

## 基于数字视网膜的低空主动感知网及其关键技术

陈轲<sup>①</sup> 吴少聪<sup>①</sup> 吴海<sup>①</sup> 张舒航<sup>②</sup> 吴哲<sup>①</sup> 邹龙坤<sup>①</sup> 刘康俊<sup>①</sup>  
曹桂平<sup>①</sup> 肖麟慧<sup>①</sup> 李鑫<sup>①</sup> 王耀威<sup>\*①③</sup>

<sup>①</sup>(鹏城实验室感知智能研究所 深圳 518000)

<sup>②</sup>(北京大学电子学院 北京 100871)

<sup>③</sup>(哈尔滨工业大学(深圳)计算机科学与技术学院 深圳 518055)

**摘要:** 低空智能感知技术旨在将低空目标与环境的物理空间转化为可计算的数字空间, 是支撑低空经济活动安全有序开展的基础。该文系统分析了当前低空场景下大规模视觉感知、目标感知以及环境感知等技术的发展现状与面临的挑战。针对现有挑战, 该文提出了基于数字视网膜端边云协同架构的低空主动感知网, 并从整体网络架构、云侧基础模型底座、端侧目标与环境感知技术等维度, 详细阐述了其核心机制与关键技术。最后, 通过初步实验, 验证了该文所提感知网络在空地带宽受限条件下实现高效协同感知计算的有效性。

**关键词:** 低空感知网; 端边云协同; 数字视网膜; 低空目标感知; 低空环境感知

**中图分类号:** TP391

**文献标识码:** A

**文章编号:** 2095-283X(2026)x-0001-28

**DOI:** 10.12000/JR26022

**CSTR:** 32380.14.JR26022

**引用格式:** 陈轲, 吴少聪, 吴海, 等. 基于数字视网膜的低空主动感知网及其关键技术[J]. 雷达学报(中英文), 待出版. doi: 10.12000/JR26022.

**Reference format:** CHEN Ke, WU Shaocong, WU Hai, *et al.* Architecture and key technologies of a low-altitude active perception network based on the digital retina[J]. *Journal of Radars*, in press. doi: 10.12000/JR26022.

## Architecture and Key Technologies of a Low-altitude Active Perception Network Based on the Digital Retina

CHEN Ke<sup>①</sup> WU Shaocong<sup>①</sup> WU Hai<sup>①</sup> ZHANG Shuhang<sup>②</sup> WU Zhe<sup>①</sup>  
ZOU Longkun<sup>①</sup> LIU Kangjun<sup>①</sup> CAO Guiping<sup>①</sup> XIAO Linhui<sup>①</sup>  
LI Xin<sup>①</sup> WANG Yaowei<sup>\*①③</sup>

<sup>①</sup>(Institute of Perceptual Intelligence, Pengcheng Laboratory, Shenzhen 518000, China)

<sup>②</sup>(School of Electronics, Peking University, Beijing 100871, China)

<sup>③</sup>(School of Computer Science and Technology, Harbin Institute of Technology (Shenzhen), Shenzhen 518055, China)

**Abstract:** Low-altitude perception technology aims to transform the physical space of low-altitude targets and environments into a computable digital space, providing a foundation for the safe and organized development of low-altitude economic activities. This paper systematically examines the current progress and challenges of technologies such as large-scale visual perception, object detection, and environmental sensing in low-altitude scenarios. To address these challenges, we introduce a low-altitude active perception network based on the digital retina featuring a collaborative architecture that integrates end, edge, and cloud computing. The key

收稿日期: 2026-01-16; 改回日期: 2026-04-18; 网络出版: 2026-xx-xx

\*通信作者: 王耀威 wangyaowei@hit.edu.cn \*Corresponding Author: WANG Yaowei, wangyaowei@hit.edu.cn

基金项目: 鹏城实验室重大攻关项目(PCL2025A14, PCL2025A02)

Foundation Items: The project of Pengcheng Laboratory (PCL2025A14, PCL2025A02)

责任编辑: 陈小龙 Corresponding Editor: CHEN Xiaolong

©The Author(s) 2026. This is an open access article under the CC-BY 4.0 License

(<https://creativecommons.org/licenses/by/4.0/>)

mechanisms and methods are outlined across various aspects, including the overall network structure, cloud-based foundation models, and end-based object and environmental perception technologies. Early applications and experimental results confirm the effectiveness of the proposed network in enabling efficient joint perception under bandwidth limitations.

**Key words:** Low-altitude perception network; Edge-cloud collaboration; Digital retina; Perception of low-altitude targets; Perception of low-altitude environment

## 1 引言

低空经济作为新质生产力的典型代表,正步入高速发展期,随着《无人驾驶航空器飞行管理暂行条例》等法规的实施以及《通用航空装备创新应用实施方案(2024—2030年)》的印发,低空经济的应用场景从传统的物流配送、应急救援拓展至城市空中交通(Urban Air Mobility, UAM)及低空文旅等多元领域。这一趋势标志着低空空域正从自然资源加速转化为具备巨大经济价值的生产要素。然而,低空空域的高密度、高动态及异构性特征,对空域的精细化管理与安全运行提出了严峻挑战<sup>[1]</sup>。

低空智能感知技术旨在将低空目标与物理环境转化为可计算的数字空间,是保障低空经济活动安全有序开展的基础。具体而言,该过程通过多模态传感探测与特征提取,将空域中连续的物理要素(如飞行器轨迹、短临气象演变、电磁波谱分布等)映射为结构化数字信息,以支持计算机的量化分析与预测<sup>[2]</sup>。鉴于低空感知涉及面极广,本文将重点探讨在受限通信带宽与边缘算力下,实现“低慢小”目标精准探测与低空环境(气象与电磁频谱)精细态势重构的感知技术,为航路规划与空域资源利用提供底层支撑。

面向低空高动态与高密度场景,现有感知技术仍面临三大挑战。首先,相比传统民航的中高空飞行空间,低空场景下环境地形、天时、天候等要素复杂,低空目标雷达散射截面积小、视觉表观信息少,且悬停、慢飞等行为的普勒效应弱,导致单一传感器(如激光/微波雷达、光电等)难以发现低空目标,多传感器智能联动协同效率低、虚警率高<sup>[3]</sup>。其次,低空环境复杂快变<sup>[4]</sup>,受限于观测设备与手段,保障飞行的环境信息存在时空稀疏性,难以为低空活动提供连续、稠密的精细环境信息。最后,低空适飞天候下全时感知覆盖需要多传感器组网,但当前组网普遍存在“脑眼分离”的问题,即前端传感器仅负责数据采集与压缩,主要依赖云侧平台进行理解分析,导致海量数据在传输过程中占据大量带宽,造成严重的通信拥塞与处理延迟,难以满足规模化高密度低空飞行保障的实时感知需求。

综上,当前的核心科学问题在于:高维、高动

态的低空物理空间感知需求,与端边云网络体系下受限的通信带宽与算力负荷之间存在深刻的矛盾。如何突破传统“先传输,后计算”的信道容量瓶颈,在端边云协同架构下实现多模态异构感知信息的最优表征与自适应资源调度,是实现低空规模化主动感知的关键所在。针对此痛点,本文提出了基于数字视网膜端边云协同架构的低空主动感知网(如图1所示)。“数字视网膜(Digital Retina)”概念受人类视觉系统机制的启发:生物视网膜并非将所有感光信号原始全量地上传大脑,而是同时向大脑视觉中枢传递用于环境重构的通用信号,以及经过初步提炼的语义特征信号<sup>[5]</sup>。将这一概念映射到机器感知与计算系统中,数字视网膜架构打破了传统感知设备仅作为“数据采集器”的局限,创新性地提出了“三流协同”(即面向人类视觉与存储的压缩数据流、面向机器视觉的紧凑特征流,以及支持模型在线演进的模型流)的全新计算范式<sup>[6]</sup>。基于这一新型的端边云协同智算架构,本文构建了一种具备全局统一时空基准、支持特征流、数据流与模型流协同工作的低空主动感知网络,该网络通过结合端侧的目标环境感知技术与云侧的基础模型底座,在大幅降低传输带宽占用的同时,实现云侧基础模型向端侧的知识蒸馏与增量更新,持续进化端侧感知性能,形成可进化的感知-传输-计算一体化闭环。

## 2 低空智能感知技术发展现状

当前,学术界与产业界围绕大规模感知计算、低空场景下目标和环境感知方面已开展了广泛研究。本节将首先从整体架构层面,梳理感知网从“云侧集中式”向“端侧分布式”演进的脉络,剖析现有架构在低空高动态场景下的局限性,在此基础上引出“数字视网膜”,并对其特性进行阐述;随后,聚焦于低空场景下的低空目标、环境感知以及低空物联网3个维度,阐述国内外的研究进展与技术瓶颈。

### 2.1 大规模感知计算范式演进及数字视网膜架构

大规模感知计算关键技术软硬件系统是支撑低空经济和低空安防等领域的关键基础设施,其核

心在于通过大范围部署的视觉、光电、雷达、气象等多源传感器组网<sup>[7,8]</sup>，实现对复杂场景的实时监控、智能分析和快速响应。相较于传统地面视频监控，低空场景需要针对目标快速机动、空域开放动态变化和感知资源受限等特点，支持对感知视场、感知频率和计算资源的主动调整<sup>[9]</sup>。目前，在学术界和工业界(如华为、海康威视以及Axis Communications、Bosch等)构建的感知网系统架构普遍采用感知端(“眼”，即摄像头等端侧感知设备)与智能分析处理的计算中心(“脑”，即云平台)分离的系统架构，主要体现为两种范式：集中式云侧处理和端侧边缘智能处理(如图2所示)。面对低空场景

“高动态、大带宽、算力受限”的严峻挑战，传统感知计算架构的演进逐渐暴露出局限性，并最终催生了以“数字视网膜”为代表的新型智算范式。

(1) 云侧处理主导的感知计算架构

集中式云侧感知计算架构通过在前端部署大量传感器完成感知数据采集，并在后端依托集中式云服务器实现数据的统一存储与分析。这一模式的优势在于云侧具备充足算力与存储，可部署复杂模型并进行统一管理，支持跨区域的感知计算需求<sup>[10,11]</sup>，适合低空场景下的全局分析。例如，ZHOU等人<sup>[12]</sup>提出的智能校园监控系统和AKLAMATI等人<sup>[13]</sup>基于云的视频监控服务方案，均强化了云侧集中化管

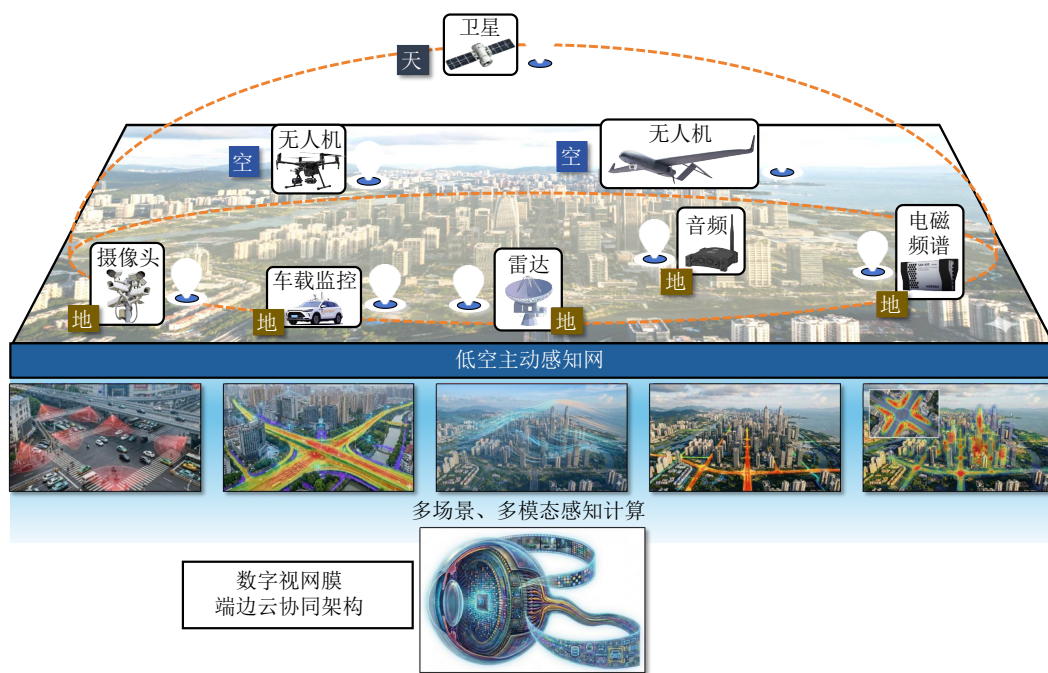


图 1 基于数字视网膜的低空主动感知网

Fig. 1 The digital retina-based active perception network for low-altitude airspace

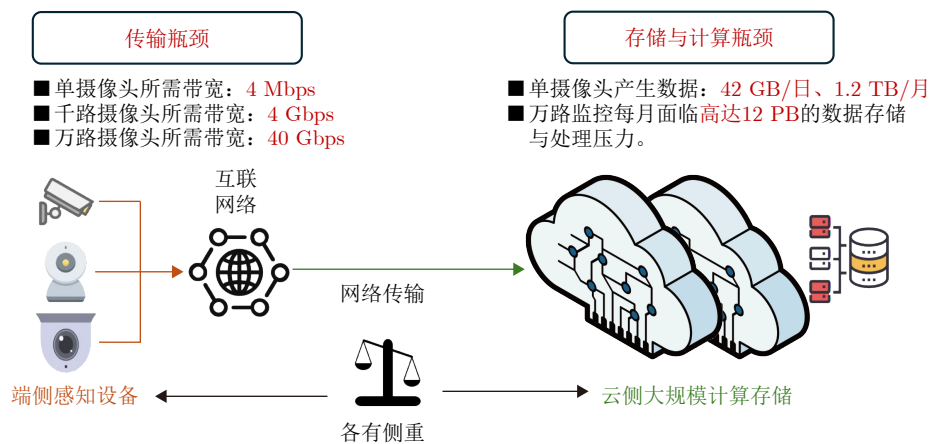


图 2 云侧感知和端侧感知两类范式及其面临的瓶颈

Fig. 2 The cloud-based and edge-based perception paradigms and their respective bottlenecks

理和可扩展性的能力,但依然沿用了“端侧只负责编码、云侧负责解码与分析”的基本模式。然而,云侧集中式处理在规模化部署下面临带宽与时延瓶颈:单路高清摄像头可能占用数Mbps至数百Mbps的带宽,数千路并发上行迅速将带宽需求推高到Gbps级,一旦带宽受限就需要通过降帧率、降分辨率、预过滤等手段“以牺牲精度换传输”<sup>[10]</sup>。此外,原始数据上云还带来隐私泄露与合规风险,使得纯云侧模式在部分高隐私需求的场景中难以落地。传统架构多以“数据中心+云存储”为核心,强调存储可靠性和访问性能<sup>[14]</sup>,而对于前端数据理解分析与系统级协同的关注较少<sup>[15,16]</sup>。

### (2) 端侧智能驱动的感知计算架构

为缓解集中式处理带来的巨大带宽和时延压力,随着高效AI芯片(如NPU,GPU,FPGA)的发展与端侧计算能力的逐步提升,将计算能力下沉到前端设备,形成依托分布式智能前端的感知计算架构,时敏的智能分析(如移动侦测、人脸抓拍、车牌识别等)直接在摄像头端侧完成。设备通常只将结构化或半结构化的特征数据(元数据)上传至云侧,而非原始数据流<sup>[17]</sup>。主流厂商均在积极布局端侧智能。例如,海康威视的“深眸”系列和Axis Communications搭载内置深度学习处理单元的ARTPEC系列芯片<sup>[18]</sup>,使得摄像头能够在端侧独立运行视频数据分析任务。这种智能前移的趋势正在改变大规模感知计算系统的算力分布模型。

