

中图法分类号: TP391.41; V279 文献标识码: A 文章编号: 1006-8961(XXXX)XX-0001-32

论文引用格式: Lyu Pin, Wu Yiquan. A Review of Low-Altitude UAV Detection and Tracking Methods Using Vision and Deep Learning[J/OL]. Journal of Image and Graphics, XXXX: 1-32. DOI: 10.11834/jig.260151. (吕品, 吴一全. 利用视觉和深度学习的低空无人机检测和跟踪方法综述[J/OL]. 中国图象图形学报, XXXX: 1-32. DOI: 10.11834/jig.260151. [DOI: 10.11834/jig.260151])

利用视觉和深度学习的低空无人机检测和跟踪方法综述

吕品, 吴一全*

南京航空航天大学电子信息工程学院, 南京 211106

摘要: 低空无人机的“黑飞”“乱飞”对公共安全构成严峻挑战, 基于视觉与深度学习的检测跟踪技术因此成为研究焦点。本文系统综述了近5年该领域进展。首先概述无人机类型及其视觉探测特点, 接着归纳检测跟踪任务面临的核心挑战。在检测方法部分, 重点从改进小目标检测、轻量化网络设计、抗恶劣天气及多模态融合四个方向, 梳理了深度学习技术的提升路径; 在跟踪方法部分, 则从鲁棒性增强、轻量化与实时性、长期跟踪与重识别、多目标跟踪四个维度总结了关键进展。此外, 整理了常用数据集与评价指标, 并评估了代表性方法性能。最后对未来研究方向进行了展望。

关键词: 反无人机; 低空经济; 无人机检测; 无人机跟踪; 小目标检测; 多目标跟踪

A Review of Low-Altitude UAV Detection and Tracking Methods Using Vision and Deep Learning

Lyu Pin, Wu Yiquan*

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

Abstract: With the rapid and unprecedented expansion of the global low-altitude economy, unmanned aerial vehicles (UAVs), commonly known as drones, have transitioned from niche technological tools into mainstream assets deeply embedded within the critical infrastructure of modern society. Their applications now span a remarkably diverse spectrum, including, but not limited to, precision logistics and last-mile delivery, emergency medical services and disaster response, automated inspection of energy grids, bridges, and wind turbines, agricultural monitoring and targeted spraying, as well as media production and aerial surveillance. This widespread integration delivers immense economic and social value by enhancing efficiency, reducing costs, and enabling access to previously unreachable areas. However, the very accessibility, affordability, and operational ease that fuel this adoption paradoxically introduce severe and escalating security vulnerabilities. Incidents involving unauthorized flights, negligent operations, and even the malicious use of drones for espionage, smuggling, or disruptive attacks pose tangible and growing threats to national security, public safety, the integrity of critical infrastructure such as airports and power plants, and individual privacy. Consequently, the development of robust,

收稿日期: 2026-03-25; 修回日期: 2026-09-16

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

基金项目: 国家自然科学基金项目(61573183); 南京航空航天大学研究生创新实践竞赛培育项目

Supported by: National Natural Science Foundation of China (61573183); Graduate Innovation Practice Competition Cultivation Project of Nanjing University of Aeronautics and Astronautics

reliable, and efficient technologies for the detection, tracking, and identification of low-altitude UAVs has emerged as a pressing global research and operational imperative across defense, security, and regulatory domains. From a computer vision perspective, low-altitude UAVs present a uniquely challenging set of characteristics, aptly summarized as "Low, Slow, and Small" (LSS) targets. The "small" signature is particularly problematic: at typical operational ranges for ground-based surveillance, a drone may occupy only a minuscule number of pixels in a captured image or video frame. This results in an extremely weak signal, with limited texture and shape information, making it easily lost amidst background noise or suppressed during the down-sampling stages of deep convolutional neural networks. The "low" and "slow" attributes further compound the difficulty. These UAVs operate within complex, cluttered, and dynamic environments—such as urban canyons, near buildings, over forests, or against varied terrain—where they are susceptible to frequent partial or complete occlusion. Furthermore, they must be reliably distinguished from a wide range of visual confusers, most notably birds, but also kites, balloons, and moving vegetation. This leads to the dual challenges of minimal inter-class differences and significant intra-class variation among UAVs themselves, which appear in diverse forms such as multi-rotors, fixed-wing aircraft, and hybrid vertical take-off and landing systems. The operational environment imposes additional stringent constraints: vision systems must maintain performance under drastic illumination changes and adverse weather conditions, while also meeting the demanding requirement for real-time, low-latency processing, often on computationally constrained and power-limited edge devices such as drones, surveillance towers, or mobile units. This comprehensive survey paper systematically reviews, analyzes, and synthesizes the state of the art in vision-based and deep learning-powered methodologies for the automated detection and tracking of low-altitude UAVs. The review begins by contextualizing the problem space, defining the operational envelope of "low altitude" as a challenging regime for sensors, and categorizing the primary types of UAVs along with their distinct visual and kinematic signatures. It then crystallizes the core technical challenges that any effective solution must address: (1) the fundamental problem of small target detection; (2) the discrimination problem arising from low inter-class variance and high intra-class variance; (3) robustness against complex, dynamic background interference and environmental degradation; and (4) the imperative for computational efficiency and real-time performance on edge platforms. The core of the survey is logically divided into two interconnected pillars: detection and tracking. For the detection task, we first provide a concise overview of traditional image processing paradigms, including background subtraction techniques, frame differencing, optical flow for motion-based detection, and hand-crafted feature descriptors such as histogram of oriented gradients (HOG) for appearance-based classification. Recognizing their limitations in generalization and robustness, the focus then shifts to modern deep learning-based detectors. This analysis is organized around four pivotal and active research thrusts. First, architectural innovations for small target detection explore specialized network designs, including enhanced feature pyramid networks for multi-scale context aggregation, the integration of attention mechanisms to amplify critical features, and the design of novel loss functions tailored for small object localization. Second, lightweight and efficient model design investigates techniques for deployment on edge devices, such as neural architecture search for optimal efficiency-accuracy trade-offs, model pruning, quantization, knowledge distillation, and the use of efficient operators like depthwise separable convolutions. Third, enhancing robustness under adverse conditions examines methods beyond standard red-green-blue (RGB) imagery, including multi-spectral fusion, the use of polarization cues, and emerging sensor modalities such as event cameras that offer high dynamic range and improved motion clarity. Fourth, multi-modal sensor fusion strategies are discussed, which synergistically combine complementary data from visual cameras, radar, light detection and ranging (LiDAR), and acoustic sensors to create a more resilient perception system. For the tracking task, following initial detection, the survey details the evolution from traditional correlation filter-based and kernelized methods to contemporary deep learning-driven trackers, which typically follow a Siamese network or transformer-based paradigm. Advancements are examined from four complementary perspectives: (1) improving robustness and accuracy, focusing on advanced motion modeling, robust feature representation learning against appearance changes, and effective scale estimation strategies; (2) enabling real-time performance on edge devices through model compression, efficient backbone architectures, and adaptive inference strategies; (3) mechanisms for long-term tracking and recovery, which are crucial for practical deployment, involving dedicated re-identification modules, global search strategies, and memory-augmented networks to handle occlusion and target disappearance; and (4) scaling to multi-UAV sce-

narios and swarms, which introduces challenges in data association, identity maintenance, and group behavior understanding. To ground the discussion and facilitate benchmarking, the survey provides a curated compilation of publicly available datasets specifically designed for UAV detection and tracking, categorized into visible-spectrum and infrared (IR) datasets, detailing their scale, attributes, and inherent challenges. It also explains the standard evaluation metrics prevalent in the field, such as average precision (AP), mean average precision (mAP), Success Rate, and Precision for tracking. A comparative analysis of representative state-of-the-art methods on key benchmarks is presented to provide a clear performance snapshot. Finally, the paper concludes by delineating several promising and critical future research directions. These frontiers include the development of more intelligent and adaptive multi-modal fusion frameworks that move beyond simple early or late fusion; the pursuit of generalizable and invariant feature learning through self-supervised learning, meta-learning, or domain adaptation techniques to handle novel drones and unseen environments; the design of environment-aware and actively adaptive algorithms capable of anticipating and compensating for perceptual degradation; the exploration of next-generation neuromorphic and efficient computing paradigms, such as spiking neural networks, for ultra-low-power deployment; and the deeper integration of novel sensing modalities, including event-based vision, to overcome the limitations of traditional frame-based imaging. This survey aims to serve as both a detailed technical reference and a strategic roadmap for researchers and engineers working to advance the field of intelligent, vision-based counter-UAV systems.

Key words: anti UAV; low altitude economy; detection of UAV; tracking of UAV; detection of small targets; tracking of multiple targets

论文引用格式: Lyu Pin, Wu Yiquan. A Review of Low-Altitude UAV Detection and Tracking Methods Using Vision and Deep Learning [J/OL]. Journal of Image and Graphics. DOI:10.11834/jig.260151. (吕品, 吴一全. 利用视觉和深度学习的低空无人机检测和跟踪方法综述 [J/OL]. 中国图象图形学报. DOI:10.11834/jig.260151.)

0 引言

随着低空经济的迅猛发展,无人机已深度融入应急救援、基础设施巡检、物流配送、农业植保、航拍摄影等诸多领域,展现出巨大的应用潜力与社会价值(Jiang等,2025)。然而,其便捷性与普适性也带来了日益严峻的“黑飞”和“乱飞”等安全隐患。未经授权或具有恶意的低空无人机可对重点区域安保、航空秩序、公共安全乃至国防安全构成现实威胁,这使得发展高效可靠的无人机监管与反制技术成为一项紧迫任务(Jie等,2021)。

与高空高速飞行器不同,低空无人机在复杂的城市或野外环境中飞行,对其探测与跟踪面临着前所未有的技术挑战。一方面,无人机目标本身具有“低、慢、小”的典型特征,其雷达散射截面积小,在光学图像中占比低,且运动轨迹多变,传统针对大型飞

行器的探测手段往往存在性能短板或盲区(Sun等,2021)。另一方面,密集建筑物遮挡、动态变化的自然背景、与飞鸟等干扰物难以区分、以及光照和天气影响等诸多复杂环境因素,极易导致探测系统出现高虚警或漏检(李栋梁等,2025)。因此,如何在复杂环境下实现对此类无人机的鲁棒感知,已成为当前反无人机技术领域研究前沿与难点。

在无人机探测技术中,基于视觉传感器的解决方案具有显著的实用优势:摄像头硬件成本低、产业链成熟;可获取包含RGB色彩、纹理细节乃至红外热成像的多维度视觉信息,为目标检测与跟踪提供丰富数据基础;易于与城市现有的安防监控网络无缝集成(Amin等,2023)。此外,视觉系统采用被动感知机制,不主动发射信号,既不会对周边电磁环境造成干扰,也避免了对无人机通信链路的依赖,适用于探测静默飞行的目标。在视觉传感器为无人机探测提供丰富数据的基础上,深度学习方法以强大的自主学习与融合优化能力成为提升检测跟踪性能的核心引擎(高广帅等,2025)。

针对低空无人机目标及其所处的复杂动态背景,如何实现快速、精准的视觉检测与稳定跟踪,受到相关研究者的普遍关注并发表了有关综述。罗俊海和王芝燕(2022)总结了无人机探测与对抗技术的国内外进展,然而该综述对技术优劣的对比多基于

定性描述,缺乏定量指标。Wang 等人(2024)回顾了基于视觉的反无人机探测与跟踪方法,然而该综述大量涉及传感器基础技术,导致视觉与非视觉方法的边界模糊。Xu 和 Luo(2025)综述了基于视觉的反无人机检测与跟踪技术,然而该研究并未梳理相关数据集。陈唯实等人(2025)对机场净空区飞鸟与无人机多源探测技术的最新进展进行了系统综述,全面分析了雷达、光电、无线电侦测和声学等传感器的原理、典型系统及基于深度学习的目标识别算法。李旻姝等人(2025)系统回顾了多模态融合技术在无人机检测与识别领域的最新进展,重点探讨了雷达、光电、无线电侦测、声学等传感器的集成方法。以上两篇综述虽然包含基于视觉的低空无人机检测和跟踪方法相关内容,但其占比很小。

鉴于此,本文将系统分类与综述基于视觉和深度学习的低空无人机检测和跟踪方法,梳理相关数据集和评价指标,并对代表性方法的性能进行评估。

1 低空无人机的种类和特点

1.1 低空的定义

在低空无人机探测与跟踪语境下,“低空”是融合空域管理与技术挑战的综合概念。从空域管理看,国际民航组织及各国机构制定低空标准,如图1

所示,此范围是无人机等“低慢小”航空器主要运行区,特征是飞行高度与地面环境紧密耦合。

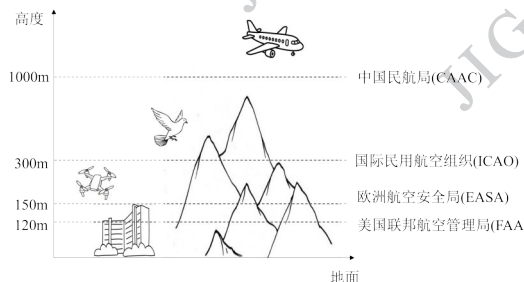


图1 全球低空标准对比图

Fig. 1 Global low-altitude standards comparison chart

技术研究更关注其感知挑战,实际高度阈值更低,聚焦300米以下超低空域——无人机在此深度融入建筑、山脉等地面背景杂波,受遮挡、电磁干扰、动态光影影响,传感器信号信噪比极低、特征微弱(Zhou 和 Wang, 2023)。因此,探测与跟踪研究中,“低空”本质是对感知系统构成严峻考验的“挑战性空域”。

1.2 各类无人机特点及在图像中的表现

低空无人机种类繁多,其不同的设计和工作原理直接决定了其外观、运动模式和在图像中的表现,从而对视觉感知系统提出不同的挑战。不同种类无人机示例如图2所示,特点如表1所示。

表1 无人机的种类、特点及在图像中的表现

Table 1 Types, characteristics, and performance of UAV in images

种类	特点	图像中的表现
旋翼机	①垂直起降与悬停:核心优势,可在空中保持静止 ②高机动性:可进行多方向及旋转的灵活运动 ③主流构型:呈“X”或“十”型框架	①运动模糊:旋翼高速旋转导致在图像中易产生模糊 ②悬停挑战:长时间静止易被误判为背景的一部分 ③外形多变:从不同视角观察,外形变化显著
固定翼机	①气动飞行:依靠机翼产生升力,类似传统飞机 ②续航长、速度快:气动效率高,但无法悬停 ③起降要求高:需要跑道或弹射/伞降	①姿态变化大:迎头、侧飞、俯视等角度下外观差异巨大 ②无静止状态:持续运动,对跟踪算法的预测能力要求高 ③背景变化快:高速飞行导致图像背景快速移动
垂直起降固定翼机	①混合布局:结合垂直起降和高效巡航 ②模式转换:通过倾转旋翼或机翼转换飞行模式	①构型突变:在起降/转换模式时,外形发生剧烈变化 ②算法需灵活:检测模型需能识别旋翼和固定翼两种状态
仿生无人机	①模仿生物:模仿鸟类或昆虫的外形和飞行方式 ②隐蔽性强:易于融入自然环境	①类间差异极小:与真实鸟类/昆虫在视觉上高度相似 ②运动模式复杂:飞行轨迹难以用常规运动模型描述



(a)旋翼机



(b)固定翼机



(c)垂直起降固定翼机



((a) rotorcraft; (b) fixed-wing aircraft; (c) vertical takeoff and landing fixed-wing aircraft; (d) bionic UAV)

图2 不同种类无人机示例

Fig. 2 Examples of different types of UAV

2 低空无人机检测与跟踪的核心挑战

2.1 低空无人机检测的核心挑战

1)小目标检测难题

无人机在远距离成像时往往仅占极少数像素,其特征信息微弱,在深层卷积神经网络中易被忽略或丢失。当无人机在图像中的尺寸小于 32×32 像素时,其可提取的视觉特征极为有限,导致检测性能显著下降(童康和吴一全,2023)。此外,复杂背景中的噪声干扰会进一步降低信噪比,使得小目标与背景杂波难以区分,造成漏检率升高。

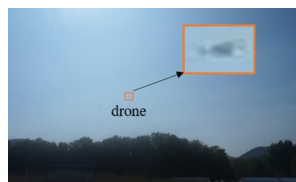
2)类间差异小与类内差异大

不同型号的无人机在外观、尺寸和材质上存在显著差异;与此同时,无人机与飞鸟(Harvey和Inman,2021)、风筝等自然干扰物在运动模式和视觉形态上又具有高度相似性。这种类间相似性和类内差异性使得传统分类器难以建立稳定的判别边界。此外,消费级无人机与工业级无人机在尺寸上可能相差数倍,而其旋翼结构又与鸟类翅膀的运动模式具有相似性。核心挑战示例图如图3所示。

2.2 低空无人机跟踪的核心挑战

1)复杂背景干扰

复杂背景带来的挑战主要体现在多个方面:动



(a)小目标检测



(b)类间差异



(c)类内差异



((a) small object detection; (b) inter-class variation; (c) intra-class variation; (d) complex background)

