棒性,并能利用全局上下文信息进行特

征学习;面对目标遮挡时,亦可通过未遮挡部分提取全局特征以提高识别率(Alzubaidi等,2021);此外,深度模型可通过大规模数据训练,学习不同姿态下的特征,对姿态变化具有较强的鲁棒性。在基于视觉与深度学习的空间目标感知领域,目标识别方法包括多模态视觉融合、目标检测与分割以及迁移学习与少样本学习。多模态融合丰富目标特征表达,检测与分割实现目标定位与形状分割,迁移学习与少样本学习提升模型在数据匮乏场景下的泛化能力。位姿估计方法则可分为4类:直接回归通过端到端预测位姿参数;关键点检测先定位语义关键点再结合几何关系推算姿态;无监督及域自适应解决标签缺乏与跨域泛化问题;多模态视觉融合整合多传感器数据以提升感知准确性与稳定性。这些方法体系相辅相成,共同推动了非合作空间目标感知技术的进步。

魏振忠等人(2023)综述了基于视觉的位姿估计原理,介绍了基于点特征、线特征和区域特征等位姿估计方法,但主要聚焦于传统的位姿估计方法。牟金震等人(2021)对空间目标检测进行了系统阐述,分析了特征提取、图像配准以及三维重建等技术在位姿估计中的应用,但侧重于传统方法并且没有给出各方法之间的定量分析。朱文山等人(2023)综述了基于深度学习的航天器位姿估计研究进展,分析传统特征提取方法的局限性,聚焦公开比赛及深度学习算法,但其在跨模态数据融合以及传统深度学习协同方面讨论不足。Pauly等人(2023)对基于深度学习的单目航天器姿态估计领域进行了系统性总结,主要将现有方法分成两类混合模块方法及端到端方法,但其对于域自适应方法讨论较少且未提及空间目标及其部组件识别。因此,有必要对基于视觉与深度学习的非合作空间目标识别和位姿估计技术进行全面的综述,以梳理当前研究进展,分析现有方法的优势与不足,并揭示未来的研究方向。

本文系统综述了基于视觉与深度学习的非合作空间目标识别与位姿估计方法。首先阐述了该领域的研究背景及其在空间任务中的重要性,阐明了目标识别与位姿估计的基本概念及二者的任务特点,并简要回顾其发展脉络。重点分析了基于视觉与深度学习的目标识别与位姿估计方法,系统比较各类方法的性能指标,总结其优势与局限性。随后,梳理了当前的空间目标公开数据集及典型自建数据集,

为后续研究提供数据支持。最后,针对现有研究中的关键技术挑战,展望了未来的发展方向。本文旨在为相关研究人员提供系统参考,推动非合作空间目标识别与位姿估计技术的持续发展与工程应用。

1 非合作空间目标识别与位姿估计的任务特点及其发展历程

1.1 任务特点

非合作空间目标识别旨在基于遥感观测信息对空间目标的基本属性进行判别与分类,分类通常按照目标的类型,如卫星、碎片和飞行器等;按照目标的类别,如通信卫星、遥感卫星等;以及按照目标的结构组成,如主体舱段、太阳翼和天线等部件(Lin等,2023)进行,如图1所示。

非合作空间目标识别任务是空间态势感知系统的核心组成部分,对于目标归属分析、威胁预警、轨道管理以及空间资源统计等具有重要意义。识别任务通常依赖于图像识别、语义分割、雷达信号处理及三维重建等技术,其中深度学习方法如CNN、Transformer等在图像分类和检测任务中表现优异,而基于雷达散射截面特征(radar cross section, RCS)的序列建模方法(Deng和Li,2023)则在雷达识别领域得到广泛采用。随着多模态观测数据的可获取性增强,结合图像、光谱和激光雷达等多源数据的融合识别方法也成为研究重点。

非合作空间目标位姿估计是指在缺乏任何通信、导航或定位协同机制的条件下,利用单目/双目相机、激光雷达和雷达成像等遥感手段对轨道空间目标进行三维位置与姿态的6自由度位姿估计,图2和图3分别为三维位置和姿态估计的可视化。

非合作空间目标位姿估计任务面向的对象通常为失效卫星、空间碎片或其他非合作航天器,其显著特征是目标无主动标识装置如反光标、光学信标或全球导航卫星系统(global navigation satellite system, GNSS)等且运动状态不确定(Federici等,2021)。位姿估计是执行在轨服务、主动碎片移除以及交会对接等任务的基础,其结果直接用于导航、路径规划与姿控系统的输入,因此对实时性、精度与鲁棒性有较高要求。主流方法包括基于几何特征的视觉估计、深度学习回归网络、点云配准技术(胡伏原等,2025)以及多模态视觉融合等。近年来结合图像与

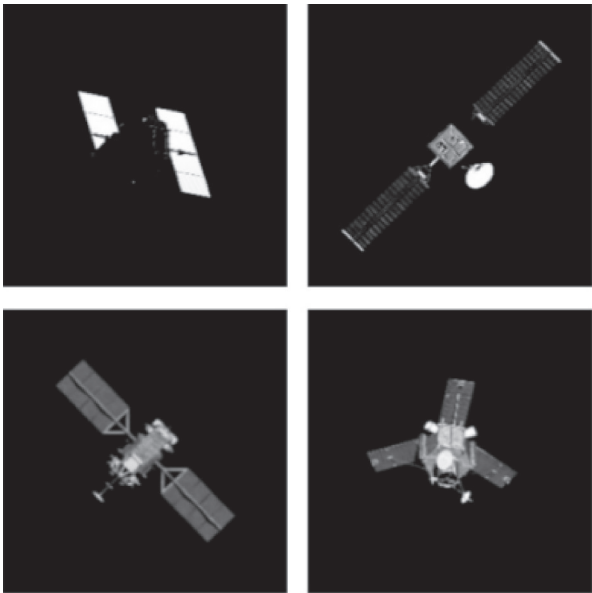


图1 目标部件示意图(郝强等,2022)
Fig. 1 Schematic diagram of the target components
(Hao et al., 2022)



图2 三维位置估计可视化
Fig. 2 Visualization of 3D position estimation

点云的端到端估计模型也逐渐成为研究热点。非合作空间目标识别与位姿估计在任务目标上各有侧重,但在空间场景应用中高度耦合(Li等,2022b)。如表1所示,两者在研究目标、输入数据及典型方法等方面存在差异:识别侧重于判定目标类型与结构属性,而位姿估计则关注目标的三维位置与姿态求解。然而在实际系统中,二者常作为前后关联的处理模块协同运行。

图4展示了两者的总体流程。图4(a)为目标识别算法,包括图像输入、预处理、特征提取、深度特征融合与目标分类,实现对目标的精确识别;图4(b)为位姿估计算法,依次进行特征提取、特征匹配、位姿求解与优化,用于获取目标的空间位姿信息。两

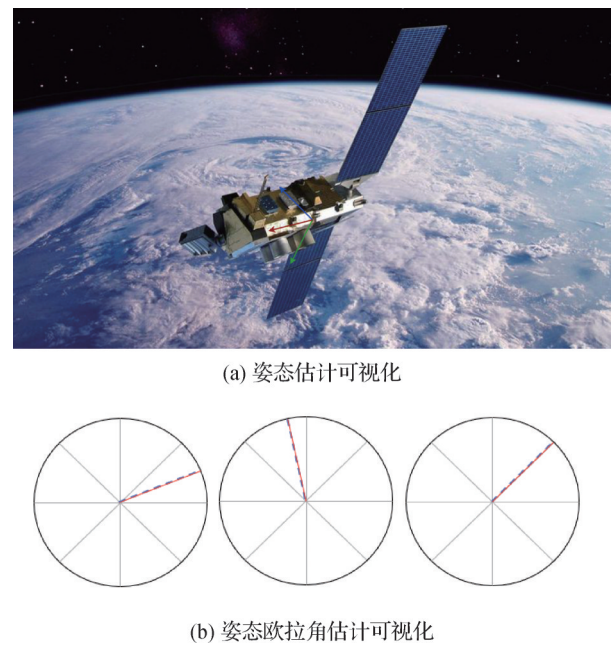


图 3 三维姿态估计可视化

Fig. 3 Visualization of 3D pose estimation ((a) pose estimation visualization; (b) pose Euler angle estimation visualization)

种流程在特征提取与后续优化阶段具有高度一致性,为识别与位姿估计的联合建模、一体化感知及深度融合提供了技术支撑。

表 1 非合作空间目标识别与位姿估计的区别与联系

Table 1 Differences and connections between non-cooperative spatial target recognition and pose estimation		
对比维度	目标识别	位姿估计
研究目标	判定目标的类型、类别与结构组成	获取目标的 3D 位置和姿态
输出形式	目标类别、语义部件等	位置向量 + 姿态四元数(袁萌萌和张泽旭, 2023)
输入数据	图像、视频、RCS 序列和三维重建结果等	图像、点云、视频流、雷达成像等
典型方法	CNN 分类器(Chu 等, 2023)、YOLO 检测(Davis 和 Pernicka, 2019)、语义分割(Du 等, 2023)	PnP 算法(Barbier 和 Gao, 2023)、ICP 配准(Chang 等, 2021)、深度回归网络
依赖条件	更依赖外观信息与语义特征提取能力	通常需要特征可见性与几何完整性

相较而言,刘福才和张晓(2021)则侧重于利用图像几何与颜色特征,提出了一种基于三级分类器级联的太阳帆板检测算法,主要通过旋转图像检测角度未知的长矩形目标、颜色分布积分图和基于积分图的特征提取方法,以提高效率。同时,使用 Adaboost 分类器结合颜色和 Haar-like 特征进行目标检测,并采用非极大值抑制(non-maximum suppression, NMS)和 Soft-NMS 处理检测框的重叠问题,但该方法仍依赖手工特征设计,在复杂干扰场景可能出现误检。由上述可知,传统方法在特征选择与参

1.2 目标识别方法的发展历程

非合作空间目标识别技术的发展经历了从依赖人工特征的传统方法向深度学习驱动的智能识别方法的转变。在非合作空间目标识别任务中,传统方法主要依赖于手工设计的图像特征与分类器的组合进行目标类型判别。这类方法通常包括边缘检测、轮廓提取、纹理分析以及颜色分布计算等低层视觉特征提取手段,结合模板匹配、统计学习或投影分析等方式进行识别判断。

杨虹等人(2020)提出了一种基于局部 Trace 变换和集成分类器的逆合成孔径雷达图像(inverse synthetic aperture radar, ISAR)空间目标识别方法。其核心是通过 Canny 边缘检测和 Hough 变换提取 ISAR 图像最长轴,在最长轴附近划分局部区域进行 Trace 变换生成低维 Trace 矩阵,结合移位对准操作消除平移影响,并引入 AdaBoost.M2-KNN(adaptive boosting multi-class version 2-k-nearest neighbors)集成分类器进行目标分类,尽管通过主成分分析(principal component analysis, PCA)降维缓解数据冗余,但其计算复杂度较高,局部区域角度划分和多 Trace 函数运算可能影响实时性。

数设定上高度依赖经验知识,难以应对复杂多变的空

间任务环境。

随着深度学习的快速发展,尤其是卷积神经网络的广泛应用,研究者开始探索利用大规模合成图像数据训练识别模型,实现端到端的特征学习与分类判别。典型方法如基于残差网络(residual network, ResNet)(He 等, 2020)、视觉几何组(Visual Geometry Group, VGG)(Su 等, 2023)和密集连接的卷积网络(densely connected convolutional network, DenseNet)(AlDahoul 等, 2022b)等结构构建的多分

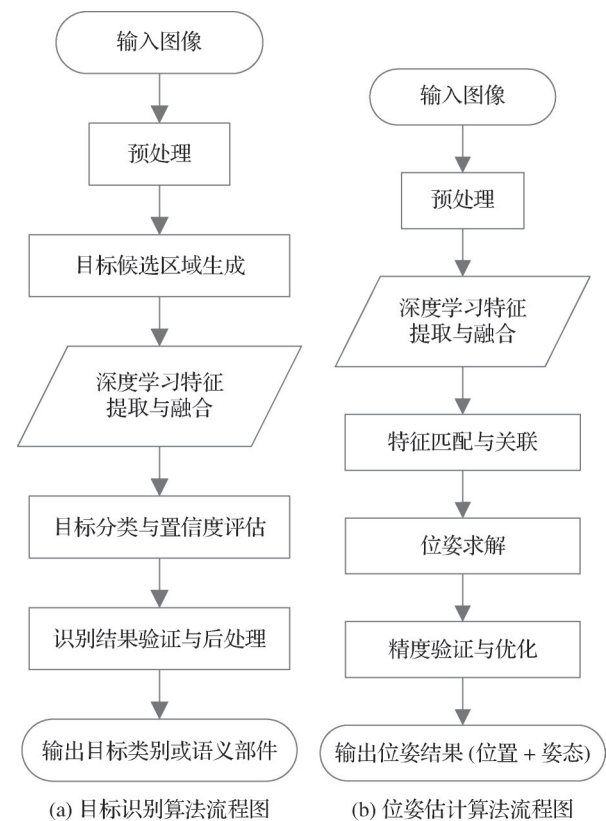


图4 目标识别与位姿估计算法流程示意图

Fig. 4 Schematic diagram of target recognition and pose estimation algorithm process ((a) target recognition algorithm flowchart; (b) pose estimation algorithm flowchart)

类识别网络,能够从图像中自动提取空间目标的纹理、形状、边缘与上下文等多层级语义特征,并应用于目标类型识别、构型识别和部组件识别等任务。例如,基于 ResNet 主干网络的识别框架常用于静态图像识别,而结合时序信息的 3D CNN 或长短期记忆网络(long short-term memory, LSTM)模型(Wang 等, 2020)则适用于视频流中的动态目标识别。此外,考虑到空间图像中目标结构复杂、语义区域分布不均,多尺度特征融合与注意力机制被广泛引入。例如压缩与激励(squeeze-and-excitation, SE)模块、卷积注意力模块(convolutional block attention module, CBAM)以及 Transformer Encoder 等结构可增强模型对关键部位如天线、太阳帆板和推进器等的感知能力,提升对细粒度结构的识别效果。在多模态融合方面,研究者逐渐采用 RGB 图、深度图和红外图等信息融合,通过 Early Fusion、Late Fusion 或 Cross-modal Attention 等机制,增强模型对复杂目标的描述能力,尤其适用于低可见度或图像不完整场景。为应对数据稀缺与仿真—现实域间差异,迁移

学习、无监督学习和少样本学习方法得以引入。典型策略包括使用在 ImageNet 等大规模数据集上预训练的网络作为特征提取器,对空间目标进行微调;采用对比学习或伪标签方法进行无监督训练;以及基于原型网络、模型正则化等方法进行少样本识别(Wang 等, 2021)。例如,使用合成图像进行模型预训练,再通过领域自适应方法对模型进行调优,有效缓解了仿真与真实图像之间的域差异问题。当前的非合作空间目标识别研究已从早期依赖显式建模与特征工程的方法,发展为融合多模态输入、自适应学习机制与高语义建模能力的深度识别框架,具备更强的适应性与推广性。

1.3 位姿估计方法的发展历程

非合作空间目标位姿估计技术的演进经历了从传统基于几何模型的计算机视觉方法向以深度学习为核心的智能感知方法的转变。早期研究主要依赖图像特征提取与几何建模,通过构建目标的 3D 先验模型,在 2D 图像中提取角点、边缘和纹理等低级视觉特征,并采用多点透视问题(perspective- n -point, PnP)、迭代最近点(iterative closest point, ICP)或高效多点透视问题(efficient perspective- n -point, EPnP)(Sharma 和 D'Amico, 2016)等几何解算算法实现姿态恢复,在目标纹理丰富、成像稳定时精度较高。但传统方法高度依赖手工特征和模型匹配,难以适应光照变化、遮挡和弱纹理等复杂环境,易出现估计失败或误差放大问题。

传统方法主要可分为基于点云配准的方法与基于特征提取的方法两类。

基于点云配准的方法,如 Li 等人(2021)提出的共形几何代数(conformal geometric algebra, CGA)跨源点云融合算法,融合 LiDAR(light detection and ranging)与 Kinect 数据并结合卡尔曼滤波优化位姿;王世昌等人(2024)结合截断最小二乘估计与半正定松弛(truncated least squares estimation and semidefinite relaxation, TEASER)与 ICP 算法,通过粗配准与精配准提升精度,但依赖调参且抗噪性差。

基于特征提取的方法,如 Hu 等人(2023)融合双目与飞行时间(time of flight, TOF)相机(孙典奇 等, 2021)数据,通过联合标定与 ICP 配准位姿,但依赖初始值;束安等人(2021)采用三目视觉与 ORB(oriented fast and rotated brief)特征匹配,结合随机采样一致性(random sample consensus, RANSAC)解算位

姿,但极端光照下鲁棒性不足。

此外,Zhu等人(2022)改进的距离加权迭代最近点算法(distance-weighted iterative closest point, DWICP)、姜丽辉等人(2022)的TOF相机位姿解算等,均因计算复杂度高或对特征完整性要求高而存在局限。在此基础上,一些研究开始探索多传感器融合与滤波方法以提升稳健性,庄海孝等人(2024)设计了结合单目视觉与激光雷达的随机点采样(random point sampling, RPS)系统,通过多传感器融合实现目标距离、方位及三维成像测量,采用特征匹配与位姿解算算法,结合深度处理单元(deep processing unit, DPU)和可变形部件模型(deformable part model, DPM)构建硬件,但该系统对目标特征依赖性强,面对非合作目标或复杂几何体时特征提取稳定性易降低,且激光雷达单点测距在动态场景下空间覆盖有限。单目视觉与滤波融合方法也被提出用于增强估计鲁棒性,Jin等人(2020)提出融合扩展卡尔曼滤波器(extended Kalman filter, EKF)与扩展卡尔曼粒子滤波器(extended Kalman particle filter, EKFPF)的单目视觉姿态估计方法,通过建立相对运动模型、以特征点平面坐标为观测值构建量测方程,采用EKFPF实现位姿估计,该方法依赖特征点稳定匹配。Deng等人(2022b)提出基于单目序列图像的运动估计方法,通过特征点匹配和极线约束分解本质矩阵实现旋转姿态估计与相对运动参数推导,但平移估计缺乏尺度信息,需依赖外部传感器初始化,且特征点提取匹配精度影响算法性能,在快速旋转或远距离观测时易失效。

随着深度学习,尤其是卷积神经网络的快速发展,研究者开始尝试将数据驱动的方法引入位姿估计任务。这一阶段的典型方法多以端到端回归方式展开,通过CNN对图像进行全局特征提取,直接输出目标的旋转向量、四元数或欧拉角等姿态参数。例如PoseNet系列(Jo和Kim,2022)方法在地面场景的初步尝试,展示了深度学习在视觉位姿回归任务中的可行性。随着网络结构优化与损失函数设计的逐步改进,该方法被进一步迁移并应用于航天器等非合作空间目标的姿态估计中。在此基础上,关键点检测与姿态优化相结合的范式逐渐成为主流。该类方法首先通过网络检测目标表面或组件上的语义关键点,随后基于这些关键点与3D模型之间的空间约束关系,通过EPnP或最小二乘优化算法(王世昌