目前,端侧设备已具备相互协同执行感知计算的能力,王秉路等人<sup>[19]</sup>研究发现多节点之间可以通过特征融合与协同处理提升系统整体感知性能与鲁棒性,Jain等人<sup>[20]</sup>发现多摄像头视频流在空间与时间上存在显著内容相关性,可用于缩小分布式分析搜索空间,实验显示协同利用多摄像头相关性,可节省约74%的运算资源并提升系统可扩展性。端侧智能驱动的感知计算通过本地运行轻量模型与控制元数据上传,可显著缓解传输压力并提升响应速度,同时具备良好的隐私保护性与局部协同潜力,为低空场景的感知数据传输与计算分析提供基础。然而,这一架构在低空场景下,面临端侧设备算力有限、缺少持续更新与全局协调机制的困境,易出现模型性能退化,难以支撑感知任务的长期稳定运行的问题。

### (3) 端云协同的感知计算架构

为同时兼顾实时性与精度,端云协同的感知计算逐渐成为新的主流方向:端侧或边缘侧承担低时延推理与快速响应,云侧承担复杂推理、全局检索以及模型的训练更新,并通过持续学习缓解模型性

能退化。例如,Neff等人<sup>[21]</sup>提出了一种基于端侧预过滤的视觉感知框架,通过在感知端运行轻量模型过滤无信息或低信息量的视频片段,将传输带宽需求减少约80%,从而把有选择地将有限的带宽留给富含信息的特定视频内容。然而,现有的端云协同多采用静态协同策略,在低空场景下,由于网络质量、目标状态、观测条件和任务优先级均可能快速变化,此类静态机制容易导致通信资源浪费或关键感知信息丢失等问题。因此,自适应任务与资源调度是端云协同感知计算这一新范式所面临的关键挑战:需要根据带宽、端侧实时算力、时延约束与感知计算任务难度与需求等信息,动态调整数据与计算任务在端、边或云上的传输与部署执行。

### (4) 应对低空感知瓶颈的数字视网膜端边云协同架构

针对现有端云协同中自适应任务与资源调度的挑战,以及低空高动态环境带来的通信带宽与端侧算力双重受限痛点,鹏城实验室高文院士团队提出了“数字视网膜”这一新型的端边云感知协同范式<sup>[5,6]</sup>,旨在借鉴生物视网膜感知世界的方式,使前端摄像头从被动采集设备演进为兼具视频编码、特征表达与模型更新能力的智能节点。其核心内涵是在统一时间与精确位置约束下,端侧感知设备能够同时输出面向存储与人工查看的压缩视频流,以及面向分析与检索的紧凑特征流,并进一步通过模型流支持端边云协同更新。在低空场景中,引入数字视网膜架构的必要性在于其从根本上打破了传统“脑眼分离”的局限。该架构通过特征流极大地缓解了低空海量数据引发的空地通信拥塞;通过模型流弥补了端侧设备算力不足且难以持续更新的短板,形成了“特征前置、数据按需、模型在线更新”的闭环。数字视网膜这一架构具有高性能、高效率、可伸缩、隐私保护和可软件定义等优势,为低空场景下实现任务驱动、按需传输和持续演化的大规模主动感知提供了不可或缺的底层逻辑与可行路径。

## 2.2 低空目标感知

低空目标感知旨在对无人机、通航航空器及飞鸟等低空目标进行检测、定位、识别与跟踪等。本节将深入探讨在低空复杂背景下,基于光电、雷达、射频信号、声学阵列等传感器数据的低空目标感知方法研究现状,并剖析各项技术在低空场景中面临的关键瓶颈。

### (1) 基于光电的低空目标感知

视觉数据如可见光三原色(Red-Green-Blue,

RGB)图像、热红外图像、视频等因其成本较低、光学成像对环境干扰小成为目前主要的低空感知手段<sup>[22,23]</sup>。近年来,针对低空目标尺寸小、对比度弱、远距离易退化等特点,文献<sup>[24,25]</sup>在通用检测框架基础上进行结构改进,引入小目标增强与多尺度特征融合,有效提升了模型在复杂天空背景和远距离场景中的检出性能。这类技术的优点是部署成本低,适用条件为光照良好或目标热对比度高的环境。为实现全时感知,可见光-热红外(Red-Green-Blue and Thermal infrared, RGB-T)融合逐渐成为研究热点。Anti-UAV数据集<sup>[26]</sup>基于可见光与热红外双模态,涵盖多种环境场景和多尺度无人机实例,为RGB-T目标检测、跟踪算法研究提供了统一的基准平台<sup>[27,28]</sup>。然而,单一视觉模态仍具有明显的局限性<sup>[29]</sup>。首先,RGB可见光与热红外均受制于成像光学系统,有效感知距离受限。其次,光学传感器对环境条件高度敏感,可见光成像易受到光照变化、阴影、反射、雨雪雾等因素影响,而红外容易受到背景热辐射、太阳加热效应或多物体同温背景的干扰,导致低质量成像,进而影响感知性能。同时,高分辨率视频流的实时解析对通信带宽与计算负载提出了严峻挑战,使得感知计算压力在端云之间的分配难以优化。

### (2) 基于雷达的低空目标感知

雷达具备探测距离远、对雨雾及低能见度环境具有天然穿透能力与良好鲁棒性、可直接获取目标径向速度等优势,长期以来雷达感知方法聚焦在基于脉内/脉间多普勒处理、恒虚警率检测、波束形成、高精度角度估计算法,以及运动目标指示与跟踪等技术<sup>[30,31]</sup>,此类传统技术路线在背景杂波较弱的高空或开阔区域具有极强的适用性。近年来,随着人工智能技术的发展,数据驱动方法逐步渗透至雷达感知研究前沿<sup>[32,33]</sup>。面向无人机与飞鸟等典型低慢小目标的探测、识别与跟踪已形成体系化研究,微动特征、多波段融合、时序运动建模与实测数据集构建等方向不断取得突破<sup>[34-37]</sup>。

然而,在低空场景中,受低空目标散射特性复杂、地物杂波强以及环境信噪比波动等因素影响,雷达回波中目标与背景难以有效区分。由于典型微型无人机的雷达散射截面(Radar Cross Section, RCS)通常仅为 $-20$  dBsm $\sim$  $-10$  dBsm,导致其在强杂波背景下的信杂比(Signal-to-Clutter Ratio, SCR)常低于 $0$  dB<sup>[38]</sup>。在这种条件下,传统恒虚警检测器的探测效能会严重退化,难以从强干扰中有效提取目标信号,使得虚警率控制与误报抑制成为雷达低空应用实践中的关键挑战。此外,雷达高采

样率原始回波的处理往往依赖后端高性能计算资源,难以在探测末端平衡实时性与算力开销;且相较光学传感器,单一雷达系统在复杂低空多径效应下的角度/距离测量精度与目标类型识别能力上仍存在差距。面向轻量化部署与端边协同的雷达信号处理与目标检测算法已经成为解决该问题的重要路径<sup>[39,40]</sup>,多波段雷达微动特征精细化提取、低慢小目标专用数据集构建与图网络特征融合分类等方法,为低空雷达感知的工程化落地提供了高效解决方案。

### (3) 基于射频信号的目标感知

该类技术通过侦收与分析目标设备在通信过程中辐射的时空电磁指纹,实现对特定目标的非合作式检测与识别。研究方法已从依赖人工特征提取(如信号频谱、小波能量熵与功率谱熵<sup>[41]</sup>)转向基于深度学习的智能识别,显著提升了识别性能<sup>[42]</sup>,但其也易受低空环境下多径效应与阴影衰落影响,在复杂城区环境下的定位精度通常仅在百米量级<sup>[43]</sup>,且由于缺少与其他传感器在物理层面的深度联动,难以实现高精度空间定位。

### (4) 基于声学阵列的目标感知

基于声学阵列的低空目标感知利用螺旋桨、电机等产生的独特声纹实现被动目标探测、定位与识别。在近场且信噪比尚可的场景中,传统的阵列信号处理路线优势明显。早期研究集中于阵列信号处理,Chen等人<sup>[44]</sup>提出平面声阵列的被动定位方法,为低空声学测向奠定基础;Tong等人<sup>[45]</sup>进一步利用单阵列并结合多普勒与时延模型估计三维运动轨迹。针对复杂环境中弱信号易被噪声淹没的问题,Song等人<sup>[46]</sup>构建多五元阵列系统并结合变分模态分解与小波阈值去噪以增强时延估计鲁棒性;Liu等人<sup>[47]</sup>在广义互相关框架下设计改进加权函数,并引入3次互相关策略,在低信噪比条件显著降低无人机声源定位误差。随着目标检测与识别需求提升,研究从传统特征工程转向机器学习与深度学习<sup>[48-51]</sup>。然而,音频模态本身存在传播衰减快、空间分辨率受限、对环境噪声高度敏感等固有不足,其感知范围通常局限于 $500$  m以内<sup>[52]</sup>,在复杂开放低空环境中,目标声纹易受风噪、交通噪声、工业背景声等非平稳干扰影响,难以在远距离、低信噪比条件下实现精确的目标感知。

如表1所示,当前低空目标感知领域仍面临严峻的挑战。首先,类似于前文所述的大规模感知计算架构,现有目标感知同样受限于“脑眼分离”的架构特征,目标感知计算的单方面集中于云侧或端侧,导致感知系统在处理海量感知数据流时,

表1 多种传感器数据的感知优势、挑战和现有研究贡献对比

Tab. 1 Comparison of perception advantages, challenges, and existing research contributions across multiple sensor modalities

传感器	核心优势	主要缺陷	现有研究贡献
雷达感知	距离远、全天候、径向速度测量	杂波干扰强、SCR极低、算力开销大	脉内/外处理、数据驱动的方法
光电感知	成本低、成像直观、目标识别力强	受环境/光照影响大、感知距离有限	小目标增强、RGB-T融合、Anti-UAV基准
射频感知	非合作式侦收、电磁指纹识别	定位精度低(百米级)、受多径影响	从人工特征转向深度学习智能识别
声学感知	被动探测、独特声纹识别	传播衰减快、距离近(<500 m)、易受噪	3D轨迹估计、加权函数改进、降噪增强

面临高带宽占用与高传输时延的矛盾，且缺乏一种根据环境复杂度与目标威胁等级进行动态调节的自适应端云协同机制。其次，现有目标感知方法研究主要围绕视音频监控数据开展，对于低空场景所特有的雷达、气象仪、频谱仪等多源异构感知设备探索深度不足，缺乏有效的物理层与特征层整合能力。这导致多模态数据之间难以实现时空精准对齐与置信度互补，成为制约低空跨场景泛化感知能力提升的核心痛点。

### 2.3 低空环境感知

低空飞航环境信息是保障低空飞行器开展经济活动的关键。本节聚焦气象与电磁两大关键维度：气象环境感知以捕捉低空复杂气象要素为核心，实现对气流波动、能见度、降水等关键参数的微观气象态势生成；电磁环境感知围绕频谱空间展开，涵盖频谱监测、干扰识别、信号溯源等核心能力，实时捕捉低空电磁信号分布与变化特征，规避电磁干扰对通信、导航系统的影响，保障低空活动的电磁环境安全。

#### (1) 气象环境感知

传统气象预报通常通过求解流体动力学控制方程实现，但由此带来了高昂的计算成本<sup>[53-55]</sup>，且预测周期通常限制在数小时内。近年来，人工智能被广泛应用于气象预报领域，但仍局限于中长尺度时空范围的预测，无法捕获长时和高分辨率的微观气象细节。其中，哈尔滨工业大学林连雷教授团队<sup>[56]</sup>创新性地构建出融合多元气象数据的风速预报网络(Multivariate meteorological data Fusion Wind Prediction Network, MFWPN)，实现了低空短时风场的高精度智能预报。上海人工智能实验室主导研发的FengWu-GHR<sup>[57]</sup>是全球首个实现10公里级分辨率(0.09°经纬网格)的AI气象大模型，将有效预报时长提升至11.25天。微软Aurora<sup>[58]</sup>气象模型基于3D Swin Transformer与Perceiver编解码器架构，实现0.1°经纬网格(约11公里)分辨率的气象预测，并能精准捕捉复杂大气的局部演变细节。同时，MetNet系列<sup>[59,60]</sup>和StormCast<sup>[61]</sup>将气象预测的

空间分辨率进一步提升至3公里以内。东北师范大学高天焱团队<sup>[62]</sup>提出了YingLong大模型，通过构建自适应傅里叶神经算子与Swin Transformer并行融合的新型网络架构和边界平滑策略，在3公里分辨率上实现了近地面风速感知，其48小时内的气象预测精度超越美国国家海洋与大气管理局(National Oceanic and Atmospheric Administration, NOAA)所提出的高分辨率气象快速预报模型(High-Resolution Rapid Refresh, HRRR)。

尽管气象环境感知技术近年来取得了显著的进展，但是围绕低空场景的应用需求，仍然存在以下两个方面的问题亟待攻克：一方面，现有方法普遍依赖ERA5等全球尺度下的气象数据，这类数据的空间分辨率通常在25公里以上，难以适配低空场景下对于高分辨率气象环境场公里级/亚公里级建模的需求；另一方面，当前方法主要采用数据驱动的模式构建范式，通过大量的历史气象数据驱动模型挖掘气象演化的季节性、周期性与地域性特征，其感知结果符合训练数据在统计层面上的规律，但并未严格遵循大气系统中热力学、流体力学等物理规律，缺乏可解释的物理约束，对于风切变、短临突发降雨等影响低空活动开展的重要气象变化的感知精度有待进一步提升。总的来看，现有气象环境感知技术在大尺度空间分辨率数据的预测问题上取得了重大成果，传统数值天气预报具备物理一致性强、可解释性高的优点，但计算成本极其高昂，数据驱动的人工智能预报在挖掘气象的中长期趋势方面表现出色，但精度不足。如表2所示，当前方法均难以满足低空环境感知中所要求的公里级甚至百米级以下的高分辨率实时气象预报。

#### (2) 电磁环境感知

随着深度学习模型体系的不断演进，电磁环境感知<sup>[63,64]</sup>的基础网络架构已由早期以卷积操作为核心的判别式模型，逐步拓展至图神经网络、对抗生成模型以及扩散生成模型等生成式范式。相关研究表明，深度学习正在推动电磁环境感知由离散测点反演向连续空间表征与生成演进<sup>[65,66]</sup>。在稀疏采样、传播机理不完备以及场景结构高度复杂等约束

下，如何充分挖掘频谱数据中的空间关联、传统特性与场景先验，实现高分辨率、连续化频谱态势地图重建，已成为低空电磁环境建模与智能感知中的关键问题，也为低空空域电磁态势可视化和智能决策提供了重要支撑。

其中，Levie等人<sup>[67]</sup>构建基于物理传播仿真的大规模数据集RadioMapSeerT，并提出基于U-Net的频谱重建模型RadioUNet，率先确立了深度学习方法生成电磁频谱地图的技术路线。随后，PM-Net<sup>[68]</sup>通过跨层连接与Atrous卷积扩大感受野，有效提升了复杂场景下的电磁态势预测精度。Chen等人<sup>[69]</sup>将频谱图建模为无向图结构，利用图神经网络(Graph Neural Network, GNN)实现图域频谱重建；DeepREM<sup>[70]</sup>结合U-Net网络与条件对抗生成网络(Conditional Generative Adversarial Network, CGAN)，通过对抗训练增强模型对场景分布变化的鲁棒性。RM-Gen<sup>[71]</sup>则将发射器位置信息和稀疏采样点作为条件输入，引入条件扩散生成机制以捕获长距离依赖与复杂多径传播模式，更适用于发射器布局不确定或采样极度稀疏的低空场景。RadioDiff<sup>[72]</sup>进一步采用变分编码器(Variational Auto-Encoder, VAE)将输入映射至隐空间，并在