图3 核心挑战示例

Fig. 3 Examples of core challenges

态变化的自然场景会产生大量虚假运动信号,干扰运动目标的准确检测(Shabaz和Soni,2024)。城市环境中的建筑物遮挡会导致目标频繁消失与重现,对跟踪算法的重识别能力提出更高要求。此外,光照条件的剧烈变化会显著改变目标的外观特征,影响颜色和纹理等视觉特征的稳定性。恶劣天气条件还会引入额外的图像退化,降低图像质量并影响特征的可区分性。

2)实时处理要求

无人机的高速机动性要求检测与跟踪算法必须在极短的时间内完成处理,通常需要在30帧/秒(frames per second,FPS)甚至更高的帧率下运行,这对计算效率提出了苛刻限制。高精度的深度学习模型通常计算复杂度高,难以在资源受限的嵌入式平台上实现实时推理(Yang等,2024)。

3 低空无人机视觉检测方法

3.1 传统图像处理检测方法

在深度学习技术广泛应用之前,基于传统图像处理方法的低空无人机检测技术已得到深入研究。这些方法主要利用目标与背景在视觉特征上的差异进行识别,具有计算复杂度低、可解释性强等优势,在资源受限的嵌入式系统中仍具有一定应用价值。

3.1.1 运动特征检测方法

运动特征是区分无人机与静态背景的关键线索。背景减除是最基础的运动检测方法,通过建立背景模型并检测与背景差异显著的像素区域来识别运动目标(Kalsotra 和 Arora, 2022)。典型算法包括高斯混合模型、ViBe 等,这些方法在静态相机场景下对缓慢运动目标具有较好检测效果。Zhang 等人(2010)利用基于马尔科夫场的背景减除法检测移动的无人机。然而,在复杂自然环境中,背景模型的建立与维护面临巨大挑战,动态背景会产生大量虚假运动区域,而光照的渐变和突变会破坏背景模型的稳定性。

帧间差分法是另一种运动检测技术,通过计算相邻帧间像素差异来检测运动区域(Srinivas 和 Ganivada, 2025)。该方法计算简单、对场景变化适应性强,但通常只能检测目标边界,难以获得完整运动区域。谢家阳等人(2021)利用改进的三帧差分法来提取无人机分类候选区域,该方法虽然能部分解决"空洞"问题,但对缓慢运动目标仍不敏感。

光流法通过计算图像序列中像素点的运动矢量场来表征目标运动。与前述方法相比,光流法不仅能检测运动目标,还能获取运动方向和速度信息(Alfarano 等, 2024)。基于稀疏特征点的光流计算效率较高,适用于实时系统;而稠密光虽然能提供更丰富的运动信息,但计算复杂度限制了其实际应用。Briese 等人(2018)通过计算连续帧间的差分图像,结合光流估计进行运动补偿,并使用时序滤波来检测无人机的运动特征,从而实现快速轻量化的检测。光流法的主要挑战在于对噪声敏感,且在均匀纹理区域和运动模糊情况下性能显著下降。

3.1.2 表观特征检测方法

除运动特征外,无人机的静态表观特征也是重要的检测依据。基于形状特征的方法利用无人机特有的几何结构进行识别,如旋翼无人机的十字形或X形结构。Hough变换可用于检测直线特征,结合轮廓分析可识别特定形状模式(Zhang 等, 2024)。Zhou 等人(2016)提出了一种基于Hough变换的方法,以提高无人机检测和跟踪性能。然而,在远距离成像条件下,无人机通常仅占少量像素,形状特征极其微弱,且易受噪声干扰。

基于纹理特征的方法通过分析图像局部模式来区分目标与背景。局部二值模式、方向梯度直方图

(Saeed, 2024)HOG 等特征描述子被广泛应用于目标表征。例如,Wang 等人(2018)提取无人机的局部HOG特征,然后将特征发送到分类器进行分类和识别。这些特征对光照变化具有一定鲁棒性,但同样受限于小目标情况下特征信息不足的问题。

3.1.3 传统方法的局限性

传统图像处理检测方法的核心局限性在于其高度依赖人工设计的特征与固定流程,难以应对复杂多变的真实监测场景。首先,无论是运动特征还是表观特征,均为针对特定先验知识设计的浅层特征,其表征能力有限,尤其在无人机目标尺寸小、背景动态复杂时,特征区分度差,导致检测率低、虚警率高。其次,传统方法通常采用“运动检测-特征提取-分类”的串联流程,各模块独立优化,误差会逐级累积,难以实现全局最优。此外,这类方法对光照剧变、目标尺度与姿态变化、以及恶劣天气等干扰因素的鲁棒性普遍较差,泛化能力不足,制约了其在开放环境下的实际应用效果。

3.2 基于深度学习的无人机目标检测方法

基于深度学习的目标检测是当前无人机视觉检测任务的主流技术。根据其技术流程,主要可以分为两大类:双阶段检测方法和单阶段检测方法(曹家乐等, 2022)。这两种方法在检测精度和速度上各有侧重,适用于不同的无人机视觉检测场景。

双阶段检测方法将检测过程明确分为两个顺序阶段:首先生成候选区域,然后对每个候选区域进行分类和精确位置回归,框架如图4所示。其基本流程是:无人机图像首先通过一个主干网络提取特征图;然后,第一阶段使用区域提议网络在这些特征图上快速生成一系列可能包含无人机目标的候选框;第二阶段将每个候选框对应的特征区域通过池化或对齐操作转化为固定大小的特征图,再送入后续的分类网络和边界框回归网络,精确判断目标类别并微调位置。

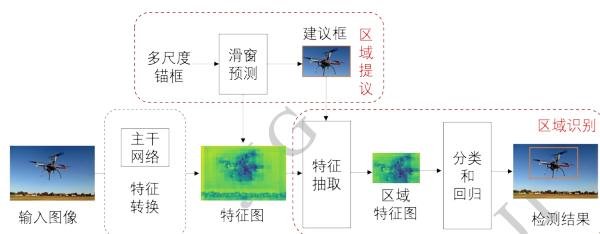


图4 双阶段无人机目标检测框架

Fig. 4 Two-stage UAV object detection framework

单阶段检测方法摒弃了生成候选区域的步骤, 实行“端到端”的检测。其核心思想是在特征图上进行密集采样, 直接预测目标的类别概率和边界框偏移量, 框架如图5所示。无人机图像经主干网络提取特征后, 通常会利用特征金字塔网络融合不同尺度的特征以应对目标尺度变化; 随后, 检测头直接在特征图的每个预设锚点上并行完成分类和回归任务。

单阶段和双阶段检测算法的优缺点对比如表2所示。目前无人机目标检测主要采用单阶段方法, 原因在于其卓越的运行速度适合实时应用, 同时通

过注意力机制有效抑制背景干扰, 增强目标特征。此外, 单阶段方法更易部署在边缘设备, 满足反无人机系统对高效性和实时性的需求, 而双阶段方法因速度慢和误报高, 在此场景中适用性较低。

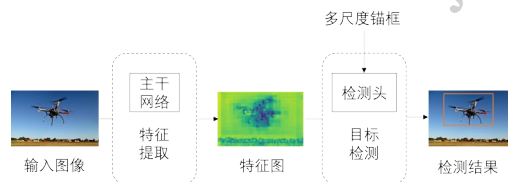


图5 单阶段无人机目标检测框架

Fig. 5 One-stage UAV object detection framework

表2 单阶段和双阶段检测算法的优缺点对比

Table 2 Comparison of advantages and disadvantages between single-stage and two-stage detection algorithms

方法类型	优点	缺点
双阶段	检测精度高、定位精确、共享计算量、参数可单独优化	运行速度慢、训练时间长、误报相对高
单阶段	运行速度快、背景干扰小、学习目标的泛化特征	精度低、小目标检测效果差、参数的学习难度大

低空无人机检测具有目标尺度小、视觉特征弱、易受复杂环境干扰以及边缘端计算资源受限等特点。为全面梳理基于深度学习的无人机检测研究, 本文并非依据单一标准对现有方法进行互斥分类, 而是从任务挑战、工程约束、环境条件和感知信息来源四个不同但相互关联的维度展开分析, 并据此归纳为小目标检测、轻量化网络设计、抗恶劣天气检测和多模态融合四个方向。

其中, 小目标检测对应无人机远距离成像时目标像素占比低、细节特征不足这一任务层面的核心难点; 轻量化网络设计对应检测模型在嵌入式设备和边缘平台上部署时面临的参数量、计算量、功耗及实时性约束; 抗恶劣天气检测对应雨、雾、雪、低照度等外部环境造成的图像质量下降和特征退化问题; 多模态融合则从感知信息来源角度, 利用可见光、红外、事件相机及雷达等传感器之间的互补信息增强目标表征。上述四个方向并非相互独立: 恶劣天气会进一步加剧小目标的特征缺失和背景混淆, 多模态融合可为小目标检测和恶劣天气条件下的检测提供信息补偿, 而轻量化网络设计则为相关高精度检测方法和多模态融合模型在资源受限平台上的实时应用提供工程支撑。因此, 本文按照“任务难点—环境影响—信息增强—部署实现”的逻辑关系, 对四类研究方向分别进行综述。

3.2.1 小目标检测改进方法

小目标检测是计算机视觉领域一项极具挑战性的任务, 特指在图像中识别和定位那些尺寸极小、往往只占据几十甚至几个像素的目标; 由于像素信息稀少, 这些目标缺乏清晰的轮廓和纹理特征, 极易淹没在复杂的背景之中(Li等, 2025)。

无人机目标检测正是小目标检测的典型场景, 这主要源于其实际应用条件: 为了保障安全, 系统必须在无人机还处于较远的距离上就及时发现它, 在这个探测距离上, 即便是体型中等的无人机在图像上也仅能呈现为一个微小的像素点(潘晓英等, 2023)。同时, 复杂环境引入了密集建筑物遮挡、动态变化的自然背景、与飞鸟等干扰, 进一步增加了识别难度, 使得无人机目标本身模糊的特征更加难以提取(冯潇等, 2025)。因此需要通过增强网络对小目标的特征提取能力来提升无人机的检测性能。针对小目标检测的改进方法汇总如表3所示, 分为检测框架与结构改进; 特征增强、融合与重建; 训练与优化策略; 数据增强与预处理4个方向。

1) 检测框架与结构改进

这类改进方法通过改进检测框架与结构, 例如添加新模块、调整层结构来增强小目标特征提取和能力。为了解决检测模型感受野较大的问题, 肖建国等人(2025)基于YOLOv11算法引入集成上下文

锚点注意力机制(contextual anchor attention, CAA)和路径聚合网络(path aggregation network, PA-Net)到颈部网络,通过引入上下文锚点提升了无人机小目标检测能力。Bo 等人(2024)通过 k-means 重新聚类锚框尺寸、引入 SPPFCSPC-SR 模块和 Inception-

NeXt 模块增强特征提取能力和感受野,显著提升小目标检测的准确性和鲁棒性。Wan 等人(2025)采用匈牙利二分匹配实现端到端检测,提高了长距离检测性能。

表3 针对小目标检测的改进方法汇总

Table 3 Summary of improvement methods for small object detection

优化方向	网络	团队	优化措施	优点	缺点
检测框架与结构改进	YOLOv7-GS	Bo 等人 (2024)	通过改进锚点聚类、引入 SPPFCSPC-SR 模块和 InceptionNeXt 模块来增强特征捕获和建模	减少漏检	FPS降低
	Vira-DSNet	Wan 等人 (2025)	采用级联和匈牙利二分匹配实现端到端检测	长距离检测性能优异	工程部署复杂度较高
	DETR	Zhu 等人 (2025)	把可学习记忆块作类别先验送入解码自注意力,提升运动模糊/尺度变化鲁棒性	对极小/密集/遮挡/模糊检测更稳	整体参量与计算量高于轻量 YOLO 系
	YOLOv11	肖建国等人 (2025)	引入 CAA 和 PA-Net,并添加额外检测头以提升小目标检测	小目标检测性能好	参数量大
特征增强、融合与重建	YOLOv8	左震等人 (2025)	通过在颈部网络构建多尺度特征融合结构,并引入 SimAM,以增强语义信息和特征提取	丰富特征信息	计算量增加
	DACG-Net	Li 等人 (2025)	通过全局上下文注意力动态加权融合多尺度特征	提升小目标特征关注度	需在精度与计算成本间权衡
	U形网络	Fang 等人 (2022)	通过引入 DSRDB 和 MFFRB 来增强特征提取和重建能力	提高检测精度	内存需求大
	DFFormer	Li 等(2026)	在进入 Encoder 前就把浅层细粒度线索以通道统计门控方式注入深层,避免传统方法导致的小目标细节丢失	小目标检测性能好	天空边缘场景易漏检
	SCINet	Fang 等人 (2024)	通过 DBFI,以及 DFCM 中的空间与通道注意力机制,以增强对比度和空间细节	增强空间细节和目标对比度	训练和推理成本高
	Magno-VTOD	Wang 等人 (2024)	通过运动引导的注意力机制增强视频中小型无人机目标的显著性	在复杂背景下鲁棒性强	对静态变化敏感
训练与优化策略	YOLOv4	Dong 等人 (2023)	使用 CIoU 损失函数替代 IoU,考虑重叠区域、中心距离和宽高比	检测精度高	检测速度略有下降
	PSO 优化网络	Lei 等人 (2022)	利用 PSO 算法动态调整网络参数,以提高检测精度和效率	提高检测性能	收敛慢

表3续表

优化方向	网络	团队	优化措施	优点	缺点
数据增强与 预处理	YOLOv5	Köksal 等人 (2022)	结合平衡焦点损失和损失排名挖掘等硬样本挖掘方法	适应不平衡数据	需平衡参数
	FENet	刘坤等人 (2025)	采用联合损失函数(SLS + Soft IoU),平衡整体监督与像素级监督	多尺度检测效果好	模型结构较复杂
	RetinaNet	梁栋等人 (2020)	通过相机运动补偿和目标重定位,以增强输入数据的质量	增强运动模糊下的鲁棒性	在纹理少场景易失效
	YOLOv4	Shi 和 Li (2020)	采用Mosaic数据增强,并用Dropblock数据增强替代Cutout	抗类不平衡	需要大量数据训练
	IRWT-YOLO	Cheng 等人 (2025)	使用SAM图像分割预处理、背景抑制增强数据	有效分离弱目标	预处理步骤增加延迟

除以 YOLO 系列为代表的卷积神经网络外,基于 Transformer 的检测变换器(detection transformer, DETR)类方法也逐渐应用于无人机小目标检测。YOLO 系列主要依靠卷积运算和多尺度特征金字塔提取局部空间信息,具有推理速度快、工程实现成熟等优势,但卷积操作的局部感受野使其对远距离区域之间关联信息的建模能力相对有限。相比之下,DETR 类方法通过自注意力机制建立图像中不同空间位置之间的全局关联,并采用目标查询与集合预测机制实现端到端检测,能够综合利用无人机目标周围的背景、轮廓和上下文信息。在远距离无人机像素占比较低、局部纹理不明显或目标受到遮挡时,全局上下文建模有助于增强弱小目标与背景之间的特征区分能力,并减少对锚框设计和非极大值抑制等人工处理环节的依赖。Zhu 等人(2025)基于 DETR 网络把可学习记忆块作类别先验送入解码自注意力,提升了运动模糊/尺度变化鲁棒性。

然而,DETR 类方法在无人机小目标检测中仍存在一定局限。一方面,标准自注意力机制需要在大量空间位置之间计算相关性,计算量和显存开销通常高于轻量化 YOLO 模型,限制了其在边缘设备和实时反无人机系统中的直接应用;另一方面,深层特征下采样可能造成小目标细节丢失,目标查询对极小目标的响应也容易受到复杂背景干扰。此外,部分 DETR 模型存在训练周期较长、收敛速度较慢以及对数据规模较为敏感等问题。针对这些不足,现有研究通常通过多尺度特征输入、可变形注意力、

查询优化、跨维度注意力及轻量化编码器等方式,在全局建模能力与计算效率之间进行平衡。总体而言,YOLO 类方法在实时性和边缘部署方面更具优势,而 DETR 类方法在全局上下文建模、复杂背景抑制及端到端检测方面具有较大潜力,二者分别适用于不同的精度与效率需求。

2)特征增强、融合与重建

此类方法主要通过特征提取与重建、多尺度信息融合、浅深层特征交互及注意力增强等方式,改善小目标特征在网络传递过程中的弱化与丢失问题,增强目标细节和语义信息的表达能力,从而提升小目标的检测与判别性能。为了解决算法难以有效提取红外无人机目标特征导致检测精度较低等问题,左震等人(2025)通过在 YOLOv8 的颈部网络构建多尺度特征融合结构,并引入三维无参注意力机制(three-dimensional parameter-free attention mechanism, SimAM),该机制可减弱背景干扰,此外融合深层与浅层特征以增强红外无人机小目标的语义信息和特征提取,从而显著提升检测精度。Li 等人(2025)提出了 DACG-Net,该网络通过全局上下文注意力动态加权融合多尺度特征,进而提升了对小目标特征关注度。Fang 等人(2022)利用深度可分离残差密集块(depthwise separable residual dense block, DSRDB)和多尺度特征融合与表示块(multi-scale feature fusion and representation block, MFFRB)构建 U 形网络,通过残差图像预测和全局残差连接有效提取和重建无人机红外小目标特征,减少背景