等,2024)进行姿态解算,显著提升了估计精度和鲁棒性,特别适用于结构清晰、形态稳定的空间目标。近年来,随着深度学习模型结构的进一步演进以及跨模态感知需求的提升,非合作空间目标位姿估计进入了多模态融合与结构感知驱动的发展阶段。研究者广泛引入GAN和Transformer结构等新型网络架构,在保持端到端训练优势的同时,增强了模型对图像上下文信息、结构拓扑关系以及跨尺度特征的建模能力。多模态数据的融合成为趋势,如结合RGB图像与深度图、红外图像,通过特征级融合或决策级融合方式,实现对目标更全面、鲁棒的感知。此外,考虑到真实空间任务中高质量标签数据的获取成本极高,越来越多研究引入无监督学习、合成数据训练与领域自适应策略,如利用仿真图像进行预训练,并通过对抗损失或风格迁移方法缩小“模拟—现实”域差异,从而增强模型的泛化能力。部分研究还探索了少样本学习、小样本适配以及动态模板匹配等方法,以适应空间目标种类多样、变化频繁的任务需求。

综上,非合作空间目标位姿估计技术正由传统几何方法向深度学习驱动的智能感知系统演进。当前研究重点已从特征提取与模型匹配,转向多模态融合、无监督学习与结构感知等方向,展现出更强的鲁棒性与泛化能力。未来,该领域将持续向高精度、强鲁棒、轻量化和实时性的综合目标迈进,为空间交会、自主服务等关键任务提供核心支撑。

2 基于视觉与深度学习的非合作空间目标识别方法

基于视觉与深度学习的空间目标识别方法主要分为多模态视觉融合、目标检测与分割、迁移学习与少样本学习3类。这些方法主要区别在于信息表达形式、学习范式及样本依赖程度等方面的不同。多模态视觉融合方法通过结合RGB、深度和红外等多源传感器信息,丰富目标特征表达,提升在复杂环境中的识别精度与鲁棒性;目标检测与分割方法侧重于实现目标的精确定位与轮廓提取,是构建高精度识别系统的基础;迁移学习与少样本学习方法则面向样本获取困难的实际问题,通过知识迁移或有限样本学习实现对新类别的识别,具有良好的泛化能力与实用价值。

2.1 多模态视觉融合

基于多模态视觉融合的空间目标识别方法通过融合来自可见光、红外、深度图像和雷达等多种传感器的数据,充分利用不同模态之间的互补性,有效提升目标识别的准确性与鲁棒性。该方法能够缓解单一模态在复杂背景、低光照或遮挡情况下的性能下降问题。通过特征对齐、跨模态信息融合及深度学习模型训练,实现对空间目标的稳定识别,表2为多模态视觉融合的非合作空间目标识别方法。

Zhang等人(2024a)针对多卫星协同观测系统的非合作空间目标识别任务,提出两阶段识别模型LCNN-DMT (lighter convolutional neural network model with distance merge threshold),阶段1基于MBRT-D (minimum bounding rectangle with threshold for distance merging)融合最小外接矩形与距离阈值合并;阶段2基于改进的SqueezeNet,引入可分离卷积+高效通道注意力机制(efficient channel attention, ECA)分类卫星与碎片,最终解决多卫星协同观测场景下的宽视场图像拼接问题,但在极端噪声环境下识别精度会下降。

而对于多模态数据特征对齐难题和远距离小目标识别精度不足问题,Wang等人(2022)提出基于多模态特征融合的渐进式注意力网络,采用双流CNN-Transformer混合架构,引入跨模态交叉注意力模块。

Aldahoul等人(2022b)提出了一种多模态深度学习方案,结合了RGB图像和深度图像的信息,使用了预训练的Vision Transformer模型和端到端CNN网络,通过全连接层后对空间目标直接分类。

此外,Chen等人(2021)提出了一种基于主动光

源检测与多源数据融合的非合作空间目标抓取位置识别方法,在高亮度环境下采用ResNet处理RGB-D多源数据,在低亮度环境下使用改进的PointNet++网络处理点云数据,运用最远点采样(farthest point sampling, FPS)算法进行点云特征提取与分类,融合二维图像与三维点云的优势,在复杂空间场景中保持毫米级定位精度。

本节综述了基于多模态视觉融合的空间目标识别方法,重点介绍了融合可见光、红外、深度图像和雷达等多源传感数据,以提升识别准确性与鲁棒性的策略。该方法充分利用模态互补性,通过特征对齐、跨模态信息融合及注意力机制,实现复杂背景下目标的稳定识别。多模态视觉融合算法通过整合多源传感器数据提升空间目标识别性能。从识别精度看,LCNN-DMT模型在场景数据集上,mAP50 (mean average precision)达97.5%,分类准确率98.83%,CNN-Transformer模型识别准确率为94.7%,ResNet和改进的PointNet++在合成数据集上分割平均交并比(mean intersection over union, mIoU)达93.2%。实时性方面,CNN-Transformer模型仅12.3帧/s,实时性较差,LCNN-DMT模型对比YOLOv8n(you only look once)提升速度17.6%,实时性更优。在适应性上,LCNN-DMT模型在极端噪声环境下识别精度下降,ResNet和改进的PointNet++未测试真实数据,CNN-Transformer模型可解决远距离小目标识别精度不足问题。整体而言,多模态视觉融合通过特征对齐与注意力机制,在复杂场景中展现出较强优势,但在极端环境适应性与实时性上仍有提升空间。

表2 多模态视觉融合方法
Table 2 Multimodal vision fusion methods

文献	网络模型	改进点	数据集	实验结果	局限
Zhang等人(2024a)	LCNN-DMT	距离合并阈值替代传统IoU阈值	自建空间目标数据集/SPARK数据集	场景数据集mAP50达97.5%,对比YOLOv8n提升速度17.6%,分类准确率为98.83%	成像差异过大,会影响识别精度
Wang等人(2022)	CNN-Transformer	设计跨模态可变形注意力模块	SPEED	参数量为48.7 M, RMSE = 1.53°, Accuracy = 94.7%, mAP50 = 82.1%	实时性较低,仅12.3帧/s
Chen等人(2021)	ResNet+改进的PointNet++	通过亮度阈值动态选择网络分支,克服单一模态在光照变化下的性能退化问题	合成数据集	mIoU为93.2%	未测试真实数据

2.2 目标检测与分割

基于视觉与深度学习的空间目标识别中,目标检测与分割方法不断演进以适应复杂空间环境。

表3展示了多种轻量化网络结构与注意力机制,如迭代交叉注意力特征融合(iterative cross-attention feature fusion, IC-AFF)、双向特征金字塔网络(bidirectional feature pyramid network, BiFPN)等以提高模型精度与效率。YOLO、单阶段多框检测器(single shot multibox detector, SSD)及其改进版本广泛应用于快速检测,结合Transformer、ResNet和DenseNet等结构增强特征提取能力。分割方面,Mask R-CNN、快速网络特征神经网络(fast grid-based neural network, FGBNN)等模型通过多尺度融合与上下文增强,有效提升空间目标的分割精度。部分研究结合点云与图像信息,实现对小目标及弱纹理目标的精细识别。

Wang等人(2025)提出了一种轻量级显著性检测模型,该模型结合Lightweight Feature Extractor、特征融合模块IC-AFF和部分一整体验证模块(part-whole verification module, PWV),参数量仅1.99 M,通过注意力机制与多尺度特征提取应对空间复杂的成像条件。

进一步地,针对多模态信息融合问题,Pensado等人(2023)提出基于多层Transformer的双分支RGB-D显著目标检测网络,采用并行编码器处理RGB和深度模态,通过跨模态互补模块(cross-modal complementary module, CMCM)、空间注意力机制(spatial attention module, SAM)和边界细化模块(boundary refinement module, BRM)实现多层次特征融合。

针对三维点云数据处理中的数据稀疏与密度不均问题,Zhao等人(2022)改进了PointNet++,增加了几何感知的特征提取(feature extraction, FE)层和新损失函数。

Li等人(2022c)提出了一种新型的非合作航天器目标检测方法SCNN-lite(lite model of spacecraft convolutional neural network),基于深度卷积神经网络(deep convolutional neural network, DCNN)框架,主要采用了深度可分离卷积结合残差网络模块改进网络的主干结构,还使用了基于K-means聚类方法,参数量比YOLOv3减少近50%,表现出高效的检测性能与轻量化优势。

在小目标和远距离目标检测方面,徐云飞等人

(2020)在SSD架构基础上构建了NCDN(non-cooperative detection network)网络,融合MobileNetV2以适应尺度变化与成像退化场景,但在距离大于40 m时仍面临识别准确率下降的问题。

为增强空间目标的细粒度识别能力,Yang等人(2020b)针对空间目标识别任务,提出了一种混合卷积神经网络HCNN-PSI(hybrid convolutional neural network with partial semantic information),结合全局特征与局部语义信息增强识别精度。该方法通过三阶段框架实现:目标检测阶段采用改进的最小包围矩形阈值法(minimum bounding rectangle with threshold, MBRT),通过合并交并比大于0.5的候选矩形框,解决深空图像中卫星组件因像素不连续导致的误分割问题,结合ResNet-50特征提取器,在仿真数据集I中实现99.94%的检测准确率;部件分割阶段利用Mask R-CNN对卫星主体结构进行像素级分割;多源融合识别阶段将全局图像与关键部件特征通过权重融合,强化决定性部件的分类贡献,使得细粒度识别准确率提升至99.70%。

面向空间碎片检测的挑战,Xiang等人(2020)提出了一种基于快速网格特征神经网络(FGBNN)的空间碎片检测方法,主要解决了传统算法在低信噪比下难以快速检测微小空间碎片的问题,核心创新在于将目标检测转换为网格级回归问题,通过参数共享和简化特征提取网络,显著提升了运算效率。

Chen等人(2020)则针对复杂姿态与光照条件下卫星部件检测困难的问题,提出融合DenseNet与ResNet-FPN(residual network-feature pyramid network)的RSD(regional-based convolutional neural network based satellite-component-detection model)模型,提升多尺度特征融合效果,实现了太阳能板与天线等部件的像素级分割,性能优于原始Mask R-CNN。

Cao等人(2024)提出了一种基于硬件在环采集的半物理仿真系统构建的航天器关键部件检测与分割数据集,通过对比实验选择了高分辨率网络(high-resolution network, HRNet)作为骨干网络,分别优化了检测模型Cascade R-CNN和分割模型FCN(fully convolutional networks),最终建立了检测任务的基准模型Cascade R-CNN + HRNet和分割任务的基准模型FCN + HRNet,平均精度提升2.42%,mIoU达0.877。

结合多阶段识别流程,Aldahoul等人(2022a)引入EfficientDet与EfficientNet-B4联合架构,实现空间

表3 目标检测与分割识别方法
Table 3 Object detection and segmentation recognition methods

文献	网络模型	改进点	数据集	实验结果	局限
Li 等人(2022c)	SCNN-lite	DSC 模块显著降低参数量	自建航天器仿真数据集	参数量为 28 M , mAP = 93.28%	当高斯噪声方差超过 2 时,检测精度显著下降
郝强等人(2022)	改进的 YOLOv3	引入 Bottleneck 模块,通过深度可分离卷积降低计算复杂度	CGRO/GPS quadsat	参数量为 105 M , mAP = 91.1%	未分析模型对动态空间碎片的泛化能力
Liu 等人(2022b)	YOLOv5-GMB	引入 MHSA 提高特征捕获,以及 BiFPN 进行多尺度特征融合	自建数据集	mAP = 81%	在极端复杂环境性能受限
Ai 等人(2022)	改进的 YOLOX_L	TOOD 头替代 YOLOX 的解耦头,通过任务对齐学习 TAL 提升分类与定位协同性	自建航天数据集	参数量为 270.9 M , mAP = 96.28%	对复杂场景中极小目标检测效果有限
Cosmas 和 Kenichi(2020)	CNN + YOLOv3	首次在航天器位姿估计中实现 CNN 的 FPGA 在轨推理,功耗仅约 1.32 W	SPEED	RMSE = 1.98	未测试真实数据
徐云飞等人(2020)	NCDN	轻量化,通道压缩策略节省 75% 计算量,精度损失仅 5%	SPEED	40 m 内, mAP = 90%	40 m 以上时检测精度明显下降
Li 等人(2023)	基于 YOLOv5 的 YSCRM 模型	在 YOLOv5 Neck 末端引入自注意力机制,增强特征表示潜力	自建数据集	mAP50 = 91.3%	未说明在极端光照条件下的鲁棒性
Yang 等人(2020b)	MBRT + Mask R-CNN	使用 MBRT 定位,改进传统最小包围矩形 MBR,通过交并比 IoU 阈值合并邻接框,提升小目标检测性能	合成数据集/ BUAA-SID 1.0	Top-1 分类准确率 = 99.94%;	计算效率低,混合网络参数多,训练速度为纯全局特征方法的 50%
Zhao 等人(2022)	改进的 Point-Net++	多源数据训练使 mIoU 提升 0.3%~0.5%;配准后特殊部件分割精度达 89.2%	自建多源 3D 航天器部件分割数据集	参数量为 20.8 M , 3DSatNet 达 84.7%,	数据集规模受限
Wang 等人(2025)	LFE + IC-AFF + PWV	ILBlock 与 IFBlock 的组合减少计算冗余,提升特征提取效率。	SwissCube	参数量为 1.99 M , Max-F = 0.956 4, MAE = 0.032, E-measure = 0.946	实验场景单一
Cao 等人(2024)	Cascade R-CNN + HRNet + FCN	针对小目标问题,通过 HRNet backbone 保持高分辨率特征图	Spacecraft-DS	AP = 82.42%, mIoU = 0.877	极端光照下小目标识别精度仍不足
Pensado 等人(2023)	多层 Trans-former	提出 CMCM、SAM 和 BRM 模块实现多层次特征融合	NJUD(1 985)、NLPR(1 000)、SIP(929)	参数量为 18.5 M , s_{α} = 0.923	对深度图质量敏感
Xiang 等人(2020)	FGBNN	通过稀疏连接和参数共享降低计算量	合成数据集	Acc = 84%, Recall = 80%, F1-score = 82%	未针对不同尺寸碎片设计多尺度检测策略
Chen 等人(2020)	RSD	融合 ResNet-FPN 与 DenseNet,增强多尺度特征融合	合成数据集	整体 F1-score 为 93%	未在真实场景图像中验证
AlDahoul 等人(2022a)	EfficientDet + EfficientNet-B4	通过结合检测模型和分类模型提升对模糊/噪声图像分类准确率	SPARK	Acc = 94%, F1-score = 86%, F2-score = 98.77%	特定类别分类效果差

目标的检测、定位与二次分类的分工协同,有效提升整体识别精度。De Vittori等人(2022)则提出了一种基于 U-Net 深度神经网络(deep neural network, DNN)(Szczuko, 2019)的创新方法,用于从望远镜观测图像中实时提取空间碎片的轨迹,但存在对复杂真实场景适应性不足问题。

为解决传统阈值方法在复杂背景下稳定性不足的问题, Lin 等人(2023)提出将全局检测任务转化为局部二分类问题,结合背景去除、区域提议生成与引导信息模块设计的 SF-CNN,实现更高鲁棒性的小目标检测。该方法首先通过背景估计与去除技术预处理图像,然后利用局部极大值点生成多尺度区域提议,最后设计了一个轻量化的 SF-CNN 网络结合引导信息模块和偏置损失来增强网络对目标分布特征的提取能力。

在 YOLO 系列改进方面, Liu 等人(2022b)提出了一种基于 YOLOv5 改进的轻量化航天器组件检测方法 YOLOv5-GMB,通过引入 Ghost 模块和通道压缩技术降低网络参数量,同时提升计算效率,针对航天器图像受光照变化、地球背景干扰等问题,采用多头注意力机制(multi-head self-attention, MHSA)增强全局特征提取能力,并结合加权双向特征金字塔网络(BiFPN)实现多尺度特征融合,显著提升了小目标检测精度,使 mAP 达到 0.81,在 CPU/GPU 上的检测速度分别达到 16/53 帧/s。

郝强等人(2022)提出了一种轻量化的 YOLOv3 改进版本,通过结合 MobileNetv2 和 Res2Net 模块显著降低模型复杂度,参数量减少 55.5%,提升目标检测速度和小目标敏感度,适用于嵌入式平台上的非合作空间目标部件识别场景。

Ai 等人(2022)通过优化特征金字塔(feature pyramid network, FPN)结构和解耦预测网络,结合 YOLOF 和任务对齐头(task-aligned one-stage object detection, TOOD),提出了改进的 YOLOX_L 网络,在自建的航天数据集上 mAP 达到 96.28%。

此外, Cosmas 和 Kenichi(2020)利用基于 CNN 的编码器—解码器架构进行航天器姿态估计中的二维关键点定位,并通过 YOLOv3 实现航天器的检测与图像裁剪以提升定位精度。

Li 等人(2023)则基于 YOLOv5 框架,通过引入特征融合层和选择性核网络增强多模态特征选择能力,加入 Transformer 编码器模块挖掘自注意力机制

的预测潜力,并采用通道稀疏训练方法实现模型轻量化,最终提出了一种用于卫星部件识别的改进轻量级模型 YSCRM(YOLOv5-based satellite components recognition model),改进后的 YSCRM 在卫星部件识别任务中 mAP 达到 84.3%,较基准模型提升 2.3%,且模型尺寸显著减小。

本节综述了当前空间目标识别中基于视觉与深度学习的目标检测与分类方法的研究进展。该方法侧重于在复杂背景下准确定位和识别目标, YOLO 系列算法如 YOLOv5-GMB 引入 Ghost 模块和 MHSA,提升小目标检测精度至 mAP 为 0.81 且检测速度达 16/53 帧/s;改进的 PointNet++ 增加几何感知特征提取层,分割精度平均 IoU 达 93.2%; EfficientDet + EfficientNet-B4 在 SPARK 数据集准确率达 94%。这些方法在精度和速度上各有优势,但对极端环境适应性和动态目标泛化能力有待提升。

2.3 迁移学习与少样本学习

基于迁移学习与少样本学习的空间目标识别方法旨在解决空间目标图像样本稀缺、标注成本高的问题。如表 4 所示,迁移学习通过将已有大规模数据训练的模型知识迁移到空间目标识别任务中,提升模型在小样本条件下的识别能力;少样本学习则通过度量学习、元学习等方法,从少量标注样本中快速泛化识别能力。两者结合可有效提高模型对新目标类别的适应性,广泛应用于空间碎片识别、在轨识别和小型航天器识别等实际场景中。