隐空间中执行扩散过程，在一定程度上兼顾了训练效率与生成质量。

表3对所列举的电磁环境感知方法进行了总结与对比。尽管现有方法在频谱地图重建与态势预测方面已取得一定进展，但面向低空应用需求时仍存在显著瓶颈。低空场景通常要求模型同时具备高精度空间刻画、快速动态响应、跨区域迁移以及决策支撑能力，而现有方法大多建立在大规模仿真或离线数据驱动基础之上，对传播机理、遮挡效应和多径衰落等物理规律缺乏显式约束，导致其在复杂环境下的结果虽具有一定数值拟合能力，却在物理一致性、可解释性与极端条件下的可靠性方面存在不足。与此同时，低空电磁环境具有开放动态、结构突变和多源耦合等特征，发射源状态、场景结构与传播条件的持续变化进一步放大了模型对训练分布的依赖，使其在连续感知、稳定泛化和实时更新方面面临挑战。

## 2.4 低空智联建设

低空智联是面向低空场景的下一代新型网络，其提出是为了弥补低空飞行活动从“少量、封闭、人工审批”走向“高频、开放、规模化运行”后所暴露出的基础设施缺口<sup>[73]</sup>。传统通航和无人机管理体系更适合低密度、低复杂度运行，但随着物流配送、城市治理、应急救援和载人低空交通等场景增长，仅靠单点通信、单机导航或局部监管已难以支撑安全、高效、连续的低空运行。因此，低空智联逐渐被定义为服务低空应用的端到端信息化系统，以网络、终端、平台和安全体系为基础，为低空飞行器提供数据传输、感知监视、定位导航、路径规划和智能计算等能力，是低空经济规模化发展的数字底座。其核心特征不是单一网络，而是通

表 2 气象感知方法的空间分辨率对比  
Tab. 2 Comparison of spatial resolutions of different meteorological sensing methods

感知方法	空间分辨率
FengWu-GHR	10公里
Aurora	11公里
MetNet	3公里以内
StormCast	3公里以内
YingLong	3公里以内

表 3 电磁频谱地图生成方法对比总结  
Tab. 3 Comparison and summary of electromagnetic spectrum map generation methods

感知方法	核心架构	关键机制	主要优势	主要局限
RadioUNet	U-Net	基于编码器-解码器的多尺度特征提取与重建	率先建立了深度学习生成电磁频谱地图的基本技术路线；结构简洁，易于训练与迁移	对复杂传播关系的显式建模能力有限
PMNet	改进型CNN U-Net变体	跨层连接+Atrous卷积扩大感受野	能更充分融合浅层细节与深层语义信息，在复杂场景下提升预测精度	仍主要依赖卷积局部建模，对超长距离依赖和全局拓扑关系刻画不足
GNN-based 方法	图神经网络(GNN)	将频谱图建模为无向图，在图域进行信息传播与重建	能更自然地建模空间节点之间的关联关系，适合不规则采样分布	图构建方式对性能敏感，计算复杂度较高，大规模场景下扩展性受限
DeepREM	U-Net + CGAN	对抗训练增强生成结果的真实性与分布鲁棒性	相比重建模型，对场景变化和分布偏移具有更强适应性	对抗训练稳定性较差，训练成本较高，易出现模式不稳定问题
RM-Gen	条件扩散模型	以发射器位置与稀疏采样点为条件输入，利用扩散过程建模复杂传播分布	能较好捕获长距离依赖和复杂多径传播模式，适合极稀疏采样和发射器布局不确定场景	扩散模型训练与推理开销较大，对条件设计和采样策略较敏感
RadioDiff	VAE + 扩散模型	先映射到隐空间，再在隐空间执行扩散生成	在一定程度上兼顾训练效率与生成质量，缓解像素空间扩散计算开销	隐空间压缩可能带来细节损失，模型设计更复杂，对潜变量质量依赖较强

信网、感知网、导航网、气象网和算力网“五网协同”，共同支撑低空活动的“通、监、导、气、算”需求<sup>[74]</sup>。低空物联网难点并不只是单项能力不足，而在于这些能力在真实场景中必须协同工作。在感知方面，低空目标小、快、轨迹变化大，城市环境中还存在遮挡、杂波和非合作目标识别难题；而在算力方面，面临机载设备算力、能耗和载荷受限，而云侧集中处理又容易引入时延和带宽瓶颈，因此必须在端、边、云之间动态分工<sup>[75]</sup>。

### (1) 传统低空感知体系

主要来源于安防、防空和通航监管需求，其基本思路是依靠专用传感器和地面站点实现区域监视，典型手段包括雷达、无线电侦测、ADS-B、Remote ID、光电设备以及人工值守平台。这类体系的优点是技术路径相对成熟，目标明确，适合重点空域、机场周边、重要设施周边等高价值区域部署，尤其在合规飞行目标识别和固定区域预警方面已有较多实践。其不足也较明显：首先，它往往偏向“点状布设+局部覆盖”，难以支撑城市级连续感知；其次，对非合作目标、低慢小目标以及复杂城市背景下的遮挡与虚警问题较为敏感<sup>[34]</sup>；再次，这类体系通常更强调“发现和告警”，而非与通信、导航、调度和智能计算形成实时闭环，因此更适合传统监管场景，不完全适应低空经济下大规模、常态化运行的需求。美国联邦航空局的无人机系统交通管理(Unmanned Aircraft System Traffic Management, UTM)和欧盟的无人机空域管理系统U-space也进一步阐明，单一传感器体系不足以支撑未来规模化低空运行，仍需依赖跨系统协同。

### (2) 新兴低空感知体系

在传统监视思路基础上的进一步演进，其突出特点是从“孤立感知”转向“网联感知”，从“被动发现”转向“感知-传输-决策-反馈”闭环协同。代表性的路线包括多源融合感知、5G-A通感一体、空地协同感知、边缘智能分析以及端边云联合推理<sup>[66]</sup>。与传统体系相比，这类体系更强调把基站、边缘节点、无人机和云平台联结起来，通过融合雷达、视觉、无线信号和飞行状态信息，提升目标识别、轨迹跟踪和异常预警能力；同时借助边缘计算和模型更新机制，在本地完成快速筛查，在云侧完成复杂分析与模型演进。近期研究开始把无人航空器(Unmanned Aerial Vehicle, UAV)视为“感知+计算+通信”一体化边缘节点，在有限带宽下通过特征流、数据流和模型流协同完成任务分配与模型更新。其优势在于更适应低空目标动态性强、业务变化快的特点，但不足在于系统复杂度高，对

跨网协同、标准接口、资源调度和工程成本提出了更高要求，距离大范围统一部署仍有一段距离。

## 3 基于数字视网膜端边云协同架构的低空主动感知网及其关键技术

本文提出了一种基于数字视网膜端边云协同架构的低空主动感知网方案(以下简称“低空主动感知网”)。其核心方法创新与贡献主要包含以下3个方面：

(1) 提出了特征驱动的“三流协同”端边云闭环机制。针对现有架构“脑眼分离”导致的通信拥塞与协同迟缓问题，突破传统的原始数据单向全量传输模式，依托数字视网膜架构，通过“特征流、数据流、模型流”的并行交互，实现了低空特征前置提取、数据按需回传与感知模型在线演化的闭环。

(2) 构建了面向低空场景的感知驱动资源调度框架。针对空地通信带宽波动大、端侧感知节点(如无人机)算力受限的痛点，引入广义速率-效用优化(Generalized Rate-Utility Optimization, GRUO)模型，实现了有限带宽下特征数据压缩、传输码率与计算任务卸载的动态自适应均衡，确保了高动态环境下系统感知效用的全局最大化。

(3) 设计了云端优势互补的多模态感知算法体系。针对低空非合作目标“低慢小”及微观气象、电磁环境复杂多变的挑战，针对性地构建了云侧感知基础模型与端侧轻量化感知模型的技术体系，形成了云侧高维稀疏数据处理分析、端侧实时敏捷响应的感知能力。

以下将详细阐述该网络的整体架构与各项代表性关键技术的细节。

### 3.1 低空主动感知网整体架构

本文提出的低空主动感知网整体架构如图3所示。这一架构以数字视网膜“端边云协同智算”架构为基座，是整个感知网络的“中枢神经”。所谓“协同智算”是指打破单一计算节点的算力与信息壁垒，依托互联网络实现端侧感知设备(各类传感器)、边侧中继计算节点与云侧(大规模计算存储平台)在数据流动、AI算法推理与算力资源分配上的全局动态统筹与深度融合。基于该协同智算基座，本方案具备“特征实时汇聚、数据按需调取、模型在线更新”三大特性，通过特征流、数据流与模型流3条链路，连接了物理空间的端侧感知设备与数字空间的云侧感知大脑。在云侧，部署“鹏城·大基”系列感知基础模型作为“感知大脑”，承担长

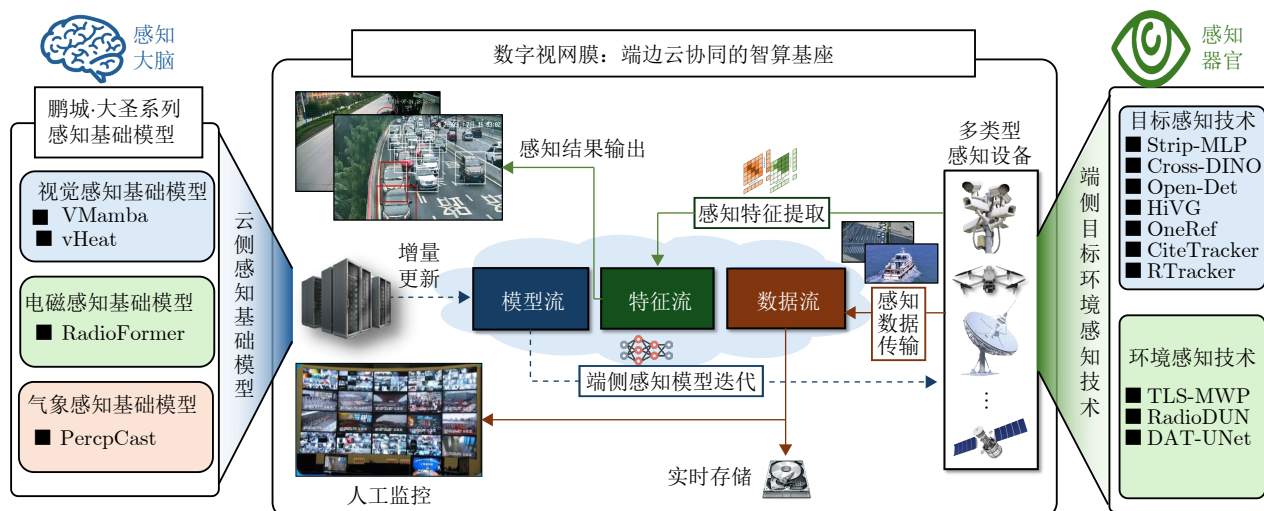


图 3 基于数字视网膜的低空主动感知网架构示意图

Fig. 3 Architecture of a digital retina-based low-altitude active perception network

时序、大空间尺度的复杂推理、跨模态关联与通用知识演化任务；在端侧，部署多种轻量化的高效目标环境感知技术，在功耗与算力受限的边缘节点实现对低空高动态目标及气象、电磁环境的实时捕捉与特征提取。

通过上述组件的有机结合，低空主动感知网实现了从“被动采集”向“主动感知”的跨越，形成了“感知-传输-计算”一体化闭环。为了系统地阐述该架构的运行机理与技术创新，以下将通过架构对比分析其在低空场景下的适应性与优势，详细介绍能够最大化发挥端侧感知设备能力的“特征-数据-模型”三流协同机制，并阐述低空主动感知网在受限资源下保障感知效用最大化的端边云协同计算机制，以及多源感知信息融合与决策方法。

### 3.1.1 架构对比与优势分析

现有的感知系统主要采用两种典型架构：一是“云侧集中式”，即将前端传感器采集的原始视频或雷达回波全量上传至云侧处理。该方式虽然能利用云侧强大算力，但在面对低空海量数据流时，极易造成通信链路拥塞，导致高延迟甚至数据丢失；二是“端侧智能驱动式”，即依赖端侧感知设备的算力进行感知任务的本地处理。受限于端侧的功耗与算力，这类架构难以部署大参数量的先进模型，且模型一旦部署便难以更新，难以适应低空环境的动态变化。

针对上述痛点，本文提出的低空主动感知网采用数字视网膜端边云协同智算架构，其核心优势在于：

(1) 通信与感知计算的全局优化。通过“特征前置提取、数据按需回传”，将繁重的感知计算任

务合理分配至端、边、云侧。端侧仅需上传极低带宽占用的特征流，仅在必要时(如目标确认、证据留存)回传原始数据，从而在有限的低空空地带宽下实现实时感知。

(2) 持续进化的主动感知能力。不同于传统传感器功能固化，低空主动感知网支持通过“模型更新信息传输”的方式更新端侧算法，使得前端传感器具备动态适应的特性，能够随着环境变化和任务需求主动进化感知能力。

### 3.1.2 融合“特征-数据-模型”的三流协同机制

为了充分发挥端侧异构感知设备(如摄像机、雷达、频谱仪)在低空场景中的能力，低空主动感知网方案建立了如图4所示的“特征流-数据流-模型流”三流协同机制，具体如下：

**特征流(Feature Stream)：**端侧设备利用轻量级深度神经网络或传统的特征描述(如CDVS/CDVA标准)，从原始感知数据中提取任务驱动的紧凑特征表达。特征流的数据量通常仅为原始数据流的1/100甚至更低，且包含了数据的高维语义信息。在低空感知网中，特征流被实时上传至云侧，用于执行大规模的快速检索与初步分析，有效缓解了上行链路的带宽压力。

**数据流(Data Stream)：**数据流指经过高效编码的数据或原始信号。在数字视网膜架构中，数据流通常存储于端侧或边缘节点，仅在特定条件下触发传输。当云侧通过特征流分析发现异常事件(如不明飞行物入侵)或需要高精度复核时，系统采用“按需调取”的策略回传数据流。此外，数据流还可包含用于数据重构的残差信息，与特征流配合以增强云侧基础模型的推理精度。

