

低空反无人机探测与感知

陈振国, 孙一铭, 朱鹏飞

(东南大学自动化学院, 江苏 南京 210096)

摘要: 近年来, 随着无人机技术的广泛应用, 未经授权的“黑飞”事件频发, 对低空空域安全构成了严重威胁, 反无人机探测与感知技术的研究极具现实紧迫性。然而, 低空场景中存在微小目标难以捕捉、密集遮挡及复杂电磁与气象干扰等问题, 传统单一传感器的探测精度与鲁棒性面临严峻挑战。针对这一问题, 本文全面综述了低空反无人机探测与感知技术的关键算法与前沿进展。首先梳理了基于视觉传感器的目标检测算法演进, 重点剖析了传统卷积神经网络、极致轻量化的 YOLO 系列, 以及具备全局上下文感知能力的 Transformer 架构在“低、慢、小”无人机检测中的创新应用。其次, 系统分析了雷达、声波和射频等非视觉传感器结合深度学习技术的检测优势与技术瓶颈。最后, 针对全天候与复杂环境下的探测需求, 深入探讨了可见光与红外以及多光谱等多模态融合检测算法的发展现状。本文总结了当前反无人机核心算法面临的感知鲁棒性与泛化能力挑战, 并展望了视觉大模型、空地多智能体协同与多源信息融合等未来重点研究方向, 以期为低空反无人机探测技术的进一步突破与工程实践提供有益参考。

关键词: 低空反无人机; 探测与感知; 多模态感知; 目标检测;

中图分类号: V279; [U8]

文献标志码: A

Low-Altitude Counter-UAV Detection and Perception

Abstract: In recent years, with the widespread application of unmanned aerial vehicle technology, frequent unauthorized “black flying” incidents have posed a severe threat to low-altitude airspace security, making the research of counter-UAV detection and perception technologies highly urgent. However, conventional single-sensor detection methods face significant challenges in accuracy and robustness due to the difficulty of capturing micro-targets, dense obstacles, and complex electromagnetic and meteorological disturbances in low-altitude scenarios. To address this issue, this paper provides a comprehensive review of the key algorithms and cutting-edge advancements in low-altitude counter-UAV detection and perception. First, it systematically summarizes the evolution of vision-based object detection algorithms, focusing on the innovative applications of traditional Convolutional Neural Networks, the ultra-lightweight YOLO series, and Transformer architectures with global context perception capabilities in detecting “low, slow, and small” UAVs. Second, it analyzes the detection advantages and technical bottlenecks of non-visual sensors, including radar, acoustic, and radio frequency, when combined with deep learning technologies. Finally, addressing the demands for all-weather detection in complex environments, the paper deeply explores the current development of multimodal fusion algorithms, such as RGB-infrared and multispectral fusion. This paper concludes by summarizing the current challenges in perception robustness and generalization capabilities of core counter-UAV algorithms, and anticipates future critical research directions including visual large models, air-ground multi-agent coordination, and multi-source information fusion, aiming to provide useful references for the technological evolution and engineering practice of low-altitude counter-UAV systems.

Key words: low-altitude counter-UAV; detection and perception; multimodal perception; object detection

收稿日期: 2026-03-29

基金项目: 国家自然科学基金重点项目(62436002); 国家自然科学基金青年科学基金项目(62506073); 国家重点研发计划雄安新区科技创新专项(2025XAGG0039); 天津市杰出青年科学基金项目(23JCJC00270); 中国博士后科学基金国家资助博士后研究人员计划(GZB20250395)

近年来,无人机(UAV)技术在智能化与自动化领域取得突破性进展,其应用已从早期的军事侦察迅速拓展至物流配送、精准农业、影视制作与灾害救援等多元化民用场景。据中商产业研究院数据显示,全球无人机市场正呈现爆发式增长:军用市场规模预计将从2023年的121.24亿美元攀升至2025年的135.43亿美元;与此同时,中国民用无人机市场规模也预计由2023年的1174.3亿元激增至2025年的1691亿元。随着“低空经济”于2021年被正式纳入《国家综合立体交通网规划纲要》的战略框架,低空空域的开发与利用正迎来前所未有的发展机遇。

伴随产业繁荣而来的无人机滥用问题日益凸显,未取得合法飞行资格或未实名注册的“黑飞”违规行为,已成为威胁低空空域安全的新型隐患。低空“黑飞”事件不仅干扰公共秩序,更对民航运行与国家军事安全构成严重威胁。在民航领域,2024年3月广东韶关丹霞机场及同年9月天津滨海国际机场,均因无人机非法侵入导致航班大面积延误,造成巨大的经济损失与极其恶劣的社会影响。在军事安全防范方面,部分违规飞行甚至直抵敏感区域,如2021年11月国家安全机关查处的“利用无人机非法拍摄军舰案”,更是敲响了防范新型技术泄密的警钟。

面对日益严峻的低空安全挑战,全球多国相继出台政策法规并加大资金投入,以构建多层次的反无人机防御体系。由于任何高效的拦截与反制手段均高度依赖于前端精准的态势感知,美国国防部联合反小型无人机系统办公室(JCO)计划在2024至2029年间投入超100亿美元,重点升级多传感器融合的早期预警与探测网络;欧盟亦通过相关通用规则进一步规范了低空空域的监视、识别与探测标准。在国内,2024年1月正式施行的《无人驾驶航空器飞行管理暂行条例》对低空目标的“早发现、早识别”提出了更高的硬性要求,而如“天穹”等综合防御系统的实战部署,则进一步确立了复杂场景下多模态探测与目标精细化识别的先决地位。这种政策法规的宏观引导与实战防御的迫切需求相互交织,共同确立了“感知层”在军民两用反无人机体系中的先决战略地位,并为低空复杂环境下的探测与感知算法研究指明了核心方向。