Deng 等人(2025)提出了一种基于光流对齐的多帧超分辨率重建算法,结合离散小波变换(discrete wavelet transform, DWT)和 CNN 提取空间与频率域特征,再通过空间金字塔网络(spatial pyramid network, SpyNet)生成三级光流金字塔进行多尺度特征对齐,自适应融合时序特征并重建高分辨率图像。

在样本稀缺的条件下, Zhang 等人(2024b)提出了一种基于迁移学习的框架,通过预训练的 ResNet-50 网络提取微多普勒特征,结合局部最大均值差异(local maximum mean discrepancy, LMMD)进行域适应,对齐源域与目标域的分布差异,在实验数据极少,提取非合作目标的微多普勒特征并完成姿态分类。

进一步地, Yang 等人(2020a)提出了一种称为 D2N4 的方法,用于少样本空间目标识别,该方法结

合了深度嵌入模块和 Image-to-Class 度量模块,并引入了类内紧凑性损失以减少类内差异,同时使用全局池化信息降低背景噪声的干扰。

此外,针对小样本问题,Zeng 和 Xia(2017)提出了一种基于数据增强的深度学习框架解决空间目标识别中的小样本问题。该方法设计了9层深度卷积神经网络 DCNN,通过融合旋转、裁剪、透视变换和噪声注入等多级数据增强技术,在训练阶段生成24倍扩增数据,并在测试阶段引入12倍增强数据,结合 softmax 分数聚合策略,该方法在400幅合成空间目标图像上实现了99.9%的分类准确率。

本节介绍基于迁移学习与少样本学习的空间目

标识别方法,以解决样本稀缺与标注成本高的问题。相关算法聚焦样本稀缺难题:迁移学习方法通过预训练网络提取特征并结合域适应技术,对齐不同域分布,提升新目标姿态分类效果;少样本学习方法如 D2N4,利用深度嵌入、特定度量模块及损失函数减少类内差异,增强识别能力;数据增强方法经多阶技术扩充数据,结合 softmax 分数聚合,分类准确率达99.9%。这些算法通过迁移大规模训练模型知识提升识别效率与泛化能力,结合元学习、图像增强和多分支网络等策略,在少样本条件下实现空间目标准确识别,已在在轨识别、小型航天器识别等场景中取得显著效果。

表4 迁移学习与少样本学习识别方法
Table 4 Identification methods of transfer learning and few-shot learning

文献	网络模型	改进点	数据集	实验结果	局限
Zhang 等人 (2024b)	ResNet-50 + LMMD	结合预训练(模拟数据)与局部域适应(LMMD提升细粒度对齐),解决数据分布差异	自建数据集	形状分类准确率为95.26%,姿态角识别79.02%	实验数据量较少
Yang 等人 (2020a)	D2N4	首次在少样本度量学习中引入中心损失,显著提升特征判别性	BUAA-SID-Share 1.0	Top-1 准确率:1-shot 93.08%,5-shot 97.85%	未测试动态环境
Zeng 和 Xia (2017)	DCNN	将数据增强同时应用到训练和测试数据,以及选择最优的组合变换	合成数据集	分类准确率 = 99.90%	未验证真实空间图像的泛化能力
Deng 等人 (2025)	MWCNN + SpyNet + FG-DCN	MWCNN 结合 DWT 提升细节保留能力,SpyNet 光流金字塔优化多尺度运动补偿	LTSPEED+	SSIM:0.912,PSNR:32.6 dB,LPIPS:0.152	依赖固定变换生成的合成数据

2.4 小结

本节主要分析了3类典型的空间目标识别方法:多模态视觉融合、目标检测与分割以及迁移学习和少样本学习。

表5对本节所讨论的非合作空间目标识别方法在典型的公开数据集和自建数据集上,从参数量、推理速度、准确率和 mAP 等维度进行了算法性能比较。例如,多模态视觉融合类的 LCNN-DMT,参数量仅0.928 M,准确率和 mAP 分别达98.83%、97.5%;目标检测与分割类的 HCNN-PSI 准确率高达99.7%,展现出强劲的识别精度;迁移学习与少样本学习类的 ResNet-50 + LMMD,在自建数据集上准确率也达到95.26%,体现出小样本场景下的良好性能。

表6进一步总结了这3类方法的核心原理与适用场景。多模态视觉融合方法通过集成 RGB 图像、激光雷达和红外成像等多源数据,在结构复杂或纹理缺失的场景中有效增强特征表达能力,提升目标

类别与部件识别的准确性与稳定性;目标检测与分割方法则以 CNN、YOLO 和 Mask R-CNN 等深度网络为基础,实现了高效的目标定位与轮廓提取,适用于多类别目标快速识别与结构判定任务;迁移学习与少样本学习方法则缓解了训练样本不足的问题,通过模型预训练、特征迁移或元学习框架实现对新类别或变化目标的识别能力提升,在实际空间任务中具备良好的实用性与扩展性。综上,空间目标识别方法正朝着多源融合、轻量高效与小样本适应方向发展,进一步推动了非合作目标从基础识别向高层语义解析转变,为后续任务提供了关键支撑。

3 基于视觉与深度学习的非合作空间目标位姿估计方法

当前基于视觉与深度学习的非合作空间目标位姿估计方法可大致划分为4类:直接回归方法、关键

表 5 目标识别算法性能比较
Table 5 Comparison of target recognition algorithm performance

类别	方法	设备	参数量	推理速度/(帧/s)	准确率/%	mAP/%	数据集
多模态视觉融合	LCNN-DMT(Zhang 等,2024a)	Intel Core i7-13700KF CPU, NVIDIA RTX 4090 GPU	0.928 M	163.9	98.83	97.5	SPARK
	CNN-Transformer (Wang 等,2022)	NVIDIA RTX2080Ti GPU, Intel i7-7700 k CPU	48.7 M	6.67	94.7	82.1	SPEED
目标检测与分割	NCDN(徐云飞 等, 2020)	Interl(R) Xeon(R) Silver 4 114 CPU@220 GHz, NVIDIA Quadro P4000 GPU	2.2 M	—	—	90	SPEED
	YOLOv5-GMB (Liu 等,2022b)	CPU AMD R7 5800 H, GPU NVIDIA RTX 3070	1 M	53	95	81	SPEED
	HCNN-PSI(Yang 等,2020b)	NVIDIA GTX1080Ti GPU , Intel Xeon E5-2609 CPU	25.5 M	—	99.7	—	BUAA-SID-1.0
	SCNN-lite(Li 等, 2022c)	Intel Core i7 9700 Processor 3.0 GHz, NVIDIA 2070s GPU	0.028 M	5.99	—	93.28	自建数据集
	改进的 YOLOv3 (郝强 等,2022)	Ubuntu16.04.6, NVIDIA GeForce RTX 2080Ti GPU	—	65	—	91.1	自建数据集
	改进的 YOLOX_L (Ai 等,2022)	i9@9900K CPU, 64 GB RAM, NVIDIA RTX 2080Ti GPU	—	50	96.28	88.86	自建数据集
迁移学习与少样本学习	D2N4(Yang 等, 2020a)	Intel Core i7-6700 CPU, NVIDIA GeForce GTX 1080Ti GPU	—	—	93.06	—	BUAA-SID-1.0
	ResNet-50+LMMD (Zhang 等,2024b)	—	—	—	95.26	—	自建数据集

注:“—”表示该项内容为空。

表 6 基于视觉与深度学习的空间目标识别方法对比
Table 6 Comparison of spatial object recognition methods based on vision and deep learning

方法类别	原理	优势	局限性	适用性
多模态视觉融合方法	集成 RGB、深度图、红外和激光雷达等多源数据,利用特征对齐与跨模态融合增强目标表征能力	融合多源信息增强特征表达,适应复杂结构、遮挡及光照变化	数据融合复杂,传感器依赖高,异常模态数据可能干扰结果	复杂结构识别、部件识别、多源遥感融合应用
目标检测与分割方法	基于 YOLO、R-CNN 等网络,通过目标定位、语义分割等实现多尺度特征融合	可定位与分割目标轮廓,适配多类别识别,精度较高	对遮挡、弱纹理敏感,训练依赖大量标注数据	舱段识别、空间碎片分类等目标明显的识别场景
迁移学习与少样本学习方法	结合元学习、数据增强等策略,从少量样本中学习新类别特征	缓解样本不足问题,适应新目标,训练成本低	对源目标领域依赖强,泛化性与适应机制设计要求高	快速部署、新目标扩展、数据稀缺情况下的识别任务

点检测方法、无监督及域自适应方法和多模态视觉融合方法。这些方法主要是在任务建模思路、监督策略以及信息利用方式等方面存在本质差异。直接回归方法将位姿估计建模为从图像到姿态参数的端到端回归问题,强调特征到位姿的直接映射,具有结

构简洁、计算高效等优点;关键点检测方法通过先定位图像中的关键点,再结合几何模型进行姿态求解,强调显式结构建模与几何约束的结合,适用于具有明确结构特征的目标;无监督及域自适应方法则致力于减少对真实姿态标签的依赖或解决跨域泛化问

题,通常通过无监督、对比学习和图像重建损失等策略实现模型训练,提升其在现实复杂环境下的适应能力;多模态视觉融合方法充分利用多源信息,通过信息互补提升位姿估计的鲁棒性与准确性,特别适用于光照变化大、纹理弱或信息缺失的场景。

3.1 直接回归

在非合作空间目标位姿估计中,直接回归类方法是指通过端到端的深度学习模型,直接从输入数据,如图像、点云等中预测目标的6自由度位姿参数而无需依赖中间步骤,如关键点检测或特征匹配。这类方法的核心目标是通过数据驱动的方式隐式学习输入与位姿之间的复杂映射关系,尤其适用于缺乏先验模型或环境高度动态的场景。表7为基于直接回归的非合作空间目标位姿估计方法。

Xiao等人(2024)提出的LSTM方法解决了单星观测下非合作目标主动段参数估计的模型精度问题,其三维位置估计误差相比时序融合与对齐校正(temporal integration and alignment correction, TIAC)算法降低65.8%,通过数据驱动方法降低对先验信息的低依赖性,但其依赖仿真生成数据会导致在实际场景中泛化性不足。

Renaut等人(2025b)将LiDAR点云投影至 224×224 像素深度图像,采用MobileNetV3-Large作为CNN主干,结合双预测头进行位姿联合估计,提出基于螺旋采样的姿态空间离散化方法,引入对称距离函数 d_{sym} 与高斯权重解码处理六边形对称航天器的姿态模糊性问题,结合改进的法线分布变换(normal distributions transform, NDT)算法进行位姿细化,替代传统ICP提高效率。

进一步地,Renaut等人(2025a)提出基于点云深度学习的航天器快速位姿估计方法,采用优化的PointNet++网络处理3D激光雷达点云,单阶段推理输出相对位姿,姿态分类模块用对称距离函数处理 120° 旋转对称性,以合成数据训练。经真实数据验证,星载计算机上平均推理速度为420 ms,位置误差为6.63 cm、姿态误差为 0.73° 。

另一方面,Ye等人(2023)提出了一种基于双通道Transformer的非合作空间目标位姿估计方法,主干网络采用EfficientNet-B0提取特征,平移通道回归空间目标的三维平移量,旋转通道回归四元数形式的旋转量,端到端网络高效输出位姿参数,但未验证真实环境适应性。

Yang等人(2024)提出了PVSPE(multi-task pyramid transformer spacecraft pose estimation network)端到端多任务Transformer网络,主要由EnPVT(enhanced pyramid vision transformer)和SlimGFPN(slimmed generalized feature pyramid network)两个模块构成,EnPVT结合了全局空间和通道注意力,SlimGFPN用于更有效地融合特征。此外,采用Matrix Fisher和多元高斯分布建模姿态回归的不确定性,提升准确性。

Qiao等人(2024)则提出了一种名为DSOAE-Net(directly space-object-attitude-estimation network)的深度学习网络,采用CNN或Transformer提取图像全局特征,直接回归非合作空间目标的实时三维旋转姿态。

针对关键点回归和姿态求解问题,Huang等人(2023)提出端到端CNN架构,主干使用HRNet-w48,采用多分支独立结构独立回归每个2D关键点提升精度,可微位姿估计子网络代替PnP求解器,实现端到端训练。

此外,为解决传统方法依赖多站雷达或稀疏数据融合,精度低于光学图像问题,Hou等人(2023)使用UNet++进行ISAR图像增强,Swin Transformer直接从增强后的图像中回归目标的三轴姿态角。

为解决传统方法需要轨道机动或多卫星协同问题,Gong等人(2023)提出了基于DNN的线性相对运动模型,该模型直接用DNN从光学相机的视线角输入回归目标相对姿态。

在网络架构创新方面,Gao等人(2023)基于U-Net架构,引入了密集残差块和多头自注意力块,提出了SU-Net(spacecraft U-shaped network)架构,旨在解决航天器ISAR成像中的特征稀疏、噪声大和数据量少的问题。Cao和Cheng(2023)首次将EfficientNet-B0与FPN结合,通过多尺度特征融合将远距离目标的平均位姿估计精度提升7.6%。

此外,Huang等人(2022)设计了一个改进的端到端姿态估计网络URSONet-Improve,基于ResNet主干,融合自适应双通道特征提取模块和双重注意力机制,提升目标特征的鲁棒性和准确性。

针对模型轻量化和计算效率优化,刘佳辉等人(2024)设计了LSPENet(lightweight space pose estimation network),结合深度可分离卷积和ECA通道注意力机制,分别回归三维平移和四元数旋转,采用

表 7 直接回归位姿估计方法
Table 7 Direct regression pose estimation methods

文献	网络模型	改进点	数据集	实验结果	局限
Renaut 等人 (2025b)	MobileNetV3-Large + 螺旋采样姿态空间离散化 + 改进 NDT	MobileNetV3-Large 在准确率 (97.9%) 与推理时间 (737 ms) 之间取得均衡	EPOS	在 EPOS 数据集上平均成功率 97.9%, 位置/姿态平均误差为 6.64 cm/0.77°; 合成测试集成功率为 99.24%	丢失 3D 空间信息特征, 对稀疏点云的适应性受限
Deng 等人 (2022a)	YOLOv5+ scSE-LHRNet	通过 DSC 和 scSE 模块实现了参数量降低至 11.82 M, 同时保持高分辨率特征	SPEED	参数量为 11.82 M, AP = 94.22%、mIoU = 86.61% 综合误差 = 0.016 9	在复杂背景和弱光条件下, 姿态预测存在偏差
Ye 等人 (2023)	双通道 Transformer	双通道设计, 解耦平移与旋转信息, 减少参数耦合	SPEED	平均姿态误差为 1.573°, 位置误差为 0.044 m, 参数量为 186 M	真实场景泛化性未充分验证
Yang 等人 (2024)	基于 Transformer 的多任务网络 PVSPE	SlimGFPN 减少跨层连接, Epose 降低 23.7%	SPEED+	在 lightbox 和 sunlamp 测试域, Epose 分别达 0.101 和 0.178, 参数量为 70.15 M	对极端反光场景仍存在误判
Huang 等人 (2021)	主干网络 ResNet-50 结合 SE 模块 + 3 个并行子网络	四元数软分类 + 误差回归, 缓解离散化空间大范围分类的误差	SPEED/URSO	URSO 数据集: 姿态误差 ≤ 5° (近距离)/20° (远距离), 位置误差 ≤ 0.25 m	真实数据泛化性受限
Qiao 等人 (2024)	DSOAE-Net	多模态 R6D 表示, 解决对称目标的多峰映射问题	SPEED/SPEED+	旋转矩阵误差为 2.14°	在真实卫星图像中性能下降
Huang 等人 (2023)	端到端 CNN 架构	多分支关键点回归提升精度, 可微位姿估计子网络实现端到端训练, 以位姿误差优化	SPEED+/SwissCube	角度精度提升 52%	计算复杂度高
Phisannupawong 等人 (2020)	改进的 GoogLeNet	损失函数创新, 对比测试指数损失和加权欧氏损失, 后者效果更优	URSO	位置误差为 1.2 m; 姿态误差为 87.93%	姿态估计误差较大
窦凯云等人 (2023b)	ResNet-50 + PSA 模块	引入 PSA 模块, 通道与空间分支的极化滤波和 HDR 机制, 增强空间特征并抑制背景噪声	URSO	位置误差为 0.85 m, 姿态误差为 9.7°	未验证真实空间场景泛化性
Xiao 等人 (2024)	LSTM	提出面向时序数据的 LSTM 架构, 替代传统轨迹模板匹配	仿真数据集	位置误差为 358.8 m, 速度误差为 0.49 m/s, 高度误差为 8 084.8 m	真实数据不足
Hou 等人 (2023)	Swin Transformer	UNet++ + Swin Transformer 组合实现图像增强与姿态回归的端到端优化	5 000 组模拟 ISAR 图像	增强后姿态估计误差 < 1°	实验未在真实雷达系统中测试
Gong 等人 (2023)	DNN	一次前向传播即可输出相对状态, 适用于实时任务。	100 万样本的轨道场景	共椭圆轨道: 位置误差 < 9.7%, 速度误差 < 48.2%; 非共面轨道: 位置误差 < 9.9%	当仰角接近 0° 或方位角 ≈ 90° 时误差显著增加
Gao 等人 (2023)	SU-Net	通过密集连接与残差结构结合, 减少特征下采样损失	SpeedPlus 数据集	平均绝对误差为 0.282°, 标准差为 0.065°	未验证其泛化性
Cao 和 Cheng (2023)	EfficientNet-B0 + FPN	通过多级特征融合增强多尺度鲁棒性	SwissCube	参数量为 10.7 M, 平均精度为 50.4%	对低光照等极端场景的鲁棒性不足
Huang 等人 (2022)	URSONet-Improve	引入双通道特征提取与注意力机制, 提升姿态估计精度, 误差降低 18.1%	Tango 卫星图像	Esyn: 0.029 6; Ereal: 0.132 8; 姿态角误差: 4.74°	真实图像训练集仅 5 幅, 模型泛化能力不足

软分配编码技术将连续姿态离散化,有效减少75.5%参数量并提升特征表达能力。Deng等人(2022a)则将轻量化的YOLOv5s和HRNet融合,并引进scSE(spatial and channel squeeze and excitation)模块和DSC(depthwise separable convolutional)模块,构建了一个轻量化的scSE-LHRNet网络,实现了高效的非合作目标姿态估计,在资源受限的航天器场景中平衡了精度与实时性。

Huang等人(2021)提出一种基于CNN的非模型单目姿态估计网络,网络由3个并行子网络构成:1)姿态预测子网络,通过软分类与误差四元数回归预测相对姿态;2)位置回归子网络,通过直接回归预测相对位置;3)目标检测子网络,基于改进的SSD网络检测航天器并控制姿态输出。