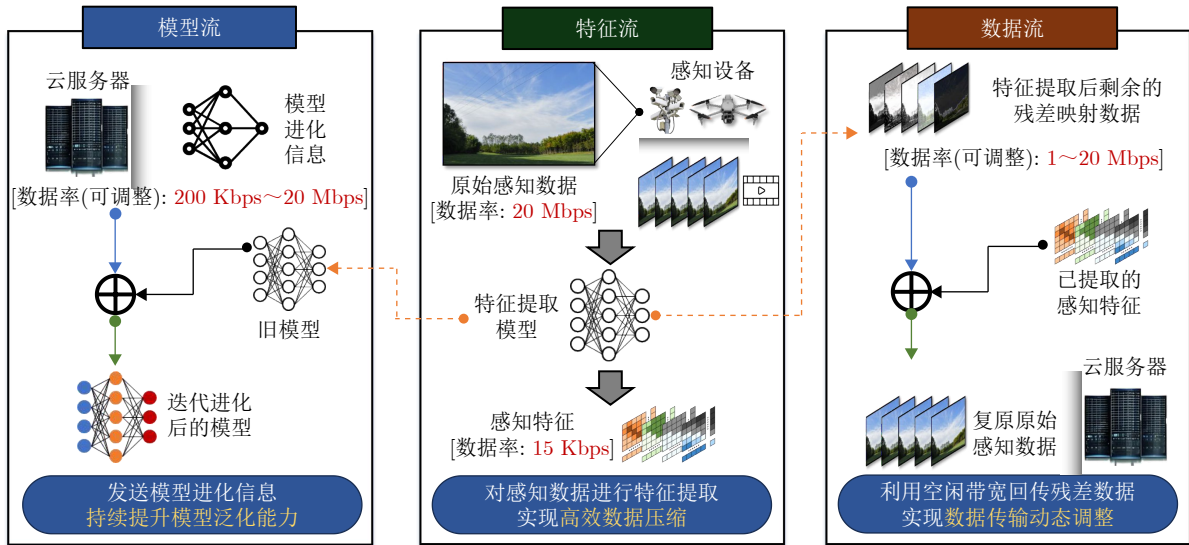


图4 数字视网膜中三流协同机制

Fig. 4 Three-Stream collaborative mechanism in digital retina

**模型流(Model Stream):** 模型流是一条自云侧向端侧传输的逆向传输信息流。云侧利用汇聚的海量数据训练或微调高精度的感知基础模型,并通过模型流将更新后的模型参数、权重差值或轻量化模型(云测模型通过量化、知识蒸馏等方式得到)下发至端侧。通过模型流的持续交互,端侧设备能够动态加载适应当前低空环境(如特定光照、气象或目标类型)的专用算法,实现了感知系统的软件定义与在线进化能力。在实际工程部署中,为保障系统运行的鲁棒性,模型更新通常采用基于模型差异(Difference of Models, DoM)的增量更新机制<sup>[6]</sup>,具体而言,云侧仅计算更新后模型与端侧现有模型之间逐层的权重差值,并对该差异矩阵进行标量量化与熵编码后下发,端侧接收后通过模型补偿即可完成参数重建。这一机制不仅降低了下行通信开销与更新延迟,端侧设备还配套内置了严格的版本兼容性校验与通信失败回滚机制,以确保在低空弱网或突发断网条件下,感知模型能够平稳运行并实现安全恢复。

### 3.1.3 端边云协同感知计算机制

低空主动感知网的核心不仅在于“三流协同”,更在于通过端边云的紧密协作,在有限的计算与通信资源下达成感知性能的最优。为此,低空主动感知网引入了广义速率-效用优化(GRUO)模型,以解决低空复杂环境下的资源分配难题<sup>[5]</sup>。不同于传统视频编码仅关注信号层面的率失真平衡,GRUO模型聚焦于机器感知的效用最大化,其目标函数定义如下:

$$S^* = \operatorname{argmin}_S (S|V, T; \Theta_L) + \lambda R(S; \Theta_R) \quad (1)$$

其中,  $V$ 表示原始感知数据,  $S$ 表示其在网络中的表征形式(例如面向存储的视频纹理流或面向分析的紧凑特征流),而  $S^*$ 代表求解得到的最优表征。 $T$ 为特定的感知任务(如目标检测或目标跟踪等),  $L$ 为效用损失函数,用于衡量在特定任务下的机器感知效用或精度损失,其被  $\Theta_L$ 参数化( $\Theta_L$ 代表决定感知效用的模型网络参数,如神经网络的权重);  $R$ 为传输码率或计算开销的代价函数,被  $\Theta_R$ 参数化( $\Theta_R$ 为决定资源消耗的编码参数,如视频编码中的步长、特征压缩模式等);  $\lambda$ 则是拉格朗日乘子,用于动态调节感知精度与传输/计算代价之间的平衡。通过求解该优化问题,低空主动感知网能够根据当前的信道状态与任务需求,动态调整特征流与数据流的编码策略。

为保障GRUO模型在低空高动态环境下的时效性以满足在线动态调整需求,低空主动感知网在实际部署时避免了极耗算力的逐帧在线最优化求解。具体而言,低空主动感知网采用“离线建模+在线查表”的策略:在离线阶段,针对不同任务(如目标检测、目标轨迹跟踪等)、不同分辨率以及不同编码量化参数或特征压缩模式,预先评估并构建“速率-效用先验关系表(R-U Lookup Table)”;在线运行阶段,端侧资源调度模块仅需根据当前探测到的无线信道状态(带宽波动)和下发的任务优先级,以极低的算力开销(通常在毫秒级)实时匹配出最优的模型提取配置参数( $\Theta_L$ )与编码调控参数( $\Theta_R$ )。这种查表或单次轻量级前向推理的机制,使得整体架构能够以几乎零延迟的代价,动态且平滑

地切换特征流与数据流的编码和传输策略，满足低空场景下的实时感知需求。

### 3.1.4 全局时空基准下的多源感知信息融合与决策

低空感知涉及雷达、光电、频谱等多源异构感知设备，单一感知设备往往存在感知盲区(如视觉传感器易受光照影响、雷达难以避免杂波干扰等)。为了实现多源信息的高效融合以支持精准决策，低空主动感知网建立了全局统一的时空基准(Unified Spatio-Temporal Reference)，从而具备以下特性：

(1) 多模态感知数据时空对齐。低空主动感知网中所有端侧设备均内置高精度授时与定位模块(北斗/GPS)，在具体实现上，系统采用硬件触发与软件授时相结合的跨模态同步机制<sup>[5]</sup>：在硬件层，通过全球导航卫星系统(Global Navigation Satellite System, GNSS)输出的秒脉冲(Pulse Per Second, PPS)信号对雷达、光电等异构传感器进行统一的底层物理触发，确保微秒级的采集对齐；在软件层，辅以精确时间协议(Precision Time Protocol, PTP)或网络时间协议(Network Time Protocol, NTP)进行系统级时钟的连续校准，消除漂移误差。通过上述软硬协同的技术手段，为每一帧视觉特征、每一个雷达点云数据赋予包含高精度绝对时间戳与地理坐标的唯一标识符，从而确保异构数据在云侧能够基于统一的时空坐标系进行像素级或特征级的精准配准。

(2) 融合多源感知的决策。基于统一时空基准，云侧感知大脑能够关联不同视角的特征流。例如，利用雷达的高精度测距与测速能力弥补视觉感知的空间深度缺失，利用视觉的丰富纹理特征弥补雷达对微小目标分类能力的不足。通过多模态特征的互补融合，系统能够剔除虚警信号，构建出包含目标位置、轨迹、类别的完整低空态势图，从而为空域管理、航路规划及威胁处置提供高置信度的决策依据。

综上，本文所提出的低空主动感知网整体感知计算流程如图5所示：首先，在真实的低空场景中，部署在端侧的多类型感知设备(如摄像头、雷达、无人机及卫星等)对“可见”的目标实体(如行人、车辆、船舶、无人机等)与“不可见”的环境场(如风、雨、温度、电磁频谱等)进行原始信号捕获。随后，端侧设备利用内置部署的轻量级感知模型对原始感知数据进行本地处理与特征编码，提取出关键的感知特征。在此协同架构下，端侧将提取的紧凑感知特征作为“特征流”实时上传至云侧，并根据实际任务需求，将原始感知数据或特征提取后的残差数据与未经进一步处理的原始感知数据合并为“数据流”进行传输。云侧在接收到多源异构数据后，执行多源特征融合与多源数据融合，并依托云侧大算力、大存储的条件，运用感知基础模型完成大范围、多粒度、多模态的复杂感知计算与态势推理。同时，云侧利用汇聚的数据进行基础模型

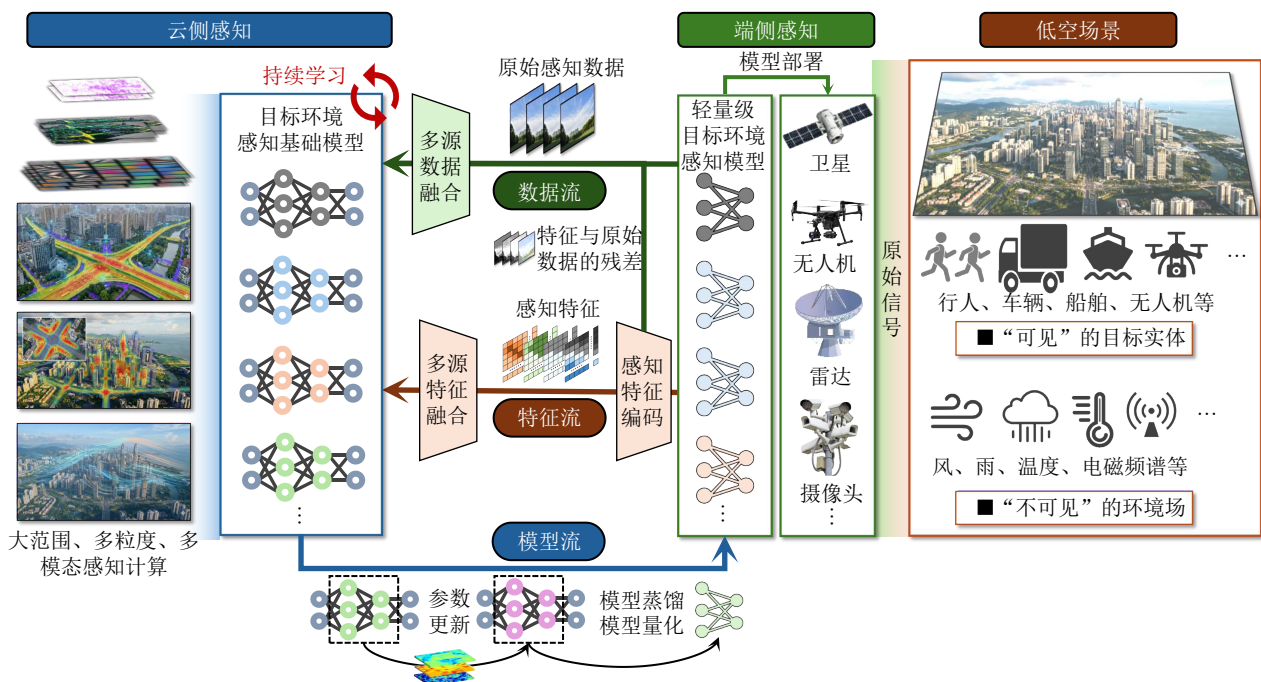


图 5 基于数字视网膜的低空主动感知网感知计算流程示意图

Fig. 5 Perception and computation pipeline of the digital retina-based low-altitude active perception network

的持续学习,通过模型蒸馏、模型量化等技术将基础模型的能力赋予新的端侧模型,并以“模型流”的形式逆向下发至端侧,实现端侧轻量级模型的在线部署与增量更新,从而形成整个感知网络自适应动态环境并持续进化的主动感知能力。

### 3.2 云侧感知基础模型

低空主动感知网的云侧基础模型主要承担全局性的时空关联和推理、为端侧模型提供通用知识的作用。为了实现对低空物理世界的全方位感知,本方案在云侧部署“鹏城·大圣”系列感知基础模型,构建了一套涵盖“可见目标实体”与“不可见环境”的完备感知体系。其中,视觉感知基础模型面向各类无人机、航空器及地物等“可见”实体,旨在解决低空目标的精准表征问题;而电磁与气象感知基础模型则面向频谱分布、大气流场等人眼“不可见”的环境场,通过数据驱动与物理约束相结合的方式,实现对环境态势的高精度反演与重构。这3类模型功能协同,互为补充,实现从物理空间到数字空间的低空要素映射,支撑低空主动感知网对低空目标、电磁频谱环境及气象环境的精准感知。3类感知基础模型具体如下:

#### (1) 视觉感知基础模型

面向低空目标感知场景中对宽幅图像高效处理的需求,针对传统视觉Transformer架构中自注意力机制带来的二次复杂度计算开销问题,本文构建了一系列低计算复杂度的视觉目标感知基础模型,以下重点阐述其中两类具有代表性的创新架构:基于状态空间模型的VMamba<sup>[76]</sup>和基于物理热传导原理的vHeat<sup>[77]</sup>(如图6所示)。前者引入了二维选择性扫描(2D-Selective-Scan, SS2D)机制,区别于传统模型的一维扫描,SS2D通过“左上至右下、右下至左上、右上至左下、左下至右上”4条交叉路径对图像特征进行展开与独立扫描,随后将特征融合,从而在保证特征具备全局二维空间感知能力的同时,实现了 $O(N)$ 的线性计算复杂度。后者受物理热传导原理启发,将图像中的特征斑块(Patch)视为热源,提出了热传导算子(Heat Conduction Operator, HCO),通过在频域内自适应预测热扩散系数 $k$ ,模拟视觉语义特征在多通道间的热能扩散过程。其核心特征传递过程可表示为

$$U^t = F^{-1} \left( F(U^0) e^{-k(\omega_x^2 + \omega_y^2)t} \right) \quad (2)$$

其中, $U^0$ 为初始特征矩阵, $F$ 和 $F^{-1}$ 分别代表二维离散余弦变换(Discrete Cosine Transform, DCT)及其逆变换, $\omega_x$ 和 $\omega_y$ 为空间频率变量。实验表明,在处理 $768 \times 768$ 高分辨率图像时,所提出的VMamba

模型在保持高精度的同时实现了远超Swin-Transformer的吞吐量;而vHeat模型不仅在ImageNet-1K数据集上实现了更高的分类精度,其显存占用亦降低了近80%,为低空场景下计算资源受限的端侧设备部署提供了强有力的高效视觉表征底座。

#### (2) 电磁环境感知基础模型

针对低空电磁环境复杂多变、监测设备部署受限导致的观测数据极端稀疏(如空间采样率低至1%)问题<sup>[78]</sup>,以及现有视觉模型在处理稀疏数据时存在的计算冗余与特征提取瓶颈,本文构建了基于多粒度(Multiple-Granularity)Transformer架构的电磁感知基础模型RadioFormer<sup>[79]</sup>。该模型创新性地提出了一种“双流自注意力(Dual-stream Self-Attention, DSA)”与“跨流交叉注意力(Cross-stream Cross-Attention, CCA)”相结合的特征提取与融合框架(如图7所示)。具体而言,DSA模块在“斑块(Patch)”级别利用视觉编码器提取建筑物等障碍物的几何拓扑特征 $F_B$ ,同时在“像素(Pixel)”级别对稀疏采样点的坐标位置与信号强度进行编码提取观测特征 $F_O$ 。随后,CCA模块将观测特征 $F_O$ 作为查询向量( $Q$ ),将建筑物特征 $F_B$ 作为键( $K$ )和值( $V$ )进行跨模态多尺度特征交互,其核心的交叉注意力(Cross-Attention, CA)计算过程可表示为

$$CA(Q, K, V) = \text{Softmax}(QK^T)V \quad (3)$$

融合后的特征序列通过轻量级解码器重构为完整的连续空间电磁频谱地图 $M'$ ,由于电磁频谱分布严格受物理衰减规律约束,模型采用均方误差(Mean Squared Error, MSE)损失函数进行端到端优化:

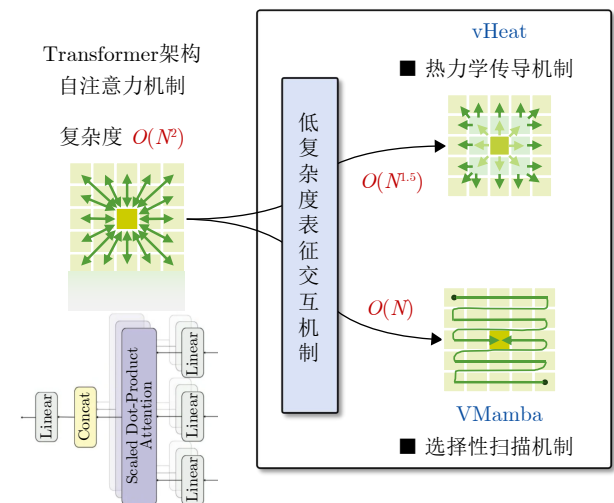


图6 VMamba与vHeat中的低复杂度计算机制

Fig. 6 Low-complexity computing mechanisms in VMamba and vHeat

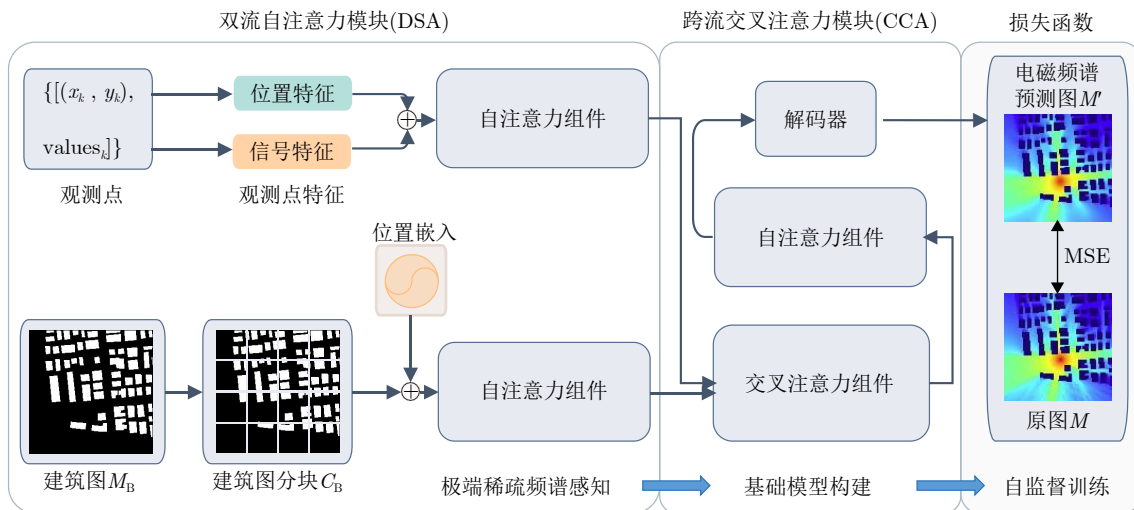


图 7 电磁感知基础模型RadioFormer的整体框架示意图

Fig. 7 Framework of the electromagnetic sensing foundation model

$$L_{\text{MSE}} = \frac{1}{N} \sum (M - M')^2 \quad (4)$$

其中,  $M$  为真实的电磁频谱地图,  $N$  为像素总数。通过这种多粒度特征融合机制, RadioFormer 有效避免了极度稀疏采样下的计算冗余。实验表明, 在仅有 5 或 9 个稀疏观测点的极端条件下, RadioFormer 的均方根误差 (Root Mean Squared Error, RMSE) 与结构相似度指数 (Structural Similarity Index Measure, SSIM) 均显著优于传统 RadioUNet 与 Pixel Transformer 等方法, 且有效消除了模型对非平稳采样的过度依赖, 实现了极稀疏空间采样下对低空全局电磁频谱态势的高精度、低延迟精细重构。

### (3) 气象环境感知基础模型

针对低空飞行对短临降水预报的高时效性与高清晰度需求, 本文构建了一种融合感知约束的降水预报模型 PercpCast<sup>[80]</sup>, 采用“先估计后约束”的两阶段框架, 以结合 ConvLSTM 与 Transformer 的时空特征融合网络为基础, 捕捉降水场的局部细节与全局依赖。其核心在于引入了感知约束机制, 从结构相似性、纹理细节和时空一致性 3 个维度优化输出, 并融合物理一致性要求, 实现了感知与物理的双重约束, 其核心损失定义为

$$E[\|\mathbf{Y} - \hat{\mathbf{Y}}^*\|^2] + E\left[\int_0^1 \|(\mathbf{Y} - \mathbf{Z}_0) - v_\theta(\mathbf{Z}_t, t)\|^2 dt\right] \quad (5)$$

其中,  $\mathbf{Y}$  和  $\hat{\mathbf{Y}}^*$  表示预测值和真实值,  $\mathbf{Z}_0$  后验均值序列,  $\mathbf{Z}_t$  表示在流程时间  $t$  的中间状态,  $v_\theta$  是一个参数化的向量场,  $E[\cdot]$  表示期望算子, 损失的前半部分用于最小化预测的均方误差, 后半部分为感知约束项, 使生成数据与真实分布对齐。如图 8 所示, 相较于 SimVP<sup>[81]</sup> 等现有基于均方误差 (MSE) 约束的方法, PercpCast 能生成更锐利、更清晰的降

水图像, 能够完整清晰地保留强降水中心的结构和边缘, 为低空环境下的短临强降水提供准确预警。

## 3.3 端侧目标环境感知

低空感知网中端侧感知设备直接面对复杂动态的低空物理环境, 但受限于端侧设备的功耗与算力限制, 本节将介绍部署于端侧的轻量化高效感知技术, 以支撑低空主动感知网络在目标与环境跨场景条件下的泛化感知能力。

### 3.3.1 低空目标感知

针对低空目标的高动态非线性运动模式以及复杂背景干扰等挑战, 本文在端侧构建了一套涵盖检测、定位与跟踪的目标感知技术体系, 旨在受限算力下实现对低空目标的精准感知。

#### (1) 面向端侧计算的高效目标检测

低空目标检测与开放场景下的未知目标检测是感知网的核心任务。传统基于卷积神经网络 (Convolutional Neural Network, CNN) 的检测器虽易于部署, 但受限于局部感受野, 难以在复杂低空背景中有效提取目标语义; 而基于 Transformer 的架构虽具备全局建模能力, 但其平方级计算复杂度使得端侧实时性大幅下降。为平衡端侧设备的计算效率与特征提取能力, 本文提出了 Strip-MLP 模型<sup>[82]</sup>, 通过引入 Strip MLP Layer 基础网络层范式, 能够自适应实现图像空间长短距离信息的交互, 克服了 Transformer 的高计算复杂度与 CNN 局部感受野的局限, 从而同步捕捉全局与局部特征。该模型能够处理任意分辨率大小的图像数据, 在保持线性计算复杂度的前提下, 克服了传统 CNN 感知范围受限与 Transformer 算力冗余的矛盾。同时, 雷达探测

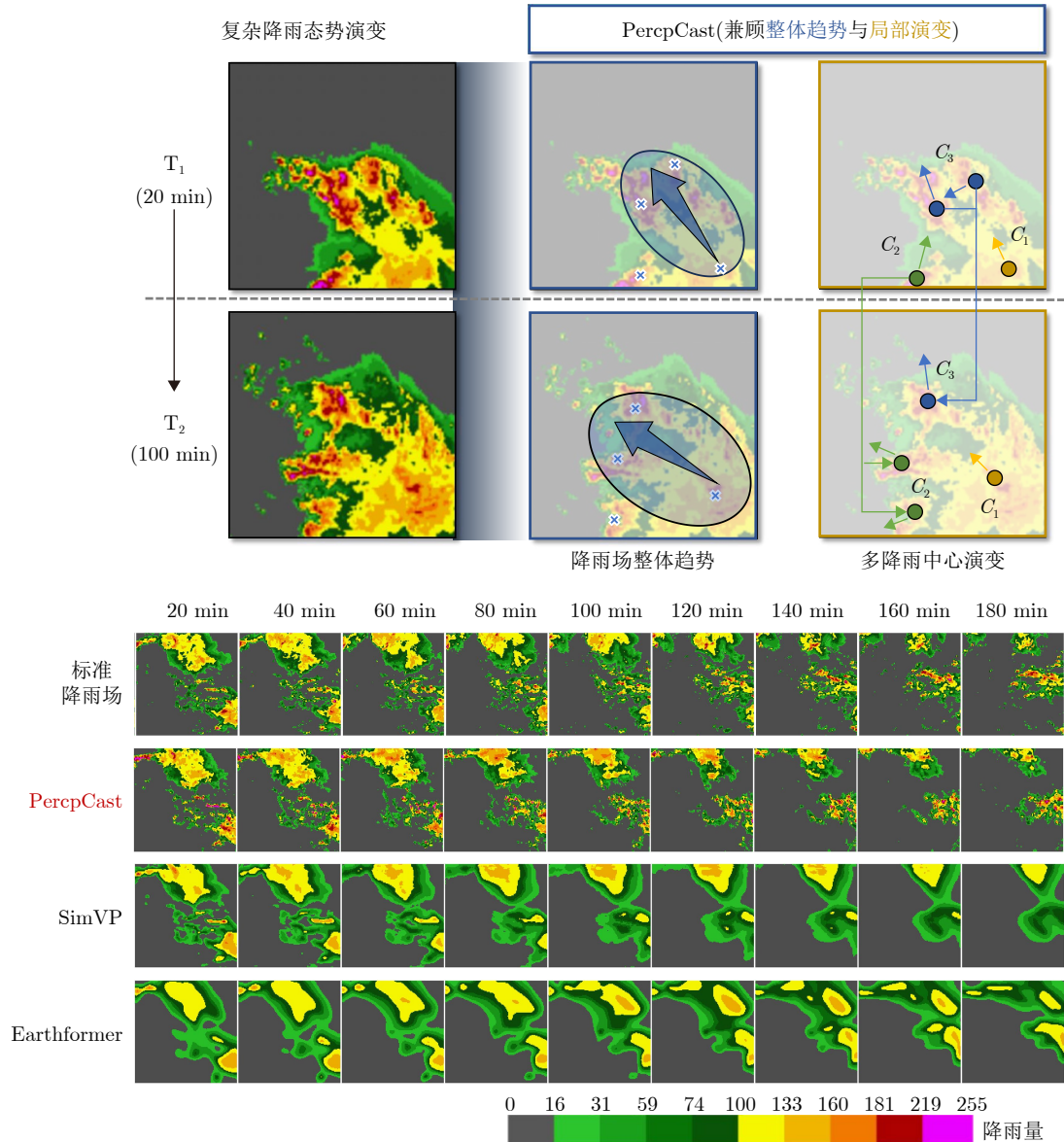


图8 降水预测模型PercpCast与现有方法的预测结果对比

Fig. 8 Comparison of prediction results between the PercpCast and existing methods

领域针对稀疏目标的超分辨重建方法<sup>[83]</sup>,也为端侧受限算力下目标检测的稳健性优化提供了一种重要的参考思路。

在此基础上,针对小目标像素占比少、特征有限及易漏检的问题,本文构建了Cross-DINO检测方法<sup>[84]</sup>,将Strip-MLP提取的特征与经Transformer编码器增强的特征进行交叉编码,融合了上下文捕捉优势与实例建模能力。同时,通过设计增强损失(Boost Loss,如式(6)所示),将目标尺寸信息融入分类预测,自适应增大对小目标的关注度,显著降低了远距离小目标的漏检率,弥补了通用检测器在特定低空场景下的表征能力不足,这一损失的设计借鉴了雷达无监督目标特征提取方法中多重对比损失<sup>[85]</sup>,为实现特征挖掘与融合提供有力支撑。

$$L_{\text{Boost}} = -\frac{1}{N} \sum_{i=1}^N \{ \alpha(1 - c\hat{s}_i^\beta)^\gamma c\hat{s}_i^\beta \log(p_i) + (1 - \alpha) p_i^\gamma (1 - y_i) \log(1 - p_i) \} \quad (6)$$

其中,  $c\hat{s}_i$  表示目标框大小与图像尺寸的比值,  $p_i$  表示预测概率,  $y_i$  为标签,  $\alpha, \beta, \gamma$  为调控因子,值均在[0,1]范围内。

此外,针对缺乏先验知识的开放世界场景,本文进一步提出了Open-Det方法<sup>[86]</sup>,通过解耦目标检测与识别过程,并将类别无关检测器与生成式语言模型融合,实现了不依赖预定义词汇表的开集目标检测,使系统具备了对未知“非合作”目标的泛化识别能力。该模型在无需文本先验的情况下,能够对输入图像中潜在的任意目标进行检测和识别,具备“万物检测”的能力,如表4所示,其性能显

表 4 开放世界场景下的目标检测方法在LVIS数据集<sup>[89]</sup>的性能对比  
Tab. 4 Comparison of open-world object detection methods on the LVIS<sup>[89]</sup> dataset

方法	骨干网络	AP <sub>r</sub>	AP <sub>c</sub>	AP <sub>f</sub>	AP
Grounding DINO	Swin-Tiny	14.4	19.6	32.2	25.6
GenerateU	Swin-Tiny	20.0	24.9	29.8	26.8
本文所提Open-Det	Swin-Tiny	21.9	25.1	30.4	27.4

著超越开放词汇检测模型Grounding-DINO<sup>[87]</sup>和同类型的开放式检测模型GenerateU<sup>[88]</sup>。

### (2) 跨模态语义引导的目标定位

视觉定位(Visual Grounding, VG)作为典型的跨模态任务,对于低空场景下的智能搜寻与人机交互至关重要。现有的视觉定位模型依赖于大语言模型,通常参数量巨大,导致视觉与文本语义融合不充分,且端侧推理延迟难以控制。针对上述问题,本文提出了两种面向端侧部署的高效方法:HiVG<sup>[90]</sup>和OneRef<sup>[91]</sup>。这两种方法均基于视觉与文本的跨模态对齐能力,能够从端侧采集的可见光(RGB)或红外(Infrared, IR)图像中,根据用户输入的自然语言描述(如“那架白色无人机”)精确锁定目标区域。具体地,在HiVG方法中,为缓解视觉Transformer自注意力机制带来的二次计算复杂度,同时避免多层感知机对全局信息建模能力的不足,提出了一种层级多模态细粒度调制框架。该方法提出层级多模态低秩适应(Hierarchical Low-Rank Adaptation, HiLoRA)机制,将预训练视觉语言模型的网路层划分为多个组,并在不同阶段逐层适配:

$$h_j^l = \begin{cases} \mathbf{W}_0^l x^l, & \text{当 } l > j \cdot L/G \\ \mathbf{W}_0^l x^l + \sum_{k=\lceil l \cdot G/L \rceil}^l \mathbf{B}_k^l \mathbf{A}_k^l x^l, & \text{当 } l \leq j \cdot L/G \end{cases} \quad (7)$$

其中,  $\mathbf{W}_0^l \in R^{d \times k}$  为第  $l$  层预训练权重,  $\mathbf{A}_k^l \in R^{r \times k}$ ,  $\mathbf{B}_k^l \in R^{d \times r}$  为低秩分解矩阵,  $r \ll \min(d, k)$  为低秩维度,  $L$  为网络总层数,  $G$  为划分的组数,  $\lceil \cdot \rceil$  表示向上取整。通过该层级适配结构,HiVG可在保持预训练知识的同时,实现视觉与文本特征的渐进式细粒度对齐,显著提升端侧设备在复杂语义场景下的目标定位精度。

在OneRef方法中,本文进一步探索了基于统一单塔架构的指代表达理解框架。该方法引入掩码指代建模(Mask Referring Modeling, MRefM),通过重构视觉-文本的指代关系增强跨模态语义对齐。其指代感知掩码图像建模和掩码语言建模分别通过点积操作融合视觉与文本聚合特征,重构被掩码的原始内容和目标关系分数,在统一特征空间中实现高精度目标定位与分割,进一步降低了端侧部署的计算开销。