1 反无人机探测感知挑战与手段

1.1 低空探测任务的本质挑战与难点

反无人机探测感知的核心难点深刻植根于低空环境下“低、慢、小”目标带来的极低检测概率与极高虚警率,其物理挑战可高度抽象为三个交织的维度。首先在极低信杂比特征提取中,微小无人机不足0.01平方米的雷达截面积及其微弱的声热信号,极易被城市复杂电磁与环境噪声彻底淹没,导致目标在特征空间中极度不显著。其次,低空探测持续面临建筑植被等严重地物杂波及飞鸟等类似物的频繁侵入,这种高度非平稳的动态监测背景迫使算法必须具备敏锐的语义判别与特征解耦能力,以有效规避实战中的“虚警爆炸”。最后雾、雨、雪及光照突变等极端气象会造成光学对比度骤降或信号严重衰减,引发跨环境维度的不可逆性能退化。目前探测感知算法在恶劣条件下的泛化与鲁棒性表现,已然成为制约系统从理论走向规模化实战部署的关键技术瓶颈。

1.2 现有反无人机检测技术

为突破上述物理限制,现有的反无人机探测技术正经历着从单一物理传感器向视觉智能化、再向多源异构深度融合的跨越演进。如图1的反无人机检测算法发展历程所示,该时间线清晰地刻画了探测手段从“单点突破”迈向“多维协同”的技术脉络。

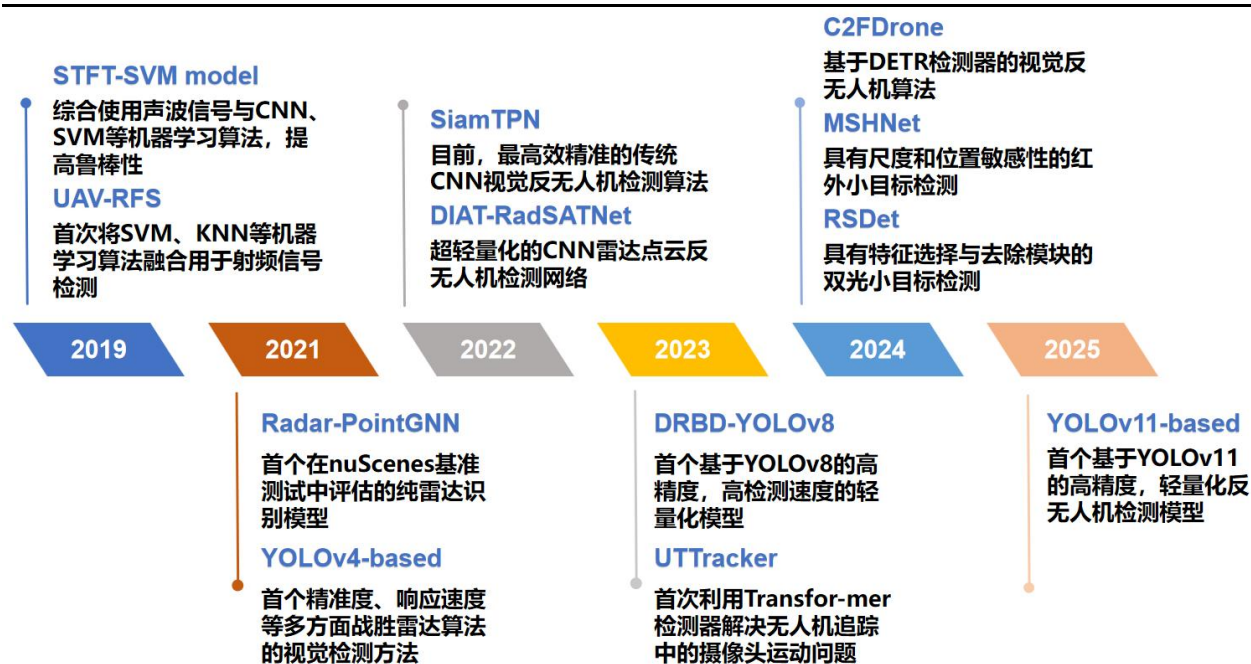


图 1 反无人机检测算法发展历程

Fig.1 Development history of anti-UAV detection algorithms

目前, 反无人机探测感知体系主要划分为以下三大类技术路径:

1) 视觉检测技术:视觉感知凭借高分辨率与精细化的语义识别能力, 已成为近年反无人机领域的研究核心。在 2021 年以前, 研究多侧重于传统卷积神经网络 (CNN) 对特定场景的适配; 而随着数据集的完善, 基于 YOLO 系列的极致轻量化检测器和具备全局上下文感知能力的 Transformer 架构, 显著提升了微小目标在复杂动态背景下的检测精度。然而, 视觉传感器对气象条件极度敏感, 在夜间、逆光或雨雾环境下面临严重的成像质量退化。

2) 其他传感器检测技术 (雷达、射频与声波):在视觉模态受限的场景下, 雷达、射频与声波等传感器提供了不可或缺的物理特征补充。雷达技术通常依赖激光或毫米波雷达获取目标的微多普勒特征, 实现远距离全天候监测^[2], 但在应对非金属材料、低速悬停目标以及城市地面杂波时存在盲区。射频 (RF) 探测与声波阵列作为被动探测手段, 虽具备良好的隐蔽性, 但在密集的通信干扰或高背景噪声下, 极易产生频率混叠与误报, 且声波探测的有效距离通常难以满足工程预警需求。

3) 多模态融合检测技术:纵观视觉及各类物理传感器的性能瓶颈可以发现, 面对复杂多变的低空环境, 单一模态技术路径在远距探测和全天候适应能力上存在明显不足, 构建多模态融合的反无人机检测体系已成为低空安防的核心发展趋势^[1]。这种演进趋势旨在通过构建异构传感器协同架构, 利用空间分辨率、穿透性及抗干扰能力上的深度互补, 有效克服单一传感器失效的风险, 从而实现全天候、全场景下的高鲁棒性探测。

基于对各类探测手段物理特性的深入剖析, 表 1 进一步确立了多维度的横向评价框架, 旨在通过对感知距离、环境鲁棒性等关键指标的定量化梳理, 揭示单一探测模态在复杂低空环境下的性能阈值。这一对比结果不仅凸显了视觉算法在目标精细化识别领域的先导地位, 更客观地揭示了跨模态协同的必要性, 多源异构信息的深度耦合才是实现全场景感知可靠性的核心演进方向。