此外,Phisannupawong等人(2020)构建了一个基于视觉的位姿估计模型,使用改进的预训练GoogLeNet,应用于非合作对接任务。采用Unreal Engine 4生成的Soyuz宇宙飞船数据集,通过CNN学习图像与6自由度参数之间的关联。

苏迪等人(2023)则提出基于CNN的两阶段非合作目标位姿估计方法,第1阶段结合目标检测模块和位置回归模块,利用ResNet50提取特征;第2阶段根据任务情况设计不同的姿态估计模型,绕飞时用分类加姿态求解器,接近时用直接回归单位四元数。

为了改进传统两阶段流程的缺陷,Hu等人(2020)提出一种单阶段的6D物体姿态估计方法,设计了一个基于CNN和多层感知器(multilayer perceptron, MLP)直接处理3D-2D对应关系的深度网络架构,采用端到端训练方式直接优化姿态估计目标。

窦凯云等人(2023b)创新性地提出了一种融合极化自注意力与残差网络的端到端位姿估计方法,通过极化自注意力(polarized self-attention, PSA)机制增强空间特征,软编码提升姿态精度,适用于复杂背景下的空间目标位姿估计,但需进一步验证真实场景适用性及降低计算复杂度。

Zhang等人(2024c)设计了多尺度注意力块(multi-scale attention block, MAB),结合多尺度卷积和通道注意力机制。然后基于MAB构建了密集多尺度注意力(dense multi-scale attention, DMA)骨干网络,采用密集连接结构。网络结构分为3个分支,分别处理平面位置、深度位置和姿态估计,对深度估

计设计了单独的分支和损失函数。此外,采用了基于欧拉角的软分类方法,将姿态回归转化为分类问题。

本节综述了基于深度学习的直接回归方法在非合作空间目标位姿估计中的应用,不同算法在网络结构和性能上各有特点,基于双通道Transformer的算法平均旋转误差 1.573° 、平移误差0.044 m,优势是端到端建模效率高,不足是真实场景泛化性待验证;改进EfficientNet-B0与FPN结合的算法平均精度50.4%,优势是多尺度特征融合,不足是对低光照场景鲁棒性欠缺;融入PSA模块的算法位置误差0.85 m、姿态误差 9.7° ,优势是增强空间特征抑制噪声,不足是未验证真实场景适用性;URSONet-Improve算法在Tango卫星图像上模拟误差0.029 6、真实误差0.132 8、姿态角误差 4.74° ,优势是引入注意力机制提升精度,不足是真实图像训练集仅5幅,泛化能力弱;结合YOLOv5与HRNet的算法姿态估计综合误差0.016 9,优势是轻量化平衡精度与实时性,不足是复杂背景下性能受限;采用ResNet-50结合SE模块的算法在URSO数据集近距离姿态误差 $\leq 5^{\circ}$ 、位置误差 ≤ 0.25 m,优势是多分支子网络提升精度,不足是真实数据泛化性受限。

3.2 关键点检测

在非合作空间目标位姿估计中,关键点检测法是一种通过识别目标表面的显著特征点,结合几何约束和优化算法来估计目标位置和姿态的技术,如表8所示。这类方法的核心是从图像或点云中检测关键点,并与目标的3D模型或先验信息匹配,最终通过数学优化求解位姿。

Zhong等人(2025)基于改进的CSPDarknet53架构提取航天器目标区域(region of interest, ROI),然后通过多分辨率并行CNN分类器,回归3D-bbox顶点投影坐标,PnP几何约束和像素重投影误差优化,输出精细化姿态,但针对域偏差和对称结构航天器的姿态估计存在多解问题。

为进一步提升关键点检测与姿态解算的准确性,Wang等人(2024b)以PVTv2-B1为骨干网络,结合多层反卷积和空间感知回归直接预测关键点坐标,通过几何约束生成可见表面伪标签,采用立体感知增强策略,结合边缘化采样一致性PnP算法(marginalized sample consensus perspective-n-point, MAGSAC PnP)求解姿态,并通过在线自训练优化跨域泛化。

针对关键点回归的精度问题,王梓等人(2024)

以 Lite-HRNet 为核心骨干网络,提升特征图分辨率为输入尺寸的 1/2,后续设计包含两个卷积层的热力图预测头,通过亚像素解码法实现超分辨率关键点预测,较传统逐像素极大值法量化误差降低,基于 EPnP + RANSAC 初值,结合 Ceres 库非线性优化提升位姿解算鲁棒性。

Chen 等人(2024)则基于 YOLO-w6-pose、YOLO-tiny-pose 等单阶段检测网络,结合 2D 关键点回归实现目标关键点检测,并利用共面特征构建单应性矩阵进行位姿分解,最后通过非线性优化进一步提升位姿估计精度。

为解决 CNN 在平移等变性和长程依赖方面的不足,Ahmed 等人(2024)采用 Transformer 网络替换传统的 CNN 以提高目标卫星的关键特征点检测的准确性,另外结合了几何优化方法,使用 EPnP 和 Gauss-Newton 方法进行位姿求解。

针对六面对称航天器因关键点视觉模糊导致的位姿混淆问题,Kajak 等人(2023)采用基于分割的关键点回归 CNN 来预测目标航天器在图像中的 2D 关键点位置,然后通过 EPnP 算法从关键点计算 6 自由度(degree of freedom, DoF)位姿。

对于对称结构空间目标,王艺诗等人(2023)提出了一种基于点云深度学习的相对位姿测量方法,设计了基于空间目标点云的特征提取网络和关键点回归网络,通过两个并行回归网络分别输出目标平移向量和带固定标签的三维边界框角点坐标,利用标签稳定的角点求解目标姿态,从而有效解决了对称结构导致的点云误匹配问题。

在单目视觉端到端估计方面,Piazza 等人(2022)将任务分为目标检测网络(spacecraft localization network, SLN)、地标回归网络(landmark regression network, LRN)和几何优化求解器。SLN 使用 YOLOv5 进行区域检测,LRN 使用 HRNet32 回归关键点,然后通过 EPnP 和 Levenberg-Marquardt 方法优化位姿。

类似地,Li 等人(2022a)提出了一种基于学习的位姿估计框架,使用两阶段网络实现,航天器检测网络(spacecraft detection network, SDN)采用 YOLOX-tiny 裁剪出航天器区域,关键点检测网络基于 CSP-Darknet 和 FPN 特征金字塔,结合多尺度锚框检测 11 个预选关键点的 2D 坐标及其不确定性,再由不确定性预测自动过滤低精度关键点,保留高置信结果

通过 EPnP 算法估计 6D 位姿。

Cassinis 等人(2021)提出了一种结合单栈 Hourglass CNN 进行特征检测,使用 CEPPnP 求解器处理带协方差的位姿,以及紧耦合(modified extended kalman filter, MEKF)联合优化位置和姿态,相比松耦合方法减少偏差。

Xu 等人(2020)则聚焦于遮挡环境下的关键点检测,提出基于 Faster R-CNN 进行目标检测,结合高分辨率网络、扩张卷积和在线困难关键点挖掘策略(online hard keypoint mining, OHKM),有效提升了复杂场景下的关键点检测与姿态估计精度。

Huo 等人(2020)结合目标检测、关键点回归和几何优化,提出了一种基于深度学习的快速准确航天器姿态估计方法,主要步骤包括使用轻量级 YOLO-like CNN 检测航天器并预测 3D 模型关键点的 2D 投影位置,设计了边界框可靠性判断和关键点存在概率模型,最后通过 PnP 算法和几何优化器估计 6 自由度姿态。

此外,王梓等人(2022)提出了一种基于 Transformer 的单目卫星位姿估计方法,该方法通过端到端的关键点回归网络直接预测卫星目标的 2D 关键点位置,并结合 PnP 算法解算 6DoF 位姿,突破了传统卷积神经网络在长距离建模和绝对距离间接描述上的局限性。

Alsabhan 和 Alotaiby (2022)提出了基于混合卷积神经网络与注意力机制的框架,通过多尺度特征融合、3D 关键点预测以及跨模态对齐等技术手段实现目标物体的 6DoF 位姿估计。

为避免传统两阶段方法将目标检测与位姿估计分离的弊端,Liu 等人(2024)提出了一种基于 Transformer 的单阶段网络 DTSE-SpaceNet,网络采用关键点检测架构直接预测 2D 关键点位置,并通过 PnP 算法求解位姿参数。

本节综述了基于关键点检测的位姿估计方法,强调通过深度学习检测目标关键点并结合 PnP 等算法实现目标位姿推理的流程。关键点检测法采用改进 YOLO、HRNet、Transformer 等网络,结合 EPnP、RANSAC 等算法,在 SPEED 等数据集上,合成数据位姿误差均值平移 ± 10.36 cm、旋转 $\pm 2.24^\circ$,真实数据 ADD 指标达 82%~94.4%。其优势是显式结构建模精度高,不足是易受遮挡光照影响、对关键点可见性依赖强,适用于结构清晰目标。

表 8 关键点检测位姿估计方法
Table 8 Pose estimation methods for key point detection

文献	网络模型	改进点	数据集	实验结果	局限
Xu 等人(2020)	Faster R-CNN-HRNet + PnP	在 HRNet 特征融合阶段引入扩张卷积,使感受野扩大 44%	SPEED	AP 99.89%,mIoU 90.03%;位姿估计误差:0.021 9	依赖预建 3D 线框模型,泛化能力受限于模型精度
Huo 等人(2020)	tiny-YOLOv3 + PnP + RANSAC	通过逻辑回归分配关键点存在的概率,加速网络收敛	SPEED	平均精度 AP = 73.2%、平均召回率 AR = 77.6%;旋转误差 0.681 2°,平移误差 0.003 20 m	未验证真实场景泛化性
王梓等人(2024)	Lite-HRNet + 亚像素解码法 + EPnP + RANSAC	解决轻量化导致特征表达能力下降时,关键点定位精度损失问题	SPEED	参数量为 1.1 M; E = 15.1×10 ³ ; A = 103.0%	遮挡条件下关键点定位存在明显偏差
Piazza 等人(2022)	SLN + LRN + EPnP + Levenberg-Marquardt	结合 HRNet32 的高分辨率热图与几何优化	SPEED	位姿误差均值为平移 ±10.36 cm,旋转 ±2.24°	目标尺寸过小时关键点检测精度显著降低
Li 等人(2022a)	SDN + KDN + EPnP	提出多尺寸锚框适应不同距离下的目标,提升小目标检测性能	SPEED	参数量为 35.2 M,合成数据:平均误差 0.017 8;真实数据:平均误差 0.132 3	真实图像泛化能力不足
Ahmed 等人(2024)	Transformer + EPnP + Gauss-Newton	提高特征点检测的准确性,模型参数减少 73%,FLOPs 下降 57%	SPEED	参数量为 7.73 M,平移误差 Et = 6.48 cm,旋转误差 Eq = 1.52°	目标距离 < 5 m 时姿态误差较大
Wang 等人(2024b)	PVSAR	基于目标几何结构确定可见表面,局部调整光照,增强模型泛化能力	SPEED+/SHIRT	参数量为 30.6 M,Lightbox 域 SPose = 0.076; Sunlamp 域 SPose = 0.112	在 sunlamp 域的图像中,强反光导致关键点模糊时误检率较高
Cassinis 等人(2020)	Hourglass CNN + CEPPnP	引入热图导出的协方差矩阵到 CEPPnP,改进姿态估计	SPEED	位置误差: 0.118 m; 姿态误差:1.33°	远距离位姿估计误差增加
Chen 等人(2024)	YOLO-w6-pose + 2D 关键点回归	融合检测与关键点回归任务,GFLOPs 仅 102.3	URSO	参数量为 80.5 M, S = 0.068, ADD 达 94.4%	对非共面航天器不适用
Zhong 等人(2025)	CSPDarknet53 + 多分辨率并行分类器	等效直径和偏移量适应不同目标尺度和姿态	URSO/SwissCube	ADD: 99.6%, ADD: 86.4%; 光照场景 S = 0.056	对称目标性能受限
Kajak 等人(2023)	CNN + EPnP	引入对称性感知损失,动态选择最佳对称位姿作为训练目标	训练集 A/B,测试集 C-G	合成数据:平均角度误差 1.6~2.3°;真实数据:召回率 71.7%	约 30% 真实图像因过曝光导致 CNN 误判对称轴方向
Alsabhan 和 Alotaiby(2022)	ResNet-50/PointNet++ Transformer	通过跨模态注意力权重分配提升特征融合效率。相较传统拼接法,参数效率提升 37%	LINEMOD +合成数据增强	准确率 = 89.7%	未考虑动态物体的运动模糊效应

注:实验结果中 A 为位姿估计成功率,E 为位姿估计总误差得分,S 为位姿综合误差得分。

3.3 无监督及域自适应

在非合作空间目标位姿估计中,无监督与域自适应方法(Fang 等,2024)有效缓解了真实与合成数据间的域偏移问题。如表 9 所示,研究者提出多任务网络,结合边界检测、热图回归等多目标任务提升

模型鲁棒性。自训练框架通过伪标签生成与领域一致性约束,增强跨域适应能力。域自适应方法中,融合图像与文本提示、多模态对齐、结构嵌入与无监督策略等,提升模型在未知域下的泛化性能。部分方法还引入自注意力机制或替代 Transformer 结构,以

提升特征提取能力。

Park 和 D’Amico (2024) 在 SPN (spacecraft pose network) 的基础上进行改进, 提出了一种新的多尺度网络 SPNv2。该网络为多任务网络架构, 包括航天器存在与否的二分类、边界框预测、目标位姿估计、目标姿态估计、关键点热图回归和图像前景的逐像素二值分割。此外, 还提出一种通过目标域图像进行在线域优化的方法, 得到了很好的效果。

进一步地, Wang 等人 (2024a) 提出了一种基于几何约束的自训练框架, 用于解决卫星姿态估计中的领域适应问题。该方法通过融合稀疏关键点预测与密集分割信息, 有效弥合成数据与真实数据之间的域差异。

在提升模型泛化性方面, Sun 等人 (2023) 提出了一个结合无监督 Transformer 网络和 CNN 的双目

视觉位姿估计模型, 整合了 SIFT (scale-invariant feature transform) 特征提取、CNN 的全局平均池化替代了 Transformer 的前馈网络。

Pérez-Villar 等人 (2023) 采用双头堆叠 Hourglass 网络, 联合预测 2D 关键点热图与深度信息, 将 3D 结构信息嵌入训练目标, 提升域适应能力, 结合鲁棒伪标签自训练策略实现无监督域适应。

Liu 等人 (2023) 则使用了视觉—语言模型, 提出了 DSF-VLM (diffusion supervised fusion-vision language model) 方法, 结合图像和文本提示, 通过双尺度融合来提升域适应能力。

此外, Cassinis 等人 (2022) 提出了一种 CNN 的单目视觉姿态估计方法, 其核心创新在于设计了一种结合纹理随机化、光照增强和背景扰动的数据增强策略, 以缓解合成数据与真实实验室图像间的域偏移问题。

表 9 无监督及域自适应位姿估计方法

Table 9 Unsupervised and domain-adaptive pose estimation methods

文献	网络模型	改进点	数据集	实验结果	局限
Park 和 D’Amico (2024)	SPNv2	姿态回归/分割/关键点联合学习	SPEED+	参数量为 52.5 M, 旋转角误差为 6.48°, 位移误差为 0.175 m	在线域细化 ODR 可能使部分样本性能恶化
Wang 等人 (2024a)	HRFormer-S	通过分割掩码提供稠密监督信号, 缓解关键点稀疏性问题	SPEED	Sunlamp 域 $S = 0.054\ 5$; Lightbox 域 $S = 0.162\ 7$	极端光照/遮挡下关键点不可见时性能下降
Pérez-Villar 等人 (2023)	Hourglass CNN + PnP	深度联合预测头支持 3D 关键点重构, 增强几何一致性	SPEED+	平均误差为 0.081 2	需预定义 3D 关键点模型
Liu 等人 (2023)	DSF-VLM	首次引入视觉—语言模型解决空间姿态估计的域适应问题	SPEED	旋转误差为 0.3°; 位移误差为 1.98 cm	极端低光照条件下平移参数估计出现偏差
Sun 等人 (2023)	STN	CNN-Transformer 混合架构, 弥补小样本下 Transformer 的局限	自建数据集	位置误差为 0.96%; 姿态误差为 4.39°	真实场景泛化性待验证
Cassinis 等人 (2022)	HRNet + EPnP	结合 SoftPosIT 算法优化关键点误匹配情景下位姿稳定性	合成数据集	位置误差为 0.2 m, 姿态误差为 0.8°	对实验室极端光照适应性有限

本节介绍了无监督与域自适应方法在非合作空间目标位姿估计中的应用, 旨在缓解合成与真实数据之间的域偏移问题。SPNv2 在 SPEED+ 数据集上的旋转角误差为 6.48°、位移误差为 0.175 m, 在线域优化可提升性能但部分样本恶化; 基于几何约束的自训练框架在 SPEED 数据集 Sunlamp 域 $S =$

0.054 5、Lightbox 域 $S = 0.162\ 7$, 极端光照下关键点不可见时性能下降; DSF-VLM 在 SPEED 数据集旋转误差 0.3°、位移误差 1.98 cm, 极端低光照平移估计出现偏差; 结合数据增强的 CNN 方法在高曝光场景 3% 图像达高精度, 位置误差 0.2 m、姿态误差 0.8°, 对极端光照适应性有限。