干扰。Fang 等人(2024)提出 SCINet,通过空间增强分支和对比增强分支的双分支特征交互(dual-branch feature inter-action, DBFI),以及双维特征校准模块(dual-dimensional feature calibration module, DFCM)中的空间与通道注意力机制,融合多尺度特征以增强红外无人机小目标的对比度和空间细节。Li 等(2026)提出 DFFormer 模型,该模型在进入 Encoder 前就把浅层细粒度线索以通道统计门控方式注入深层,避免传统方法导致的小目标细节丢失。Wang 等人(2024)受生物视觉系统启发,利用巨细胞计算模型提取运动特征,并将其与深度网络中的静态特征融合,通过运动引导的注意力机制增强视频中小型无人机目标的显著性,从而提升检测鲁棒性。

3) 训练与优化策略

此类方法聚焦于改进训练与优化策略,以更好地处理小目标的尺度敏感性和定位精度问题。Dong 等人(2023)在 YOLOv4 的基础上,引入了完整交并比(complete intersection over union, CIoU)损失来替代传统的交并比(intersection over union, IoU)损失,以更好地优化边界框回归。其位置损失函数的具体公式如下:

$$L_w = \lambda_w \sum_{i=0}^{S^2} \sum_{j=0}^B I_{ij}^{obj} \left[(1 - IoU) + \frac{\rho^2(b, b^{gt})}{c^2} + \alpha v \right] \quad (1)$$

式中: λ_w 为带权重的定位损失值; S^2 为特征图网格单元的总数; B 为每个网格单元预测的边界框的数量; I_{ij}^{obj} 为指示函数,当第 i 个网格单元包含第 j 个预测框所负责的某个真实目标的中心点时值为 1,否则值为 0; IoU 为预测框与真实框的交并比; $\rho^2(b, b^{gt})$ 为预测框与真实框中心点的欧氏距离; v 为用于衡量长宽比一致性的参数; α 为一个权衡参数。该损失函数综合考虑了重叠面积、中心点距离和长宽比,使得边界框的回归更加精确,尤其有利于小目标的定位。刘坤等人(2025)采用了 SLS 损失函数,此外该研究强调将 SLS 损失与 SoftIoU 损失联合使用,以平衡整体监督与像素级监督之间的权衡,共同优化模型。这种联合损失策略旨在同时提升目标的定位精度和形状分割能力,有效降低小目标的漏检和误检。针对目前反无人机检测方法不能快速处理传统反无人机和反无人机结构复杂的情况,Lei 等人(2022)通过粒子群优化(particle swarm optimization, PSO)算法动态调整神经网络参数和结构,结合特征捕获和建

模模块,优化无人机小目标检测性能。Köksal 等人(2022)通过结合平衡焦点损失和损失排名挖掘等硬样本挖掘方法,优化 YOLOv5 模型以增强对难样本的关注,间接通过特征层面的注意力机制提升小目标检测性能。

4) 数据增强与预处理

此类方法通过图像处理技术提升输入数据质量,如超分辨率重建或背景抑制。梁栋等人(2020)针对无人机检测中小目标的运动模糊、外观信息缺失和背景复杂问题,提出了一系列预处理改进策略。这些改进主要集中在相机运动补偿和目标重定位方面,旨在增强输入数据的质量,从而提升检测模型的鲁棒性和准确性。Shi 和 Li(2020)采用 Mosaic 数据增强方法扩展数据集,并结合 YOLOv4 的 CSPDarknet-53 主干网络,提升小目标检测的泛化能力和精度。Cheng 等人(2025)利用通用分割模型(Segment Anything Model, SAM)进行背景减除预处理,通过分割弱目标区域并提取最小边界框,减少背景干扰,显著提升红外小目标检测的准确性。

3.2.2 轻量化网络设计方法

由于无人机检测系统常需部署于算力有限的边缘设备,模型必须满足严格的实时性和功耗约束(Sun 等, 2025)。庞大的参数量和复杂的计算图不仅会耗尽设备资源,更会直接导致检测延迟,无法满足实际应用中对瞬时响应的要求。无人机目标的动态特性进一步加剧了这一挑战,其高速移动和背景多变要求检测模型必须具备高效的特征提取与推理能力。传统的深层网络虽然精度较高,但其巨大的计算成本难以在机载或便携式设备上稳定运行。因此,需要通过轻量化技术,在模型大小、推理速度与检测精度之间寻求最佳平衡,从而确保无人机检测系统在资源受限环境下依然可靠工作(李强等, 2025)。针对轻量化的改进方法汇总如表 4 所示,包括轻量化主干网络设计、高效卷积与注意力机制和特征融合优化 3 个方面。

1) 轻量化主干网络设计

这类方法是轻量化模型的基础,涉及替换或优化主干网络,以直接减少参数和计算量。Fang 等人(2023)采用空间感知通道注意力和通道调制可变形空间注意力等差异化注意力机制,增强了多尺度特征提取能力,有效降低了模型复杂度并提高了在复杂背景下的检测速度。针对现有反无人机无人机检

测算法检测精度低、计算资源消耗大的问题, Wang 等人(2022)设计了一种基于YOLOX的轻量化模型, 采用深度可分离卷积方法对模型进行压缩, 构建纳米网络, 大幅减少了参数数量和模型大小。同时集成了深度可分离卷积、挤压与激励(squeeze-and-excitation, SE)模块和卷积块注意力模块(convolutional block attention module, CBAM), 使无人机群检测在保持高精度的同时适用于各种边缘计算设备。Huang 等人(2024)改进的EDGS-YOLOv8模型融合了Ghost卷积、C3Ghost模块和EMA注意力机制, 在压缩模型体积的基础上强化了小目标检测能力, 实现了实时性和准确性的平衡, 显著降低了计算开销。张瑞鑫等人(2022)优化的CenterNet方法引入了联合注意力和轻量化主干网络DLAX28, 结合了通道分流策略, 该策略通过减少网络宽度与重复梯度信息的学习, 降低计算复杂度和内存成本, 大幅减少了参数数量和推理时间, 提升了低空无人机检测的实时性和定位质量。Zhang 等人(2026)提出了Anti-DETR, 该模型结合全局多通道注意力机制增强空间位置敏感性, 在显著降低参数数量与十亿次浮点运算(giga floating-point operations, GFLOPs)的同时提升小目标特征提取能力。

2) 高效卷积与注意力机制

这类技术通过优化卷积操作和增强注意力来提升特征提取效率, 但相对主干设计, 它更侧重于局部优化。Tian 等人(2024)通过利用分组卷积和空间自适应融合, 在增强多尺度特征表达的同时大幅减少参数和计算量, 提升无人机检测的实时性。Wan 等人(2025)通过SCDown轻量化下采样和自适应空间特征融合, 在quad-head结构中集成二次特征提取层, 使用部分卷积和通道注意力机制, 有效压缩模型大小并提升小目标检测精度。高仪等人(2025)通过采用部分多尺度特征融合和通道分割和协调注意力机制, 减少单通道灰度影像的冗余计算, 实现模型参数量降低至0.5M。邵剑飞等人(2025)通过重参数

化卷积, 利用多分支拓扑和注意力优化, 将参数量压缩73%, 显著提升检测速度。仲元昌等人(2025)集成大型可分离卷积核注意力(large separable kernel attention, LSKA), 结合WIoUv3损失函数, 在添加极小目标检测头的同时剔除冗余层, 使模型大小减少65.1%, 提升FPS至126.58帧/秒。Ahmad 等人(2025)提出了VM-QCNN, 使用量子卷积层替代经典卷积以降低计算复杂度。随着量子图像表示、量子图像处理及量子神经网络等技术的发展, 量子计算与传统视觉模型的结合为降低复杂视觉任务的计算开销提供了新的研究思路(吕品和吴一全, 2026)

3) 特征融合优化

在轻量化设计中, 特征融合优化虽然不直接减少参数或计算量, 但通过以下方式间接提高效率或降低复杂度: 更好的特征融合可以减少因尺度变化导致的重复检测或漏检, 从而避免不必要的计算开销; 通过高效路径聚合, 模型能以更少的层数实现更丰富的特征表示, 从而在保持精度的同时降低网络深度或宽度, 间接支持轻量化。Gao 等人(2024)采用跨特征融合模块和P2head小目标检测头来优化特征融合过程, 在增强小目标检测能力的同时, 将模型参数降低了38.4%, 实现了轻量化和效率的平衡。Yang 等人(2025)设计增强自适应特征金字塔网络进行渐进式特征融合, 通过减少非相邻层间的语义差距, 将模型大小压缩了75%。Munir 等人(2025)利用Scale-blended Feature Aggregation颈部和无参数SimAM注意力模块强化特征融合, 这是一种基于特征的注意力机制, 没有任何参数, 允许调整特征图的3D注意力权重的能力, 以增强模型的能力提取特征, 在恶劣天气条件下维持多尺度检测性能, 并通过优化网络结构降低了计算复杂度, 提升了模型的鲁棒性。Qin 和 Li(2026)提出了DCAM-DETR, 该模型通过跨维度注意力与自适应特征融合在保持精度的同时降低计算负担。

表4 针对网络轻量化的改进方法汇总

Table 4 Summary of improvement methods for network lightweighting					
优化方向	网络	团队	优化措施	参数量	模型大小
轻量化主干网络设计	DAGNet	Fang 等人(2023)	采用空间感知通道注意力和通道调制可变形空间注意力等差异化注意力机制	64.20 M	-

表4续表

优化方向	网络	团队	优化措施	参数量	模型大小
高效卷积与注意力机制	YOLOX	Wang 等人(2022)	通过深度可分离卷积、SE模块和CBAM注意力机制的集成,大幅减少了参数数量和模型大小	0.93 M	3.85 MB
	Anti-DETR	Zhang 等人(2026)	结合全局多通道注意力机制增强空间位置敏感性,在显著降低参数量与GFLOPs的同时提升小目标特征提取能力	13.90 M	-
	YOLOv8	Huang 等人(2024)	融合了Ghost卷积、C3 Ghost模块和EMA注意力机制	-	4.23 MB
	S-CenterNet	张瑞鑫等人(2022)	引入联合注意力和轻量化主干网络DLAX28,结合通道分流策略,大幅减少了参数量和推理时间	-	8.3 MB
	YOLOv8n	Tian 等人(2024)	通过利用分组卷积和空间自适应融合,在增强多尺度特征表达的同时大幅减少参数和计算量	-	3.72MB
	YOLOv8	Wan 等人(2025)	通过SCDown轻量化下采样和自适应空间特征融合,在quad-head结构中集成二次特征提取层	17.9 M	-
	VM-QCNN	Ahmad 等人(2025)	用量子卷积层替代经典卷积以降低计算复杂度	-	-
	YOLOv8n	高仪等人(2025)	通过采用部分多尺度特征融合和通道分割和协调注意力机制,减少单通道灰度影像的冗余计算	0.5M	1.4MB
	YOLOv5n	邵剑飞等人(2025)	通过重参数化卷积,利用多分支拓扑和注意力优化,将参数量压缩73%	0.46M	2.2MB
	YOLOv8n	仲元昌等人(2025)	集成LSKA,结合WIoUv3损失函数,在添加极小目标检测头的同时剔除冗余层	0.93M	2.2MB
特征融合优化	YOLOv11n	Gao 等人(2024)	采用跨特征融合模块和P2head小目标检测头来优化特征融合过程,将模型参数降低了38.4%	1.59M	-
	DCAM-DETR	Qin 和 Li(2026)	通过跨维度注意力与自适应特征融合在保持精度的同时降低计算负担	47.6 M	-

表4续表

优化方向	网络	团队	优化措施	参数量	模型大小
	YOLOv8s	Yang 等人(2025)	设计增强自适应特征金字塔网络进行渐进式特征融合,通过减少非相邻层间的语义差距,将模型大小压缩了75%	2.4M	5.3MB
	YOLOv5	Munir 等人(2025)	利用Scale-blended Feature Aggregation颈部和SimAM参数免费注意力模块强化特征融合	81.2M	-

3.2.3 抗恶劣天气方法

将探测视距延伸至恶劣天气和复杂光照条件下,无人机目标检测任务的难度显著增加。其根本原因在于,远距离探测本身已使无人机在图像中仅占据极小的像素区域,而雨、雾、雪及异常光照等因素会进一步破坏本已有限的目标特征信息。不同环境条件造成的图像退化机理存在明显差异:雨、雾、雪主要通过大气散射、光线衰减和局部遮挡降低目标对比度,并引入与小目标相似的动态噪声;夜间及低照度环境主要表现为有效光照不足、信噪比降低和细节缺失(Qiu 等,2025);强光、逆光及过曝则容易造成局部像素饱和、目标轮廓和纹理信息丢失。因此,现有抗恶劣天气无人机检测方法可根据主要退化类型,从雨雾雪等低能见度环境、夜间与低照度环境以及强光和高动态范围环境三个方面进行分析,相关代表性方法汇总如表5所示。

1)雨雾雪等低能见度环境检测

在雨、雾和雪等低能见度条件下,空气中的水滴、雾粒子和雪花会造成光线散射、衰减及局部遮挡,使远距离无人机的轮廓、纹理及目标—背景对比度进一步减弱,同时雨雪形成的动态噪声还可能与尺寸较小的无人机目标产生较高的像素级相似性,从而增加误检和漏检风险。针对这类问题,Wu 等人(2025)引入基于不确定性量化的红外无人机检测方法,利用贝叶斯卷积神经网络(Bayesian convolutional neural network,BCNN)并通过蒙特卡洛Drop-out生成不确定性图,对分割过程进行优化,从而提高低能见度条件下无人机目标检测的稳定性。Liu 等人(2025)则从检测模型本身入手,对YOLOv8的损失函数和注意力机制进行改进,以增强模型在雨、雪及光照变化等复杂环境中的目标特征提取能力。与前者侧重于利用预测不确定性提高低能见度场景

下检测可靠性不同,后者主要通过网络特征表达和优化机制增强模型对多种环境变化的适应能力,但相应的模型改进也会带来一定的检测速度下降。

2)夜间与低照度环境下的检测方法

夜间及低照度环境的主要问题是有效光照信息不足,由此造成图像亮度降低、信噪比下降以及无人机细节和边缘信息弱化。对于本身像素占比较低的远距离无人机,简单提高图像亮度还可能同步放大背景噪声,因此仅依赖可见光图像增强难以稳定恢复目标判别特征。针对复杂光照条件,李森等人(2024)提出PVNet,将事件相机产生的数据建模为三维时空点云,采用双视图融合分割网络提取不同视角下的目标信息,并通过门控融合机制对特征进行自适应加权。事件相机能够利用亮度变化信息进行异步响应,使该方法能够适应低光照以及高动态范围等传统帧式相机较难处理的成像条件。另一方面,Liu 等人(2025)的改进YOLOv8方法通过注意力机制和损失函数优化增强复杂光照变化下的目标表征能力,体现了从网络鲁棒性增强与成像方式改变两个不同层面对低照度无人机检测问题的解决思路。然而,事件相机方案依赖专用硬件,而基于传统视觉检测网络的改进方法仍会受到极端低照度条件下原始信息不足的限制。

3)强光过曝与高动态范围环境下的检测方法

与低照度条件下的信息不足不同,强光、逆光及过曝主要会造成局部区域像素饱和,使目标轮廓、纹理及灰度层次被高亮背景淹没,耀光还会形成较大范围的干扰区域。针对这类问题,乔新博等人(2025)提出ANDD,采用分焦平面红外偏振成像获取目标的偏振信息,并结合偏振解码蒸馏网络提取受背景干扰较小的特征,以改善烟雾、扬尘、耀光及复杂背景下的无人机检测性能;该方法利用偏振信

表 5 针对抗恶劣天气的改进方法汇总

退化类型	网络	团队	优化措施	适用场景	缺点
雨雾雪/低能见度	BCNN	Wu 等人(2025)	不确定性量化框架,蒙特卡洛 Dropout 生成不确定性图,优化分割精度	雨、雾等低能见度天气	模型泛化能力有限
多种天气退化	YOLOv8	Liu 等人(2025)	改进损失函数和注意力机制	恶劣天气如雨、雪和光照变化	检测速度略有下降
烟雾/耀光等复杂退化	ANDD	乔新博等人(2025)	采用分焦平面红外偏振成像,结合偏振解码蒸馏网络提取噪声更少的偏振信息	烟雾、扬尘、耀光及复杂背景干扰下的无人机检测	轻量化程度较低
高动态范围光照	PVNet	李森等人(2024)	将事件相机数据建模为三维时空点云,采用双视图融合分割网络,并设计门控融合机制自适应加权特征	强光过曝、低光照等高动态范围复杂光照条件	依赖专用事件相机硬件