表 1 典型反无人机探测技术多维度性能评价框架

Tab.1 Multi-dimensional performance evaluation framework for typical anti-UAV detection technologies						
探测模态	有效感知距离	目标敏感性(RCS)	全天候能力	遮挡/背景鲁棒性	实时性	部署成本
雷达检测	远 (>3 km)	较低	极强	中等 (受杂波干扰)	高	极高
视觉检测	近/中程	极高	弱	弱 (易受遮挡)	中	较低
射频检测	远 (非视距)	较低	强	中等 (受电磁干扰)	极高	中等
声波检测	极近 (<200 m)	中等	中等	弱 (易被噪声淹没)	高	极低
多模态融合	远	极高	极强	极强 (互补优势)	中等	较高

2 视觉检测算法

2.1 传统 CNN 方法

CNN 模型起始应用于手写数字识别，而后又逐渐拓展到图像分类，图像识别与目标检测以及一系列后续的任务。CNN 模型在图像处理，视觉识别领域具有得天独厚的优势，同时在科研工作者的不断推进下，CNN 在视觉领域具有不可撼动的地位，在反无人机检测这个任务上，基于 CNN 模型的经典工作也有许多。

在 2020 年，Tong 等提出的 DAGNet^[3]是一种新颖的差异化注意力引导网络，旨在提高复杂背景下红外无人机目标的检测性能。DAGNet 通过精心设计的注意力机制，自适应地增强不同尺度无人机目标的特征并抑制复杂背景干扰。该网络包含空间感知通道注意力(SCA)、通道调制变形空间注意力(CDSA)以及结合前两者的双维度联合注意力(DCA)。其设计基于对不同层次特征贡献的理解：SCA 通过保留关键空间特征并利用通道间依赖关系聚焦于大尺度目标；CDSA 通过细化通道上下文和动态感知空间特征关注小尺度目标；DCA 则用于中等尺度目标的检测。此外，网络在检测分支中嵌入了特征聚合器(FA)，通过自下而上的上下文调制路径，指导高低层特征图之间的信息交换，将低层空间细节传递到深层特征图中，从而有效增强了网络对目标位置的感知与分类意识，为红外无人机监控系统提供了一种高效的检测方案。

在 2024 年 Cao 等提出的 SR-TOD^[4] (self-reconstructed tiny object detection) 框架则创新性地解决了微小目标检测中的信息丢失问题。该框架通过引入自重建机制，构建差异图作为微小目标的先验信息，使特征对检测器更加可见。具体而言，图像输入骨干网络后，先利用特征金字塔网络(FPN)创建多尺度特征，随后将 P2 级别特征图输入重建头，生成与原始输入尺寸一致的重建图像。通过计算重建图像与原图的差异图，网络能够精准定位并增强微小目标特征。其核心创新点在于，自重建机制在不改变特征图分辨率的前提下增强了目标特征；同时，基于差异图引导的特征增强模块(DGFE)利用先验信息，通过元素级注意力机制使微小目标特征更加清晰。该方法在作者提出的高难度微小无人机数据集(DroneSwarms)及其他公开数据集上均表现优异，有效提升了系统的鲁棒性。

同年，Liu 等针对复杂环境提出了 MSHNet^[5] (Multi-Scale Head Net) 红外小目标检测模型。无人机检测常面临目标尺寸小、背景复杂及光照噪声干扰等挑战。MSHNet 在保持模型结构轻量化的同时，引入了新颖的尺度和位置敏感损失函数(SLS Loss)以及多尺度预测头。在训练过程中，SLS Loss 通过为不同尺度的目标分配自适应权重，使模型更关注尺度差异较大的目标；同时，位置敏感损失利用极坐标系下的距离和角度计算，显著增强了对目标中心位置的定位能力，有效减少了误报和漏报。此外，其基于 U-Net 结构的多尺度预测头在不同特征层上分别预测并融合结果，不仅提高了对无人机多尺度变化的

适应性，还降低了计算复杂度，使其在实时性要求较高的场景中极具优势。

总而言之，这一阶段的改进方法主要沿着特征表达增强与定位与损失优化两条主线演进。虽然这些方法在特定场景下取得了显著成效，但随着低空对抗环境的日益复杂，传统网络在全局上下文建模能力和计算资源消耗上的局限性逐渐显现。这也促使反无人机视觉检测技术开始向注重极致轻量化的 YOLO 系列，以及具备全局感知能力的 Transformer 架构加速演进。

2.2 YOLO-based 方法

YOLO (You Only Look Once) 是一种基于视觉的端到端深度学习目标检测系统^[6]。YOLO 系列模型以其检测速度快、实时性高而闻名，在反无人机领域，随着从早期的 YOLOv3^[7]到 YOLOv5、YOLOv7^[8]及 YOLOv8^[9]的迭代，对低慢小无人机的识别能力不断提升。特别是 YOLOv8，作为 Ultralytics 公司在 2023 年推出的重大更新版本，同时支持图像分类、物体检测和实例分割任务，为反无人机检测提供了更强劲的基线。

针对小目标特征极易在卷积中丢失的难题，Fang 等^[10]在 YOLOv8 的基础上设计了 SEB-YOLOv8s 检测方法。该方法首先引入了 SPD-Conv 卷积模块^[11]，利用空间到深度 (SPD) 层和非跨步卷积有效提取并保留小目标特征。同时，为更高效地利用空间信息，作者对原有的 C2f 模块进行改进，设计出融合 EMA 模块^[12]的 AttC2f 结构，增强了网络对浅层特征图的关注，显著提升了复杂背景下的检测抗干扰能力。网络中还嵌入了双层路由注意力机制^[13]，通过逐层过滤最不相关的区域，以极低的计算成本实现了对无人机目标的精准聚焦，有效解决了小目标的错检与漏检问题，实现了高精度与低计算消耗的平衡。类似地，ES-YOLO^[14]也通过细节特征增强与冗余特征抑制的策略，有效提升了小目标检测精度。