3.4 多模态视觉融合

在非合作空间目标位姿估计中,多模态视觉融合方法是指通过整合来自多种传感器或数据源,如可见光相机、红外相机、激光雷达和雷达等的信息,

提升位姿估计的鲁棒性和精度的技术,如表 10 所示。其核心在于利用不同模态数据的互补性,克服单一传感器在复杂太空环境,如极端光照、遮挡以及高速运动下的局限性。

表 10 多模态视觉融合位姿估计方法
Table 10 Pose estimation methods of multimodal visual fusion

文献	网络模型	改进点	数据集	实验结果	局限
Qiao 等人(2022)	ResNet18 + PointNet++	首次将点云数据与单目图像结合,通过卷积提取 3D 结构特征	BUAA-SID-POSE 1.0	参数量为 43 M, MAE: 偏航角 9.85°、俯仰角 2.88°、滚转角 2.56°; 精度 ACC: 偏航 71.4%、俯仰 94.2%、滚转 95.1%	对称卫星姿态估计精度下降
Zhang 等人(2025)	SegFormer-MiT-B5	提出局部类区域学习模块提升估计精度	自建数据集	参数量 98.2 M, mAPE = 0.63°	低分辨率图像表现受到限制
王金聪等人(2025)	改进的 Dense-Fusion 网络	引入 PConv 模块,降低冗余计算,提升对光照变化的鲁棒性	LineMOD/自建数据集	LineMOD: 平均 ADD-S 达 97.6%; 自制数据集: 平均 ADD-S 为 89.2%	未验证真实太空场景下的性能
Lei 等人(2024)	MSFT	引入时序感知的跨模态注意力机制	自建数据集	参数量为 48.7 M, 平均精度提升 15.6%	对高频运动场景鲁棒性不足
Gavilanez 和 Moncayo(2024)	Mask R-CNN + PCR-GAN + GICP	结合 MSN 补全遮挡与 PU-GAN 提升点云密度, MSE 误差降低 35.59%	自建数据集	点云质量: PCR-GAN 将 CD 从 7.806×10^{-3} 降至 7.159×10^{-4} , MSE 降低 35.59%; 位姿精度: 平移误差: y 轴提升 45.7%, 姿态角误差: Yaw 角提升 61.4%	PCR-GAN 生成的内部点密度高于表面,可能影响复杂几何精度
陈炜等人(2023)	YOLOv5 + ICP	结合 RGB 图像与深度图的异步数据流	自建数据集	Precision = 99.5%, AP50 = 99.5%	对遮挡敏感,目标被遮挡超过 30% 面积时点云配准失败
Guthrie 等人(2022)	Siamese CNN + RANSAC	首次在端到端可训练框架中集成并行 RANSAC, 提升异常值鲁棒性	合成数据集	在模拟数据集上,相较于 ORB、SIFT 与 ICP 算法,其平均姿态误差降低 40% 以上,实时处理速度达 23 Hz	真实场景泛化性待验证
Rondao 等人(2023)	CNN + LSTM	首次引入 LSTM 处理航天器序列图像中运动连续性	Astos 合成数据集/City 实验数据集	ASTOS/06 序列达到 0.49% 相对位置误差、1.85° 平均姿态误差	真实数据验证局限于实验室环境

Zhang 等人(2025)提出了一种基于多通道匹配和动态关键帧优化的无模型三维姿态估计方法,分为基于 SuperPoint 架构关键点提取、多通道匹配和姿态图优化三阶段,旨在解决非合作目标的快速旋转导致特征匹配失效问题。

为提升稠密点云下的姿态估计精度,Gavilanez 和 Moncayo(2024)设计了一种双阶段视觉架构,基于改进的 Mask R-CNN 对立体相机 RGB 图像进行目标分割和掩膜生成,结合点云重建生成对抗网络

(point cloud reconstruction generative adversarial network, PCR-GAN)与广义迭代最近点算法(generalized iterative closest point, GICP),通过对抗训练提升点云密度,利用 GICP 算法求解位姿参数,但 PCR-GAN 生成的内部点密度高于表面,可能影响复杂几何精度。

Lei 等人(2024)采用双分支网络架构,视觉编码分支采用 ResNet-50 骨干提取图像特征,跨模态融合模块通过 Transformer 层融合视觉与惯性测量单元

(inertial measurement unit, IMU)传感器数据以及引入层级注意力机制整合多尺度特征,解决单目视觉尺度模糊性问题,但对高频运动场景鲁棒性不足。

此外,Qiao等人(2022)提出了基于CNN的卫星相对姿态估计方法,主要分为3个模块:特征提取、3D先验信息融合和位姿求解模块。ResNet18提取图像特征,3D融合模块使用点云卷积,最后全连接层回归三维姿态角。

陈炜等人(2023)基于YOLOv5神经网络构建目标检测框架,结合深度相机数据通过几何反投影生成目标点云,采用ICP算法与CAD模型点云进行配准,最终实现空间位姿估计和AR(augmented reality)虚实融合,但该方法对遮挡敏感,目标被遮挡超过30%面积时点云配准失败。

为进一步提升多模态输入下的关键点稳定性,Guthrie等人(2022)采用连体卷积神经网络Siamese CNN从多模态传感器数据中提取关键点,并结合改进的RANSAC算法滤除异常值,最终通过均方根误差(root mean square deviation, RMSD)优化算法或全连接神经网络(fully connected neural network, FCNN)估计目标姿态。

在处理复杂光照和模态差异方面,Rondao等人(2023)提出了一种名为ChiNet的深度递归卷积神经网络,结合了CNN和LSTM模块,处理多模态的可见光和红外图像,并通过3个阶段的分层训练策略优化模型,解决光照变化、强反射等的鲁棒性问题。

王金聪等人(2025)通过改进DenseFusion网络,结合部分卷积和大核注意力来提高卫星6D位姿估计的准确性,基础框架是DenseFusion,主要做了3方面改进:改进了Blender渲染数据集的方法,利用多模态数据进行训练以提升模型的表现,加入部分卷积(partial convolution, PConv)模块来降低对光照和背景噪声的敏感度,以及设计了LKA-PSPNet(large kernel attention pyramid scene parsing network)来捕捉多尺度特征。

本节介绍了多模态视觉融合方法在非合作空间目标位姿估计中的应用,通过融合可见光、红外、激光雷达和深度图等多源异构传感器信息,提升鲁棒性与精度。基于SuperPoint的方法可解决快速旋转特征匹配失效问题;MSFT在SPEED数据集解决单目尺度模糊,对高频运动鲁棒性不足;改进DenseFusion在LineMOD数据集ADD-S达89.2%~97.6%,

提升光照鲁棒性;结合PCR-GAN与GICP的方法使点云MSE降低35.59%,但影响复杂几何精度;YOLOv5+ICP在自建数据集上AP@0.5=99.5%,遮挡超30%配准失败;ChiNet处理多模态图像,在Astos数据集获0.49%相对位置误差,1.85°平均姿态误差,真实数据验证有限。

3.5 小结

本节围绕深度学习在非合作空间目标位姿估计中的应用展开,系统分析了4类主流技术路径:直接回归、关键点检测、无监督与域自适应以及多模态视觉融合方法。表11对本节所讨论的非合作空间目标位姿估计方法在典型的公开数据集上,从参数量、推理速度、姿态误差和位置误差等维度进行了算法性能比较。

表12进一步总结了4类方法的核心原理和适用场景等。直接回归方法通过构建端到端网络模型,绕过传统的特征提取和几何解算流程,具备计算效率高、部署简便等优势,适合实时性要求高的任务;关键点检测方法强调结构建模与几何约束结合,借助PnP等算法提升姿态精度,适用于具有清晰结构先验的目标场景;无监督与域自适应方法通过无监督、多任务联合训练等机制,降低对标注数据的依赖,提升跨域泛化能力,适合缺少真实数据的任务环境;多模态视觉融合方法则整合来自RGB图像、深度图、红外以及激光雷达等多源信息,有效提升弱纹理、遮挡与复杂光照条件下的鲁棒性,适用于复杂空间环境中的精确位姿估计。总体来看,基于视觉与深度学习的位姿估计方法呈现出从端到端建模到多模态融合、无监督增强的演进趋势,逐步突破传统方法在光照变化、目标遮挡以及结构模糊等场景下的适用瓶颈,展现出更强的适应性与智能化水平。

4 数据集

数据集在基于视觉与深度学习的非合作空间目标识别与位姿估计研究中具有基础性作用。高质量的数据集不仅为深度学习模型提供了充足的训练样本,支持模型学习目标在不同视角、姿态和光照条件下的特征,还为各类算法提供了统一的测试平台,用于评估性能与对比优劣。

4.1 公开数据集

非合作空间目标识别与位姿估计任务常用的公

表 11 位姿估计算法性能比较
Table 11 Comparison of pose estimation algorithms performance

类别	方法	设备	参数量	推理速度/ (帧/s)	姿态误差	位置误差	数据集
直接 回归	双通道 Transformer (Ye 等, 2023)	i9-11900K CPU, NVIDIA RTX3090 GPU	186 M	29.9	1.573°	0.044 m	SPEED
	PVSPE: 基于 Transformer 的多任务网络 (Yang 等, 2024)	NVIDIA GeForce RTX 3090 GPU	70.15 M	31.25	4.81°	0.017 m	SPEED+
	基于 CNN 的端到端网络 (Huang 等, 2023)	Intel Xeon Silver 4210 CPU, NVIDIA Quadro RTX 6000 GPU	—	10	1.47°	0.011 9 m	SPEED+
	YOLOv5 + scSE-LHRNet (Deng 等, 2022a)	64 GB Intel Core i9-10900K CPU, 32 GB NVIDIA 2080Ti GPU	11.82 M	—	0.016 9°	—	SPEED
	改进的 GoogLeNet (Phisannupawong 等, 2020)	NVIDIA GeForce RTX 2080 GPU	—	—	13.70°	1.2 m	URSO
关键 点 检测	Lite-HRNet + 亚像素解码法 + EPnP + RANSAC (王梓 等, 2024)	—	1.1 M	35.21	0.62°	0.030 9 m	SPEED
	Transformer + EPnP + Gauss-Newton (Ahmed 等, 2024)	NVIDIA-TITAN XP GPU	7.73 M	—	1.52°	0.064 8 m	SPEED
	SDN + KDN + EPnP (Li 等, 2022a)	—	35.2 M	—	0.739°	0.004 9 m	SPEED
	Hourglass CNN + CEPPnP (Cassinis 等, 2021)	—	26.8 M	—	1.33°	0.118 m	SPEED
	Faster R-CNN-HRNet + PnP (Xu 等, 2020)	Intel Core i7-6700 CPU, NVIDIA GeForce GTX 1080Ti GPU	26.2 M	—	0.972 9°	0.007 4 m	SPEED
	tiny-YOLOv3 + PnP + RANSAC (Huo 等, 2020)	—	—	50	0.681 2°	0.003 2 m	SPEED
无监 督及 域自 适应	SPNv2 (Park 和 D'Amico, 2024)	—	52.5 M	—	6.48°	0.175 m	SPEED+
	双头堆叠 Hourglass CNN 网络 + PnP (Pérez-Villar 等, 2023)	—	—	—	0.049 3°	0.081 2 m	SPEED+
	DSF-VLM (Liu 等, 2023)	NVIDIA Geforce RTX 2080 SUPER GPU, Intel i7-9700 CPU	—	21.93	0.3°	0.019 8 m	SPEED
	HRFormer-S (Sun 等, 2023)	NVIDIA GTX 3090 GPU	7.8 M	17.45	4.696°	0.096 3 m	SPEED
多模 态视 觉融 合	ResNet18+ PointNet++ (Qiao 等, 2022)	Intel(R) Xeon(R) Silver 4114 CPU, NVIDIA GTX 2080Ti GPU	43 M	—	2.56°	—	BUAA-SID-POSE 1.0
	CNN+LSTM (Rondao 等, 2023)	NVIDIA Turing V100 Tensor Core GPU	—	—	1.85°	0.23 m	Astos

注：“—”表示该项内容为空。

开数据集涵盖多种典型航天器目标,提供多视角、高分辨率图像及精确的 6D 位姿标注,常用的公开数据集如表 13 所示。图 5 为部分 URSO 数据集的示例。

这些数据集支持深度学习模型的训练与评估。
AKM (apogee kick motor) 数据集是由 Rathinam 等人 (2022) 提出的一个专门针对无纹理空间目标的

表 12 基于视觉与深度学习的位姿估计方法对比

Table 12 Comparison of pose estimation methods based on vision and deep learning

方法类别	原理	优势	局限性	适用性
直接回归方法	通过深度神经网络直接回归目标的平移和旋转参数,实现端到端位姿预测	端到端建模,计算效率高	泛化能力受限,对对称结构存在多解问题,真实环境适应性需验证	缺乏几何先验、需快速响应的动态任务场景
关键点检测方法	检测目标关键点或特征点,并结合 PnP 算法计算位姿,显式利用几何关系	显式结构建模,精度高,适配几何先验	易受遮挡与光照影响,对关键点可见性依赖强,处理流程复杂	结构清晰、具有模型先验的目标
无监督及域自适应方法	通过重投影损失、对比学习等无监督方式或域适应技术进行训练,减少对真实标签依赖	弱标签依赖,支持无监督与伪标签,跨域泛化强	训练复杂,依赖伪标签质量,在极端场景下鲁棒性不足	缺少真实标签、任务需从仿真数据过渡到真实环境的应用
多模态视觉融合方法	融合可见光、红外、深度图、LiDAR 等多源感知信息,通过特征级或决策级融合提升位姿估计精度	多源互补,抗遮挡、抗光照变化能力强,精度高	系统复杂度高,传感器标定与数据对齐困难,硬件成本大	多源异构环境、高遮挡或弱纹理条件下的高精度感知任务

姿态估计数据集,主要研究对象是固体火箭发动机。该数据集包含 10 000 幅 $1\,024 \times 1\,024$ 像素的合成图像,通过 Blender 模拟器在类轨道光照条件下生成。目标物体建模为直径 0.62 m、长度 1.066 8 m 的轴对称无纹理圆柱体,模拟了 STAR 24 型发动机的外观特性。数据集中目标位置在 1~10 m 范围内均匀随机分布,姿态角也采用随机分布,每幅图像都标注有精确的相对位姿信息,但不足在于缺乏真实观测图像,目标单一,限制了算法在多目标任务中的泛化能力。

BUAA-SID-POSE 1.0 数据集是由北京航空航天大学于 2022 年发布,涵盖 LRO、ICECube 等航天器,每类模型生成 56 656 幅 256×256 像素的灰度图像,拍摄距离 2~50 m,使用 BUAA-SID 1.0 仿真工具生成。其优势在于航天器类型多样,视角与距离覆盖广,适用于多目标识别与位姿估计,但不足在于全部为合成数据,缺少真实成像噪声与光照变化,影响模型对实际环境的适应性。

SHIRT (satellite hardware-in-the-loop rendezvous trajectories) 数据集是由 Park 和 D’Amico (2022) 提出的首个包含真实硬件在环图像与合成图像的航天器交会轨迹数据集,用于验证跨域差距的导航算法性能。该数据集基于 PRISMA 任务中的 Tango 航天器模型,通过 OpenGL 渲染器和美国斯坦福大学 TRON 机器人测试床生成两种轨道交会场景 (ROE1 和 ROE2) 的序列图像,其中,ROE1 模拟沿轨保持,ROE2 模拟缓慢接近并伴有目标翻滚。每 5 s 采集的

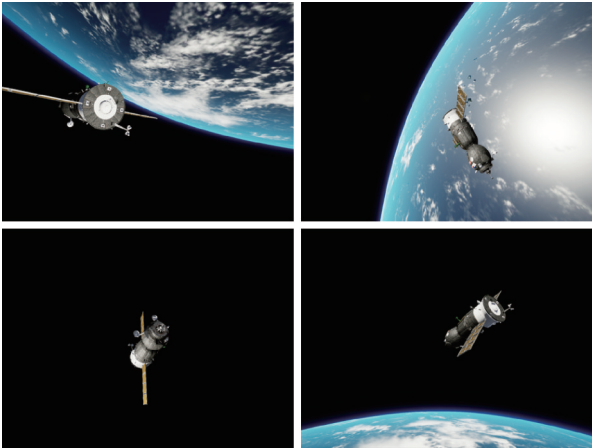


图 5 URSO 数据集示例 (Proença 和 Gao, 2020)
Fig. 5 Examples of URSO datasets
(Proença and Gao, 2020)

图像配有毫米级精度的位姿标签,为单目视觉导航滤波器的跨域性能评估提供了基准。

SwissCube 数据集采用了高精度卫星 CAD (computer-aided design) 模型与光谱级渲染技术,真实模拟太空环境中太阳、地球及星体反光等复杂光照效应,通过轨道动力学仿真生成目标深度范围 1 d—10 d (d 为卫星直径) 的图像序列,弥补了 SPEED 数据集在远距离样本不足的缺陷;包含 50 000 幅 $1\,024 \times 1\,024$ 像素分辨率图像,覆盖近/中/远距离场景。但该数据集在地面实验室拍摄的真实图像仅 20 幅用于微调,无法完全反映太空极端光学条件。

SPEED+ (spacecraft pose estimation dataset enhanced) 数据集由美国斯坦福大学与 NASA 于

2021 年联合发布, 包含约 60 000 幅合成图像与 10 000 幅真实图像, 目标为 Tango 卫星, 分辨率 $1\,920 \times 1\,200$ 像素, 拍摄距离 3~40.5 m, 使用 OpenGL 渲染。其优势在于增加真实图像比例, 提升跨域泛化能力, 适用于 Sim2Real 姿态回归研究, 但不足在于航天器类型单一, 仅涵盖 Tango 卫星, 场景多样性有限。

Cygnus 数据集由 NASA 于 2021 年发布, 包含约 20 000 幅合成图像和 540 幅太空实拍图像, 分辨率为 $1\,024 \times 1\,024$ 像素, 目标为 Cygnus 飞船, 拍摄距离 35~75m, 采用 Blender 渲染。其优势在于包含真实太空图像, 适用于验证交会对接算法, 但不足在于图像总量较少, 且仅覆盖 Cygnus, 缺少多目标与多场景样本。

Minerva-II2 数据集 (Hematulin 等, 2025) 是由 Andrew Price 和 Kazuya Yoshida 为训练航天器姿态估计神经网络而构建的合成数据集, 专门用于 Hayabusa2 任务中部署 Minerva-II2 巡视器时拍摄的真实图像特性。该数据集通过 Solidworks Photoview 360 软件构建了 Minerva-II2 的 3D 模型, 并采用 12 个定制化噪声过滤步骤, 逼真再现了原始图像中由 CCD 传感器饱和、高动态范围和非均匀光照造成的复杂噪声特征。数据集包含两组主要数据: SetA 包含 10 000 幅真实模型渲染图, 用于实际部署图像分析;

SetB 包含 10 000 幅虚构非对称模型渲染图, 用于概念验证。首次针对深空任务特有的图像噪声模式进行了系统建模, 为解决对称目标姿态估计这一核心挑战提供了训练基础。

SPARK (spacecraft recognition leveraging knowledge of space environment) 数据集是一个专为空间目标识别任务设计的合成数据集, 由卢森堡大学 (University of Luxembourg) 和 LMO 团队开发, 包含 150 000 幅 RGB 图像和 150 000 幅深度图像, 覆盖 10 类卫星和 1 类太空碎片。该数据通过 Unity3D 引擎在真实空间仿真环境中生成, 模拟了多样化的轨道场景和复杂背景, 包括地球海洋云层、高对比度光照、物体随机位置、多尺寸目标及高噪声等挑战性条件。

SPEED 数据集是由美国斯坦福大学 Space Rendezvous Laboratory 与欧洲空间局合作开发的航天器姿态估计基准数据集, 主要用于非合作航天器单目视觉导航研究。该数据集包含基于计算机建模生成的 15 000 幅 PRISMA 任务中 Tango 航天器的合成图像和通过测试台 TRON 设施获取的 300 幅硬件在环图像。

URSO (University of Rome Spacecraft Orientation) 数据集是一款基于 Unreal Engine 4 构建的空间目标图像数据集, 专为航天器在轨服务任务中的识别、定位与位姿估计研究设计。该数据集通过模拟