息提高了复杂成像条件下目标与背景的可区分性,但轻量化程度相对有限。李森等人(2024)提出的PVNet则利用事件相机具有高动态范围的特点,并通过三维时空建模、双视图融合及门控特征加权,提高强光过曝条件下无人机目标的感知能力,但需要额外的事件相机硬件。由此可见,针对恶劣天气和复杂光照的无人机检测目前主要存在两类技术思路:一类通过改进检测网络直接学习具有环境鲁棒性的判别特征,另一类则借助红外偏振、事件视觉等成像方式缓解传统RGB图像在极端环境中的信息退化。

除此之外,图像恢复与目标检测的协同优化也是值得关注的技术路线。传统“先增强、后检测”的串联方式通常分别优化去雾、去雨或低照度增强模型与目标检测器,图像视觉质量的改善并不必然带来检测性能提升;相比之下,将图像增强模块与检测网络进行端到端联合训练,可利用检测任务损失反向约束前端增强过程,使其更加关注无人机轮廓、局部对比度及判别特征的恢复,从而实现图像质量改善与目标检测性能提升之间的协同优化。但这类方法也会增加网络复杂度和训练难度,且面对多种退化同时存在的开放环境时,其跨天气泛化能力仍有待进一步提高。

3.2.4 多模态融合方法

与检测地面车辆或行人不同,无人机目标本身具有体积小、材质多样、飞行轨迹灵活多变等特点,

这给单一传感器带来了巨大挑战。例如,光学相机在白天视野良好,但极易受天气、云层和光线突变的影响,且对伪装成小鸟或融入背景的无人机分辨能力有限;雷达对运动金属目标敏感,但对低速或悬停的塑料无人机探测能力弱,且在城市环境中容易受到高楼等静止杂波的干扰。因此,融合可见光、红外、雷达甚至声学等多种模态信息,成为了构建可靠反无人机体系的必然选择,如图6所示。这种融合能够发挥各自优势:雷达可以实现大范围、全天候的粗搜索与预警,引导光学和红外系统进行精确识别与跟踪,从而有效克服单一传感器的盲区,实现对“低慢小”无人机的高概率探测与低虚警辨识。针对多模态融合的改进方法汇总如表6所示。

Diamantidou 等人(2019)提出了一种多模态深度学习框架,通过融合热红外、视觉和2D雷达传感器的特征,使用多层感知器神经网络进行特征融合,从而显著提升了无人机检测的准确性和鲁棒性。Xie 等人(2024)提出了一种RF-visual定向融合框架,用于精确的无人机定位,这是典型的多模态工作。通过联合校准阵列天线和相机,将射频(radio frequency, RF)信号与图像对齐,从而利用RF信号的环境适应性和视觉数据的精细粒度。框架包括一个基于分割的去噪方法来去除频谱噪声,并通过空间校准将RF方位角投影到图像平面上,以缩小感兴趣区域,提升检测精度和鲁棒性。这种方法在复杂电磁环境下能有效区分无人机和类似物体。



图6 多模态无人机目标检测
Fig. 6 Multimodal UAV object detection

He 等人(2025)开发了一种声学-光学混合多模态检测方法,利用Capon波束成形进行声学成像和YOLOv11进行视觉检测,并通过空间一致性验证,在低光、遮挡或多目标场景中实现声学视觉的互补,从而大幅提升检测鲁棒性。Wang 等人(2024)设计了RTM-UAVDet实时多模态检测器,采用动态卷积模块和多尺度特征融合编码器,自适应处理RGB和热红外(thermal infrared, TIR)模态的差异,无需严格图像对齐即可在静态图像中实现高精度检测,从而增强在夜间、雾霾等多变成像条件下的鲁棒性。

表6 针对多模态融合的改进方法汇总

Table 6 Summary of improvement methods for multimodal fusion

优化方向	团队	采用网络或方法	融合模态	缺点
多模态融合	Diamantidou E 等人(2019)	- 多层感知器融合网络 - Faster R-CNN 用于TIR 和视觉特征提取 - CNN 用于2D 雷达特征提取	TIR、视觉和2D 雷达融合	时间注册要求精确同步,实践中难以实现
	Xie W 等人(2024)	- RF 模态: 使用均匀圆形阵列天线和多重信号分类算法进行方位角估计 - 视觉模态: YOLOv11 进行目标检测	RF 和视觉融合	RF 模态易受环境噪声影响
	He T 等人(2025)	- 声学模态: Capon 波束成形进行声源成像 - 视觉模态: YOLOv11 进行目标检测	声学 and 光学融合	融合系统复杂度高,实时性可能不足
	Wang G 等人(2024)	- 多尺度特征动态融合编码器 - 基于CSPDarknet53的backbone	RGB 和 TIR 融合	依赖高质量标注数据,可能过拟合

4 低空无人机视觉跟踪方法

4.1 传统图像处理跟踪方法

在基于深度学习的视觉跟踪方法成为主流之前,研究者们主要依靠传统图像处理技术实现对低空无人机的持续定位。这类方法的核心在于利用目标外观特征在视频序列中的时空相关性,通过构建动态模型来预测其位置。它们通常不依赖大规模数据训练,具有较低的计算开销和良好的可解释性,因此在计算能力受限的机载嵌入式平台上,此类方法依然拥有特定的应用场景,并为后续方法的发展奠定了重要基础。

4.1.1 生成式跟踪方法

生成式跟踪方法的核心思想是将跟踪视为一个搜索问题,即在当前帧中寻找一个区域,使其外观模型与在初始帧或上一帧中建立的目标模型最为匹配。

均值漂移算法因其计算效率高,早期被广泛用于地对空无人机跟踪(王旭东等,2025)。其通过迭代优化巴氏系数等相似性度量,使搜索窗口向目标位置收敛。然而,在动态的地对空场景中,无人机目标的尺度与姿态会因其自身的高速机动而剧烈变化,该算法的固定窗口机制成为其主要瓶颈。为解决此问题,一些研究尝试引入尺度估计策略,例如通过在不同尺度空间进行均值漂移迭代,以自适应跟

踪目标的大小变化(Kemsaram等,2016)。

基于光流的跟踪方法尤其适用于应对无人机目标的快速机动。稀疏光流法通过跟踪图像序列中的角点或特征点,能够有效解算目标的相对运动。在地面平台的视觉里程计与目标跟踪结合的研究中, Lucas-Kanade 光流法及其改进版本被广泛应用。例如,在对高速无人机的持续跟踪任务中,可通过光流场估计来分离或补偿背景运动,从而实现对机动目标的稳定跟踪。然而,该方法对图像噪声、运动模糊以及光照突变非常敏感,在目标高速运动和复杂背景下性能不稳定(Granillo和Beltran,2018)。

4.1.2 判别式跟踪方法

判别式方法将跟踪视为一个在线学习的分割问题,通过持续区分目标与其周围背景来提升跟踪的鲁棒性。

相关滤波框架是传统方法中在精度与速度权衡上的一个高峰,其思想被后续众多无人机视觉跟踪研究作为基线。最小输出误差平方和滤波器展示了惊人的速度潜力(Fu等,2021)。而核相关滤波算法(kernelized correlation filters algorithm, KCF)通过利用循环矩阵和核技巧,在保持高速的同时显著提升了精度,使其非常适合在无人机上处理常规场景的跟踪任务。

4.1.3 传统方法的局限性

传统图像处理跟踪方法虽在计算效率上具有优势,但其性能受限于人工设计特征与模型更新机制的固有缺陷。生成式方法依赖于固定的表观模板,难以适应目标剧烈的尺度、姿态变化与光照波动,容易在目标外观改变时发生跟踪漂移或丢失。判别式方法通过在线学习区分目标与背景,对形变和部分遮挡有更好鲁棒性,但其模型更新通常基于单一的近期外观,在目标经历完全遮挡、快速运动模糊或背景杂波干扰时,容易学习到错误信息,导致模型污染和后续跟踪失败。此外,绝大多数传统跟踪器采用手工特征,其特征表达能力有限,在复杂空域背景、小目标及低对比度场景下判别力不足。最后,这些方法通常采用简单的线性运动模型,难以准确预测高机动性无人机的不规则运动轨迹。上述局限性使得传统方法在动态、开放的低空复杂环境中,其稳定性、精度和长期跟踪能力面临严峻挑战。

4.2 基于深度学习的无人机跟踪方法

基于深度学习的无人机目标跟踪流程始于对初

始目标的标注,随后通过一个高效的“搜索-匹配-更新”闭环实现持续跟踪,如图7所示。系统首先在首帧中确定无人机目标位置;在后续每一帧中,以上一帧目标位置为中心划定搜索区域,并利用轻量级深度神经网络分别提取无人机模板和搜索区域的多尺度特征;通过相似度匹配计算,在搜索区域内定位最可能的无人机位置,并输出当前帧的跟踪结果;同时,系统会根据跟踪结果自适应更新无人机外观模板以应对姿态变化和尺度缩放。

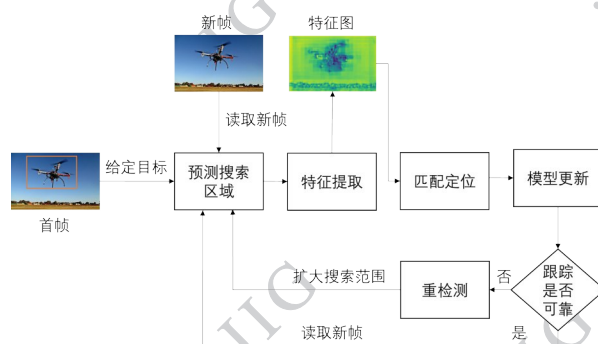


图7 基于深度学习的无人机目标跟踪流程

Fig. 7 Deep learning-based UAV object tracking process

当跟踪置信度较低时,系统会启动重检测机制,在更大范围内重新捕获无人机目标,从而形成一个稳定、自适应的端到端跟踪流程。该流程充分发挥了深度学习在特征学习和匹配方面的优势,使其对无人机目标的尺度变化、快速运动和复杂背景等挑战具有出色的鲁棒性。

基于深度学习的无人机视觉跟踪研究,围绕其核心任务流程,主要从提升跟踪性能、满足部署约束、完善跟踪机制、扩展跟踪场景四个关键维度进行深化与创新。首先,鲁棒性增强聚焦于解决跟踪算法本身在面对目标快速运动、尺度变化、光照突变及复杂背景干扰时的稳定性与精度问题,是提升跟踪性能的核心。在此基础上,研究进一步向三个方向延伸:其一,轻量化与实时性旨在优化模型与策略,以满足嵌入式平台或地面站对计算效率、功耗和实时响应的严苛部署约束;其二,长期跟踪与重新检测针对目标因严重遮挡、飞出视场等原因消失的现实场景,致力于设计包含记忆、全局检索等模块的完整跟踪机制,以确保任务的连续性;其三,多目标跟踪则将问题扩展到无人机集群或密集空域监控等复杂应用场景,核心挑战在于跨帧的稳定数据关联与身

份维持。下文将依此框架,对各个方向的主流方法进行系统综述。

4.2.1 鲁棒性增强方法

无人机视觉跟踪场景中的目标常面临快速运动、尺度变换、光照突变及复杂背景干扰等严峻挑战,这对跟踪算法的鲁棒性提出了极高要求。为应

对此问题,基于深度学习的跟踪方法致力于通过多种策略提升模型在复杂环境下的稳定性与容错能力。针对鲁棒性增强的改进方法汇总如表7所示,包含运动估计与时空建模、尺度估计与特征融合和环境适应与图像预处理3个方向。

表7 针对鲁棒性增强的改进方法汇总
Table 7 Summary of improvement methods for robustness enhancement

优化方向	网络	团队	优化措施	优点	缺点
运动估计与时空建模	Siamese	Xie 等人 (2023)	提出联合建模关联和运动线索,使用运动估计模块处理遮挡和背景杂波	鲁棒于快速运动、热交叉和相似干扰	模板更新不自适应
	Keep-Track	Xu 等人 (2024)	利用运动估计模块进行尺度补偿,并融合关联匹配和运动线索	抗遮挡能力强	跟踪速度较慢
	YOLOv7	Zhang 等人 (2025)	通过动态相似性尺度估计和自适应遮挡感知机制,结合运动线索优化模型更新	高实时性,适合动态环境	超参数调优依赖场景
	Transformer	Yu 等人 (2025)	利用局部跟踪和全局检测的注意力模块,改善前景-背景区分	提升跟踪速度和准确性	框架复杂度高,需大量训练数据
尺度估计与特征融合	SiamR-CNN	Huang 等人 (2021)	通过时空约束估计目标尺度变化,并融合运动特征和变化检测相关滤波器	处理小尺度和快速运动效果好	计算复杂度较高
	EfficientNet-B0	谢家阳等人 (2021)	通过运动背景补偿估计尺度,并融合深度特征和运动重定位策略	处理动态背景效果好	对极端运动或遮挡可能敏感
环境适应与图像预处理	YOLOv5	刘璐琦等人 (2024)	基于 HSV 色彩空间的阴影识别算法排除目标投影干扰,并结合二次定位技术融合目标框中心点和结构关键点	提升稳定性和安全性	目标快速脱离视野时难以重新识别

1)运动估计与时空建模

这类方法通过整合运动线索、时空信息或背景补偿来处理相机运动、目标快速移动、遮挡和场景变化,确保跟踪的连续性和稳定性。Xie 等人(2023)提出联合建模关联和运动线索,使用运动估计模块处理遮挡和背景杂波,并结合 YOLOv7 应对目标消失和重现问题,提升在复杂运动下的鲁棒性。为了解决了背景干扰和遮挡导致的跟踪不稳定问题,Xu 等人(2024)利用运动估计模块进行尺度补偿,并融合关联匹配和运动线索,增强在遮挡和背景杂波下无人机跟踪的稳定性。Zhang 等人(2025)基于 KCF 框

架,通过动态相似性尺度估计和自适应遮挡感知机制,结合运动线索优化模型更新,从而增强无人机跟踪对尺度变化、遮挡和重现的稳定性,同时保持高实时性能。Yu 等人(2025)通过场景补偿收集器捕获时空变化信息,并注入到 Transformer-based 跟踪框架中,设计了局部跟踪和全局检测的注意力模块,该模块将场景补偿信息无缝集成到跟踪框架的全局检测器中,增强前景特征,抑制背景特征,改善前景-背景区分,从而提升无人机跟踪在复杂场景下的鲁棒性和准确性。

2)尺度估计与特征融合

此类方法通过多特征融合、尺度自适应机制或聚焦策略来处理目标尺度变化、相似物干扰和特征判别问题,提高定位精度。Huang等人(2021)提出SiamSTA,通过时空约束估计目标尺度变化,并融合运动特征和变化检测相关滤波器,从而在复杂背景中提升对小目标和快速移动无人机的跟踪鲁棒性。谢家阳等人(2021)开发动态云台摄像机算法,通过运动背景补偿估计尺度,并融合深度特征和运动重定位策略,提高在动态场景中无人机跟踪的实时性和鲁棒性。

3)环境适应与图像预处理

这类方法通过图像增强、阴影识别或自定位技术来应对暗光、阴影或动态环境等外部因素,提升跟踪在恶劣条件下的可靠性。刘璐琦等人(2024)基于色调-饱和度-明度色彩空间(hue-saturation-value color space, HSV)的阴影识别算法排除目标投影干扰,并结合二次定位技术融合目标框中心点和结构关键点,减少定位误差。

4.2.2 轻量化与实时性改进方法

面向地对空无人机检测的地面计算单元,通常存在严格的功耗、算力与存储限制。因此,在保证跟踪精度的前提下,实现模型的轻量化与推理的实时性,是相关技术得以实际部署的关键。针对轻量化与实时性的改进方法汇总如表8所示,分为模型压缩与结构优化、动态策略与自适应调整和硬件与框架优化3部分。

1)模型压缩与结构优化

该类方法通过减少参数数量和计算复杂度来实现轻量化,如使用Ghost卷积、轻量主干网络或注意力机制。Wang等人(2025)提出FocusTrack框架,通过自适应搜索区域调整策略和注意力到掩码模块优化Transformer-based跟踪器,动态调整计算资源,在保持高精度的同时实现高效实时跟踪。针对模型参数量大导致模型难以实时跟踪的问题,程鲲等人(2025)改进YOLOv8网络,采用空间到深度分组卷积和深度可分离混洗网络结构进行模型压缩,该卷积将特征图空间信息转化为深度信息,减少卷积时小目标的空间信息损失,同时对原始特征图分组,降低计算量的同时进一步增加通道保留小目标的空间信息。

2)动态策略与自适应调整

该优化方向通过智能策略动态调整搜索区域、

尺度或采样率,减少不必要的计算开销。Wu和Liu(2022)通过引入帧差分动态机制,在语义跟踪器不自信时利用运动信息辅助跟踪,自适应调整搜索区域,提升了红外无人机跟踪的鲁棒性和轻量化。Wu等人(2020)通过集成特征注意力模块和扩展搜索策略,动态调整模型参数和搜索中心,增强了跟踪器对突然运动的适应性,实现了高实时性跟踪,同时保持了轻量化设计。Feng等人(2025)通过尺度自适应增强模块隐式调整模板特征,动态响应目标尺度变化,优化了全局跟踪器的轻量化和实时性,但在复杂背景中抗干扰能力不足。