除了致力于提升检测精度，部分研究则侧重于模型的极致轻量化，以适应边缘计算设备（如无人机载算力或便携式防空设备）的实时部署需求。例如，Jiang 等^[15]提出的 DRBD-YOLOv8 模型，通过在网络中大规模引入深度可分离卷积，显著降低了模型参数量。在特征提取与融合阶段，该模型设计了 RELAN 模块替换默认的 C2f 模块以优化梯度路径，并引入 BiFPN 结构加强多尺度信息（特别是 P2 层小目标特征）的深度融合。在损失函数层面，作者基于 ShapeIoU^[16]创新性地提出了 DN-ShapeIoU 损失，能够根据样本质量动态调整梯度增益，进一步优化了边界框回归的精度。

总结而言，基于 YOLO 系列的反无人机目标检测方法凭借模型规模小、运行效率高的天然优势，极大地满足了无高性能算力支持的实战场景需求。当前 YOLO 模型在反无人机领域的改进主线清晰地聚焦于“微小目标特征保留”与“面向边缘计算的轻量化”，这使其成为目前低空防护体系中应用最广泛、落地最成熟的视觉算法分支。

2.3 Transformer-based 方法

Transformer^[17]最初应用于自然语言处理，其核心的自注意力机制随后在计算机视觉领域 (Vision Transformer) 引发了重大变革^[18]。在视觉任务中，Transformer 将图像划分为序列元素，通过计算每个元素与其他所有元素的关联程度，生成包含全局上下文信息的表示。这一机制打破了传统卷积神经网络 (CNN) 局部感受野的限制，能够对图像全局信息进行有效建模。例如，在包含无人机的图像中，自注意力机制可让模型精准捕捉无人机各个部件以及周围环境元素之间的关系。

在小目标检测 (SOD) 领域，随着 DETR^[19]的提出，基于 Transformer 的端到端目标检测方法快速发展。研究早期，为缓解计算压力，DETR 采用了 CNN 进行特征提取后接入 Transformer 编码器的混合

系统。但随着研究深入，从YOLOS^[20]开始，纯Transformer架构逐渐占据主导，CNN架构被ViT及其变体(如Swin Transformer^[21])全面取代。然而，小目标检测需要依赖高分辨率特征图，这导致Transformer的计算复杂度急剧上升。为此，ViDT^[22]作为第一个专为小目标检测优化的模型，创新性地采用无编码器(Encoder-free)架构，仅利用一个Transformer解码器接收Swin Transformer主干输出的多尺度特征，在保证检测性能的同时大幅减小了计算开销。

在反无人机实战场景中，目标常面临严重失真与密集遮挡。传统CNN的下采样过程极易加重噪声并导致目标信息丢失，而Transformer凭借卓越的上下文捕捉能力能显著缓解这一问题。以现代反无人机检测模型C2FDrone^[23]为例，该方法基于DAB DETR架构进行了创新设计。模型首先采用Swin Transformer作为骨干网络，利用补丁合并技术生成既保留细节又富含上下文的多尺度特征，随后通过特征金字塔网络(FPN)进行融合。更为关键的是，针对实际视频帧中的噪声问题，文章创新性地提出了目标增强网络(OEN)。该网络以Swin Transformer末端的特征作为输入，通过抑制背景噪声并增强前景细节，生成目标增强特征图；进而应用阈值生成“粗检测”掩模，以此初始化DAB DETR解码器，大幅缩小了目标的搜索范围并提升了最终的定位性能。

总结而言，基于Transformer的视觉模型(如C2FDrone由粗到细的检测策略)充分利用了自注意力机制的全局感受能力，从根本上改善了传统CNN因池化操作导致小目标细节丢失的缺陷。在树木、云朵等复杂环境干扰下，Transformer能够更精准地捕捉微小无人机的关键特征。可以预见，在未来的低空反无人机视觉感知系统中，Transformer架构将发挥愈发核心的作用。

3 其他传感器方法

除视觉探测技术外，雷达、声波和射频等非视觉传感器在反无人机系统中同样扮演着至关重要的角色，尤其在应对非视距、复杂遮挡、恶劣天气或夜间环境时展现出不可替代的优势。如图2所示的典型非视觉传感器反无人机架构，当前基于不同传感器的探测方法均已形成了一套标准化的技术链路，主要包含“物理层”、“信号预处理与映射”以及“模型层”三个核心阶段。下文将分别针对雷达、声波和射频这三种主流传感探测技术的发展现状与核心算法进行详细阐述。