表 13 公开数据集
Table 13 Public datasets

数据集	合成/实验/太空	航天器	分辨率/像素	距离/m	图片类型	仿真工具
AKM (Rathinam 等, 2025)	10 k/-/-	AKM	$1\,024 \times 1\,024$	1~10	灰度图	Blender
BUAA-SID-POSE 1.0 (Qiao 等, 2024)	60 k/-/-	LRO、ICECube 等	$1\,280 \times 720$	2~50	灰度图	BUAA-SID 1.0
SHIRT (Park 和 D'Amico, 2023)	5 k/5 k/-	Tango	$1\,920 \times 1\,200$	< 40	灰度图	OpenGL
SwissCube (Wang 等, 2025)	50 k/-/-	SwissCube	$1\,024 \times 1\,024$	0.1~1	RGB	Mitsuba 2
SPEED+ (Park 等, 2022)	60 k/10 k/-	Tango	$1\,920 \times 1\,200$	3~40.5	灰度图	OpenGL
Cygnus (Black 等, 2021)	20 k/-/540	Cygnus	$1\,024 \times 1\,024$	35~75	RGB	Blender
Minerva-II2 (Hematulin 等, 2025)	3 k/-/-	Minerva-II2	$1\,024 \times 1\,024$	< 40	RGB	Photoview 360
SPEED (Park 和 D'Amico, 2024)	15 k/305/-	Tango	$1\,920 \times 1\,200$	3~40.5	灰度图	OpenGL
SPARK (Aldahoul 等, 2022a)	150 k/-/-	AcrimSat、Aquarius、空间碎片等	$1\,280 \times 720$	1.5~10	RGB	Unity3D
URSO (Proença 和 Gao, 2020)	15 k/-/-	Soyuz、Dragon	$1\,080 \times 960$	10~40	RGB	Unreal Engine 4

注:“-”表示该项内容为空。

近地轨道环境,高度10~20 m,随机生成5 000个相机视点,并包含地球自转、相机视角变化以及空间目标随机姿态等多样化场景。数据集包含训练集、验证集和测试集,图像分辨率为1 280×960像素,覆盖了复杂的地球背景与太空背景干扰场景。

4.2 自建数据集

在基于视觉与深度学习的非合作空间目标识别与位姿估计研究中,自建数据集具有重要意义。首先,现有公开数据集数量有限,且大多缺乏覆盖多姿态、多光照和多尺度等真实复杂环境下的图像,难以满足实际任务需求;其次,非合作空间目标具有形态多样、纹理缺失以及强光照干扰等特点,公开数据集往往无法充分反映这些挑战。通过自建数据集,研究人员可根据具体应用场景模拟目标种类、成像条件和相机参数等,提高算法训练的针对性和泛化能力,如表14所示。同样地,位姿估计任务的自建数据集多依托仿真手段构建,以适配复杂空间场景需求。如Cassinis等人(2021)构建的数据集,借助Cinema 4D、Blender、Unity3D等仿真工具,模拟航天器多姿态角、不同距离范围、轨道运动及光照条件,生成涵盖RGB图像、多模态序列等数据形式的数据集,覆盖从常规姿态到复杂轨道运动、从单一成像到多模态融合的位姿估计研究场景,为航天器位姿解算模型训练提供多样化数据支撑。但是部分数据集在模拟真实物理规律、与实际任务场景契合度上,仍需结合真实数据优化。

5 评价指标

为全面评估基于深度学习的非合作空间目标识别与位姿估计系统性能,需要选择合适的性能指标(Ahmed等,2024)。

5.1 识别算法评价指标

空间目标识别任务包括目标分类、检测定位和分割等子任务。为了全面评估识别算法性能,通常使用以下指标:精确率(precision)、召回率(recall)、F1值(F1-score)、交并比(intersection of union, IoU)和平均精度均值(mean average precision, mAP)等(胡佳等,2025)。

精确率和召回率是两个关键指标。精确率反映被模型检测为目标样本中有多少是真实目标;而召回率衡量所有真实目标中被正确检测出的比例。

在精确率和召回率存在权衡关系时,常使用其调和平均值,即F1值进行综合评价,具体为

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (1)$$

F1值越高,说明模型在保持高精度的同时也能覆盖大多数目标。

交并比(IoU)衡量预测区域与真实区域之间的重叠程度,适用于定位于分割任务,具体为

$$IoU = \frac{A_p \cap A_{gt}}{A_p \cup A_{gt}} \quad (2)$$

式中, A_p 为预测区域, A_{gt} 为真实区域,通常设定IoU阈值为0.5或0.75,以判断目标是否被成功检测。

平均精度(mAP)用于多类别识别场景,先计算每类目标的Precision-Recall曲线下面积(average precision, AP),再取所有类别的均值,具体为

$$mAP = \frac{1}{C} \sum_{i=1}^C AP_i \quad (3)$$

式中, C 表示目标类别总数, AP_i 为第*i*类的平均精度。

5.2 位姿估计评价指标

位置误差(金泽明等,2021)是衡量空间目标位置预测值与真实值之间偏差的重要指标,通常采用欧氏距离计算目标质心在三维坐标中的直线距离,具体为

$$E_T = \|\mathbf{t}_{pred} - \mathbf{t}_{gt}\|_2 \quad (4)$$

式中, $\mathbf{t}_{pred}$ 表示预测的平移向量, $\mathbf{t}_{gt}$ 表示真实的平移向量, $\|\cdot\|_2$ 表示欧氏范数,为距离向量。

姿态误差表示预测姿态与真实姿态之间的旋转偏差(窦凯云等,2023b),常用两种方式计算。

基于旋转矩阵的表示方法为

$$E_R = \cos^{-1} \left(\frac{\text{tr}(\mathbf{R}_{gt}^T \mathbf{R}_{pred}) - 1}{2} \right) \quad (5)$$

式中, $\mathbf{R}_{pred}$ 和 $\mathbf{R}_{gt}$ 分别表示预测和真实的旋转矩阵。

基于四元数的表示方法为

$$E_R = 2 \cos^{-1} \left(\left| q_{gt} \cdot q_{pred} \right| \right) \quad (6)$$

式中, q_{gt} 和 q_{pred} 分别表示真实和预测的单位四元数。

ADD(average distance and direction)适用于非对称目标。计算物体3D模型中所有点在预测位姿和真实位姿下投影位置之间的平均欧氏距离,衡量整体位姿估计的误差,具体为

$$ADD = \frac{1}{|M|} \sum_{x \in M} \left\| (\mathbf{R}_{pred} \mathbf{x} + \mathbf{t}_{pred}) - (\mathbf{R}_{gt} \mathbf{x} + \mathbf{t}_{gt}) \right\| \quad (7)$$

表 14 自建数据集
Table 14 Self-managed datasets

任务分类	来源	类型	样本数	覆盖范围	数据集内容	仿真工具
目标识别	Zhang 等人 (2024a)	RGB	训练/验证/测试: 17 k/-/3 k	—	15 类失效卫星以及 39 类太空垃圾	3D MAX
	Zhang 等人 (2024b)	灰度图	源域: 8 485; 目标域: 66 095	目标类型: 圆锥体、立方体、椭球体 (涵盖不同几何对称性目标); 姿态角 $0^{\circ} \sim 90^{\circ}$	源域: 包含 3 种目标, 圆锥体、立方体、椭球体, 在旋转状态下的微多普勒特征 ($0^{\circ} \sim 90^{\circ}$ 姿态角, 间隔 2°); 目标域: 激光雷达实测的微多普勒时频谱, 目标为 3D 打印的毫米级目标 (边长 6 mm 立方体、半径 4 mm 圆锥体、长短轴 4/12 mm 椭球体)	—
	Pensado 等人 (2023)	LiDAR 深度图像	训练/验证/测试: 40 k/10 k/-	姿态角 $0^{\circ} \sim 360^{\circ}$	基于 Envisat 卫星的 .stl 模型生成不同姿态的 LiDAR 深度图像	—
	Zhao 等人 (2022)	点云	792 个点云	—	本体、面板、天线, 共 32 个不同航天器模型	Blender
	Liu 等人 (2022b)	RGB	训练/验证/测试: 800/200/-	—	航天器主要部件, 如主体、太阳能板、天线等	—
	Chen 等人 (2021)	RGB	训练/验证/测试: 900/100/-	—	包括不同角度和颜色的目标物体及标签数据	V-rep
	郝强 等人 (2022)	RGB	训练/验证/测试: 6 k/1 k/3 k	—	包括卫星本体和部件, 如太阳能板、天线、传感器等	—
	Ai 等人 (2022)	RGB	3 117	—	卫星、太阳帆板、舱段	—
	Cassinis 等人 (2021)	RGB	训练/验证/测试: 18 k/6 k/6 k	$90\text{ m} \sim 180\text{ m}$	多姿态角场景 (偏航/俯仰/横滚角按 10° 离散化); 固定太阳高度角 (30°) 和方位角以生成不同光照条件	Cinema 4D
	Guthrie 等人 (2022)	多模态数据集	训练/验证/测试: 60 k/-/-	相对旋转速度: $0.5^{\circ}/\text{步} \sim 7^{\circ}/\text{步}$; 相对轨道距离: $20\text{ m} \sim 500\text{ m}$	模拟主动碎片清除任务中追踪卫星与目标卫星的相对轨道运动和翻滚运动	Blender
位姿估计	Rondao 等人 (2023)	RGB/RGBT	7 个序列/-/7 个序列	$> 3.8\text{ m}$	Envisat 仿真序列: 3 种不同制导模式, 3 种旋转状态, 2 种接近方向	Astos 相机模拟器
	Sun 等人 (2023)	RGB	373 248	姿态范围: $360^{\circ} \times 3$ 旋转轴; 距离范围: $> 2.8\text{ m}$	三维卫星模型模拟, 3 个旋转轴 (横滚/俯仰/偏航) 以 5° 间隔全周旋转; 4 种位置运动轨迹 (近远/水平/远近/水平交替)	Unity3D
	Hu 等人 (2023)	灰度图	90 帧连续数据	低/高速旋转航天器的姿态测量 ($X/Y/Z$ 轴位移 $\pm 500\text{ mm}$, 角度 $0^{\circ} \sim 360^{\circ}$)	航天器模型在旋转状态下的双目光学图像; TOF 相机采集的 3D 点云数据; 模拟空间光照条件下的传感器输出	ABB 机械臂仿真系统
	Gao 等人 (2023)	RGB	训练/验证/测试: 14 195/-/3 549	姿态角度覆盖范围 $0^{\circ} \sim 35^{\circ}$	低分辨率、模糊边界及高噪声的航天器在轨 ISAR 成像数据	地面 ISAR 设备采集
	Huang 等人 (2022)	RGB	训练/验证/测试: 12 005/-/3 298	—	Tango 卫星	KPEC
	Han 等人 (2025)	多模态数据集	训练/验证/测试: 140 k/20 k/40 k	$> 10\text{ m}$ 及 $< 10\text{ m}$	PEU-cloud: 100 种航天器三维点云; PEU-track: 模拟轨道交会、编队飞行等任务	Blender

注: “—”表示该项内容为空。

ADD-S(average distance and direction with scale)是ADD指标的扩展,适用于对称目标。通过对每个预测点匹配真实点集中最近点消除对称性干扰,具体为

$$ADD-S = \frac{1}{|M|} \sum_{x_1 \in M} \min_{x_2 \in M} \|(R_{\text{pred}} x_1 + t_{\text{pred}}) - (R_{\text{gt}} x_2 + t_{\text{gt}})\| \quad (8)$$

式中, M 为3D模型点集, x 为点集中的单个3D点, x_1 为预测位姿下的模型点, x_2 为真实位姿下与 x_1 最近的模型点。

ADD-S准确率衡量位姿估计误差(以ADD-S度量)在预设阈值 τ 内的样本比例,反映模型整体预测成功率,具体为

$$Accuracy_{\text{ADD}} = \frac{N_{\text{pred}}}{N} \times 100\% \quad (9)$$

式中, N_{pred} 为ADD-S误差小于 τ 的正确预测样本数, N 为待评估的总样本数, τ 常设为目标物体直径的10%。

6 面临的挑战与发展趋势展望

6.1 面临的挑战

尽管基于视觉与深度学习的非合作空间目标识别与位姿估计方法取得了显著进展,但在实际应用与工程部署中,仍面临多方面的挑战,主要包括:

1)空间目标识别的相似类别混淆与微小目标检测瓶颈。在轨识别任务中,不同航天器在形态结构上可能高度相似,在低分辨率成像、光照不均或局部遮挡情况下,易导致特征分布重叠,引发类别混淆。Jin等人(2020)在URSO数据集的相似目标识别实验中表明,高相似度类别的误检率较普通类别高出24.8%。同时,小目标检测仍是识别算法的性能短板,当目标在图像中宽度低于20 px时,YOLOv5与Faster R-CNN的mAP平均下降超过35%,对在轨微小空间碎片或远距离航天器的识别带来限制。

2)空间目标识别在复杂背景下的误检问题。在轨拍摄背景往往包含地球边缘、太阳眩光及星空背景,这些结构与航天器轮廓在低信噪比情况下会产生显著混淆。部分检测器在SPEED+的混合背景测试集上误检率上升超过20%,主要集中在光照过渡区与高反差边缘区域,影响实际任务的稳定性与可靠性。

3)位姿估计的多解歧义与旋转精度受限。非合作空间目标的几何结构常存在高度对称性,在单视角成像下会产生多个几何等价解,使位姿估计任务存在固有的多解歧义问题。当目标部分遮挡或视角单一时,该问题尤为突出。Deng和Li(2023)基于SPEED+数据集实验发现,在目标存在30%遮挡时,旋转角均方误差较无遮挡场景上升31.7%,甚至出现翻转解的错误匹配。此外,在深远距离成像条件下,目标仅占图像的极小像素比例(<2%),关键点检测与特征匹配精度大幅下降,导致姿态解算发生明显漂移。

4)位姿估计的关键点定位稳定性不足。许多基于关键点检测的位姿估计方法,如PVNet、EPnP+RANSAC在面对弱纹理区域时,特征点的重复检测与定位精度显著下降,进而影响姿态求解的稳定性。特别是在URSO、AKM等合成数据集向真实在轨图像迁移时,受传感器噪声、光学畸变与压缩伪影等因素影响,关键点检测的亚像素精度损失严重,姿态误差在跨域测试中最高增加42%。

5)目标识别与位姿估计缺乏统一评测体系。目前两类任务各自使用不同的评价指标:位姿估计常用ADD、ADD-S以及旋转/平移误差,目标识别则以mAP、IoU为主,且阈值设定与容忍范围不统一。例如,不同研究在位姿估计的误差阈值上有0.1 d、0.05 d等差异,在识别中IoU阈值设置从0.3到0.7不等,导致跨方法、跨数据集的性能对比缺乏可比性,也限制了位姿估计与目标识别协同优化的验证。

6.2 未来发展趋势展望

针对上述挑战,未来研究可从以下几个方面着力突破,以推动非合作空间目标感知技术从理论验证向实际在轨应用演进:

1)目标识别的小目标增强与判别特征优化。针对远距离小目标与相似类别混淆,可在检测网络中引入特征金字塔与多尺度检测分支,提升不同尺度目标的检测能力;结合注意力机制、增强类别间判别性,并利用对比学习优化特征空间,使相似类别在特征分布上更易分离。

2)复杂背景与极端环境下的识别鲁棒性提升。未来研究可通过背景抑制网络与多任务学习框架提升复杂背景下的识别鲁棒性,例如在检测过程中同时预测目标显著性图,降低背景特征干扰。此外,结合多模态信息可在光照剧烈变化时保持识别精度。

3)基于多模态融合与几何约束的姿态优化。未来可引入CAD模型先验、形状约束与物理可行性判别,将深度回归与几何优化结合,过滤不合理姿态解;同时融合可见光、深度图、点云及惯性测量单元等多模态信息,弥补单模态在弱纹理或遮挡场景下的特征不足。

4)跨域稳定性与在轨跟踪连续性优化。未来可结合无监督域自适应与特征增强策略,在仿真预训练模型中引入真实成像噪声模拟,提高关键点检测的跨域稳定性。同时利用时序信息,通过CNN、LSTM等深度时序网络与图优化方法实现位姿的连续优化,抑制瞬时检测误差并提升在轨跟踪稳定性。

5)建立统一评测体系。未来应建立覆盖两类任务的统一评测体系,明确各指标的适用范围与标准化阈值,并在测试集中纳入多光照、遮挡以及噪声条件下的综合场景。以SPEED+、URSO、AKM等数据集为基础扩展混合测试集,构建权威测试平台与跨机构竞赛,推动算法在多任务、多域条件下的权威性能对比与持续优化。

7 结 语

本文系统综述了基于视觉与深度学习的非合作空间目标识别与位姿估计方法,全面梳理了该领域的研究进展、技术体系、数据支撑、存在挑战与未来发展方向。首先,介绍了非合作空间目标识别与位姿估计的任务特点与应用背景,阐明了两者在空间在轨服务与态势感知中的重要性及其紧密联系,并梳理了从传统几何方法到深度学习模型的演进脉络。在技术体系方面,系统总结了空间目标识别的三大技术路径,包括多模态视觉融合、目标检测与分割以及迁移学习与少样本学习,分析了各类方法在特征表达、精度、适应性和样本依赖方面的差异;位姿估计方法则涵盖直接回归、关键点检测、无监督及域自适应以及多模态视觉融合4类技术路径,比较了不同方法在精度、鲁棒性与适用场景上的优势与局限。在数据集与评价指标方面,系统整理了当前非合作空间目标识别与位姿估计的公开数据集,以及典型的自建数据集,涵盖数据来源、分辨率、任务类型、距离范围和合成与实拍等特征;同时总结了目标识别与位姿估计任务常用的评价指标体系,如分类准确率、mAP、位置误差和姿态误差等,明确了当

前评估方法的标准与不足,为后续研究提供了必要的支撑与评估参考。最后分析了当前研究所面临的主要挑战,包括:空间目标识别的相似类别混淆与微小目标检测瓶颈;空间目标识别在复杂背景下的误检问题;位姿估计的多解歧义与旋转精度受限;位姿估计的关键点定位稳定性不足;目标识别与位姿估计缺乏统一评测体系。针对上述挑战,提出未来的发展方向:引入特征金字塔和注意力机制等增强小目标类别间特征判别性,通过背景抑制网络与多任务学习框架提升复杂背景下的识别鲁棒性,基于多模态融合与几何约束优化姿态精度,结合无监督域自适应与特征增强策略提升姿态求解稳定性,并构建标准化测试体系与权威测试平台,推动任务驱动的工程化应用落地。通过以上的总结和分析,希望可以提供有参考的价值。

参考文献(References)