3)硬件与框架优化

此类方法针对边缘设备进行部署优化,如使用专用框架或硬件加速。陶磊等人(2020)采用改进的YOLOv3模型,结合比例积分微分算法控制摄像头,在GTX1060硬件上实现了实时跟踪。

4.2.3 长期跟踪与重新检测改进方法

在实际应用中,目标无人机可能因长时间严重遮挡、飞出视场或识别失败等原因导致跟踪中断,因此,具备长期跟踪能力是现代跟踪系统的核心指标之一。长期跟踪机制不仅需要高效的短期定位模块,更关键的是要集成一个基于重检测的恢复能力。这通常通过构建一个全局的或基于关键帧的时序记忆模块,并在跟踪置信度低时,触发一个轻量级的全局检测器来重新扫描图像,从而快速找回丢失的目标,形成一个“跟踪-检测-恢复”的闭环系统,确保任务的连续性(Fang等,2024)。针对长期跟踪与重新检测的改进方法汇总如表9所示,包含多一致性融合、注意力增强与记忆网络和判别器与滤波器优化3个优化方向。

1)多一致性融合

这类方法融合多源信息或施加一致性约束,以提升跟踪在复杂场景下的鲁棒性。重新检测依赖于一致性检查或模态切换。Zhang等人(2023)采用多模态特征融合和决策融合机制,通过局部-全局转换器分析红外与可见光数据的一致性,动态触发重新检测以应对目标消失或遮挡问题。Zhang等人(2025)设计运动一致性约束和空间重新映射模块,结合自适应模板更新策略,利用背景相关性进行重新检测,在目标位移或变形时维持长期跟踪稳定性。Bu等人(2025)集成大型选择性卷积核注意力机制和时空一致性重新检测,通过背景增强模块提升初

表 8 针对轻量化与实时性的改进方法汇总
Table 8 Summary of improvement methods for lightweight and real-time performance

优化方向	网络	团队	优化措施	缺点	FPS
模型压缩与结构优化	FocusTrack	Wang 等人 (2025)	自适应搜索区域调整策略和注意力到掩码模块,动态优化计算资源,减少冗余计算	实时性不足于高速或资源受限场景	143
	YOLOv8	程鲲等人 (2025)	空间到深度分组卷积和深度可分离混洗网络,压缩模型参数,增强小目标特征提取	鲁棒性有待提升	146.3
动态策略与自适应调整	SuperDiMP	Wu 和 Liu (2022)	使用帧差分辅助语义跟踪器,当置信度低时启动决策模块自适应融合结果	速度略有下降	7.3
	ResNet18	Wu 等人 (2020)	特征注意力模块动态重校准特征图,扩展搜索策略自适应调整搜索中心以处理突然运动	无注意力时 FPS 更高但精度下降	55.6
	ResNet50	Feng 等人 (2025)	尺度自适应增强模块隐式调整模板特征,自适应阈值判断目标消失	跟踪易受相似物体干扰	30.9
硬件与框架优化	YOLOv3	陶磊等人 (2020)	改进 YOLOv3 模型结构,结合比例积分微分算法控制摄像头,迁移学习训练	无法摆脱对显卡的依赖	20

始帧信息利用,在低信杂比环境下优化长期跟踪的重新检测效率。

2)注意力增强与记忆网络

这类方法通过注意力机制增强特征判别力,并结合记忆模块动态维护历史特征,以处理长期跟踪中的外观变化和遮挡。重新检测通常由内部判别器触发全局搜索。Wang 等人(2024)设计了对比增强记忆网络,为了避免模型退化,设计了对比学习和自适应动态记忆模块,该模块应用了注意力更新策略,以更好地提取初始模板和跟踪结果的共同特征,提高了热红外视频中无人机跟踪的实例判别能力和长期稳定性。林淑彬等人(2025)提出目标感知 Transformer 跟踪器,通过自适应关系建模和特征聚合,聚焦关键信息,减少了背景干扰,提升了无人机跟踪的精度和实时性能。

3)判别器与滤波器优化

这类方法改进传统算法,通过判别器或梯度特征优化实时性和遮挡处理。重新检测基于简单阈值或模型切换机制。王文胜等人(2025)提出融合梯度改进的 YOLOv10 和 KCF 模型,通过特征级联和多尺度检测优化判别器与滤波器,以增强小目标和遮挡情况下的跟踪鲁棒性和重新检测精度。

4. 2. 4 多目标跟踪改进方法

在集群协同、空域监控等场景中,需要对多架无人机进行同时跟踪与轨迹维持,即多目标跟踪任务。该任务的核心挑战在于如何实现跨帧的、准确的目标-轨迹数据关联。基于深度学习的方法通常采用“跟踪-检测”范式,即首先在每一帧中检测出所有目标,然后通过深度学习模型学习的外观特征表示或运动模型预测,利用匈牙利算法等数据关联技术,将不同帧中的检测框关联到同一个轨迹编号上。针对多目标跟踪的改进方法汇总如表 10 所示。

如何学习更具判别力的特征以解决目标间的相似性问题与交互遮挡,是多目标跟踪研究中的重点与难点。Yu 等人(2024)提出基于局部点匹配的协作图像注册和反无人机跟踪框架,通过动态多模态注册、分层多模态图像融合、关键点轨迹跟踪和在线自适应校准模块,提升多目标跟踪在挑战性场景下的性能。Zhu 等人(2024)提出 Z 函数分解基于强化学习方法,通过优化主动无人机的传输功率和所有无人机的轨迹来提升 3D 无人机跟踪的多目标跟踪精度,减少定位误差。

表 9 针对长期跟踪与重新检测的改进方法汇总

优化方向	网络	团队	优化措施	缺点
多一致性融合	Siamese	Zhang 等人(2023)	多模态特征融合、局部-全局转换器决策融合	依赖多模态数据对齐
	OTrack	Zhang 等人(2025)	运动约束模块、空间重新映射模块、自适应模板更新	模板更新可能引入噪声
	Super-DiMP	Bu 等人(2025)	大型选择性卷积核注意力机制、时空一致性重新检测、背景增强模块	计算资源需求高
注意力增强与记忆网络	CAMTracker	Wang 等人(2024)	对比增强匹配模块和自适应动态记忆模块,用于实例判别和模板更新	在通用TIR跟踪任务中性能一般
	Transformer	林淑彬等人(2025)	自适应关系建模机制和特征聚合模块,聚焦关键令牌并增强特征判别力	对剧烈外观变化敏感
判别器与滤波器优化	YOLOv10	王文胜等人(2025)	引入梯度特征级联和多尺度检测,优化判别器特征提取和滤波器尺度自适应	模型复杂度较高,训练数据需求大

表 10 针对多目标跟踪的改进方法汇总

优化方向	网络	团队	优化措施	优点	缺点
多目标跟踪	基于点匹配的协作框架	Yu 等人(2024)	通过动态多模态注册、分层多模态图像融合、关键点轨迹跟踪和在线自适应校准模块,提升多目标跟踪	有效处理目标消失、微小目标等挑战性场景	对点检测敏感
	强化学习	Zhu 等人(2024)	联合优化主动无人机的传输功率和所有无人机的轨迹;使用分布式训练和概率分布估计未来奖励	适应雨、雾等低能见度天气	对快速运动目标可能敏感

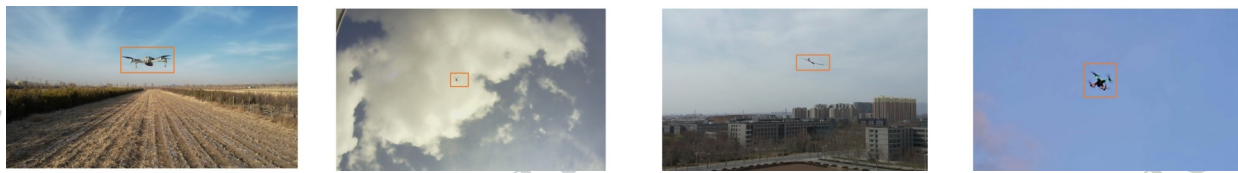
5 数据集与性能评估

5.1 无人机图像数据集

高质量数据集的可用性对于推进反无人机场景中无人机检测和跟踪的研究至关重要。对于基于视

觉的反无人机系统,数据集必须反映这些系统遇到的环境、无人机类型、光照条件和操作环境的多样性。本节回顾了公开可用的用于无人机检测和跟踪的数据集。这些数据集在表 11 和表 12 中根据它们的图像类型进行了分类。图 8 和图 9 为它们的示例。





(e)UAVfly (f)MMAUD (g)MMFW-UAV (h)UETT4K Anti-UAV((a) Drone-vs-Bird; (b) DUT Anti-UAV;(c) LRDD;(d) ICG-Drone;(e)UAVfly;(f)MMAUD;(g)MMFW-UAV;(h)UETT4K Anti-UAV)

图8 无人机RGB图像数据集示例

Fig. 8 Examples of UAV RGB image datasets

表 11 无人机 RGB 图像数据集汇总
Table 11 Summary of UAV RGB image datasets

数据集	数量	尺寸	环境	无人机种类	团队
Drone-vs-Bird	24,804	-	静态和移动摄像头多种背景	8 种不同类型	Coluccia 等人 (2021)
DUT Anti-UAV	22,500	1080×1920	天空、建筑、自然环境等	35+种类型	Zhao 等人 (2022)
LRDD	1,823	1920×1080	城市、乡村、自然环境	DJI Mavic Pro 等	Rouhi 等人 (2024)
ICG-Drone	10,281	1920×1080	室内实验室、室外开阔场地	Tello、F450、QAV250	Xiao 等人 (2024)
UAVfly	33,601	1280×720	城市街区、郊区、沙漠、田野、湖泊、天空、山脉	40+种类型	Cheng 等人 (2024)
MMAUD	50 个视频	2560×960	真实室外、机械噪声	6 种商用无人机	Yuan 等人 (2024)
MMFW-UAV	147,417	3840×2160	多视角、全天候条件	12 种固定翼无人机	Liu 等人 (2025)
UETT4K Anti-UAV	308,144	3840×2160	天空、建筑、树木等多背景	6 种类型	Awan 等人 (2025)

Drone-vs-Bird: <https://github.com/wosdetc/challenge>
DUT Anti-UAV: <https://github.com/wangdongdut/DUT-Anti-UAV>
LRDD: <https://research.coe.drexel.edu/ece/imaple/long-range-drone-detection-dataset/>
ICG-Drone: <https://github.com/NTU-ICG/multidrone-detection-tracking>
UAVfly: <https://github.com/lucien22588/UAVfly>
MMAUD: <https://ntu-aris.github.io/MMAUD>
MMFW-UAV: <https://www.scidb.cn/en/detail?dataSetId=edb443cfcf4142019499bbfd68542cc9&version=V4>
UETT4K Anti-UAV: <https://github.com/mugheessarwarawan/UETT4k-Anti-UAV>

在完成对现有数据集的梳理后,本节对各数据集的核心特性、优缺点进行横向比较,比较如表 13 所示。并分析其在评估小目标检测、抗恶劣天气、多模态融合等特定任务时的适用性,为研究者选择合适的数据集提供明确指导。

1) 小目标检测任务

研究无人机小目标检测,需按传感器模态选择数据集。可见光模态下,LRDD 专攻极限远距离微小目标检测,UETT4K Anti-UAV 则凭借 4K 分辨率利于研究上下文信息的利用。红外模态下,IST-A 与 IRDST 提供了海量复杂背景下的像素级小目标样本,是训练和评估模型的基石。DUT Anti-UAV 和 Drone-vs-Bird 可作为包含多样背景与混淆目标的通用补充基准。

2) 抗恶劣天气与夜间任务

在恶劣天气及夜间条件下,红外数据集因成像原理而成为研究首选。IST-A、IRDST、SIDD 及 Anti-UAV (IR) 等均能有效检验模型在低能见度环境下的鲁棒性。可见光模态下,可选用包含全天候条件的 MMFW-UAV 进行探索。需注意,目前仍极度缺乏专门针对极端恶劣天气的大规模可见光数据集。

3) 多模态融合任务

多模态融合研究可根据深度选择数据集。MMAUD 提供了最完备的同步多模态数据,是前沿探索的核心资源。Drone Detection Dataset 和 Anti-UAV 分别提供了易获取的可见光-红外-音频数据和高质量的可见光-红外视频对,是算法验证与原型开发的理想起点。

4) 跟踪与重识别任务

无人机跟踪研究首推权威基准 Anti-UAV 数据

表 12 无人机 IR 图像数据集汇总

Table 12 Summary of UAV IR image datasets

数据集	数量	尺寸	环境	无人机种类	团队
AllSeqs	396	384×288	多种真实环境条件	—	Wang 等人(2019)
SIRST	427	—	城市、天空、海洋	—	Dai 等人(2021)
SIDD	4,737	640×512	城市、天空、山地、海洋	多种四旋翼无人机	Yuan 等人(2023)
IST-A	68,311	288×384	复杂背景, 包括云层、建筑物、植被	鸟类、无人机、直升机	Xu 等人(2023)
IRDST	142,727	720×480	天空、云彩、树木、电线、 建筑物、山脉、湖泊	无人机、直升机、飞机	Sun 等人(2023)
Anti-UAV	318 个视频	640×512	天空、城市、山地	6 种型号无人机	Jiang 等人(2021)
USC Drone Dataset	10 个视频	1920×1080	大学校园、室外	单一无人机	Wang 等人(2019)
Drone Detection	365 个视频	320×256	瑞典机场	3 种型号无人机	Svanström 等人(2021)

AllSeqs: https://github.com/wanghuanphd/MDvsFA_cGAN

SIRST: <https://github.com/YimianDai/open-acm>

SIDD: <https://github.com/Dang-zy/SIDD>

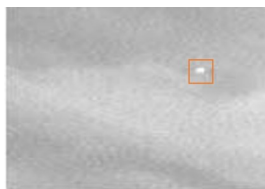
IST-A: <https://github.com/SeaHifly/IST-A>

IRDST: <https://github.com/sun11999/RDIAN?tab=readme-ov-file>

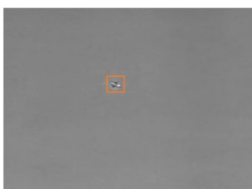
Anti-UAV: <https://github.com/ZhaoJ9014/Anti-UAV>

USC Drone Dataset: <https://github.com/chelicynly/A-Deep-Learning-Approach-to-Drone-Monitoring>

Drone Detection: <https://github.com/DroneDetectionThesis/Drone-detection-dataset>



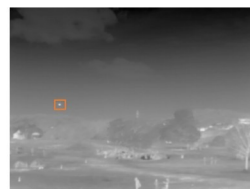
(a) AllSeqs



(b) SIRST



(c) SIDD



(d) IST-A

集,其专为评估长时跟踪与重检测性能而设计。ICG-Drone 包含多无人机交互场景,适用于多目标跟踪研究。IRDST 的序列版本也可跟踪算法提供额外的数据支持。

5.2 评价指标

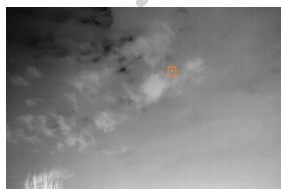
5.2.1 无人机目标检测评价指标

精确率(precision):衡量模型预测为正例的样本中真正为正例的比例,高精确率表示误报较少,计算公式为:

$$\text{Precision} = \frac{TP}{TP + FP} \tag{2}$$

式中:TP 为真正例;FP 为假正例。

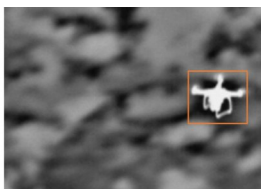
召回率(recall):衡量模型正确识别出的正例占



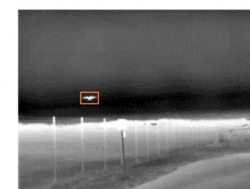
(e) IRDST



(f) Anti-UAV



(g) USC Drone Dataset



(h) Drone Detection

(a) AllSeqs; (b) SIRST; (c) SIDD; (d) IST-A; (e) IRDST; (f) Anti-UAV; (g) USC Drone Dataset; (h) Drone Detection