### (3) 复杂场景下的鲁棒目标跟踪

低空目标常伴随剧烈姿态变化、遮挡及视觉信息退化,导致纯视觉模板匹配易发生漂移。现有的视觉跟踪方法高度依赖模板匹配,当低空目标发生姿态剧变、光照波动或被短暂遮挡时,易因像素级匹配失效而产生路径漂移。为解决弱纹理与外观不稳定性问题,本文提出了图像-文本关联的跟踪新范式CiteTracker<sup>[92]</sup>(如图9所示),该范式基于采用对比学习的文本-图像预训练模型(Contrastive Language-Image Pre-training, CLIP)的跨模态对齐能力,将目标图像映射至语义属性空间。

为应对目标外观的动态变化,我们设计了语义引导模块,通过计算目标特征  $\mathbf{F}$  与语义属性空间  $\mathbf{T}$  的属性差异度  $\mathbf{D} = \mathbf{F}_{\text{cur}} - \mathbf{F}_{\text{ref}}$  (其中  $\mathbf{F}_{\text{cur}}$  为当前帧特征,  $\mathbf{F}_{\text{ref}}$  为初始参考特征),并通过Softmax函数生成属性注意力权重:

$$\text{Weight} = \text{Softmax}(-\mathbf{D}_{\text{color}}, -\mathbf{D}_{\text{texture}}, -\mathbf{D}_{\text{material}}) \\ \mathbf{T}_{\text{att}} = \mathbf{W} * \mathbf{T}_{\text{attributes}} \quad (8)$$

其中,  $\mathbf{T}_{\text{att}}$  为最终输出的动态文本特征。该语义引导机制显著抑制背景干扰,增强小尺寸、高动态目标表征鲁棒性。CiteTracker可同时处理可见光图像与文本数据,这种从“像素级模板匹配”到“语义级属性关联”的范式转变,使系统在纹理模糊、光照剧变等极端环境下仍能依托语义属性抑制干扰,稳定性显著优于传统算法;同时,雷达领域针对低空“走-停-走”目标的鲁棒跟踪技术<sup>[93]</sup>,可进一步为复杂场景下低空目标跟踪的抗漂移设计提供支撑。

此外,现有跟踪方法通常缺乏丢失找寻机制,一旦目标丢失即导致任务执行失败,针对全遮挡、出视野等导致的目标跟踪丢失痛点,本文进一步提出了可恢复跟踪方法RTracker<sup>[94]</sup>。如图10所示,采用正负样本决策树存储可见光图像中的目标外观特征(正分支)与干扰特征(负分支)。核心通过相对相似度判断目标状态:

$$\text{State} = \cos(\mathbf{F}_{\text{real}}, \mathbf{F}_{\text{node}}) \quad (9)$$

其中,  $\mathbf{F}_{\text{real}}$  为当前帧实时特征,  $\mathbf{F}_{\text{node}}$  为树节点中存储的特征。当正分支最高相似度大于负分支时,判定目标“存在”并持续跟踪;反之判定为“丢

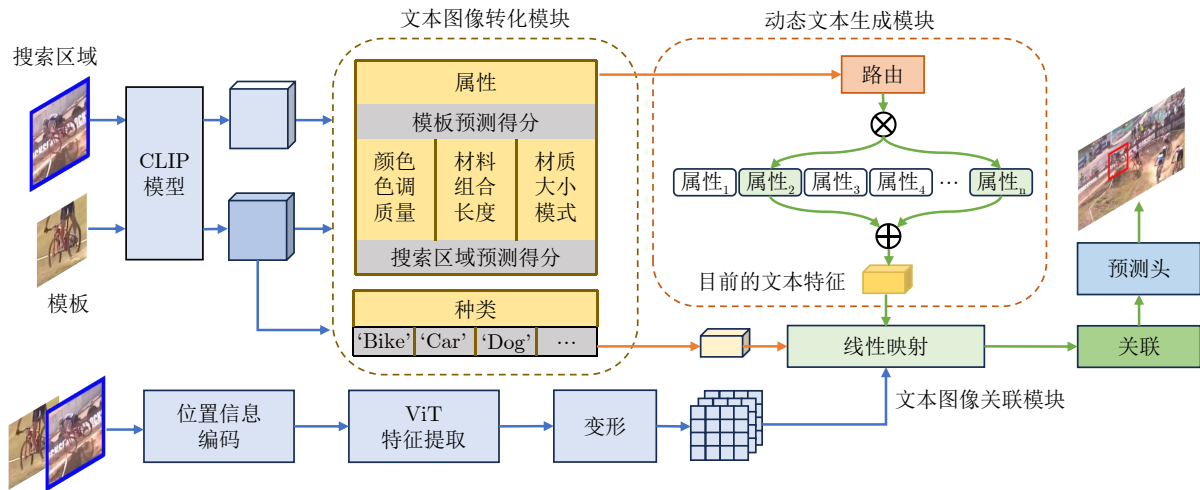


图9 基于图像与文本关联的视觉跟踪框架CiteTracker

Fig. 9 The proposed visual tracking framework CiteTracker

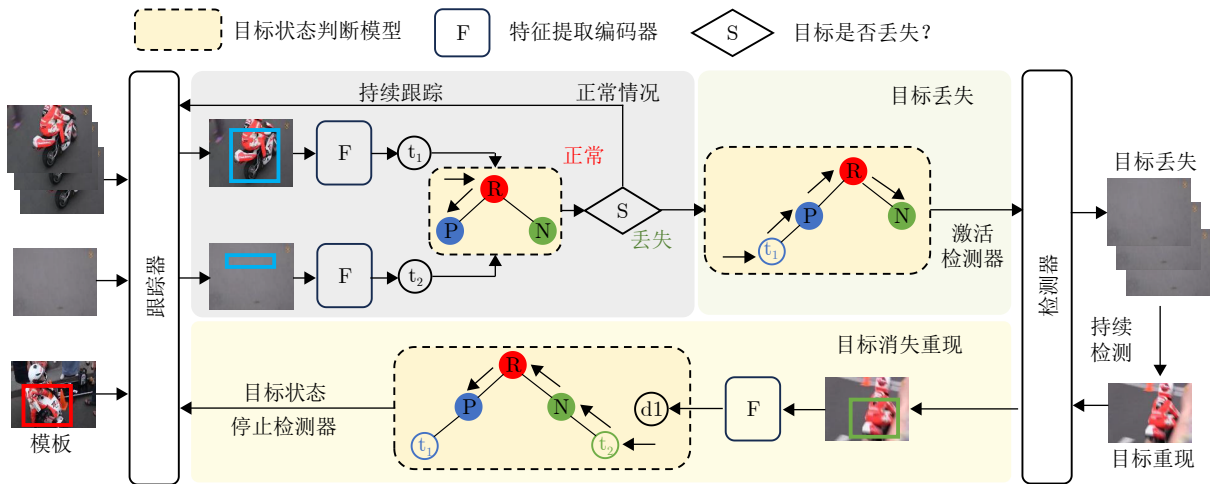


图10 可恢复跟踪框架Rtracker流程示意图

Fig. 10 The pipeline of the proposed recoverable tracking framework Rtracker

失”，系统将自动激活检测器进行全局搜索，待目标重现后将新特征存入树中并恢复跟踪。该框架结合“跟踪-检测”协同机制，实现了可靠的长时连续跟踪能力。CiteTracker与RTracker共同构成了本文的多模态感知跟踪体系，CiteTracker负责语义层面的抗干扰，RTracker负责状态层面的鲁棒恢复。

### 3.3.2 环境感知关键技术

低空环境具有高度的动态性与复杂性，气象条件的突变与电磁频谱的干扰直接影响低空飞行器的通信链路与飞行安全。针对端侧设备计算资源受限的特点，本研究重点突破了气象环境的高精度短临预报与电磁频谱态势的快速重构技术。

#### (1) 结合物理约束的气象感知

为了应对现有模型在气象感知中难以兼顾长程

时空关联与短程局部细节的问题，本文提出了一种基于张量的长短程卷积融合模型TLS-MWP<sup>[95]</sup>，突破了单维度特征拆分处理的局限，将气象数据的多维度特征以张量形式统一表征。通过设计长短程卷积模块，利用扩张卷积捕捉大气环流等大尺度规律，同时结合注意力机制提取局地突变特征，并通过张量融合层实现特征互补，显著降低了短期预测误差并提升了中期预测的时空一致性。相比于现有基于ConvLSTM或Transformer架构的气象感知方法(如SimVP<sup>[81]</sup>和TAU<sup>[96]</sup>等)，TLS-MWP在捕捉复杂气象演化方面更具优势(如表5所示)，其计算复杂度更低，高时空分辨率下的气象预测精准度更高。

#### (2) 物理信息启发与动态适配的电磁感知

面向端侧受限算力与稀疏观测条件下的频谱地图构建需求，相较于传统依赖密集采样或纯数据驱动的黑盒模型，本文研究聚焦物理先验引导与结构

表 5 TLS-MWP在“风速”这一气象要素上的预测结果对比  
Tab. 5 Comparison of prediction results on the “wind speed”

气象感知模型	MAE	MSE	RMSE
SimVP	1.03±0.02	2.20±0.08	1.46±0.03
LS-NTP	0.93±0.00	1.72±0.02	1.30±0.01
MMVP	1.17±0.01	2.72±0.03	1.65±0.01
TAU	1.02±0.00	2.15±0.02	1.45±0.01
所提TLS-MWP	<b>0.89±0.00</b>	<b>1.60±0.01</b>	<b>1.25±0.00</b>

注：表内加粗数值表示最优。

化建模在提升电磁环境感知精度与计算效率方面的关键作用。为提高端侧部署的频谱地图构建精度与效率，本文提出了基于物理传播模型增强的无线电深度展开网络RadioDUN<sup>[97]</sup>，该方法将电磁频谱态势生成问题转化为稀疏信号恢复任务，如式(10)所示：

$$\hat{x} = \underset{x}{\operatorname{argmin}} \frac{1}{2} \|\Phi x - y\|_2^2 + \lambda R(x), \quad (10)$$

其中， $\hat{x}$ 表示生成的频谱地图， $\Phi$ 表示采样矩阵， $x$ 表示理想的频谱地图， $y$ 表示采样信号， $\lambda$ 为缩放权重， $R(x)$ 为正则化系数。进一步，本文充分考虑频谱信号的物理先验，使用路径损耗模型对频谱信号进行表征，如式(11)所示：

$$x - I = 10\alpha \lg\left(\frac{d}{d_0}\right) + \eta + x_\delta \quad (11)$$

其中， $I$ 表示信源处信号强度， $d$ 表示信源与接收机之间的距离， $d_0$ 表示参考距离， $\eta$ 表示参考距离处的固定损耗， $x_\delta$ 表示阴影衰落因子，用于量化障碍物引起的随机信号衰落， $\alpha$ 为路径损耗指数。将式(10)与式(11)进行结合，并利用深度展开技术将迭代优化过程映射为可学习的层级网络结构，在显著减少推理复杂度的同时，保持了与传统优化方法一致的收敛结构与物理可解释性。

另一方面，针对低空动态环境中发射源位置变化频繁、传播路径高度非平稳等现实挑战，本文构建了DAT-UNet模型<sup>[98]</sup>，通过在U-Net架构中引入可变形注意力机制，突破了传统插值方法与固定感受野CNN在空间建模上的固有限制。该方法能够根据稀疏观测点分布与环境结构动态调整特征聚合位置，实现对观测点、遮挡结构与潜在传播路径之间关系的自适应建模。通过“动态适配稀疏观测-融合空间先验-多尺度特征建模”的协同策略，DAT-UNet在信号遮挡严重或反射路径快速变化的低空场景下，仍可自动聚焦高信息密度区域，实现对频谱态势的高精度重构与稳健感知，展现出相较于传统端侧感知方法在动态适应性与复杂场景泛化能力方面的显著提升。

综上所述，为了更清晰地展示从宏观架构到具体算法的映射关系，本文所提出的低空主动感知网方案的核心关键技术总结如表6所示。

### 3.4 低空主动感知网协同工作流程：以低空非合作目标感知为例

为了更直观地阐明上述各类基础模型与端侧感知技术在低空主动感知网中的作用机制，本节以“复杂环境下的低空非合作无人机感知”这一典型功能为例，详细解析“功能需求-传感器特性-算法模型”之间的映射与协同链路。在该功能场景下，整个低空主动感知网络的工作流可划分为环境预判、目标捕获与全局识别3个阶段：

(1) 环境态势预判与传感器调度(气象/频谱传感器→云侧基础模型)

功能需求：提前规避恶劣天气与强电磁干扰对光电/雷达探测的影响。

传感器与模型协同：云侧平台汇聚气象雷达与分布式频谱仪的数据，利用PercepCast气象模型进行高分辨率的局地降水预报，利用RadioFormer模型基于极稀疏采样重构低空电磁干扰拓扑图。云侧大脑基于预判结果，提前向受雨雾或电磁干扰区域的端侧设备下发指令，动态增加红外传感器的置信度权重。

(2) 低慢小目标的捕获与抗遮挡跟踪(雷达/光电传感器→端侧目标感知模型)

功能需求：在远距离、背景杂波强以及楼宇遮挡条件下，实现无人机的实时检测与不间断跟踪。

传感器与模型协同：在端侧，毫米波雷达首先提供具有穿透性的径向速度与距离引导，光电云台联动对准目标空域。针对无人机在画面中像素极少(“小”特性)，端侧调用Cross-DINO检测模型，利用其增强损失(Boost Loss)自适应放大微小目标的视觉响应。当无人机穿梭于楼宇间发生短暂遮挡时，传统像素匹配失效，端侧切换至CiteTracker与RTracker模型，将雷达点云的几何先验与视觉的跨模态语义属性(如“白色、四旋翼”)对齐，实现基于高维语义的抗遮挡寻回与轨迹维持。

(3) 低带宽下的目标精准识别与研判(特征流传输→云侧视觉感知基础模型)

功能需求：在空地带宽受限的前提下，利用云侧大算力对截获的无人机进行具体型号识别与全局轨迹关联。

传感器与模型协同：端侧感知节点不回传占用大量带宽的高清视频流，而是将提取的目标紧凑特征(特征流)上传至云侧。云侧接收到高维特征后，将其输入至具备全局大感受野和线性计算复杂度的

表6 低空主动感知网方案关键技术体系

Tab. 6 Key technological framework of the low-altitude active sensing network

层级	任务	代表性技术/模型	应对的低空挑战与核心机制
系统协同机制	架构与通信计算	三流协同机制	挑战: 低空空地带宽受限与通信拥塞。 机制: 单节点输出特征流、数据流与模型流, 实现“特征前置、数据按需、模型在线更新”。
		GRUO优化模型	挑战: 端云动态环境下的资源分配。 机制: 基于广义速率-效用优化理论, 自适应调节特征流/数据流的编码策略与带宽配比。
		统一时空基准	挑战: 多源异构传感器的置信度互补。 机制: 赋予异构数据带时间戳与地理坐标的唯一标识, 实现多模态物理层与特征层的精准对齐。
云侧感知基础模型	视觉目标感知	VMamba vHeat	挑战: 高分辨率宽幅图像带来的算力冗余与显存瓶颈。 机制: VMamba引入二维选择性扫描(SS2D)实现O(N)复杂度; vHeat基于热传导算子(HCO)在频域内进行自适应特征扩散。
	电磁环境感知	RadioFormer	挑战: 低空复杂环境下的电磁空间极度稀疏观测(如1%000采样率)。 机制: 采用双流自注意力与跨流交叉注意力, 实现建筑几何特征与稀疏观测信号的多粒度融合。
	气象环境感知	PercepCast	挑战: 低空突发强降水的精细化高分辨率预报。 机制: 结合流体力学物理先验, 引入MSE与感知双重约束, 保留强降水中心的锐利结构与时空一致性。
端侧目标与环境感知技术	低空目标检测	Strip-MLP Cross-DINO Open-Det	挑战: 远距离小目标像素少、易漏检及开放场景未知目标。 机制: 利用增强损失(Boost Loss)提升小目标响应; 通过脱离预定义词汇表的视觉-语言生成式框架实现开放词汇检测。
	目标定位与跟踪	HiVG OneRef CiteTracker RTracker	挑战: 目标遮挡、姿态剧变及弱纹理导致定位匹配漂移。 机制: 引入层级多模态低秩适应实现跨模态细粒度对齐; 通过“图像-文本”属性语义映射抗干扰, 并基于正负决策树实现丢失目标重捕获。
	气象与电磁感知	TLS-MWP RadioDUN DAT-UNet	挑战: 端侧算力受限下的短临气象预测与高动态电磁态势重构。 机制: 以张量融合长短期卷积捕获局地气象突变; 基于无线电深度展开网络与可变形注意力, 动态适配频谱遮挡与多径效应。