- Aglietti G S, Taylor B, Fellowes S, Salmon T, Retat I and Hall A. 2020. The active space debris removal mission RemoveDebris. Part 2: in orbit operations. *Acta Astronautica*. 168: 310-322 [DOI: 10.1016/j.actaastro.2019.09.001]
- Ahmed J, Arshad A, Bang H and Choi Y. 2024. Transformer network-aided relative pose estimation for non-cooperative spacecraft using vision sensor. *International Journal of Aeronautical and Space Sciences*, 25(3): 1146-1165 [DOI: 10.1007/s42405-023-00703-3]
- Ai H, Zhang H F, Ren L, Feng J and Geng S N. 2022. Detection and recognition of spatial non-cooperative objects based on improved YOLOX_L. *Electronics*, 11(21): #3433 [DOI: 10.3390/electronics11213433]
- Al Najar M, Thoumyre G, Bergsma E W J, Almar R, Benshila R and Wilson D G. 2023. Satellite derived bathymetry using deep learning. *Machine Learning*, 112(4): 1107-1130 [DOI: 10.1007/s10994-021-05977-w]
- AlDahoul N, Abdul Karim H, De Castro A and Toledo Tan M J. 2022a. Localization and classification of space objects using EfficientDet detector for space situational awareness. *Scientific Reports*, 12(1): #21896 [DOI: 10.1038/s41598-022-25859-y]
- AlDahoul N, Karim H A and Momo M A. 2022b. RGB-D based multi-modal deep learning for spacecraft and debris recognition. *Scientific Reports*, 12(1): #3924 [DOI: 10.1038/s41598-022-07846-5]
- Alsabhan W and Alotaiby T. 2022. Automatic building extraction on satellite images using unet and ResNet50. *Computational Intelligence and Neuroscience*, 2022: #5008854 [DOI: 10.1155/2022/5008854]
- Alzubaidi L, Zhang J L, Humaidi A J, Al-Dujaili A, Duan Y, Al-

- Shamma O, et al. 2021. Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. *Journal of Big Data*, 8(1): #53 [DOI: 10.1186/s40537-021-00444-8]
- Barbier T and Gao Y. 2023. Relative visual navigation around an unknown and uncooperative spacecraft. *Acta Astronautica*, 206: 144-155 [DOI: 10.1016/j.actaastro.2023.02.022]
- Black K, Shankar S, Fonseka D, Deutsch J, Dhir A and Akella M R. 2021. Real-time, flight-ready, non-cooperative spacecraft pose estimation using monocular imagery [EB/OL]. [2025-09-17]. <http://arxiv.org/pdf/2101.09553.pdf>
- Bonnal C, Ruault J M and Desjean M C. 2013. Active debris removal: recent progress and current trends. *Acta Astronautica*, 85: 51-60 [DOI: 10.1016/j.actaastro.2012.11.009]
- Cao Y and Cheng X H. 2023. An end-to-end pose estimation network for multiscale space non-cooperative objects//Proceedings of the 8th International Conference on Control and Robotics Engineering (ICCRE). Niigata, Japan: IEEE: 28-33 [DOI: 10.1109/ICCRE57112.2023.10155575]
- Cao Y, Mu J Z, Cheng X H and Liu F Y. 2024. Spacecraft-DS: a spacecraft dataset for key components detection and segmentation via hardware-in-the-loop capture. *IEEE Sensors Journal*, 24(4): 5347-5358 [DOI: 10.1109/JSEN.2023.3347584]
- Cassinis L P, Menicucci A, Gill E, Ahrns I and Sanchez-Gestido M. 2022. On-ground validation of a CNN-based monocular pose estimation system for uncooperative spacecraft: bridging domain shift in rendezvous scenarios. *Acta Astronautica*, 196: 123-138 [DOI: 10.1016/j.actaastro.2022.04.002]
- Cassinis L P, Fonod R, Gill E, Ahrns I and Gil-Fernández J. 2021. Evaluation of tightly- and loosely-coupled approaches in CNN-based pose estimation systems for uncooperative spacecraft. *Acta Astronautica*, 182: 189-202 [DOI: 10.1016/j.actaastro.2021.01.035]
- Chang L, Liu J X, Chen Z, Bai J and Shu L Z. 2021. Stereo vision-based relative position and attitude estimation of non-cooperative spacecraft. *Aerospace*, 8(8): #230 [DOI: 10.3390/aerospace8080230]
- Chen S P, Yang W Y, Wang W, Mai J, Liang J and Zhang X H. 2024. Spacecraft homography pose estimation with single-stage deep convolutional neural network. *Sensors*, 24(6): #1828 [DOI: 10.3390/s24061828]
- Chen W, Sun Q W, Hu F C and Chao J G. 2023. Scene understanding technology for astronaut training based on deep learning object detection. *Manned Spaceflight*, 29(2): 143-149 (陈炜, 孙庆伟, 胡福超, 晁建刚. 2023. 基于深度学习目标识别的航天员训练场景理解技术. *载人航天*, 29(2): 143-149) [DOI: 10.16329/j.cnki.zrht.2023.02.006]
- Chen Y L, Gao J M and Zhang K B. 2020. R-CNN-based satellite components detection in optical images. *International Journal of Aerospace Engineering*, 2020: #8816187 [DOI: 10.1155/2020/8816187]
- Chen Z H, Lin J Q, Wang Y B, Liang B Y and Wang Y. 2021. Grasping position recognition of spatial non-cooperative targets based on active light detection. *Journal of Physics: Conference Series*, 2029(1): #012130 [DOI: 10.1088/1742-6596/2029/1/012130]
- Chu S H, Duan Y X, Schilling K and Wu S F. 2023. Monocular 6-DoF pose estimation for non-cooperative spacecrafts using riemannian regression network//Karlinsky L, Michaeli T, Nishino K, eds. *Computer Vision — ECCV 2022 Workshops*. Cham, Germany: Springer: 186-198 [DOI: 10.1007/978-3-031-25056-9_13]
- Cosmas K and Kenichi A. 2020. Utilization of FPGA for onboard inference of landmark localization in CNN-based spacecraft pose estimation. *Aerospace*, 7(11): #159 [DOI: 10.3390/aerospace7110159]
- Dang T, Tranzatto M, Khattak S, Mascarich F, Alexis K and Hutter M. 2020. Graph-based subterranean exploration path planning using aerial and legged robots. *Journal of Field Robotics*, 37(8): 1363-1388 [DOI: 10.1002/rob.21993]
- Davis J and Pernicka H. 2019. Proximity operations about and identification of non-cooperative resident space objects using stereo imaging. *Acta Astronautica*, 155: 418-425 [DOI: 10.1016/j.actaastro.2018.10.033]
- De Vittori A, Cipollone R, Di Lizia P and Massari M. 2022. Real-time space object tracklet extraction from telescope survey images with machine learning. *Astrodynamics*, 6(2): 205-218 [DOI: 10.1007/s42064-022-0134-4]
- Deng H, Zhang Z, Zhou D and Ma L Y. 2025. A multi-frame super-resolution reconstruction method based on optical flow for space noncooperative objects. *IEEE Access*, 13: 32918-32926 [DOI: 10.1109/ACCESS.2025.3541297]
- Deng L W, Suo H F, Jia Y Q and Huang C. 2022a. Pose estimation method for non-cooperative target based on deep learning. *Aerospace*, 9(12): #770 [DOI: 10.3390/aerospace9120770]
- Deng R R, Wang D Y, E W, Ge D M and Zhu W H. 2022b. Motion estimation of non-cooperative space objects based on monocular sequence images. *Applied Sciences*, 12(24): #12625 [DOI: 10.3390/app122412625]
- Deng Z J and Li A H. 2023. A large depth-of-field virtual measurement network for non-cooperative 6DOF pose estimation in occlusion scenes. *Measurement*, 218: #113111 [DOI: 10.1016/j.measurement.2023.113111]
- Dou K Y, Fan Y S and Wang T. 2023b. Pose estimation algorithm of spatial combined with polarized self-attention mechanism. *Aerospace Control*, 41(4): 48-53 (窦凯云, 樊永生, 王涛. 2023b. 结合极化自注意力机制的空间目标位姿估计. *航天控制*, 41(4): 48-53) [DOI: 10.16804/j.cnki.issn1006-3242.2023.04.011]
- Du H, Hu H D, Wang D Y, Xie X R, Xu S F, Wei C L, et al. 2023. Autonomous measurement and semantic segmentation of non-cooperative targets with deep convolutional neural networks. *Journal of Ambient Intelligence and Humanized Computing*, 14(6): 6959-6973 [DOI: 10.1007/s12652-021-03553-7]

- Fang Y Q, Yap P T, Lin W L, Zhu H T and Liu M X. 2024. Source-free unsupervised domain adaptation: a survey. *Neural Networks*, 174: #106230 [DOI: 10.1016/j.neunet.2024.106230]
- Federici L, Benedikter B and Zavoli A. 2021. Deep learning techniques for autonomous spacecraft guidance during proximity operations. *Journal of Spacecraft and Rockets*, 58(6): 1774-1785 [DOI: 10.2514/1.A35076]
- Gao H, Li Z H, Wang N, Yang J F and Dang D P. 2023. SU-Net: pose estimation network for non-cooperative spacecraft on-orbit. *Scientific Reports*, 13(1): #11780 [DOI: 10.1038/s41598-023-38974-1]
- Gavilanez G and Moncayo H. 2024. Vision-based relative position and attitude determination of non-cooperative spacecraft using a generative model architecture. *Acta Astronautica*, 225: 131-140 [DOI: 10.1016/j.actaastro.2024.07.039]
- Gong B C, Ma Y Q, Zhang W F, Li S and Li X Q. 2023. Deep-neural-network-based angles-only relative orbit determination for space non-cooperative target. *Acta Astronautica*, 204: 552-567 [DOI: 10.1016/j.actaastro.2022.09.024]
- Guthrie B, Kim M, Urrutxua H and Hare J. 2022. Image-based attitude determination of co-orbiting satellites using deep learning technologies. *Aerospace Science and Technology*, 120: #107232 [DOI: 10.1016/j.ast.2021.107232]
- Han B, Wang C X, Zhang X Y, Zhao Z B, Zhai Z, Liu J X, et al. 2025. Pose estimation and neural implicit reconstruction toward noncooperative spacecraft without offline prior information. *IEEE Transactions on Aerospace and Electronic Systems*, 61(2): 2612-2630 [DOI: 10.1109/TAES.2024.3479199]
- Hao Q, Li J, Zhang M and Wang L. 2022. Spatial non-cooperative target components recognition algorithm based on improved YOLOv3. *Computer Science*, 49(S1): 358-362 (郝强, 李杰, 张曼, 王路). 2022. 基于改进YOLOv3的空间非合作目标部件识别算法. *计算机科学*, 49(S1): 358-362 [DOI: 10.11896/jsjx.210700048]
- He F X, Liu T L and Tao D C. 2020. Why ResNet works? Residuals generalize. *IEEE Transactions on Neural Networks and Learning Systems*, 31(12): 5349-5362 [DOI: 10.1109/TNNLS.2020.2966319]
- Hematulin W, Kamsing P, Phisannupawong T, Panyalert T, Manut-hasna S, Torteeka P, et al. 2025. Generating large-scale datasets for spacecraft pose estimation via a high-resolution synthetic image renderer. *Aerospace*, 12(4): #334 [DOI: 10.3390/aerospace12040334]
- Hou C Y, Zhang R Z, Yang K Z, Li X Y, Yang Y, Ma X, et al. 2023. Non-cooperative target attitude estimation method based on deep learning of ground and space access scene radar images. *Mathematics*, 11(3): #745 [DOI: 10.3390/math11030745]
- Hu F Y, Li C L, Zhou T, Cheng H F and Gu M M. 2025. A survey on point cloud completion algorithms for deep learning. *Journal of Image and Graphics*, 30(2): 309-333 (胡伏原, 李晨露, 周涛, 程洪福, 顾敏明). 2025. 面向深度学习的三维点云补全算法综述. *中国图象图形学报*, 30(2): 309-333 [DOI: 10.11834/jig.240124]
- Hu J, Liu J H, Liu S L, Liu J S, Wang J X and Sun L S. 2025. Pose estimation method for spacecraft pipe based on adaptive convolution and latent representation. *Journal of Mechanical Engineering*, 61(14): 150-165 (胡佳, 刘检华, 刘少丽, 刘金山, 王家修, 孙连胜). 2025. 基于自适应卷积和潜在表征的航天器管路位姿估计方法. *机械工程学报*, 61(14): 150-165 [DOI: 10.3901/JME.2025.14.150]
- Hu L, Sun D Q, Duan H X, Shu A, Zhou S S and Pei H D. 2023. Non-cooperative spacecraft pose measurement with binocular camera and TOF camera collaboration. *Applied Sciences*, 13(3): #1420 [DOI: 10.3390/app13031420]
- Hu Y L, Fua P, Wang W and Salzmann M. 2020. Single-stage 6D object pose estimation//*Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Seattle, USA: IEEE: 2927-2936 [DOI: 10.1109/CVPR42600.2020.00300]
- Huang C, Cao T Z and Huang J L. 2022. Autonomous control of the large-angle spacecraft maneuvers in a non-cooperative mission. *Sensors*, 22(22): #8586 [DOI: 10.3390/s22228586]
- Huang H R, Song B, Zhao G P and Bo Y M. 2023. End-to-end monocular pose estimation for uncooperative spacecraft based on direct regression network. *IEEE Transactions on Aerospace and Electronic Systems*, 59(5): 5378-5389 [DOI: 10.1109/TAES.2023.3256971]
- Huang H R, Zhao G P, Gu D Q and Bo Y M. 2021. Non-model-based monocular pose estimation network for uncooperative spacecraft using convolutional neural network. *IEEE Sensors Journal*, 21(21): 24579-24590 [DOI: 10.1109/JSEN.2021.3115844]
- Huo Y R, Li Z and Zhang F. 2020. Fast and accurate spacecraft pose estimation from single shot space imagery using box reliability and keypoints existence judgments. *IEEE Access*, 8: 216283-216297 [DOI: 10.1109/ACCESS.2020.3041415]
- Jiang L H, Zheng X J, Yang Y F, Zhao Y, Zhang X W and Sun S D. 2022. Pose estimation of ultra-short range space non-cooperative target based on TOF camera. *Aerospace Control and Application*, 48(3): 63-71 (姜丽辉, 郑循江, 杨逸峰, 赵畅, 张徐玮, 孙朔冬). 2022. 基于TOF相机的超近程空间非合作目标的位姿估计. *空间控制技术与应用*, 48(3): 63-71 [DOI: 10.3969/j.issn.1674-1579.2022.03.008]
- Jin Z M, Wang L, Liu H H, Du R H and Zhang X. 2020. Monocular-based pose estimation of non-cooperative space targets using EKF and EKPF//*Proceedings of the 13th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI)*. Chengdu, China: IEEE: 46-51 [DOI: 10.1109/CISP-BMEI51763.2020.9263615]
- Jin Z M, Wang L, Liu K, Du R H and Zhang X. 2021. Monocular-based pose estimation of non-cooperative space targets using EKF and EKPF. *Journal of Astronautics*, 42(7): 907-916 (金泽明, 汪玲,

- 刘柯, 杜荣华, 张翔. 2021. 联合EKF和EKPF的空间非合作目标单目位姿估计. 宇航学报, 42(7): 907-916 [DOI: 10.3873/j.issn.1000-1328.2021.07.011]
- Jo B and Kim S. 2022. Comparative analysis of OpenPose, PoseNet, and MoveNet models for pose estimation in mobile devices. *Traitement Du Signal*, 39(1): 119-124 [DOI: 10.18280/ts.390111]
- Kajak K M, Maddock C, Frei H and Schwenk K. 2023. Domain randomisation and CNN-based keypoint-regressing pose initialisation for relative navigation with uncooperative finite-symmetric spacecraft targets using monocular camera images. *Advances in Space Research*, 72(7): 2824-2844 [DOI: 10.1016/j.asr.2023.02.024]
- Lei J Q, Wang J P, Shi J C, Xu G L and Cheng Y H. 2024. Configuration optimization method of cooperative target for pose estimation with monocular vision. *Optical Engineering*, 63(2): #023102 [DOI: 10.1117/1.OE.63.2.023102]
- Li C, Zhao G P, Gu D Q and Wang Z B. 2023. Improved lightweight YOLOv5 using attention mechanism for satellite components recognition. *IEEE Sensors Journal*, 23(1): 514-526 [DOI: 10.1109/JSEN.2022.3222868]
- Li J, Zhuang Y Q, Peng Q and Zhao L. 2021. Pose estimation of non-cooperative space targets based on cross-source point cloud fusion. *Remote Sensing*, 13(21): #4239 [DOI: 10.3390/rs13214239]
- Li K C, Zhang H P and Hu C Y. 2022a. Learning-based pose estimation of non-cooperative spacecrafts with uncertainty prediction. *Aerospace*, 9(10): #592 [DOI: 10.3390/aerospace9100592]
- Li Y H, Yan Y H, Xiu X C and Miao Z H. 2022b. An uncertainty weighted non-cooperative target pose estimation algorithm, based on intersecting vectors. *Aerospace*, 9(11): #681 [DOI: 10.3390/aerospace9110681]
- Li Y X, Huo J, Ma P and Jiang R Y. 2022c. Target localization method of non-cooperative spacecraft on on-orbit service. *Chinese Journal of Aeronautics*, 35(11): 336-348 [DOI: 10.1016/j.cja.2022.04.001]
- Lin B, Wang J, Wang H, Zhong L J, Yang X and Zhang X H. 2023. Small space target detection based on a convolutional neural network and guidance information. *Aerospace*, 10(5): #426 [DOI: 10.3390/aerospace10050426]
- Liu F C and Zhang X. 2021. Research on solar panel recognition method of space non-cooperative target spacecraft. *Chinese High Technology Letters*, 31(6): 628-638 (刘福才, 张晓. 2021. 空间非合作目标航天器太阳帆板识别方法研究. 高技术通讯, 31(6): 628-638) [DOI: 10.3772/j.issn.1002-0470.2021.06.007]
- Liu F Y, Zhang Z J and Li S J. 2024. DTSE-SpaceNet: deformable-transformer-based single-stage end-to-end network for 6-D pose estimation in space. *IEEE Transactions on Aerospace and Electronic Systems*, 60(3): 2555-2571 [DOI: 10.1109/TAES.2023.3332075]
- Liu J H, Zhang Y H and Zhang W X. 2024. Lightweight network-based end-to-end pose estimation for noncooperative targets. *Laser and Optoelectronics Progress*, 61(14): #1437010 (刘佳辉, 张永合, 张文秀. 2024. 轻量化的端到端非合作目标位姿估计方法. 激光与光电子学进展, 61(14): #1437010 [DOI: 10.3788/LOP232418])
- Liu S L, Zhu X W, Cao Z W and Wang G. 2022a. Deep 1D landmark representation learning for space target pose estimation. *Remote Sensing*, 14(16): #4035 [DOI: 10.3390/rs14164035]
- Liu X, Wang H Y, Chen X L, Chen W C and Xie Z Y. 2023. Position awareness network for noncooperative spacecraft pose estimation based on point cloud. *IEEE Transactions on Aerospace and Electronic Systems*, 59(1): 507-518 [DOI: 10.1109/TAES.2022.3182307]
- Liu Y P, Zhou X Y and Han H W. 2022b. Lightweight CNN-based method for spacecraft component detection. *Aerospace*, 9(12): #761 [DOI: 10.3390/aerospace9120761]
- Lopatka A. 2018. Satellite repair and refueling demos may soon spur a new industry. *Physics Today*, 71(11): 26-29 [DOI: 10.1063/PT.3.4066]
- Mu J Z, Hao X L, Zhu W S and Li S. 2021. Review and prospect of intelligent perception for non-cooperative targets. *Chinese Space Science and Technology*, 41(6): 1-16 (牟金震, 郝晓龙, 朱文山, 李爽. 2021. 非合作目标智能感知技术研究进展与展望. 中国空间科学技术, 41(6): 1-16) [DOI: 10.16708/j.cnki.1000-758X.2021.0076]
- Park T H and D'Amico S. 2023. Adaptive neural-network-based unscented Kalman filter for robust pose tracking of noncooperative spacecraft. *Journal of Guidance, Control, and Dynamics*, 46(9): 1671-1688 [DOI: 10.2514/1.G007387]
- Park T H and D'Amico S. 2024. Robust multi-task learning and online refinement for spacecraft pose estimation across domain gap. *Advances in Space Research*, 73(11): 5726-5740 [DOI: 10.1016/j.asr.2023.03.036]
- Park T H, Märtens M, Lecuyer G, Izzo D and D'Amico S. 2022. SPEED+: next-generation dataset for spacecraft pose estimation across domain gap//*Proceedings of 2022 IEEE Aerospace Conference (AERO)*. Big Sky, USA: IEEE: 1-15 [DOI: 10.1109/AERO53065.2022.9843439]
- Pauly L, Rharbaoui W, Shneider C, Rathinam A, Gaudillière V and Aouada D. 2023. A survey on deep learning-based monocular spacecraft pose estimation: current state, limitations and prospects. *Acta Astronautica*, 212: 339-360 [DOI: 10.1016/j.actaastro.2023.08.001]
- Pensado E A, De Santos L M G, Jorge H G and Sanjurjo-Rivo M. 2023. Deep learning-based target pose estimation using LiDAR measurements in active debris removal operations. *IEEE Transactions on Aerospace and Electronic Systems*, 59(5): 5658-5670 [DOI: 10.1109/TAES.2023.3262505]
- Pérez-Villar J I B, García-Martín Á and Bescós J. 2023. Spacecraft pose estimation based on unsupervised domain adaptation and on a 3D-