图9 无人机 IR 图像数据集示例

Fig. 9 Examples of UAV IR image datasets

表 13 无人机数据集优缺点对比

Table 13 Comparison of advantages and disadvantages of UAV datasets			
模态	数据集	核心优势	主要局限
可见光	Drone-vs-Bird	数据规模较大,专门针对“无人机 vs. 鸟类”这一高混淆度任务设计,极具现实意义。	图像尺寸未统一标注,缺乏极端天气和复杂夜间场景
	DUT Anti-UAV	规模大、分辨率高、背景多样、无人机类型极其丰富,是最全面的可见光检测基准之一	恶劣天气样本较少,极远距离小目标样本比例可提升
	LRDD	专注于长距离无人机检测任务,目标像素占比小,贴合实际安防场景中对早期预警的需求	数据总量相对较小,无人机种类和飞行模式不够丰富
	ICG-Drone	包含室内、室外场景,提供了可用于跟踪研究的视频序列	整体规模中等,室外场景的背景复杂度和天气变化相对有限
	UAVfly	数据量庞大,覆盖场景极为广泛(从城市到荒漠),有助于提升模型泛化能力	图像分辨率相对较低,可能限制其对超小目标特征的刻画
红外	UETT4K Anti-UAV	当前规模最大的4K数据集,超高分辨率为小目标提供了丰富上下文信息,场景多样	标注仅为边界框,缺乏更细粒度的属性信息;无人机种类相对较少
	AllSeqs	早期针对红外图像中小目标分割的基准数据集,推动了对抗学习在该领域的应用	规模较小,目标类别不专一于无人机,且图像分辨率较低
	SIRST	是红外小目标检测领域的经典基准数据集之一,包含了多种典型背景下的目标	数据规模中等,且并非专为无人机设计,目标类别包含飞机、鸟类等
	SIDD	专为红外反无人机任务构建,覆盖城市、天空、山地、海洋四类场景,场景区分度高	无人机种类信息未明确,且图像尺寸固定为640×512,分辨率适中
	IST-A	大规模红外小目标数据集,目标绝大多数小于6×6像素,背景复杂,贴近真实非合作侦查场景	目标不限于无人机,包含鸟类、直升机等;图像分辨率较低
可见光+热红外	IRDST	同时提供单帧和序列数据,背景多样,适用于检测与跟踪研究	目标类别包含无人机、直升机、飞机,非纯无人机数据集
	MMFW-UAV	数据规模巨大,4K分辨率,包含固定翼无人机和多视角、全天候条件,稀缺性强	数据集尚未完全公开,获取存在一定壁垒
	Anti-UAV	大规模、高质量的无人机跟踪专用数据集,提供同步的红外与可见光视频,标注密集	作为跟踪数据集,其静态检测任务评测需从视频中提取帧,且数据获取需申请
多模态	USC Drone Dataset	较早公开的无人机数据集,包含可见光与热红外视频,推动了基于深度学习的无人机监测研究	规模较小,视频数量有限,场景相对单一
	MMAUD	真正的多模态同步数据集,包含可见光、红外、事件相机、雷达、激光雷达等,并记录机械噪声	以视频为主,静态图像数量未明确标注;数据采集平台复杂,复现成本高
	Drone Detection	包含同步的可见光视频、红外视频和音频数据,支持真正的多传感器融合研究,且标注了无人机、直升机、鸟类、飞机等多类别	视频分辨率较低,且部分数据为机场环境,背景相对特定

实际正例的比例,高召回率表示漏检较少,计算公式为:

$$Recall = \frac{TP}{TP + FN} \tag{3}$$

式中:FN为假负例。混淆矩阵如图10所示,在该矩

阵中,TP、FP、FN和TN分别表示真正例、假正例、假负例和真负例。

F1-score:精确率和召回率的调和平均数,用于平衡两者,公式为:

		预测	
		正	负
标签	正	TP	FN
	负	FP	TN

图 10 混淆矩阵

Fig. 10 Confusion matrix

$$F_1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (4)$$

交并比 (Intersection over Union, IoU): 预测框与真实框的重叠度, 常用于分割任务中评估像素级准确性, 如图 11 所示。

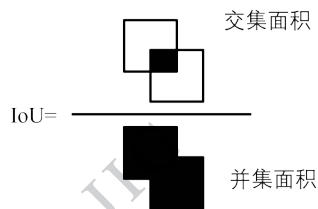


图 11 交并比示意图

Fig. 11 IoU diagram

平均精度 (average precision, AP): 基于精确率-召回率曲线下的面积计算, 反映模型在不同置信度阈值下的综合性能。常用变体包括: AP50: IoU 阈值为 0.5 时的平均精度, 侧重于宽松定位要求。AP50: 95: IoU 阈值从 0.5 到 0.95 的平均精度, 更严格地评估定位准确性。

平均精度均值 (mean average precision, mAP): 是所有类别 AP 的平均值。若任务有 N 个类别, 则公式为:

$$mAP = \frac{1}{N} \sum_{i=1}^N AP_i \quad (5)$$

5.2.2 无人机目标跟踪评价指标

成功率 (success): 在视频的所有帧中, 预测目标边界框与真实标注边界框的 IoU 超过设定阈值的帧所占比例。IoU 衡量预测框与真实框的重叠程度, 阈值通常设为 0.5, 即 $IoU > 0.5$ 时视为成功跟踪, 计算公式如下:

$$\text{Success} = \frac{1}{M} \sum_{i=1}^M \delta(IoU_i > T_0) \quad (6)$$

式中: M 是视频总帧数; IoU_i 是第 i 帧的预测框与真实框的 IoU 值; T_0 是阈值 (通常为 0.5); δ 是指示函

数, 当条件满足时值为 1, 否则为 0。

曲线下面积 (area under the curve, AUC): 表示成功率曲线下的面积, 用于综合衡量跟踪器在不同重叠阈值下的性能。成功率曲线绘制了 IoU 大于给定阈值的帧百分比, AUC 值越高, 表明跟踪器整体性能越稳健, 公式为:

$$AUC = \int_0^1 \frac{1}{M} \sum_{j=1}^M \frac{\sum_{i=1}^{N_j} I(IoU_{ji} > \tau)}{N_j} d\tau \quad (7)$$

式中: M 是视频序列数; N_j 是第 j 个序列的帧数; IoU_{ji} 是预测框与真实框的重叠率; τ 是阈值; I 是指示函数。

平均交并比 (mean intersection over union, mIoU): 表示平均重叠精度指标, 用于评估跟踪框与真实框的平均重叠程度, 特别适用于长时跟踪场景, 其中目标可能暂时消失或重现, 公式为:

$$mIoU = \frac{1}{T} \sum_t (IoU_t \cdot \delta(v_t > 0) + p_t (1 - \delta(v_t > 0))) \quad (8)$$

式中: T 是总帧数; IoU_t 是第 t 帧的重叠率; v_t 是真实可见性标志; p_t 是预测标志; $\delta(\cdot)$ 为指示函数, 条件成立时值为 1, 否则值为 0。

5.3 代表性方法性能评估

1) 无人机目标检测

依据 Precision、Recall、F1-score 和 mAP50 客观评价指标评估了 10 种无人机目标检测代表性方法在 DUT Anti-UAV 数据集上的检测性能, 结果如表 14 所示。综合分析表明, EDGS-YOLOv8 在 mAP50 上达到 0.971 的最高值, 同时 F1-score 为 0.952, 表明它在精度和召回率之间取得了良好平衡; YOLOv11+CAA+PA-Net 以 0.974 的精确率领先, 而 YOLO-G-KCF 则在召回率上表现最佳为 0.952。相比之下, LA-YOLO 和 DACG-Net 的召回率相对较低, 分别只有 0.867 和 0.851, 这限制了它们的 F1-score, 而 QHNet 在所有指标中均处于较低水平, 可能意味着其在复杂场景下的检测能力有待提升。

这些方法的性能差异反映了不同优化策略的侧重点, 例如高精度方法如 YOLOv11+CAA+PA-Net 更注重减少误报, 适合对准确性要求极高的应用, 而高召回方法如 YOLO-G-KCF 则能更好地捕捉目标, 但可能在精确性上略有妥协。综合来看, EDGS-YOLOv8 在多个指标上表现均衡且突出, 可作为该数据集上的一个强基线。

表 14 无人机目标检测代表性方法在 DUT Anti-UAV 数据集上的性能评估

Table 14 Performance evaluation of representative UAV object detection methods on the DUT Anti-UAV dataset

方法	团队	Precision	Recall	F1-score	mAP ₅₀
YOLOv7-GS	Bo C 等人 (2024)	0.968	0.903	0.934	0.932
EDGS-YOLOv8	Huang M 等人 (2024)	0.970	0.935	0.952	0.971
L-YOLO-AUAV	Gao Y 等人 (2024)	0.969	0.898	0.932	0.952
YOLOv11+CAA+PA-Ne	肖建国等人 (2025)	0.974	0.925	0.949	0.963
YOLO-G-KCF	Zhao J 等人 (2022)	0.953	0.952	0.952	0.942
LA-YOLO	Ma J 等人 (2024)	0.944	0.867	0.904	0.929
DACG-Net	Li W 等人 (2025)	0.963	0.851	0.904	0.928
YOLOv6 LERFNet	Elsayed M 等人 (2025)	0.958	0.912	0.934	0.935
QHNet	Wan Q 等人 (2025)	0.941	0.854	0.895	0.891
ESMS-YOLOv7	Wei Y 等人 (2025)	0.957	0.850	0.900	0.909

注:加粗字体为每行最优值。

2)无人机目标跟踪
依据 AUC、Precision、mIoU 和 FPS 客观评价指标评估了 10 种无人机目标跟踪代表性方法在 Anti-UAV 数据集上的跟踪性能,结果如表 15 所示。SCAR-T 在各项指标上均表现卓越,其 0.789 的

AUC、0.985 的 Precision、0.812 的 mIoU 及 63FPS 的综合成绩,使其成为当前数据集上性能最为全面的方法。紧随其后的 OSGT 与 SiamFusion 在跟踪精度上同样出色,而 EECF 虽在精度上略有妥协,却以 86.3FPS 的速度展现了突出的实时处理潜力。

表 15 无人机目标跟踪代表性方法在 Anti-UAV 数据集上的性能评估

Table 15 Performance evaluation of representative UAV object tracking methods on the Anti-UAV dataset

方法	团队	AUC	Precision	mIoU	FPS
STFTrack	Xie X 等人 (2023)	0.666	0.912	0.677	12.4
MAMC	Xu B 等人 (2024)	0.690	0.810	0.650	35
SCAR-T	Yu Q 等人 (2025)	0.789	0.985	0.812	63
SiamSTA	Huang B 等人 (2021)	0.729	0.969	0.673	-
FocusTrack	Wang Y 等人 (2025)	0.677	0.689	0.440	-
OSGT	Feng Z 等人 (2025)	0.764	0.979	-	27.3
SiamFusion	Zhang Z 等人 (2023)	0.734	0.977	-	14
MCJT	Zhang J 等人 (2025)	0.699	0.914	0.696	44
MACTrack	Xu W 等人 (2026)	0.620	0.889	0.679	19
EECF	Wu T 等人 (2024)	0.601	0.882	-	86.3

注:加粗字体为每行最优值。

进一步观察可见,跟踪性能的优化往往需要在速度与精度之间权衡:例如 MAMC 与 MCJT 在保持较高 AUC 的同时实现了较快的处理速度,而部分高精度方法如 SiamSTA 则未提供速度数据。在实际应用中,对跟踪稳定性与重叠度要求高的任务应优先

考虑 SCAR-T 等综合性能强的方法;在需要高速响应的监控场景中,EECF 或 MAMC 等高帧率方法可能更为适用。这些差异为不同应用场景下的方法选择提供了明确依据。

6 未来展望

尽管现有无人机视觉检测与跟踪方法在精度和效率方面取得了显著进展,但面对开放环境中的复杂干扰、多样化目标以及资源受限部署需求,仍存在诸多亟待突破的关键问题。未来研究应聚焦环境自适应感知、无人机本质特征建模、边缘智能协同优化以及大模型驱动的开放世界感知等方向,探索更加鲁棒、高效且具备泛化与语义理解能力的视觉感知框架,为无人机探测系统的实战化和智能化应用提供支撑。

1)实现环境自适应的无人机智能感知体系:当前无人机检测模型大多采用固定网络结构和静态参数配置,在光照变化、天气扰动、目标遮挡及背景复杂度变化时容易出现性能下降。未来研究应突破传统静态感知框架,构建具备环境认知与动态调节能力的自适应感知系统。通过实时感知场景状态并动态调整特征提取、信息融合及决策策略,实现模型对复杂环境的主动适应,从而提升无人机探测系统在开放场景中的鲁棒性和泛化能力。

2)构建面向无人机本质属性的鲁棒表征模型:当前深度学习方法更多依赖外观特征进行识别,容易受到视角变化、背景干扰以及目标伪装的影响。未来研究需要从无人机的几何结构、运动规律及动力学特征出发,探索融合视觉、时序和物理先验信息的统一表征框架,建立能够反映无人机本质属性的特征表达机制,提高模型对未知机型和相似干扰目标的区分能力。

3)实现高精度与低功耗协同的边缘智能探测:随着无人机反制系统向便携化和集群化发展,未来探测模型不仅需要具备高精度识别能力,还需满足实时性、低功耗和长期自主运行要求(王中天和吴一全,2025)。未来可探索算法—硬件协同优化机制,结合神经架构搜索、模型动态推理以及新型脉冲神经网络,实现感知、计算与资源调度的一体化设计,推动无人机探测系统从“高性能算法”向“高效智能系统”演进。

4)探索大模型驱动的开放世界无人机视觉感知:探索大模型驱动的开放世界无人机感知,随着基础模型和视觉语言模型的发展,其通用表征和跨模态语义理解能力为无人机检测提供了新的研究思

路。未来可利用大规模预训练获得的视觉与语义先验,探索零样本、少样本及开放词汇无人机检测,提高模型对未知机型、复杂场景和相似干扰目标的泛化能力。同时,应关注大模型参数量大、计算开销高以及领域适配不足等问题,通过参数高效微调、知识蒸馏和大小模型协同等方式,推动其在低空无人机智能感知中的实际应用。

参考文献(References)

- Ahmad I, Narmeen R, Mughal U A, Yang L, Almadhor A, and Dhahbi S, et al. 2025. Quantum CNN for detection and identification of UAV-enabled non-terrestrial networks. *IEEE Wireless Communications*, 32(3): 28-36. [DOI: 10.1109/MWC.001.2400419]
- Amin J, Anjum M A, Ibrar K, Sharif M, Kadry S and Crespo RG. 2023. Detection of anomaly in surveillance videos using quantum convolutional neural networks. *Image and Vision Computing*, 135: 104710. [DOI: 10.1016/j.imavis.2023.104710]
- Alfarano A, Maiano L, Papa L and Amerini I. 2024. Estimating optical flow: A comprehensive review of the state of the art. *Computer Vision and Image Understanding*, 249: 104160. [DOI: 10.1016/j.cviu.2024.104160]
- Awan M S, Zaidi S A A and Mir J. 2025. UETT4K anti-UAV: A large scale 4K benchmark dataset for vision-based drone detection in high-resolution imagery. *IEEE Access*, 13: 73553-73568. [DOI: 10.1109/ACCESS.2025.3562890]
- Briese C, Seel A and Andert F. 2018. Vision-based detection of non-cooperative UAVs using frame differencing and temporal filter//*Proceedings of the 2018 International Conference on Unmanned Aircraft Systems*. Dallas, USA: IEEE: 606-613. [DOI: 10.1109/ICUAS.2018.8453372]
- Bo C, Wei Y, Wang X, Shi Z and Xiao Y. 2024. Vision-based anti-UAV detection based on YOLOv7-GS in complex backgrounds. *Drones*, 8(7): 331. [DOI: 10.3390/drones8070331]
- Bu D, Ding B, Tong X, Sun B, Sun X and Guo R, et al. 2025. FSTC-DiMP: Advanced feature processing and spatio-temporal consistency for anti-UAV tracking. *Remote Sensing*, 17(16): 2902. [DOI: 10.3390/rs17162902]
- Cao J L, Li Y L, Sun H Q, Xie J, Huang K Q and Pang Y W. 2022. A survey on deep learning based visual object detection. *Journal of Image and Graphics*, 27(6): 1697-1722. (曹家乐, 李亚利, 孙汉卿, 谢今, 黄凯奇, 庞彦伟. 2022. 基于深度学习的视觉目标检测技术综述. *中国图象图形学报*, 27(6): 1697-1722). [DOI: 10.11834/jig.220069]
- Chen W S, Niu H C, Wang X, Wan J, Lu X F and Zhang J, et al. 2025. A survey of multi-source detection technologies for birds and UAVs in airport clearance area. *Acta Aeronautica et Astronautica*