VMamba或vHeat视觉基础模型中。由于这些模型有效突破了宽幅高分辨率数据的算力瓶颈, 能够在海量特征库中迅速完成该“非合作”无人机的细粒度型号比对与跨区域时空轨迹拼接, 最终为管控平台提供完整的决策闭环。

通过上述典型功能流转可以看出, 低空主动感知网中的多类模型并非孤立存在, 而是深度结合了各类传感器的物理特性, 在数字视网膜“三流协同”架构下各司其职, 共同化解了低空场景中“探不到、跟不上、传不回”的痛点。

#### 4 低空主动感知网的初步实践

针对低空环境中通信带宽波动大、机载计算资源受限、复杂背景下无人机等小目标定位检测难度高的特点, 本节首先介绍了基于数字视网膜架构的低空主动感知网在空地协同感知计算任务中的应用实践, 通过建立了无人机边缘节点与地面云服务器之间的动态协同感知闭环, 实现了低空场景下计算资源与通信资源的联合优化<sup>[99]</sup>。随后, 介绍低空主动感知网利用数字视网膜三流协同机制, 关联视觉、红外、雷达等多源多模态感知数据, 实现高精度的多模态协同目标感知, 并通过仿真环境验证其在提升低空小目标感知精度方面的效用。

##### 4.1 空地协同感知计算问题建模

在低空安防巡检及灾害救援等典型应用场景中, 无人机作为空中边缘节点, 需对采集的高清视

频进行实时处理与目标感知。针对该问题, 采用云侧集中式处理范式, 海量高清视频回传将造成巨大的上行带宽压力, 且在低空复杂电磁环境下, 视频流传输的可靠性难以保障; 完全依赖前端分布式处理, 受限于无人机的载荷与能耗, 机载轻量化模型的推理精度通常低于云侧模型, 难以有效应对复杂多变的低空环境和灵活的任务需求。

针对上述挑战, 我们提出的低空主动感知网改变了传统单向数据传输模式, 依托数字视网膜的三流协同特性实现空地协同视觉感知计算: 特征流指无人机端侧模型提取紧凑的视觉特征, 上传至云侧, 在保留关键语义信息的同时显著降低上行通信开销; 数据流作为特征流的补充, 按需上传残差数据或关键帧, 供云侧模型进行高精度分析与挖掘; 模型流通过云侧模型对上传数据进行微调, 生成适配无人机端侧环境的模型参数更新量, 并通过下行链路推送至无人机, 实现端侧模型的实时在线更新。

具体来说, 在有限的空地无线通信资源约束下, 本文将端云协同感知计算过程建模为一个联合资源优化问题, 旨在最大化高清视频流中目标的平均检测精度(Mean Average Precision, mAP)。假设无人机采集帧率为 $N$ , 其中比例为 $\beta$ 的关键帧卸载至云侧模型处理, 剩余 $1-\beta$ 的帧在本地通过端侧模型处理。整体感知性能mAP由云侧模型检测精度 $mAP_L$ 、边侧模型检测精度 $mAP_S$ 以及任务卸载比例 $\beta$ 共同决定, 可近似表示为<sup>[99]</sup>

$$mAP \approx \frac{mAP_L \cdot mAP_S}{(1-\beta)mAP_L + \beta mAP_S} \quad (12)$$

其中，云侧模型检测精度 $mAP_L$ 取决于上行传输的数据质量。在低空主动感知网中，通过调节上传残差数据的量化比特数 $b$ ，在带宽消耗与识别精度之间寻求平衡。边侧模型检测精度 $mAP_S$ 取决于下行模型流的更新频率。云侧通过下发模型更新量 $M$ ，持续提升边侧模型的泛化能力。在通信资源层面，设系统总物理带宽为 $B$ ，上行带宽为 $B_u$ ，下行带宽为 $B_d$ ，则资源约束条件表示为

$$\begin{cases} R_{\text{Feature}} + R_{\text{Data}}(\beta, b) \leq B_u \cdot S_u \\ R_{\text{Model}}(M) \leq B_d \cdot S_d \\ B_u + B_d \leq B \end{cases} \quad (13)$$

其中， $R$ 代表各数据流的通信传输速率， $S_u$ 和 $S_d$ 分别为上行和下行链路的频谱效率。该数学模型刻画了低空场景下的资源博弈关系：即通过消耗上行带宽传输数据以利用云侧算力(提升 $mAP_L$ )，与消耗下行带宽下载参数以增强边缘推理能力(提升 $mAP_S$ )之间的权衡。

#### 4.2 感知驱动的资源自适应调度应用验证

针对上述建模中的计算与通信资源博弈问题，采用感知驱动的资源自适应调度策略<sup>[99]</sup>，通过实时探测信道状态信息，动态调整“特征-数据-模型”三流的带宽配比，实现空地协同感知的全局最优，仿真实验在MATLAB平台上运行。面向视觉分类任务，在公开低空无人机测试数据集VisDrone<sup>[100]</sup>进行了实验。实验设置端侧部署Yolov8s模型(模拟机载资源受限环境)，云侧则为Yolov8x2模型(模拟云侧高性能环境)，无线链路的上行有效信息传输频谱效率设为2.55 bit/(Hz·s)，下行有效信息传输频谱效率为5 bit/(Hz·s)，并在不同的无线传输带

宽(单位：MHz)条件下进行了测试分析。为保证对比的严谨性与可复现性，各类资源调度策略与对比基线方案均在相同的VisDrone测试集数据序列、一致的初始预训练模型权重(Yolov8s/Yolov8x2)及统一的评价标准下进行独立测试。此外，以上实验参数均为低空通信典型场景下的数值，在此参数设置下的仿真实验与真实场景差异较小，当真实场景存在参数值差异时，仍可采用本文所提方法进行一般化求解。

##### (1) 不同通信带宽下的感知精度

如图11所示，实验对比了不同总带宽 $B$ 下，基于数字视网膜的低空主动感知网方案与传统的“仅云侧处理”、“仅端侧处理”、“穷尽搜索”、“动态带宽均衡”方案的平均检测精度(mAP)。其中，“仅云侧处理”方案下，所有数据统一上传至中心云服务器进行处理；“仅端侧处理”方案下，由各个边缘节点进行独立数据处理；“穷尽搜索”方案下，资源分配设计是通过多种候选变量组合进行枚举来选择性能最优方案，作为近似最优解参考；“动态带宽均衡”方案下，上行和下行的带宽分配比例和实际上下行信道的频谱效率成反比关系，从而进行传输数据量均衡，其余的优化变量均与本文所提出方案相同。为保证性能对比的公平性与实验的可复现性，所提出方案与各对比方案的训练轮次均设为26次，学习率均为0.005，数据训练批量大小为4，验证批量大小为1，输入数据均采用VisDrone数据集中的相同样本序列。

此外，对于“穷尽搜索”方案，需要进一步说明的是：由于本节所研究的空地协同资源优化问题具有较高复杂性，其严格意义上的全局性能上限难以直接解析求得。因此，本文采用穷举搜索的方式对该性能上限进行数值逼近。由于真实性能上限未

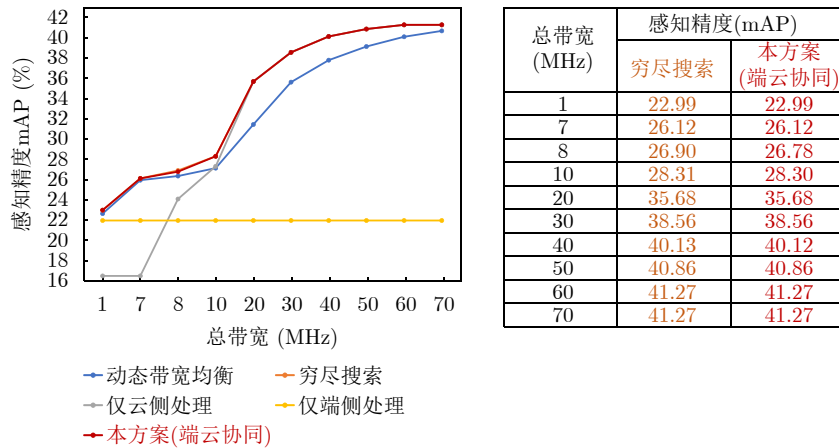


图 11 不同方案在不同带宽下的感知精度对比

Fig. 11 Comparison of different schemes under varying bandwidths

知,故无法精确量化穷举搜索结果与真实上限之间的误差。一般而言,穷举搜索的离散粒度越细,其所得结果越接近实际性能上限;但受限于计算复杂度与算力开销,搜索粒度无法无限细化。基于此,在本文的设定中,当进一步减小搜索步长后系统性能已无显著提升时,即认为当前穷举搜索结果已较为充分地逼近性能上限,并将其作为评估其他方案的近似最优解参考。

结果显示,随着传输带宽的增加,mAP性能也随之增加,并在总带宽超过60 MHz时收敛到一个稳定值。当带宽小于5 MHz时,所提出的架构的mAP与分布式边缘模型相当,此时由于数据传输能力有限,大部分任务在边缘模型上执行。当带宽大于20 MHz时,所提出架构与云侧模型处理的mAP接近,此时大部分数据被发送到云服务器进行分析。作为集中式云侧模型计算和分布式边缘模型计算的动态组合,所提出架构通过动态调整通信资源以获得最佳的任务分配与资源分配比例,其推理平均精度始终优于“仅云侧处理”、“仅端侧处理”和“动态带宽均衡”3类对比方案。实验结果显示,所提出方案与穷尽搜索之间的平均精度性能差距在所有情况下均在0~0.5%之间,证明所提出方案推理精度可以逼近理论最优上界。

(2)“特征-数据-模型”三流通信开销的动态调整

对不同带宽下各数据流的吞吐量占比分析显示(如图12),特征流由于采用了高效的特征提取技术,开销始终保持在较低水平。关键的资源分配发生在数据流与模型流之间:当带宽较小(如2 MHz以下)时,模型更新数据量占据主导地位,表明低空主动感知网采取了“以计算换通信”的策略,通过下行数据换取端侧智能的提升;当带宽超过10 MHz后,数据流的开销呈指数级增长并占据主导,这验

证了在宽带条件下,系统减少了对端侧模型更新的依赖,转而直接利用性能更强的云侧模型处理高保真数据,实现了空地计算资源的高效协同。

通过空地协同感知计算的初步应用实践,展示了低空主动感知网能够有效突破低空场景下通信带宽波动与端侧算力受限的双重瓶颈,是构建未来高可靠、高智能低空感知网的可行方案,为后续实现更复杂的多源多模态协同感知奠定了坚实基础。

### 4.3 复杂低空场景下的小目标定位与检测初步实践

低空感知网不仅需要解决受限带宽下的资源调度问题,更需应对低空小目标在复杂环境下的精准感知挑战。低空无人机等小目标通常具有雷达散射截面积小、红外特征弱、视觉纹理少等特点,尤其在远距离探测、夜间低照度或雾霾等视觉退化场景下,仅依赖单一模态传感器(如可见光相机)往往难以提取有效特征,导致虚警率高、目标定位漂移严重。针对这一难题,本文所提的低空主动感知网方案利用数字视网膜的“特征流”机制,支持多源异构数据的时空对齐与特征级融合,构建了高鲁棒性的多模态协同目标感知能力。本节将通过构建高逼真度多模态仿真环境,探讨并验证所提方案在复杂条件下的感知效能。

#### 4.3.1 仿真实验环境与低空小目标定义

受限于现有的空域管制法规、安全约束以及极端天气数据捕获的困难,在真实世界中大规模采集带有精准三维位姿及跨模态对齐标注的无人机数据面临极大的现实瓶颈。因此,本文依托虚幻引擎(Unreal Engine 4, UE4)与CARLA平台,构建了一套高逼真多模态低空感知仿真实验环境。该设置不仅能实现极端环境的可控复现,更为多模态融合提供了统一、精准的评估基准。

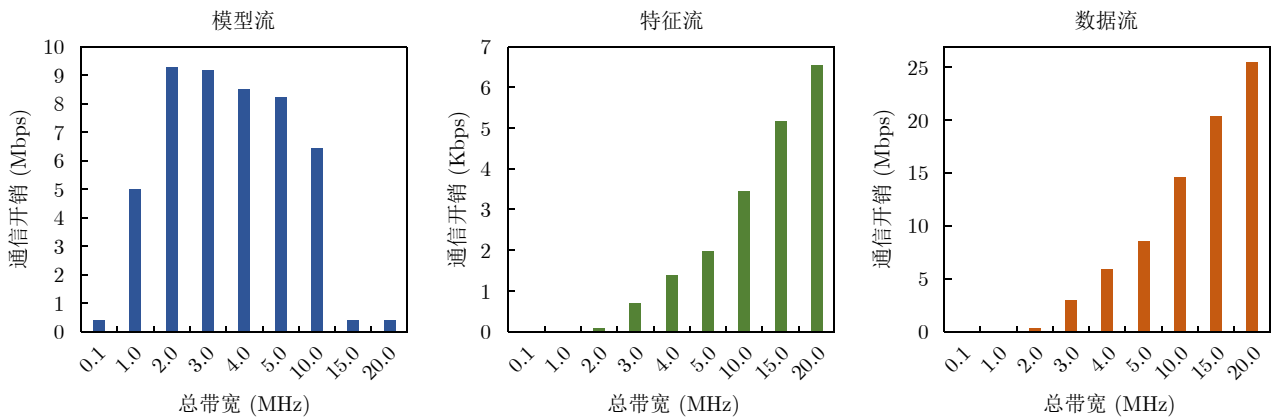


图 12 不同带宽下“三流”的通信开销

Fig. 12 Communication overhead of the “three flows” with varying bandwidths

(1) 低空小目标界定：本实验将物理尺寸较小、飞行机动性强且在观测设备中像素占比较低的目标界定为“低空小目标”。具体涵盖了从微型无人机(如折叠尺寸130×70×55 mm、质量0.12 kg)到中型无人机(尺寸1668×1518×727 mm、质量10.5 kg)等7种典型真实无人机模型，这些无人机目标(如大疆Mavic Mini、Phantom 4等)的选取具备良好的工程代表性，它们不仅是当前低空经济中普及率最高的飞行器，也是安防监控与设施巡检等典型应用的核心监测对象。在20米至100米以上的典型观测距离下，这些无人机目标雷达散射截面积小、热辐射与视觉纹理微弱，在复杂背景或恶劣天气下极易发生漏检与跟丢，充分代表了当前低空主动感知任务中最具挑战性的典型难点。