- guided loss combination//Proceedings of ECCV 2022 Workshops. Cham, Germany: Springer: 37-52 [DOI: 10.1007/978-3-031-25056-9_3]
- Phisannupawong T, Kamsing P, Torteeka P, Channumsin S, Sawangwit U, Hematulin W, et al. 2020. Vision-based spacecraft pose estimation via a deep convolutional neural network for noncooperative docking operations. *Aerospace*, 7(9): #126 [DOI: 10.3390/aerospace7090126]
- Piazza M, Maestrini M and Di Lizia P. 2022. Monocular relative pose estimation pipeline for uncooperative resident space objects. *Journal of Aerospace Information Systems*, 19(9): 613-632 [DOI: 10.2514/1.1011064]
- Proença P F and Gao Y. 2020. Deep learning for spacecraft pose estimation from photorealistic rendering//Proceedings of 2020 IEEE International Conference on Robotics and Automation (ICRA). Paris, France: IEEE: 6007-6013 [DOI: 10.1109/ICRA40945.2020.9197244]
- Qiao S J, Zhang H P, Meng G, An M, Xie F Y and Jiang Z G. 2022. Deep-learning-based satellite relative pose estimation using monocular optical images and 3D structural information. *Aerospace*, 9(12): #768 [DOI: 10.3390/aerospace9120768]
- Qiao S J, Zhang H P, Xie F Y and Jiang Z G. 2024. Deep-learning-based direct attitude estimation for uncooperative known space objects. *IEEE Transactions on Aerospace and Electronic Systems*, 60(3): 2526-2541 [DOI: 10.1109/TAES.2023.3325801]
- Rathinam A, Gaudilliere V, Pauly L and Djamil A. 2022. AKM dataset: textureless space target dataset. [s.l.]: [s.n.] [DOI: 10.5281/zenodo.7043325]
- Renaut L, Frei H and Nüchter A. 2025a. Deep learning on 3D point clouds for fast pose estimation during satellite rendezvous. *Acta Astronautica*, 232: 231-243 [DOI: 10.1016/j.actaastro.2025.03.009]
- Renaut L, Frei H and Nüchter A. 2025b. CNN-based pose estimation of a noncooperative spacecraft with symmetries from LiDAR point clouds. *IEEE Transactions on Aerospace and Electronic Systems*, 61(2): 5002-5016 [DOI: 10.1109/TAES.2024.3517574]
- Rondao D, Aouf N and Richardson M A. 2023. ChiNet: deep recurrent convolutional learning for multimodal spacecraft pose estimation. *IEEE Transactions on Aerospace and Electronic Systems*, 59(2): 937-949 [DOI: 10.1109/TAES.2022.3193085]
- Sharma S and D'Amico S. 2019. Pose estimation for non-cooperative rendezvous using neural networks [EB/OL]. [2025-09-17]. <http://arxiv.org/pdf/1906.09868.pdf>
- Sharma S and D'Amico S. 2016. Comparative assessment of techniques for initial pose estimation using monocular vision. *Acta Astronautica*, 123: 435-445 [DOI: 10.1016/j.actaastro.2015.12.032]
- Shu A, Pei H D and Duan H X. 2021. Trinocular stereo visual measurement method for spatial non-cooperative target. *Acta Optica Sinica*, 41(6): #0615001 (束安, 裴浩东, 段慧仙. 2021. 空间非合作目标的三目立体视觉测量方法. *光学学报*, 41(6): #0615001) [DOI: 10.3788/AOS202141.0615001]
- Su D, Zhang C, Wang K and Sun K. 2023. A two-stage pose estimation method for noncooperative targets based on convolution neural network. *Transactions of Beijing Institute of Technology*, 43(7): 734-743 (苏迪, 张成, 王柯, 孙凯. 2023. 基于卷积神经网络的非合作目标两阶段位姿估计方法. *北京理工大学学报*, 43(7): 734-743) [DOI: 10.15918/j.tbit1001-0645.2022.185]
- Su Y, Zhang Z X, Yuan M M and Wang Y S. 2023. Flow-based 6D pose tracking of uncooperative spacecrafts. *International Journal of Aerospace Engineering*, 2023: #9631895 [DOI: 10.1155/2023/9631895]
- Sun D Q, Duan H X, Pei H D and Hu L. 2021. Pose measurement method of space non-Cooperative targets based on TOF camera. *Acta Optica Sinica*, 41(22): #2212003 (孙典奇, 段慧仙, 裴浩东, 胡亮. 2021. 基于TOF相机的空间非合作目标位姿测量方法. *光学学报*, 41(22): #2212003) [DOI: 10.3788/AOS202141.2212003]
- Sun Q, Pan X H, Ling X, Wang B, Sheng Q H, Li J, et al. 2023. A vision-based pose estimation of a non-cooperative target based on a self-supervised transformer network. *Aerospace*, 10(12): #997 [DOI: 10.3390/aerospace10120997]
- Szczuko P. 2019. Deep neural networks for human pose estimation from a very low resolution depth image. *Multimedia Tools and Applications*, 78(20): 29357-29377 [DOI: 10.1007/s11042-019-7433-7]
- Wang J C, Yang H F, Song W L, Tang P R and Yu Z C. 2025. Satellite 6D position estimation method based on improved DenseFusion. *Laser Journal*, 46(3): 161-168 (王金聪, 杨海峰, 宋文龙, 汤普然, 于志超. 2025. 基于改进DenseFusion的卫星6D位姿估计方法. *激光杂志*, 46(3): 161-168) [DOI: 10.14016/j.cnki.jgz.2025.03.161]
- Wang L, Bai X R, Xue R H and Zhou F. 2021. Few-shot SAR automatic target recognition based on Conv-BiLSTM prototypical network. *Neurocomputing*, 443: 235-246 [DOI: 10.1016/j.neucom.2021.03.03]
- Wang M H, Feng D Z, Wang H, Guo H H, Dai P, Cui J S, et al. 2025. Improved lightweight saliency model based on neural network for noncooperative spacecraft detection. *IEEE Transactions on Aerospace and Electronic Systems*, 61(1): 642-654 [DOI: 10.1109/TAES.2024.3453218]
- Wang R R, Luo H Y, Wang Q, Li Z H, Zhao F and Huang J Y. 2020. A spatial-temporal positioning algorithm using residual network and LSTM. *IEEE Transactions on Instrumentation and Measurement*, 69(11): 9251-9261 [DOI: 10.1109/TIM.2020.2998645]
- Wang S C, Hua B C, Zhou Y E and Li X L. 2024. Pose estimation of space non-cooperative target based on TEASER algorithm. *Aerospace Control and Application*, 50(1): 25-34 (王世昌, 华宝成, 周依尔, 李小路. 2024. 基于TEASER算法的空间非合作目标位姿估计. *空间控制技术与应用*, 50(1): 25-34) [DOI: 10.3969/j.issn.1674-1579.2024.01.004]

- Wang Y S, Xu T L, Zhang Z X and Su Y. 2023. Relative pose measurement of space target with symmetrical structure based on point cloud deep learning. *Journal of Astronautics*, 44(2): 294-302 (王艺诗, 徐田来, 张泽旭, 苏宇. 2023. 基于点云深度学习的对称结构空间目标相对位姿测量. *宇航学报*, 44(2): 294-302) [DOI: 10.3873/j.issn.1000-1328.2023.02.013]
- Wang Z, Chen M L, Guo Y L, Li Z and Yu Q F. 2024a. Bridging the domain gap in satellite pose estimation: a self-training approach based on geometrical constraints. *IEEE Transactions on Aerospace and Electronic Systems*, 60(3): 2500-2514 [DOI: 10.1109/TAES.2023.3250385]
- Wang Z, Sun X L, Li Z, Cheng Z L and Yu Q F. 2022. Transformer based monocular satellite pose estimation. *Acta Aeronautica et Astronautica Sinica*, 43(5): #325298 (王梓, 孙晓亮, 李璋, 程子龙, 于起峰. 2022. 基于Transformer模型的卫星单目位姿估计方法. *航空学报*, 43(5): #325298) [DOI: 10.7527/S1000-6893.2021.25298]
- Wang Z, Wang J H, Li Y, Li Z and Yu Q F. 2024. Non-cooperative target pose estimation from monocular images based on lightweight neural network. *Acta Aeronautica et Astronautica Sinica*, 45(22): #330248 (王梓, 王靖皓, 李杨, 李璋, 于起峰. 2024. 基于轻量级神经网络的非合作目标位姿单目测量. *航空学报*, 45(22): #330248) [DOI: 10.7527/S1000-6893.2024.30248]
- Wang Z, Zhang Z, Sun X L, Li Z and Yu Q F. 2022. Revisiting monocular satellite pose estimation with transformer. *IEEE Transactions on Aerospace and Electronic Systems*, 58(5): 4279-4294 [DOI: 10.1109/TAES.2022.3161605]
- Wang Z H, Liu Y M and Zhang E. 2024b. Pose estimation for cross-domain non-cooperative spacecraft based on spatial-aware key-points regression. *Aerospace*, 11(11): #948 [DOI: 10.3390/aerospace11110948]
- Wei Z Z, Feng G K, Zhou D Y, Ma Y M, Liu M K, Luo Q F, et al. 2023. A review of position and orientation visual measurement methods and applications. *Laser and Optoelectronics Progress*, 60(3): #0312010 (魏振忠, 冯广堃, 周丹雅, 马岳鸣, 刘明坤, 罗启峰, 等. 2023. 位姿视觉测量方法及应用综述. *激光与光电子学进展*, 60(3): #0312010 [DOI: 10.3788/LOP223420])
- Xiang Y B, Xi J B, Cong M, Yang Y, Ren C F and Han L. 2020. Space debris detection with fast grid-based learning//*Proceedings of the 3rd IEEE International Conference of Safe Production and Informationization (ICSPI)*. Chongqing, China: IEEE: 205-209 [DOI: 10.1109/ICSPI51290.2020.9332372]
- Xiao H, Zhu C R, Sheng Q H, Wang B, Li J, Ling X, et al. 2024. Parameter precise estimation technology of active segment of non-cooperative targets based on long short-term memory. *Space: Science and Technology*, 4: #0150 [DOI: 10.34133/space.0150]
- Xu J J, Song B, Yang X and Nan X T. 2020. An improved deep key-point detection network for space targets pose estimation. *Remote Sensing*, 12(23): #3857 [DOI: 10.3390/rs12233857]
- Xu Y F, Zhang D Z, Wang L and Hua B C. 2020. Lightweight feature fusion network design for local feature recognition of non-cooperative target. *Infrared and Laser Engineering*, 49(7): #20200170 (徐云飞, 张笃周, 王立, 华宝成. 2020. 非合作目标局部特征识别轻量化特征融合网络设计. *红外与激光工程*, 49(7): #20200170) [DOI: 10.3788/IRLA20200170]
- Yang H, Xiao X M, Yao M B, Xiong Y G, Cui H T and Fu Y G. 2024. PVSPE: a pyramid vision multitask transformer network for spacecraft pose estimation. *Advances in Space Research*, 74(3): 1327-1342 [DOI: 10.1016/j.asr.2024.05.011]
- Yang H, Zhang Y S and Xu C. 2020. Space target recognition based on trace feature of ISAR image. *Acta Electronica Sinica*, 48(3): 431-441 (杨虹, 张雅声, 徐灿. 2020. 基于trace特征的ISAR像空间目标识别. *电子学报*, 48(3): 431-441) [DOI: 10.3969/j.issn.0372-2112.2020.03.003]
- Yang X, Nan X T and Song B. 2020a. D2N4: a discriminative deep nearest neighbor neural network for few-shot space target recognition. *IEEE Transactions on Geoscience and Remote Sensing*, 58(5): 3667-3676 [DOI: 10.1109/TGRS.2019.2959838]
- Yang X, Wu T, Wang N N, Huang Y, Song B and Gao X B. 2020b. HCNN-PSI: a hybrid CNN with partial semantic information for space target recognition. *Pattern Recognition*, 108: #107531 [DOI: 10.1016/j.patcog.2020.107531]
- Ye R D, Ren Y, Zhu X Y, Wang Y J, Liu M Y and Wang L F. 2023. An efficient pose estimation algorithm for non-cooperative space objects based on dual-channel transformer. *Remote Sensing*, 15(22): #5278 [DOI: 10.3390/rs15225278]
- Yuan M M and Zhang Z X. 2023. A recognition method for components of space target based on bi-modal fusion of 3D point cloud and image. *Journal of Astronautics*, 44(5): 796-804 (袁萌萌, 张泽旭. 2023. 三维点云与图像双模态融合的空间目标部件识别方法. *宇航学报*, 44(5): 796-804) [DOI: 10.3873/j.issn.1000-1328.2023.05.014]
- Zeng H Y and Xia Y. 2017. Space target recognition based on deep learning//*Proceedings of the 20th International Conference on Information Fusion (Fusion)*. Xi'an, China: IEEE: 1-5. [DOI: 10.23919/ICIF.2017.8009786]
- Zhang Y, Wang J Y, Chen J B, Shi D H and Chen X T. 2024a. A space non-cooperative target recognition method for multi-satellite cooperative observation systems. *Remote Sensing*, 16(18): #3368 [DOI: 10.3390/rs16183368]
- Zhang Y W, Zhang Y, Ding Z H and Wang Z J. 2024b. Classification and recognition method of non-cooperative object based on transfer learning. *Optics and Laser Technology*, 169: #110005 [DOI: 10.1016/j.optlastec.2023.110005]
- Zhang Z, Hu Y H, Zhou D, Wu L G, Yao W R and Li P. 2024c. DMA-Net: dense multi-scale attention network for space non-cooperative object pose estimation. *Transactions of Nanjing University of Aeronautics and Astronautics*, 41(1): 122-134 [DOI: 10.16356/j.

1005-1120.2024.01.010]

Zhang Z X, Xu Y L and Song J N. 2025. Robust pose estimation for non-cooperative space objects based on multichannel matching method. Scientific Reports, 15 (1) : #4940 [DOI: 10.1038/s41598-025-89544-6]

Zhao G Y, Wan X, Tian Y L, Shao Y D and Li S Y. 2022. 3D component segmentation network and dataset for non-cooperative spacecraft. Aerospace, 9(5): #248 [DOI: 10.3390/aerospace9050248]

Zhong L J, Chen S P, Wang W, Yang W Y, Yuan G X, Guo P Y, et al. 2025. Uncooperative spacecraft pose estimation with normalized segmentation coordinate space. IEEE/ASME Transactions on Mechatronics, 30 (3) : 2293-2304 [DOI: 10.1109/TMECH.2024.3442570]

Zhu W S, Mu J Z, Li S and Han F. 2023. Review and prospect of spacecraft pose estimation based on deep learning. Journal of Astronautics, 44(11): 1633-1644 (朱文山, 牟金震, 李爽, 韩飞. 2023. 基于深度学习的航天器位姿估计研究进展. 宇航学报, 44(11): 1633-1644). [DOI: 10.3873/j.issn.1000-1328.2023.11.002]

Zhu Z J, Xiang W H, Huo J, Yang M, Zhang G Y and Wei L. 2022. Non-cooperative target pose estimation based on improved iterative closest point algorithm. Journal of Systems Engineering and Electronics, 33(1): 1-10 [DOI: 10.23919/JSEE.2022.000001]

Zhuang H X, Shi W, Luo Y F, Zhang Q, Wang Y, Wu R F, et al. 2024. Design on relative pose measurement system for close range manipulation of space target. Spacecraft Engineering, 33(5): 22-28 (庄海孝, 石伟, 罗毓芳, 张强, 王瑶, 吴茹菲, 等. 2024. 面向空间目标近距离操作的相对位姿测量系统设计. 航天器工程, 33(5): 22-28) [DOI: 10.3969/j.issn.1673-8748.2024.05.004]

作者简介

姜建湘,女,硕士研究生,主要研究方向为图像处理与目标检测。E-mail: jianxiang@nuaa.edu.cn

吴一全,通信作者,男,教授,博士生导师,主要研究方向为遥感图像处理与理解、目标检测与识别、视觉检测与图像测量、视频处理与智能分析。E-mail: nuaaimage@163.com