- Sinica, 46(10): 233-259. (陈唯实, 牛红闯, 王鑫, 万健, 卢贤锋, 张洁, 等. 2025. 机场净空区飞鸟与无人机多源探测技术综述. 航空学报, 2025, 46(10): 233-259). [DOI: 10.7527/S1000-6893.2024.31251]
- Cheng X, Wang F, Hu X, Wu X and Nuo M. 2025. IRWT-YOLO: A background subtraction-based method for anti-drone detection. Drones, 9(4): 297. [DOI: 10.3390/drones9040297]
- Cheng K, Lei H T, and Lyu Z X. 2025. UAV tracking method based on improved YOLOv8 lightweight network. [EB/OL] Systems Engineering and Electronics: 1-15[2026-03-24]. (程鲲, 雷洪涛, 吕志轩. 2025. 改进 YOLOv8 轻量级网络的无人机跟踪方法. [EB/OL] 系统工程与电子技术: 1-15[2026-03-24].) <https://link.cnki.net/urlid/11.2422.TN.20250521.1742.009>.
- Coluccia A, Fascista A, Schumann A, Sommer L, Dimou A and Zarpalas D, et al. 2021. Drone vs. bird detection: Deep learning algorithms and results from a grand challenge. Sensors, 21(8): 2824. [DOI: 10.3390/s21082824]
- Cheng Q, Wang Y, He W and Bai Y. 2024. Lightweight air-to-air unmanned aerial vehicle target detection model. Scientific Reports, 14(1): 2609. [DOI: 10.1038/s41598-024-53181-2]
- Dong Y, Ma Y, Li Y and Li Z. 2023. High-precision real-time UAV target recognition based on improved YOLOv4. Computer Communications, 206: 124-132. [DOI: 10.1016/j.comcom.2023.04.019]
- Diamantidou E, Lalas A, Votis K and Tzovaras D. 2019. Multimodal deep learning framework for enhanced accuracy of UAV detection// Proceedings of International Conference on Computer Vision Systems. Cham, Switzerland: Springer International Publishing: 768-777. [DOI: 10.1007/978-3-030-34995-0_70]
- Dai Y, Wu Y, Zhou F and K Barnard. 2021. Asymmetric contextual modulation for infrared small target detection//Proceedings of Winter Conference on Applications of Computer Vision. Hawaii, USA: IEEE: 950-959. [DOI: 10.1109/WACV48630.2021.00099]
- Elsayed M, Reda M, Mashaly A S and Amein AS. 2025. LERFNet: An enlarged effective receptive field backbone network for enhancing visual drone detection. The Visual Computer, 41(4): 2219-2232. [DOI: 10.1007/s00371-024-03527-8]
- Feng X, Gong W W and Sun X E. 2025. Small drone target detection based on dual-modal image fusion. Laser & Optoelectronics Progress, 62(22): 325-333. (冯潇, 龚维维, 孙祥娥. 2025. 基于双模态图像融合的无人机小目标检测. 激光与光电子学进展, 2025, 62(22): 325-333). [DOI: 10.3788/LOP250854]
- Fang H, Ding L, Wang L, Chang Y, Yan L and Han J. 2022. Infrared small UAV target detection based on depthwise separable residual dense network and multiscale feature fusion. IEEE Transactions on Instrumentation and Measurement, 71: 1-20. [DOI: 10.1109/TIM.2022.3198490]
- Fang H, Ding L, Wang X, Chang Y, Yan L and Liu L, et al. 2024. SCI-Net: Spatial and contrast interactive super-resolution assisted infrared UAV target detection. IEEE Transactions on Geoscience and Remote Sensing, 62: 1-22. [DOI: 10.1109/TGRS.2024.3471786]
- Fang H, Liao Z, Wang X, Chang Y and Yan L. 2023. Differentiated attention guided network over hierarchical and aggregated features for intelligent UAV surveillance. IEEE Transactions on Industrial Informatics, 19(9): 9909-9920. [DOI: 10.1109/TH.2022.3232777]
- Fu C, Li B, Ding F, Lin F and Lu G. 2021. Correlation filters for unmanned aerial vehicle-based aerial tracking: A review and experimental evaluation. IEEE Geoscience and Remote Sensing Magazine, 10(1): 125-160. [DOI: 10.1109/MGRS.2021.3072992]
- Feng Z, Zhang W, Pan E, Liu D and Yu Q. 2025. Enhancing the scale adaptation of global trackers for infrared UAV tracking. Drones, 9(7): 469. [DOI: 10.3390/drones9070469]
- Fang H, Liao G, Liu Y, Zeng C, He X and Meng Q. 2024. A dual-mode framework for robust long-term tracking in video SAR. IEEE Sensors Journal, 24(8): 13028-13042. [DOI: 10.1109/JSEN.2024.3373396]
- Gao G S, Shang Y Q, and Dong Y. 2025. Review of deep learning algorithms for small object detection in optical remote sensing images. Journal of Image and Graphics, 2025, 30(11): 3479-3505. (高广帅, 尚云琦, 董燕. 2025. 光学遥感图像小目标检测深度学习算法综述. 中国图象图形学报, 2025, 30(11): 3479-3505). [DOI: 10.11834/jig.240740]
- Gao Y, Chen P, Liu Z Q, Tong X H, Yuan J, and Wang X. 2025. GUAV-YOLO: A lightweight detection model for gray-scale image UAV small targets. Acta Optica Sinica, 45(23): 98-113. (高仪, 陈鹏, 刘子奇, 童小华, 袁杰, 王响. 2025. GUAV-YOLO: 面向灰度影像无人机小目标的轻量化检测模型. 光学学报, 2025, 45(23): 98-113). [DOI: 10.3788/AOS250911]
- Gao Y, Xin Y, Yang H and Wang Y. 2024. A lightweight anti-unmanned aerial vehicle detection method based on improved YOLOv11. Drones, 9(1): 11. [DOI: 10.3390/drones9010011]
- Granillo O D M and Beltran Z Z. 2018. Real-time drone (UAV) trajectory generation and tracking by optical flow//Proceedings of the 2018 International Conference on Mechatronics, Electronics and Automotive Engineering (ICMEAE). Cuernavaca, Mexico: IEEE: 38-43. [DOI: 10.1109/ICMEAE.2018.00014]
- Harvey C and Inman D J. 2021. Aerodynamic efficiency of gliding birds vs comparable UAVs: a review. Bioinspiration & Biomimetics, 16(3): 031001. [DOI: 10.1088/1748-3190/abc86a]
- Huang M, Mi W and Wang Y. 2024. EDGS-YOLOv8: An improved YOLOv8 lightweight UAV detection model. Drones, 8(7): 337. [DOI: 10.3390/drones8070337]
- He T, Hou J and Chen D. 2025. Multimodal UAV target detection method based on acousto-optical hybridization. Drones, 9(9): 627. [DOI: 10.3390/drones9090627]
- Huang B, Chen J, Xu T, Wang Y, Jiang S and Y Wang, et al. 2021. SiamSTA: Spatio-temporal attention based Siamese tracker for

- tracking UAVs//Proceedings of the International Conference on Computer Vision. Montreal, Canada: IEEE: 1204-1212. [DOI: 10.1109/ICCVW54120.2021.00140]
- Jiang Y, Li X, Zhu G, Li H, Deng J and Han K. 2025. Integrated sensing and communication for low altitude economy: Opportunities and challenges. *IEEE Communications Magazine*, 63(12): 72-78. [DOI: 10.1109/MCOM.001.2400685]
- Jie L, Liu CF, Cheng D, Tong F, Nimin Z and Hang Z. 2021. Necessity analysis and scheme of constructing ultra-low-altitude defense system in megacities. *IEEE Aerospace and Electronic Systems Magazine*, 36(1): 14-21. [DOI: 10.1109/MAES.2020.3021303]
- Jiang N, Wang K, Peng X, Yu X, Wang Q and Xing J, et al. 2021. Anti-UAV: A large-scale benchmark for vision-based UAV tracking. *IEEE Transactions on Multimedia*, 25: 486-500. [DOI: 10.1109/TMM.2021.3128047]
- Kalsotra R and Arora S. 2022. Background subtraction for moving object detection: Explorations of recent developments and challenges. *The Visual Computer*, 38(12): 4151-4178. [DOI: 10.1007/s00371-021-02286-0]
- Köksal A, Tuzcuoğlu Ö, İnce K G, Ataseven Y and Alatan AA. 2022. Improved hard example mining approach for single shot object detectors//Proceedings of the 2022 IEEE International Conference on Image Processing (ICIP). Bordeaux, France: IEEE: 3536-3540. [DOI: 10.1109/ICIP46576.2022.9897806]
- Kemsaram N, Kanth T V R and Guntupalli D R. 2016. A mean-shift algorithm based autonomous visual tracking for micro aerial vehicles. *International Journal of Recent Trends in Engineering & Research*, 2(4): 362-369.
- Li D L, Cai H X, Wang T T, Zhao J X, Hua Y Y, and Li T, et al. 2025. Application analysis of low-altitude UAV small target detection and snapshot spectral imaging (invited). *Acta Optica Sinica*, 45(17): 376-388. (李栋梁, 蔡红星, 王婷婷, 赵俊霞, 花扬扬, 李腾, 等. 2025. 低空无人机小目标探测及快照式光谱成像应用分析(特邀). *光学学报*, 45(17): 376-388). [DOI: 10.3788/AOS251247]
- Luo J H and Wang Z Y. 2022. Review on development and application of UAV detection and countermeasure technology. *Control and Decision*, 37(3): 530-544. (罗俊海, 王芝燕. 2022. 无人机探测与对抗技术发展及应用综述. *控制与决策*, 2022, 37(3): 530-544). [DOI: 10.13195/j.kzyjc.2020.1507]
- Li M S, Zhou M H, and Zhi R C. 2025. A survey on UAV recognition based on multimodal fusion. *Computer Engineering and Applications*, 2025, 61(21): 1-14. (李旻姝, 周莫涵, 支瑞聪. 2025. 基于多模态融合的无人机识别研究综述. *计算机工程与应用*, 2025, 61(21): 1-14). [DOI: 10.3778/j.issn.1002-8331.2501-0014]
- Li Q, Zhang W, Lu W and Q Wang. 2025. Multi-branch mutual-guiding learning for infrared small target detection. *IEEE Transactions on Geoscience and Remote Sensing*, 63: 1-10. [DOI: 10.1109/TGRS.2025.3526754]
- Lei Y, Yao H, Jiang B, Tian T and Xing P. 2022. Anti-UAV high-performance computing early warning neural network based on PSO algorithm. *Scientific Programming*, 2022(1): 7150128. [DOI: 10.1155/2022/7150128]
- Liu K, Li J, Yan X W, and Zou K L. 2025. Infrared small UAV target detection based on multi-dimensional feature optimization. *Acta Photonica Sinica*, 54(8): 182-194. (刘坤, 李基, 晏行伟, 邹昆霖. 2025. 基于多维特征优化的红外小型无人机目标检测. *光子学报*, 2025, 54(8): 182-194). [DOI: 10.3788/gzxb20255408.0810001]
- Liang D, Gao S, Sun H, and Liu N Z. 2020. UAV target detection in moving camera by combining kernelized correlation filters and deep learning. *Acta Aeronautica et Astronautica Sinica*, 41(9): 159-171. (梁栋, 高赛, 孙涵, 刘宁钟. 2020. 结合核相关滤波器和深度学习的运动相机中无人机目标检测. *航空学报*, 2020, 41(9): 159-171). [DOI: 10.7527/S1000-6893.2020.23733]
- Li Q, Wang Z, Cui S W and He M Y. 2025. Lightweight spaceborne remote sensing object detection algorithm with multi-attention mechanism. *Journal of Image and Graphics*, 30(12): 3955-3968 (李强, 王智, 崔书玮, 何明一. 2025. 融合多注意力机制的轻量化天基遥感目标检测算法. *中国图象图形学报*, 30(12): 3955-3968)[DOI: 10.11834/jig.250002]
- Liu J, Wen B, Xiao J and Sun M. 2025. Design of UAV target detection network based on deep feature fusion and optimization with small targets in complex contexts. *Neurocomputing*, 639: 130207. [DOI: 10.1016/j.neucom.2025.130207]
- Li H, Xiao L, Cao L, Yao S, Wang M and Li Y. 2026. DFFormer: UAV object detection via feature scaling and interaction. *IEEE Transactions on Geoscience and Remote Sensing*, 64: 1-16. [DOI: 10.1109/TGRS.2026.3658082]
- Li M, Chen N, An W, Li B Y, Ling Q, and Li W X. 2024. Dual-view fusion detection method for event-camera-based UAV detection. *Opto-Electronic Engineering*, 51(11): 49-59. (李森, 陈诺, 安玮, 李博扬, 凌强, 李卫星. 2024. 面向事件相机探测无人机的双视图融合检测方法. *光电工程*, 2024, 51(11): 49-59). [DOI: 10.12086/oe.2024.240208]
- Liu R Q, Wang H Y and Han J Z. 2024. Detection and real-time following for UAV targets. *Computer Engineering and Applications*, 60(11): 319-327. (刘璐琦, 王红雨, 韩佼志. 2024. 面向无人机目标的检测与实时跟随. *计算机工程与应用*, 2024, 60(11): 319-327). [DOI: 10.3778/j.issn.1002-8331.2302-0281]
- Lin S B, Wu G S and Yang W Y. 2025. Target-aware transformer UAV tracking focusing on key information. *CAAI Transactions on Intelligent Systems*, 20(6): 1483-1492. (林淑彬, 吴贵山, 杨文元. 2025. 聚焦关键信息的目标感知Transformer无人机跟踪. *智能系统学报*, 2025, 20(6): 1483-1492). [DOI: 10.11992/tis.202506030]
- Liu Y, Sun Z, Xi L, Dong W, Chen C and Lu M, et al. 2025. MMFW-

- UAV dataset: Multi-sensor and multi-view fixed-wing UAV dataset for air-to-air vision tasks. *Scientific Data*, 12(1): 185. [DOI: 10.1038/s41597-025-04482-2]
- Li W, Xiao L, Li H, Yao S, Wan B and Ren D. 2025. DACG-Net: A dual-backbone and context-guided fusion network for aerial UAV detection. *IEEE Transactions on Aerospace and Electronic Systems*, 61(6): 17109-17123. [DOI: 10.1109/TAES.2025.3601109]
- Lyu P and Wu Y Q. 2026. Research progress of quantum image information processing methods. *Journal of Image and Graphics*: 1-30. (吕品, 吴一全. 2026. 量子图像信息处理方法研究进展. *中国图象图形学报*: 1-30). [DOI: 10.11834/jig.250566]
- Munir A, Siddiqui A J, Hossain M S and El-Maleh A. 2025. YOLO-RAW: Advancing UAV detection with robustness to adverse weather conditions. *IEEE Transactions on Intelligent Transportation Systems*, 26(6): 7857-7873. [DOI: 10.1109/ITS.2025.3560792]
- Ma J, Huang S, Jin D, Wang X, Li L and Guo Y, et al. 2024. LA-YOLO: An effective detection model for multi-UAV under low altitude background. *Measurement Science and Technology*, 35(5): 055401. [DOI: 10.1088/1361-6501/ad23c6]
- Pan X Y, Jia N X, Mu Y Z and Gao X R. 2023. Survey of small object detection. *Journal of Image and Graphics*, 28(9): 2587-2615. (潘晓英, 贾凝心, 穆元震, 高炫蓉. 2023. 小目标检测研究综述. *中国图象图形学报*, 28(9): 2587-2615). [DOI: 10.11834/jig.220455]
- Qin Z and Li Y. 2026. DCAM-DETR: Dual Cross-Attention Mamba Detection Transformer for RGB - Infrared Anti-UAV Detection. *Information*, 17(1):103. [DOI: 10.3390/info17010103]
- Qiu Y, Lu Y, Wang Y and Jiang H. 2023. ARODNet: Adaptive rain image enhancement object detection network for autonomous driving in adverse weather conditions. *Optical Engineering*, 62(11): 118101-118101. [DOI: 10.1117/1.OE.62.11.118101]
- Qiao X B, Guo Y, Zhao Y Q, and Liu L. 2025. UAV target detection method in complex interference environment based on infrared polarization imaging. *Acta Electronica Sinica*: 1-12. (乔新博, 郭阳, 赵永强, 刘亮. 2025. 基于红外偏振成像的复杂干扰环境下无人机目标检测方法. *电子学报*: 1-12). [DOI: 10.12263/DZXB.20250496]
- Rouhi A, Umare H, Patal S, Kapoor R, Deshpande N and Arezooman-dan S, et al. 2024. Long-range drone detection dataset//Proceedings of 2024 IEEE International Conference on Consumer Electronics (ICCE). Las Vegas, USA: IEEE: 1-6. [DOI: 10.1109/ICCE59016.2024.10444135]
- Sun H, Liu Q, Wang J, Ren J, Wu Y and Zhao H, et al. 2021. Fusion of infrared and visible images for remote detection of low-altitude slow-speed small targets. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 14: 2971-2983. [DOI: 10.1109/JSTARS.2021.3061496]
- Shabaz M and Soni M. 2024. Adaptive enhancement of spatial information in adverse weather. *Spatial Information Research*, 32(4): 473-484. [DOI: 10.1007/s41324-024-00577-x]
- Srinivas Y and Ganivada A. 2025. A modified inter-frame difference method for detection of moving objects in videos. *International Journal of Information Technology*, 17(2): 749-754. [DOI: 10.1007/s41870-024-02355-2]
- Saeed V A. 2024. A framework for recognition of facial expression using HOG features. *International Journal of Mathematics, Statistics, and Computer Science*, 2: 1-8. [DOI: 10.59543/ijmscs.v2i.7815]
- Shi Q and Li J. 2020. Objects detection of UAV for anti-UAV based on YOLOv4//Proceedings of the 2020 IEEE 2nd International Conference on Civil Aviation Safety and Information Technology (ICCASIT). Weihai, China: IEEE: 1048-1052. [DOI: 10.1109/ICCASIT50869.2020.9368788]
- Sun K, Wang X, Miao X and Q Zhao. 2025. A review of AI edge devices and lightweight CNN and LLM deployment. *Neurocomputing*, 614: 128791. [DOI: 10.1016/j.neucom.2024.128791]
- Shao J F, Cai S J and Liu J. 2025. YOLO-LDD: Lightweight drone detection algorithm. *Journal of Jilin University (Science Edition)*, 63(3): 867-877. (邵剑飞, 蔡世军, 刘杰. 2025. YOLO-LDD:轻量级无人机检测算法. *吉林大学学报(理学版)*, 2025, 63(3): 867-877). [DOI: 10.13413/j.cnki.jdxblxb.2024085]
- Sun H, Bai J, Yang F and Bai X. 2023. Receptive-field and direction induced attention network for infrared dim small target detection with a large-scale dataset IRDST. *IEEE Transactions on Geoscience and Remote Sensing*, 61: 1-13. [DOI: 10.1109/TGRS.2023.3235150]
- Svanström F, Alonso-Fernandez F and Englund C. 2021. A dataset for multi-sensor drone detection. *Data in Brief*, 39: 107521. [DOI: 10.1016/j.dib.2021.107521]
- Tong K and Wu Y Q. 2024. Progress in benchmark of small object detection based on deep learning. *Acta Electronica Sinica*, 52(3): 1016-1040. (童康, 吴一全. 2024. 基于深度学习的小目标检测基准研究进展. *电子学报*, 2024, 52(3): 1016-1040). [DOI: 10.12263/DZXB.20230624]
- Tian D, Wang C, Zhou D, Yan X and Zeng L. 2024. NFE-YOLO: A lightweight and efficient detection network for low, slow, and small drones. *IEEE Access*, 12: 175458-175471. [DOI: 10.1109/ACCESS.2024.3492058]
- Tao L, Hong T and Chao X. 2020. UAV recognition and positioning tracking based on YOLOv3. *Journal of University of Science and Technology Beijing*, 42(4): 463-468. (陶磊, 洪韬, 钞旭. 2020. 基于YOLOv3的无人机识别与定位追踪. *工程科学学报*, 2020, 42(4): 463-468). [DOI: 10.13374/j.issn2095-9389.2019.09.10.002]
- Wang B, Li Q, Mao Q, Wang J, Chen CLP and Shangguan A, et al. 2024. A survey on vision-based anti unmanned aerial vehicles methods. *Drones*, 8(9): 518. [DOI: 10.3390/drones8090518]
- Wang Z, Qi L, Tie Y, Ding Y and Bai Y. 2018. Drone detection based on FD-HOG descriptor//Proceedings of the 2018 International Con-