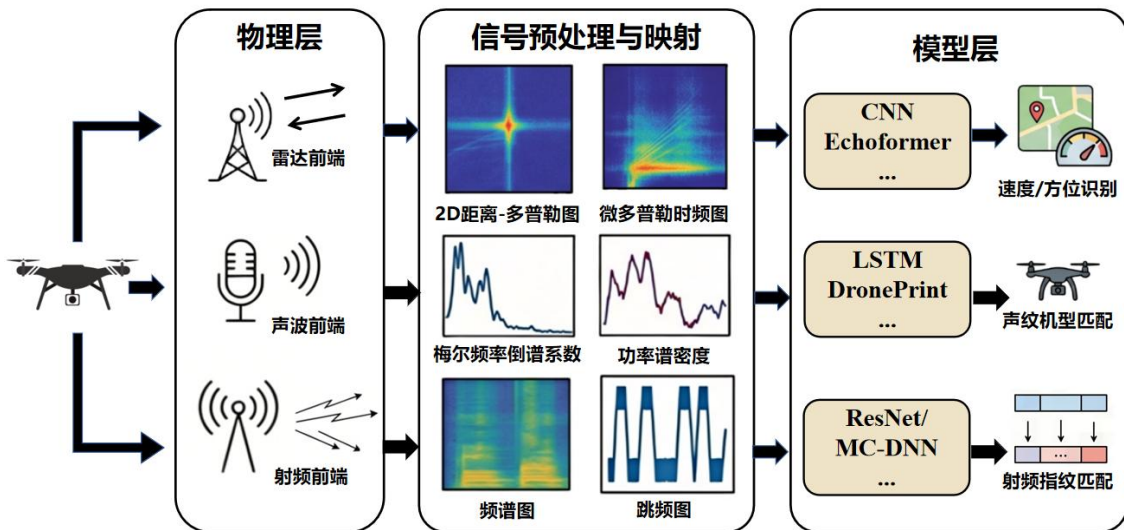


图2 典型非视觉反无人机探测与识别框架

Fig.2 Detection and Recognition Framework of Typical Non-visual Anti-UAV Systems

3.1 雷达检测

雷达传感器凭借全天候探测能力与微多普勒特征解析优势,已经成为反无人机系统的核心感知单元。近年来,基于深度学习的雷达信号处理方法通过端到端特征学习与时空建模能力突破,显著提升了系统性能。

针对小型无人机由于其尺寸微小、结构多样且飞行状态复杂,导致传统特征工程方法难以实现鲁棒检测与精确分类的挑战,Kumawat等^[24]提出了一种轻量化深度卷积神经网络架构DIAT-RadSATNet。该模型直接以彩色微多普勒时频图为输入,通过构建包含三条并行路径的特征重构机制,在有效压缩特征通道冗余的同时,精准提取了目标的局部微多普勒特征,从而在保持网络轻量化的前提下,实现了高效、高精度的微小型无人机分类识别。

此外,为了进一步提升微多普勒特征与多尺度速度特征的提取能力,Yang等^[25]提出了一种专为雷达原始回波信号设计的Transformer架构Echoformer。该模型突破了传统深度学习在雷达检测中对相干积分或时频图像映射的依赖,直接以脉冲-距离二维复数数据为输入构建端到端网络。该方法在保留完整相位信息的前提下,充分利用Transformer的全局特征提取优势,显著增强了对低空弱小目标及低速目标的探测与区分能力。

3.2 声波检测

声学检测是一种典型的被动式探测技术,主要依赖无人机旋翼气动噪声或电机运转产生的独特声频特征(通常集中在40 kHz以下)。通过高灵敏度麦克风阵列捕获这些信号并结合智能算法分析,可实现目标的非视觉检测与定位。该方法在光照受限的复杂环境中优势显著,其核心在于声学特征的提取,常用指标包括功率谱密度、梅尔频率倒谱系数及线性预测倒谱系数等^[26]。

在特征分类阶段,传统的机器学习算法(如SVM、KNN)因高度依赖手工特征质量而遭遇性能瓶颈。随着深度学习的引入,声学检测精度得到了显著跃升。例如,Kolamunna等^[27]提出了一种基于长短期记忆网络(LSTM)的开放集检测模型DronePrint。该模型巧妙利用振幅缩放与频率扭曲等数据增强策略,有效缓解了多普勒效应的干扰,并在包含20种无人机声学指纹的数据集上实现了高达95.7%的分类准确率。

尽管声学检测具备部署成本低、全天候抗恶劣天气等优势,但其规模化应用仍受制于三大技术瓶颈:首先,有效探测距离通常局限于200米以内,且极易受城市交通等复杂背景噪声的干扰导致性能骤降;其次,新型静音无人机或降噪技术的应用使其探测难度陡增;最后,缺乏大规模、标准化的声纹特征数据库严重限制了现有算法的泛化能力。未来,该领域的研究重点将转向复杂环境下的鲁棒性提升,如引入盲源分离技术抑制背景噪声干扰,以及借助迁移学习增强模型对未知机型的泛化探测能力。

3.3 射频检测

基于射频(RF)的探测方法主要通过监测无人机与遥控端之间的通信链路实现目标预警与识别。由于大多数无人机依赖2.4 GHz至5 GHz频段(如Wi-Fi、蓝牙或专用协议)进行数据与指令传输,其信号的调制方式、脉冲重复频率及强度波动等特征模式均可作为关键检测依据。例如,通过频谱分析提取控制器与设备间约30 Hz的周期性心跳信号,即可构建出独特的射频指纹^[28]。

特征提取与分类算法是射频检测的核心。早期的机器学习方法高度依赖信号频谱(SFS)、小波能量熵(WEE)及功率谱熵(PSE)等统计特征^[29]。随着深度学习的发展,基于数据驱动的分类模型逐渐

成为主流。例如, Al-Sa'd 等^[30]基于其构建的 DroneRF 开源数据集, 验证了深度残差网络在射频分类任务中相比传统 SVM 提升了约 15% 的准确率。为进一步在复杂电磁环境中提取稳健特征, Yang 等^[31]提出了一种融合特征工程生成器 (FEG) 与多通道深度网络 (MC-DNN) 的高效框架。该方法先通过 FEG 对原始信号进行频段分离与滤波降噪, 随后利用双通道网络独立学习高低频特征以抑制频间干扰。实验表明, 该方法在多机型、多飞行模式的数据集上实现了高达 98.4% 的分类准确率, 显著优于传统单一网络模型。

射频检测具备探测距离远、支持非视距预警且不受气象条件限制等显著优势。然而, 其在实战应用中也面临严峻的局限性: 首先, 该技术对复杂电磁环境极其敏感, 城市密集的 Wi-Fi 网络易导致频谱重叠, 大功率基站也极易掩盖微弱的目标信号; 其次, 对于开启无线电静默或采用“离线自主导航”模式的无人机, 射频检测将彻底失效; 此外, 高级加密通信技术与动态跳频协议的普及, 也为特征提取与信号解译带来了巨大阻碍。因此, 未来研究需重点结合自适应滤波与深度学习网络, 提升强干扰背景下的弱信号分离能力, 并积极探索加密信号的特征逆向解析方法。

4 多模态检测方法

多模态反无人机检测算法通过融合雷达、光电、无线电、声学等多种传感器数据, 有效突破了单一传感器的性能瓶颈。在复杂电磁与气象环境下, 多模态融合能够显著提高系统检测的准确性、可靠性和鲁棒性^[32], 是当前反无人机领域最具潜力的研究方向。近年来, 随着多模态反无人机数据集的相继发布, 基于异构数据融合的深度学习算法迎来了快速发展, 其中尤以视觉可见光与红外/多光谱的融合最为典型。如图 3 所示, 该类模型通过引入空间配准与时间同步机制解决异构数据异步难题, 并利用交叉注意力实现跨模态特征的深度耦合; 代表性算法如 RTM-UAVDet 通过动态卷积规避了模态未对齐的限制, 而 MSF-Net 则通过双流结构实现了多光谱信息的高效协同, 显著增强了复杂工况下的探测能力。

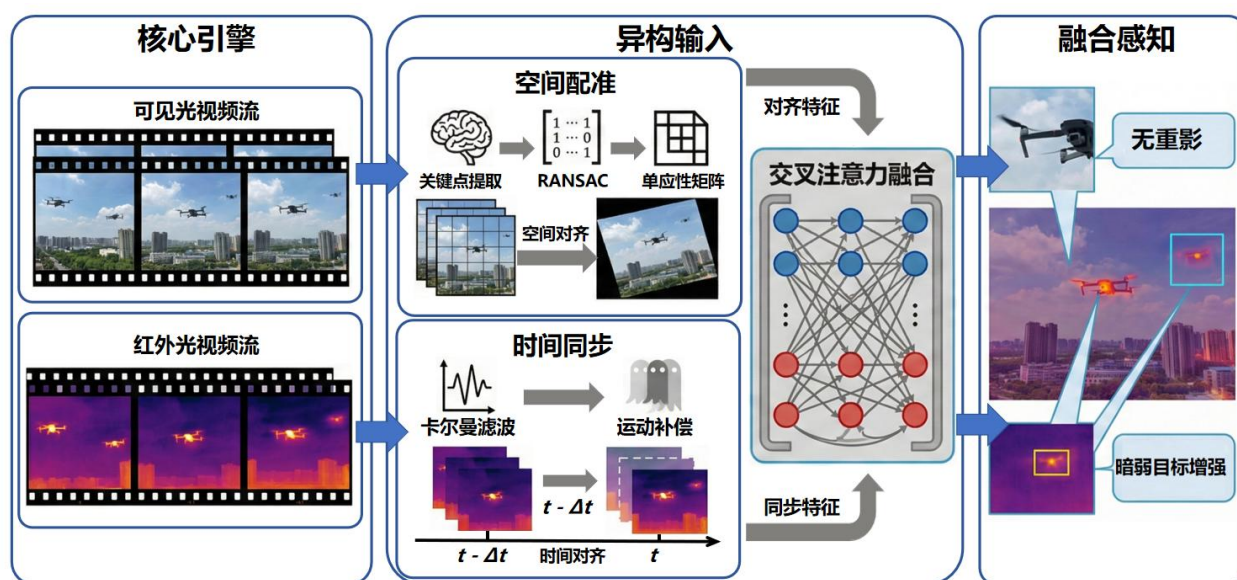


图 3 可见光-红外双模态时空对齐与交叉融合网络架构图

Fig.3 Network architecture of visible-infrared dual-modal spatiotemporal alignment and cross-fusion

基于可见光和热红外双模态的联合探测是实现全天候监控的关键。然而在实际应用中, 实时处理双

模态数据并解决异构传感器空间未对齐的问题是一项重大挑战。为此, Wang 等(2024)提出了 RTM-UAVDet^[33], 这是一种新型的实时双模态检测方法。有别于依赖复杂数据融合对齐的传统思路, RTM-UAVDet 创新性地提出了一种多模态动态卷积主干网络。该网络能够基于 RGB 和 TIR 特征的显著差异动态调整卷积核, 利用参数可变的架构自适应多种成像模态, 从而巧妙规避了未对齐数据集带来的限制。此外, 作者还专门设计了多模态预训练方法, 有效缓解了红外样本匮乏导致的过拟合问题。在 Anti-UAV^[34]数据集上的评估表明, 该模型在增加极少量参数的前提下, 检测精度大幅超越了基线模型, 完美兼顾了复杂环境下的高鲁棒性与实时性要求。

另一方面, 与传统的 RGB 图像相比, 包含红边 (REG) 和近红外 (NIR) 的多光谱图像能够提供更丰富的空间和光谱维度的信息, 在观察目标细精微特征方面具有天然优势。为此, 基于多光谱传感技术的 MSF-Net^[35]被提出。该模型构建了一种灵活的双流结构: 主干分支用于处理标准的三通道 RGB 图像, 辅助分支则专门处理 REG 和/或 NIR 光谱通道。为了实现跨模态的高效协同, MSF-Net 内部设计了专属的多模态融合模块, 并引入联合注意力机制对微小目标特征进行了深度优化。广泛的消融实验表明, 在结合了深层网络架构以及 RGB+REG+NIR 的全光谱配置下, 该模型在检测召回率和平均精度上均展现出远超单一可见光检测器的优异性能, 充分证实了多光谱信息融合在低空隐蔽目标探测中的巨大潜力。

5 结论

5.1 工作总结

通过本文上述研究可以得出, 反无人机技术在多模态感知、数据集构建以及核心算法研发等方面取得了显著进展, 为复杂环境下的反无人机系统提供了有力支持, 但目前反无人机技术仍面临着诸多挑战, 下文将对此进行总结。

1) 核心算法发展: 当前, 反无人机领域的核心算法蓬勃发展, 基于多模态感知的算法设计成为研究热点。基于图神经网络的模型 (如 Radar-PointGNN), 可以有效解析雷达点云的空间拓扑关系, 在目标识别任务中展现优越性能; 多模态融合算法 (如 RSDet) 采用跨域特征对齐策略, 在复杂光照条件下提高定位稳定性。