- ference on Cyber-Enabled Distributed Computing and Knowledge Discovery (CyberC). Zhengzhou, China: IEEE: 433-433. [DOI: 10.1109/CyberC.2018.00084]
- Wang G, Yang X, Li L, Gao K, Gao J and Zhang J, et al. 2024. Tiny drone object detection in videos guided by the bio-inspired magnocellular computation model. *Applied Soft Computing*, 163: 111892. [DOI: 10.1016/j.asoc.2024.111892]
- Wang C, Meng L, Gao Q, Wang J, Wang T and Liu X, et al. 2022. A lightweight UAV swarm detection method integrated attention mechanism. *Drones*, 7(1): 13. [DOI: 10.3390/drones7010013]
- Wan Q, Feng L, Xiao Z, Zhu Z, Xing H and Tian Y, et al. 2025. QHNet: A novel quad-head network for real-time detection of intruding drones. *IEEE Transactions on Geoscience and Remote Sensing*, 64: 1-18. [DOI: 10.1109/TGRS.2025.3567751]
- Wu C, Tang W, Rao Y, Chen Y, Ding H and Zhu S, et al. 2025. An uncertainty quantification-based method for anti-UAV detection in infrared images. *Computers, Materials & Continua*, 83(1): 1415. [DOI: 10.32604/cmc.2025.059797]
- Wan Y, Ji J, Zhou F, Wu Q and Quek T Q S. 2025. From static dense to dynamic sparse: vision-radar fusion-based UAV detection. *IEEE Transactions on Information Forensics and Security*, 20: 4454-4468. [DOI: 10.1109/TIFS.2025.3560551]
- Wang G, Jiang Q, Jin X, Wozniak M, Wang P and Yao S. 2024. RTM-UAVDET: A real-time multimodal UAV detector. *IEEE Transactions on Aerospace and Electronic Systems*, 61(1): 473-489. [DOI: 10.1109/TAES.2024.3446460]
- Wang X D, Wang Y W and Yan H. 2019. Continuously adaptive mean shift tracking algorithm with background-suppressed histogram model. *Journal of Electronics & Information Technology*, 41(6): 1480-1487. (王旭东, 王屹炜, 闫贺. 2019. 背景抑制直方图模型的连续自适应均值漂移跟踪算法. *电子与信息学报*, 2019, 41(6): 1480-1487). [DOI: 10.11999/JEIT180588]
- Wang Y, Xu T and Li J. 2025. FocusTrack: A self-adaptive local sampling algorithm for efficient anti-UAV tracking. *IEEE Transactions on Geoscience and Remote Sensing*, 63: 1-14. [DOI: 10.1109/TGRS.2025.3562958]
- Wu H and Liu G. 2022. A dynamic infrared object tracking algorithm by frame differencing. *Infrared Physics & Technology*, 127: 104384. [DOI: 10.1016/j.infrared.2022.104384]
- Wu H, Li W, Li W and Liu G. 2020. A real-time robust approach for tracking UAVs in infrared videos//*Proceedings of Conference on Computer Vision and Pattern Recognition Workshops*. Seattle, USA: IEEE: 1032-1033. [DOI: 10.1109/CVPRW50498.2020.00524]
- Wang Z, Hu Y, Yang J, Zhou G, Liu F and Liu Y. 2024. A contrastive-augmented memory network for anti-UAV tracking in TIR videos. *Remote Sensing*, 16(24): 4775. [DOI: 10.3390/rs16244775]
- Wang W S, He J Y, Huang M, and Wu G X. 2025. UAV target recognition and tracking algorithm by fusing gradient improved YOLO and KCF model. *Chinese Journal of Scientific Instrument*, 46(2): 221-233. (王文胜, 何君尧, 黄民, 吴国新. 2025. 融合梯度改进 YOLO 和 KCF 模型的无人机目标识别跟踪算法. *仪器仪表学报*, 2025, 46(2): 221-233). [DOI: 10.19650/j.cnki.cjsi. J2413509]
- Wang H, Zhou L and Wang L. 2019. Miss detection vs. false alarm: Adversarial learning for small object segmentation in infrared images//*Proceedings of International Conference on Computer Vision*. Seoul, South Korea: IEEE: 950-959. [DOI: 10.1109/ICCV.2019.00860]
- Wang Y, Chen Y, Choi J and Kuo C C J. 2019. Towards visible and thermal drone monitoring with convolutional neural networks. *APSIPA Transactions on Signal and Information Processing*, 8: e5. [DOI: 10.1017/ATSIP.2018.30]
- Wei Y, Wang X, Bo C and Shi Z. 2025. Small target drone algorithm in low-altitude complex urban scenarios based on ESMS-YOLOv7. *Cognitive Robotics*, 5: 14-25. [DOI: 10.1016/j.cogr.2024.11.004]
- Wu T, Duan H and Zeng Z. 2024. Biological eagle-eye-based correlation filter learning for fast UAV tracking. *IEEE Transactions on Instrumentation and Measurement*, 73: 1-12. [DOI: 10.1109/TIM.2024.3436132]
- Wang Z T and Wu Y Q. 2025. Research progress on vision and deep learning based scene perception methods for UAV autonomous landing. *Acta Electronica Sinica*, 53(11): 4171-4198. (王中天, 吴一全. 2025. 基于视觉与深度学习的无人机自主着陆场景感知方法研究进展. *电子学报*, 2025, 53(11): 4171-4198). [DOI: 10.12263/DZXB.20250775]
- Xu L and Luo Z. 2025. Anti-UAV detection and identification technology: Fundamentals, methods and challenges. *Physical Communication*, 2025: 102676. [DOI: 10.1016/j.phycom.2025.102676]
- Xie J Y, Wang X J, Shi Z G, Wu J F, Chen J M, and Chen Q, et al. 2021. Drone detection and tracking algorithm for dynamic pan-tilt-zoom camera. *CAAI Transactions on Intelligent Systems*, 16(5): 858-869. (谢家阳, 王行健, 史治国, 吴均峰, 陈积明, 陈潜, 等. 2021. 动态云台摄像机无人机检测与跟踪算法. *智能系统学报*, 2021, 16(5): 858-869). [DOI: 10.11992/tis.202103032]
- Xiao J G, Xiong Y H, Zhang F K, He W, and Shu M L. 2025. Research on UAV target detection technology based on YOLOv11 algorithm. *Aero Weaponry*, 32(4): 80-87. (肖建国, 熊宇虹, 张方凯, 何伟, 舒明磊. 2025. 基于 YOLOv11 算法的无人机目标检测技术的研究. *航空兵器*, 2025, 32(4): 80-87). [DOI: 10.12132/ISSN.1673-5048.2025.0027]
- Xie W, Wan Y, Wu G, Li Y, Zhou F and Wu Q. 2024. A RF-visual directional fusion framework for precise UAV positioning. *IEEE Internet of Things Journal*, 11(22): 36736-36747. [DOI: 10.1109/JIOT.2024.3424271]
- Xie X, Xi J, Yang X, Lu R and Xia W. 2023. STTTrack: Spatio-temporal-focused Siamese network for infrared UAV tracking. *Drones*, 7(5): 296. [DOI: 10.3390/drones7050296]

- Xu B, Hou R, Bei J, Ren T and Wu G. 2024. Jointly modeling association and motion cues for robust infrared UAV tracking. *The Visual Computer*, 40 (12) : 8413-8424. [DOI: 10.1007/s00371-023-03245-7]
- Xiao J, Chee J H and Feroskhan M. 2024. Real-time multi-drone detection and tracking for pursuit-evasion with parameter search. *IEEE Transactions on Intelligent Vehicles*, 9 (10) : 6027-6037. [DOI: 10.1109/TIV.2024.3360433]
- Xu H, Zhong S, Zhang T and Zou X. 2023. Multiscale multilevel residual feature fusion for real-time infrared small target detection. *IEEE Transactions on Geoscience and Remote Sensing*, 61: 1-16. [DOI: 10.1109/TGRS.2023.3324502]
- Xu W, Sun H, Wang S, Zhang X and Ji Y. 2026. MACTrack: Spatio-temporal context propagation with motion compensation for anti-UAV tracking. *Neural Networks*: 108725. [DOI: 10.1016/j.neunet.2026.108725]
- Yang X, del Rey Castillo E, Zou Y and Wotherspoon L. 2024. UAV-deployed deep learning network for real-time multi-class damage detection using model quantization techniques. *Automation in Construction*, 159: 105254. [DOI: 10.1016/j.autcon.2023.105254]
- Yang H, Liang B, Feng S, Jiang J, Fang A and Li C. 2025. Lightweight UAV detection method based on IASL-YOLO. *Drones*, 9(5) : 325. [DOI: 10.3390/drones9050325]
- Yu Q, Wang J, Yang D, Yang D, Yang W and Zhang T. 2025. A framework strengthened by scenario compensation for anti-UAV tracking. *Signal, Image and Video Processing*, 19 (12) : 977. [DOI: 10.1007/s11760-025-04465-1]
- Yu J, Xu T, Zhu X and Wu XJ. 2024. Local point matching for collaborative image registration and RGBT anti-UAV tracking//*Proceedings of Chinese Conference on Pattern Recognition and Computer Vision (PRCV)*. Singapore: Springer Nature Singapore: 418-432. [DOI: 10.1007/978-981-97-8858-3_29]
- Yuan S, Yang Y, Nguyen T H, Nguyen TM, Yang J and Liu F, et al. 2024. MMAUD: A comprehensive multi-modal anti-UAV dataset for modern miniature drone threats//*Proceedings of 2024 IEEE International Conference on Robotics and Automation (ICRA)*. Yokohama, Japan: IEEE: 2745-2751. [DOI: 10.1109/ICRA57147.2024.10610957]
- Yuan S, Sun B, Zuo Z, Huang H, Wu P and Li C, et al. 2023. IRSDD-YOLOv5: Focusing on the infrared detection of small drones. *Drones*, 7(6) : 393. [DOI: 10.3390/drones7060393]
- Zhou D and Wang X. 2023. Research on high robust infrared small target detection method in complex background. *IEEE Geoscience and Remote Sensing Letters*, 20: 1-5. [DOI: 10.1109/LGRS.2023.3297523]
- Zhang C, Wang D B and Farooq M. 2010. Real-time tracking of fast moving objects in complex background. *Transactions of Nanjing University of Aeronautics & Astronautics*, 27(4) : 321-325. [DOI: 10.16356/j.1005-1120.2010.04.012]
- Zhang B, Hu S M, Du J, Yang X, Chen X and Jiang H, et al. 2024. Detecting moving objects in photometric images using 3D Hough transform. *Publications of the Astronomical Society of the Pacific*, 136(5) : 054502. [DOI: 10.1088/1538-3873/ad481f]
- Zhou C, Liu Y and Song Y. 2016. Detection and tracking of a UAV via Hough transform//*Proceedings of the 2016 CIE International Conference on Radar (RADAR)*. Guangzhou, China: IEEE: 1-4. [DOI: 10.1109/RADAR.2016.8059453]
- Zhu C, Xie X, Xi J and Yang X. 2025. GM-DETR: Infrared detection of small UAV swarm targets based on detection transformer. *Remote Sensing*, 17(19) : 3379. [DOI: 10.3390/rs17193379]
- Zuo Z, Yuan S D, Li C, Li C, and Huang H H. 2025. Multi-scale learning based infrared UAV target detection algorithm. *Journal of National University of Defense Technology*, 47(6) : 224-234. (左震, 袁书东, 李灿, 李灿, 黄泓赫. 2025. 多尺度学习的红外无人机目标检测算法. *国防科技大学学报*, 2025, 47(6) : 224-234). [DOI: 10.1887/j.issn.1001-2486.2409004]
- Zhang R X, Li N, Zhang X X and Zhang H Y. 2022. Low-altitude UAV detection method based on optimized CenterNet. *Journal of Beijing University of Aeronautics and Astronautics*, 48 (11) : 2335-2344. (张瑞鑫, 黎宁, 张夏夏, ZHOU Huiyu. 2022. 基于优化 CenterNet 的低空无人机检测方法. *北京航空航天大学学报*, 2022, 48 (11) : 2335-2344). [DOI: 10.13700/j.bh.1001-5965.2021.0108]
- Zhong Y C, Chen Y, Yang Z C and Li D L. 2025. Improved YOLOv8 algorithm for small UAV detection application. *Journal of Hunan University (Natural Sciences)*, 52(4) : 57-67. (仲元昌, 陈宇, 杨子楚, 李大林. 2025. 面向小型无人机检测应用的改进 YOLOv8 算法. *湖南大学学报(自然科学版)*, 2025, 52(4) : 57-67). [DOI: 10.16339/j.cnki.hdxzbk.2025266]
- Zhang H, Liu Z, Ji M, Li K and Yu Y. 2025. A novel robust tracking algorithm for anti-UAV based on dynamic similarity scale estimation and adaptive occlusion-aware. *Physica Scripta*, 100 (5) : 056004. [DOI: 10.1088/1402-4896/adc3d4]
- Zhang Z, Jin L, Li S, Xia JQ, Wang J and Li Z, et al. 2023. Modality meets long-term tracker: A Siamese dual fusion framework for tracking UAV//*Proceedings of the 2023 IEEE International Conference on Image Processing (ICIP)*. Kuala Lumpur, Malaysia: IEEE: 1975-1979. [DOI: 10.1109/ICIP49359.2023.10222679]
- Zhang J, Lin Y, Zhou X, Zhu X and Zeng D. 2025. Precision in pursuit: A multi-consistency joint approach for infrared anti-UAV tracking. *The Visual Computer*, 41 (4) : 2187-2202. [DOI: 10.1007/s00371-024-03525-w]
- Zhang J, Chen Z, Zheng C, Yi W, Yan G and Wang Y. 2026. Anti-DETR: End-to-end anti-drone visual detection network based on wavelet convolution. *Neurocomputing*, 674: 132935. [DOI: 10.1016/j.neucom.2026.132935]
- Zhu Y, Chen M, Wang S, Hu Y, Liu Y and Yin C. 2024. Collaborative reinforcement learning based unmanned aerial vehicle (UAV) trajectory design for 3D UAV tracking. *IEEE Transactions on Mobile*

Computing, 23 (12) : 10787-10802. [DOI: 10.1109/TMC. 2024. 3382913]

Zhao J, Zhang J, Li D and Wang D. 2022. Vision-based anti-UAV detection and tracking. IEEE Transactions on Intelligent Transportation Systems, 23(12): 25323-25334. [DOI: 10.1109/TITS.2022. 3177627]