2) 数据集建设: 反无人机数据集已从单一的视觉模态逐步扩展到多模态融合。例如 MMAUD 数据集融合了立体视觉、激光雷达、毫米波雷达以及音频阵列, 覆盖了无人机检测、分类以及三维轨迹估计等任务。然而, 现有数据集在场景多样性、目标类型的覆盖以及极端环境下的数据采集方面仍存在局限性。例如, MMAUD 存在缺乏红外图像, 低光照下的检测能力受限、复杂环境覆盖度不高等问题。

3) 关键挑战: 尽管当前反无人机技术取得了显著进展, 但仍面临一些挑战。①感知鲁棒性不足, 复杂背景、动态行为复杂性及电磁干扰导致虚警率高、遮挡问题频发; ②多源数据协同难题, 多传感器融合延迟、时空校准精度不足限制了全天候场景下的实时响应能力; ③模型泛化能力弱, 现有数据集对隐蔽机型、极端天气 (如雨雪、浓雾) 覆盖不足, 深度学习模型在未知场景适应性上存在瓶颈。

5.2 研究展望

为改进现有反无人机技术的不足, 结合低空应用场景的特殊性和技术发展趋势, 未来研究还需要关注以下几个关键方向。

1) 视觉大模型与跨模态学习: 随着深度学习技术的不断发展, 大模型在视觉算法中的应用越来越

广泛，基于视觉的方法已成为研究热点。基于 Siamese 的方法、基于 Transformer 的方法和基于 YOLO 的方法这三类方法正受到越来越多的重视。未来反无人机领域可以基于 Swin Transformer 等视觉大模型架构构建多模态预训练框架，通过自注意力机制增强小目标特征提取能力，结合对比学习优化跨模态特征对齐，提升低照度环境下的目标判别精度，进一步强化反无人机系统在低空复杂光照条件下的目标判别能力。

2) 多智能体空地协同：由于无人机技术的迅速发展，特别是无人机集群带来的新型挑战，单个反无人机系统往往不足以应对这种复杂多目标的威胁。因此，开发能够协同作战的多智能体系统变得至关重要。未来反无人机技术的研究可以建立无人机-地面移动平台（UGV）协同感知网络，通过引入多智能体强化学习（MARL）等方法，基于博弈论模型优化群体决策效率，构建自适应环境的三维态势感知体系，从而提升系统在复杂场景下的实时响应能力。

3) 多源传感信息融合增强：多传感器协同是反无人机技术发展的重要方向。面向低空防御的多传感协同系统需要深化跨模态数据融合机制，将多维度感知数据进行整合，如可见光、红外热成像、射频定位（RF）、激光雷达（LIDAR）等，实现对数据的高感知。未来，反无人机技术研究者可以建立多源数据冲突消解机制，开发一体化决策框架，使之具备在线校准能力。

4) 自主学习与对抗博弈：无人机技术的智能化对反无人机探测和跟踪算法提出了更高的要求。在未来反无人机技术的发展中，自主学习与对抗博弈将逐渐成为焦点。为了有效应对挑战，反无人机探测和跟踪算法需要具备自主学习能力，并能够实时做出决策，以适应和对抗无人机不断进化的智能行为。

参考文献：

- [1] 王栋, 赵洁, 刘洋, 等. 多模态反无人机检测系统与技术[J]. 中国科学基金, 2026, 40(1): 73-84.
- [2] RAI P K, IDSØE H, YAKKATI R R, et al. Localization and Activity Classification of Unmanned Aerial Vehicle Using mmWave FMCW Radars[J]. IEEE Sensors Journal, 2021, 21(14): 16043-16053.
- [3] FANG H, LIAO Z, WANG X, et al. Differentiated Attention Guided Network Over Hierarchical and Aggregated Features for Intelligent UAV Surveillance[J]. IEEE Transactions on Industrial Informatics, 2023, 19(9): 9909-9920.
- [4] CAO B, YAO H, ZHU P, et al. Visible and Clear: Finding Tiny Objects in Difference Map[C]//Proceedings of the European Conference on Computer Vision. Cham: Springer Nature Switzerland, 2024: 1-18.
- [5] LIU Q, LIU R, ZHENG B, et al. Infrared Small Target Detection With Scale and Location Sensitivity[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle, WA, USA: IEEE, 2024: 17490-17499.
- [6] REDMON J, DIVVALA S, GIRSHICK R, et al. You Only Look Once: Unified, Real-Time Object Detection[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, NV, USA: IEEE, 2016: 779-788.
- [7] REDMON J, FARHADI A. YOLOv3: An Incremental Improvement[EB/OL]. (2018-04-09). <https://arxiv.org/abs/1804.02767>.
- [8] WANG C Y, BOCHKOVSKIY A, LIAO H Y M. YOLOv7: Trainable Bag-of-Freebies Sets New State-of-the-Art for Real-Time Object Detectors[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Vancouver, BC, Canada: IEEE, 2023: 7464-7475.
- [9] JOCHER G, CHAURASIA A, QIU J. Ultralytics/Ultralytics: NEW-YOLOv8[EB/OL]. (2023-01-10). <https://github.com/ultralytics/ultralytics>.
- [10] FANG A, FENG S, LIANG B, et al. Real-Time Detection of Unauthorized Unmanned Aerial Vehicles Using SEB-YOLOv8s[J]. Sensors, 2024, 24(12): 3915.
- [11] SUNKARA R, LUO T. No More Strided Convolutions or Pooling: A New CNN Building Block for Low-Resolution Images and Small Objects[C]//Proceedings of the Joint European Conference on Machine Learning and Knowledge Discovery in

- Databases. Berlin: Springer, 2022: 443-459.
- [12] OUYANG D, HE S, ZHANG G, et al. Efficient Multi-Scale Attention Module With Cross-Spatial Learning[C]//Proceedings of the ICASSP 2023 - 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Rhodes Island, Greece: IEEE, 2023: 1-5.
- [13] ZHU L, WANG X, KE Z, et al. BiFormer: Vision Transformer With Bi-Level Routing Attention[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Vancouver, BC, Canada: IEEE, 2023: 10323-10333.
- [14] 朱志亮, 黄欣荣, 刘怡, 等. ES-YOLO: 基于细节特征增强与冗余特征抑制的小目标检测方法[J]. 华东交通大学学报, 2025, 42(6).
- [15] JIANG P, YANG X, WAN Y, et al. DRBD-YOLOv8: A Lightweight and Efficient Anti-UAV Detection Model[J]. Sensors, 2024, 24(22): 7148.
- [16] ZHANG H, ZHANG S. Shape-IoU: More Accurate Metric Considering Bounding Box Shape and Scale[EB/OL]. (2023-12-29). <https://arxiv.org/abs/2312.17663>.
- [17] VASWANI A, SHAZEER N, PARMAR N, et al. Attention Is All You Need[C]//Proceedings of the 31st International Conference on Neural Information Processing Systems (NIPS'17). Long Beach, CA, USA: Curran Associates Inc., 2017: 6000-6010.
- [18] DOSOVITSKIY A, BEYER L, KOLESNIKOV A, et al. An Image Is Worth 16x16 Words: Transformers for Image Recognition at Scale[C]//Proceedings of the 9th International Conference on Learning Representations (ICLR 2021). Virtual: OpenReview.net, 2021.
- [19] CARION N, MASSA F, SYNNAEVE G, et al. End-to-End Object Detection With Transformers[C]//Proceedings of the European Conference on Computer Vision. Online Berlin: Springer International Publishing, 2020: 213-229.
- [20] FANG Y, LIAO B, WANG X, et al. You Only Look at One Sequence: Rethinking Transformer in Vision Through Object Detection[C]//Proceedings of the International Conference on Neural Information Processing Systems. Vancouver, BC, Canada: Curran Associates Inc., 2021: 26183-26197.
- [21] LIU Z, LIN Y, CAO Y, et al. Swin Transformer: Hierarchical Vision Transformer Using Shifted Windows[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. Montreal, QC, Canada: IEEE, 2021: 10012-10022.
- [22] SONG H, SUN D, CHUN S, et al. ViDT: An Efficient and Effective Fully Transformer-Based Object Detector[C]//Proceedings of the 10th International Conference on Learning Representations (ICLR 2022). Virtual: OpenReview.net, 2022.
- [23] REBBAPRAGADA S V C, PANDA P, BALASUBRAMANIAN V N. C2FDrone: Coarse-to-Fine Drone-to-Drone Detection Using Vision Transformer Networks[C]//Proceedings of 2024 IEEE International Conference on Robotics and Automation (ICRA). Yokohama, Japan: IEEE, 2024: 6627-6633.
- [24] KUMAWAT H C, CHAKRABORTY M, RAJ A A B. DIAT-RadSATNet—A Novel Lightweight DCNN Architecture for Micro-Doppler-Based Small Unmanned Aerial Vehicle (SUAV) Targets' Detection and Classification[J]. IEEE Transactions on Instrumentation and Measurement, 2022, 71: 1-11.
- [25] YANG Y, YANG F, SUN L, et al. Echoformer: Transformer Architecture Based on Radar Echo Characteristics for UAV Detection[J]. IEEE Sensors Journal, 2023, 23(8): 8639-8653.
- [26] UDDIN Z, ALTAF M, BILAL M, et al. Amateur Drones Detection: A Machine Learning Approach Utilizing the Acoustic Signals in the Presence of Strong Interference[J]. Computer Communications, 2020, 154: 236-245.
- [27] KOLAMUNNA H, DAHANAYAKA T, LI J, et al. DronePrint: Acoustic Signatures for Open-Set Drone Detection and Identification With Online Data[J]. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, 2021, 5(1): 1-31.
- [28] XIAO Y, ZHANG X. Micro-UAV Detection and Identification Based on Radio Frequency Signature[C]//Proceedings of the 2019 6th International Conference on Systems and Informatics (ICSAI). Shanghai, China: IEEE, 2019: 1056-1062.
- [29] NIE W, HAN Z C, LI Y, et al. UAV Detection and Localization Based on Multi-Dimensional Signal Features[J]. IEEE Sensors Journal, 2021, 22(6): 5150-5162.

- [30] AL-SA'D M F, AL-ALI A, MOHAMED A, et al. RF-Based Drone Detection and Identification Using Deep Learning Approaches: An Initiative Towards a Large Open Source Drone Database[J]. *Future Generation Computer Systems*, 2019, 100: 86-97.
- [31] YANG S, LUO Y, MIAO W, et al. RF Signal-Based UAV Detection and Mode Classification: A Joint Feature Engineering Generator and Multi-Channel Deep Neural Network Approach[J]. *Entropy*, 2021, 23(12): 1678.
- [32] DENG T C, ZHOU Y, WU W H, et al. Multi-Modal UAV Detection, Classification and Tracking Algorithm — Technical Report for CVPR 2024 UG2 Challenge[C]//*Proceedings of 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*. Seattle, WA, USA: IEEE, 2024: 1-12.
- [33] WANG G, JIANG Q, JIN X, et al. RTM-UAVDet: A Real-Time Multimodal UAV Detector[J]. *IEEE Transactions on Aerospace and Electronic Systems*, 2025, 61(1): 473-489.
- [34] JIANG N, WANG K, PENG X, et al. Anti-UAV: A Large Multi-Modal Benchmark for UAV Tracking[EB/OL]. (2025-04-14). <https://arxiv.org/abs/2101.08466v3>.
- [35] XU F, LIN Z, CHEN C, et al. Anti-Drone Detection in Aerial Multispectral Images[C]//*Proceedings of IGARSS 2024 - 2024 IEEE International Geoscience and Remote Sensing Symposium*. Athens, Greece: IEEE, 2024: 9388-9391.



通信作者: 朱鹏飞 (1986—), 男, 教授, 博士生导师, 国家优秀青年基金获得者和天津市杰青获得者, 获吴文俊人工智能科技进步一等奖、黑龙江省自然科学一等奖等。研究方向为低空智能感知、低空具身智能。