(2) 多模态传感器与环境配置：仿真系统在相同物理节点集成了5种异构传感器流：可见光(RGB)、红外(IR)和动态视觉传感器(Dynamic Vision Sensor, DVS)(分辨率均为1280×720，视场角90°)，以及256线激光雷达(15 Hz采样率)和毫米波雷达(有效探测距离30 m)。通过全局时间戳(world-tick)实现各模态的帧级严格同步。实验场景包含城市与郊区等多种建筑密度的拓扑环境，并覆盖了晴、雨、雾、雪与昼夜交替等8种基于物理模型渲染的复杂气象与光照条件。目标轨迹采用参数化椭圆运动模型生成。每条轨迹由均匀分布在该参数曲线上的15~25个离散航路点组成。

$$x = a \cos(\theta), y = b \sin(\theta), z = k_0 + \Delta k(\theta) \quad (14)$$

其中， $\theta$ 为曲线的角度参数， $k_0$ 为基础高度，参数 $a$ 和 $b$ 分别控制椭圆轨迹的纵向和横向跨度， $\Delta k(\theta)$ 则在预设范围内引入随机的高度波动。此外，系统对不同无人机的高度和速度曲线进行了随机化配置，从而准确模拟无人机多样的加速、减速与转向行为。

### 4.3.2 多模态目标特征空间对齐与融合机制

为了解决单一模态在恶劣天气或远距离探测中失效的问题，低空主动感知网在端侧计算节点建立了跨模态空间对齐与特征融合机制。针对异构传感器数据形式不一致的难点，我们采用了一种基于激光雷达几何引导的双分支处理架构。

具体而言，网络分别设立可见光(RGB)与红外(IR)双分支提取外观与热辐射特征。为了在视角不同的双分支间建立严格的几何一致性，系统将激光雷达的三维点云作为几何锚点。对于任意空间点 $\mathbf{P}_k = (X_k Y_k Z_k)^T$ ，通过相机内参 $K_c$ 和外参 $[R_c | t_c]$ 投影至不同传感器平面：

$$p_{c,k} \sim K_c [R_c | t_c] \mathbf{P}_k \quad (15)$$

其中， $c \in \text{RGB, IR}$ 。通过该投影模型，在RGB特征图与IR特征图之间寻找空间对应关系，构建包含深度与几何结构的先验特征图 $\mathbf{D}$ 。随后，在特征提取阶段，将视觉模态的外观特征 $\mathbf{F}_{\text{vis}}$ 与几何模态的结构特征 $\mathbf{F}_{\text{geo}}$ 进行深度融合。该融合过程可表示为

$$\mathbf{F} = \phi(\mathbf{F}_{\text{vis}}(I^{\text{RGB}}, I^{\text{IR}}), \mathbf{F}_{\text{geo}}(\mathbf{D})), \quad (16)$$

其中， $\phi$ 为特征融合算子， $I^{\text{RGB}}, I^{\text{IR}}$ 分别为可见光RGB与红外IR的图像，通过注意力机制动态加权不同模态的特征响应。基于融合后的增强特征 $\mathbf{F}$ ，感知网络能够进一步预测目标的类别概率及其在三维空间中的精确位置。

这种多模态协同机制利用“特征流”在不同的端侧传感器之间建立了互补关系，利用雷达/点云的测距精度弥补视觉在深度感知上的不足，利用红外的热特征弥补可见光在夜间或伪装场景下的失效。

### 4.3.3 复杂场景下的目标感知性能验证

为了量化验证上述多模态主动感知方案的有效性，我们在所构建的仿真环境下进行了对比实验。为保证对比实验的公平性与结果的严谨性，本节中所有参与评估的方法均在相同的软硬件环境下运行，并采用严格一致的输入设置与训练协议。具体而言，各对比方法的视觉输入数据均统一缩放至1280×720分辨率，并应用了相同的数据增强策略；在训练机制上，所有方法均使用Adam优化器，初始学习率设为0.001，并采用余弦衰减(Cosine Decay)策略，最大训练轮次统一设定为50个Epochs。所有方法均采用不同随机数种子开展3次独立重复实验，最终取平均值作为结果。评估指标涵盖了反映定位精度的位置误差(Position Error)、旋转误差(Rotation Error)、尺寸误差(Size Error)，以及综合检测性能的3D平均精度(3D AP)。参与对比的方法包括基于几何推断的检测器YOLO-6D<sup>[101]</sup>、基于中心回归的单模态基线CenterNet<sup>[102]</sup>以及本文所提的激光雷达引导多模态融合方案。

表7展示了不同方案在复杂环境下的量化性能对比结果。可以看出，简单的“决策级得分平均”(即单模态结果后融合)所带来的提升极为有限，这表明缺乏特征级空间对齐难以真正发挥多源数据的互补优势。而采用本方案的三维空间对齐与融合后，各项精度指标均实现了显著跃升：相较于性能较优的中心回归基线(Late Fusion配置)，本方案将低空目标的三维位置定位误差降低了约26%(从4.92 m降至3.64 m)，旋转误差降低了约40%(从

17.84°降至10.57°),同时3D AP从17.38大幅提升至26.72。这一量化结果证实,基于几何先验的特征级融合机制能有效校正远距离小目标的深度估算偏差。

进一步的可视化结果(图13)直观印证了上述量化分析。在白天的远距离观测中,单一视觉模型常因小目标像素过少而产生严重的定位框漂移与尺寸收缩。引入特征融合后,目标特征响应集中度大幅提高。而在夜间与雾天场景下,可见光图像信噪比急剧恶化,单模态模型极易出现漏检与虚警;得益于对红外热辐射与雷达深度点云的有效空间对齐,本系统在严重视觉退化环境中依然维持了紧凑、稳定的三维包围框追踪,展现出在复杂飞航场景下卓越的环境鲁棒性。

尽管高逼真仿真环境提供了可控的极限测试条件与精准的三维位姿基准,但当前场景与实际低空物理环境仍存在一定的差距。具体而言,物理引擎难以完美复现城市峡谷中复杂的电磁多径效应、动态频谱干扰,以及真实恶劣天气(如非均匀团雾、

突发阵风)带来的极端非平稳随机性。面向未来在真实低空场景的规模化部署,本方案的演进路径将重点聚焦于跨域自适应技术的探索。后续研究将利用在仿真环境预训练的基础感知模型,结合真实复杂天气下采集的实测数据进行知识迁移与微调,以有效弥合虚实数据间的分布差异,逐步构建由纯数据驱动向物理环境自适应演进的混合感知网络。

## 5 结语

本文提出了一种基于数字视网膜端边云协同架构的低空主动感知网方案,通过“特征流、数据流、模型流”的三流协同机制,有效突破了低空场景下通信带宽受限、端侧算力不足以及低空小目标难以精确定位等瓶颈。本方案在多项关键性能指标上实现了显著提升:在空地通信与计算协同方面,特征流机制将上传数据量压缩至原始视频的1%以下,且在动态带宽受限条件下,端云协同感知精度逼近理论最优上界;在低空目标感知方面,面对复杂气象与远距离观测条件,基于几何引导的多模态

表 7 仿真环境下不同目标感知方案的6-DoF位姿与检测性能对比

Tab. 7 Comparison of different object perception schemes in a simulated environment

方法	模态与融合策略	旋转误差(°)	位置误差(m)	尺寸误差(m)	三维目标检测精度(AP)
YOLO-6D	RGB-only(单模态)	26.47	7.21	0.43	11.78
YOLO-6D	IR-only(单模态)	24.15	6.83	0.41	12.26
YOLO-6D	决策级融合(Late Fusion)	22.73	5.95	0.40	13.89
Center-based	RGB-only(单模态)	21.65	6.48	0.42	13.07
Center-based	IR-only(单模态)	19.29	5.76	0.40	14.64
Center-based	决策级融合(Late Fusion)	17.84	4.92	0.39	17.38
本文所提出的低空感知网方案	几何引导特征融合	<b>10.57</b>	<b>3.64</b>	<b>0.35</b>	<b>26.72</b>

注:加粗数值表示最优。

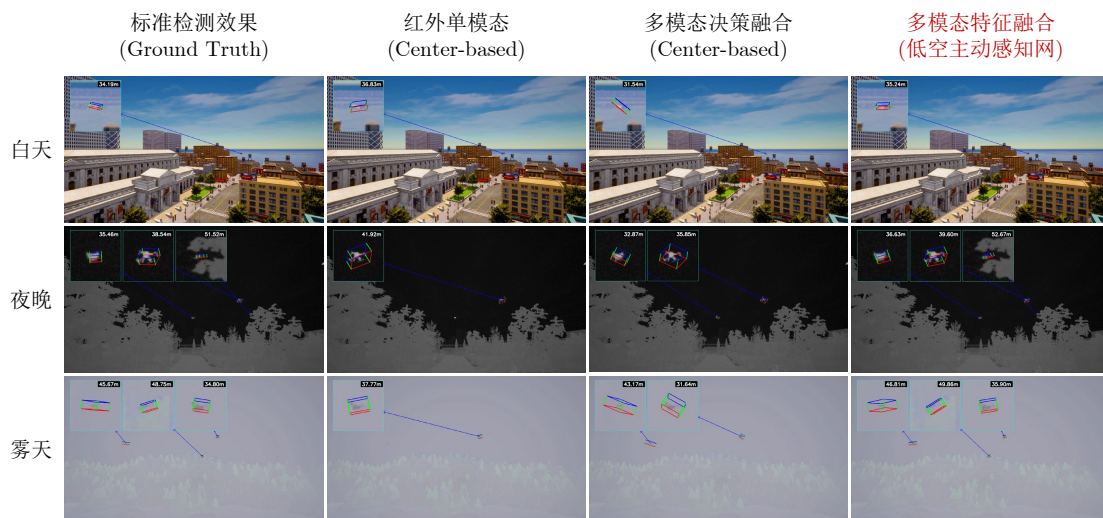


图 13 不同气象与光照条件下的低空小目标感知效果可视化对比

Fig. 13 Visual comparison of low-altitude small target perception performance with varying meteorological and illumination conditions

特征融合机制将低空小目标的三维位置误差降低约 26%，旋转误差降低约 40%，3D 平均检测精度由 17.38 跃升至 26.72；在云侧感知基础模型方面，新型视觉感知模型在保持高精度的同时降低了近 80% 的显存占用，电磁感知模型在空间采样率低至万分之一的极端稀疏条件下仍能实现电磁态势的高精度重构。

未来，低空感知仍然面临多源多模态数据融合的挑战，具体包括视场异构、尺度差异及数据不同步等。亟需进一步突破多视角特征建模与跨模态语义对齐等关键技术，在多智能体架构下建立分布式的协同感知与任务分配机制，支撑异构智能体间的信息共享与协作决策，以持续提升低空感知的协同效率和整体感知效能。

**利益冲突** 所有作者均声明不存在利益冲突

**Conflict of Interests** The authors declare that there is no conflict of interests

### 参 考 文 献

- [1] TANG Hualong, ZHANG Yu, MOHMOODIAN V, *et al.* Automated flight planning of high-density urban air mobility[J]. *Transportation Research Part C: Emerging Technologies*, 2021, 131: 103324. doi: [10.1016/j.trc.2021.103324](https://doi.org/10.1016/j.trc.2021.103324).
- [2] 蒲钊, 陈志杰, 刘杨, 等. 数字低空融合运行空中交通管理技术[J]. *航空学报*, 2025, 46(11): 531331. doi: [10.7527/S1000-6893.2025.31331](https://doi.org/10.7527/S1000-6893.2025.31331).  
PU Fan, CHEN Zhijie, LIU Yang, *et al.* Air traffic management technologies for digital low-altitude integrated operations[J]. *Acta Aeronautica et Astronautica Sinica*, 2025, 46(11): 531331. doi: [10.7527/S1000-6893.2025.31331](https://doi.org/10.7527/S1000-6893.2025.31331).
- [3] VITIELLO F, CAUSA F, OPROMOLLA R, *et al.* Radar/visual fusion with fuse-before-track strategy for low altitude non-cooperative sense and avoid[J]. *Aerospace Science and Technology*, 2024, 146: 108946. doi: [10.1016/j.ast.2024.108946](https://doi.org/10.1016/j.ast.2024.108946).
- [4] CHAN Y Y, NG K K H, LEE C K M, *et al.* Wind dynamic and energy-efficiency path planning for unmanned aerial vehicles in the lower-level airspace and urban air mobility context[J]. *Sustainable Energy Technologies and Assessments*, 2023, 57: 103202. doi: [10.1016/j.seta.2023.103202](https://doi.org/10.1016/j.seta.2023.103202).
- [5] GAO Wen, MA Siwei, DUAN Lingyu, *et al.* Digital retina: A way to make the city brain more efficient by visual coding[J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2021, 31(11): 4147–4161. doi: [10.1109/TCSVT.2021.3104305](https://doi.org/10.1109/TCSVT.2021.3104305).
- [6] 高文, 田永鸿, 王坚. 数字视网膜: 智慧城市系统演进的关键环节[J]. *中国科学: 信息科学*, 2018, 48(8): 1076–1082. doi: [10.1360/N112018-00025](https://doi.org/10.1360/N112018-00025).
- [7] GAO Wen, TIAN Yonghong, and WANG Jian. Digital retina: Revolutionizing camera systems for the smart city[J]. *Scientia Sinica Informationis*, 2018, 48(8): 1076–1082. doi: [10.1360/N112018-00025](https://doi.org/10.1360/N112018-00025).  
易伟, 袁野, 刘光宏, 等. 多雷达协同探测技术研究进展: 认知跟踪与资源调度算法[J]. *雷达学报*, 2023, 12(3): 471–499. doi: [10.12000/JR23036](https://doi.org/10.12000/JR23036).  
YI Wei, YUAN Ye, LIU Guanghong, *et al.* Recent advances in multi-radar collaborative surveillance: Cognitive tracking and resource scheduling algorithms[J]. *Journal of Radars*, 2023, 12(3): 471–499. doi: [10.12000/JR23036](https://doi.org/10.12000/JR23036).
- [8] 葛建军, 唐思琦, 李明强, 等. 复杂感知系统信息理论与构建方法[J]. *雷达学报(中英文)*, 2025, 14(3): 651–663. doi: [10.12000/JR25078](https://doi.org/10.12000/JR25078).  
GE Jianjun, TANG Siqi, LI Mingqiang, *et al.* Information theory and construction methods of complex perception systems[J]. *Journal of Radars*, 2025, 14(3): 651–663. doi: [10.12000/JR25078](https://doi.org/10.12000/JR25078).
- [9] 宋晓程, 李陟, 任海伟, 等. 目标动态威胁度驱动的分布式组网相控阵雷达资源优化分配算法[J]. *雷达学报*, 2023, 12(3): 629–641. doi: [10.12000/JR22240](https://doi.org/10.12000/JR22240).  
SONG Xiaocheng, LI Zhi, REN Haiwei, *et al.* Threat-driven resource allocation algorithm for distributed netted phased array radars[J]. *Journal of Radars*, 2023, 12(3): 629–641. doi: [10.12000/JR22240](https://doi.org/10.12000/JR22240).
- [10] ANANTHANARAYANAN G, BAHL P, BODÍK P, *et al.* Real-time video analytics: The killer app for edge computing[J]. *Computer*, 2017, 50(10): 58–67. doi: [10.1109/MC.2017.3641638](https://doi.org/10.1109/MC.2017.3641638).
- [11] NAN Ya, JIANG Shiqi, and LI Mo. Large-scale video analytics with cloud-edge collaborative continuous learning[J]. *ACM Transactions on Sensor Networks*, 2024, 20(1): 14. doi: [10.1145/3624478](https://doi.org/10.1145/3624478).
- [12] ZHOU Zhiqing, YU Heng, and SHI Hesheng. Optimization of wireless video surveillance system for smart campus based on internet of things[J]. *IEEE Access*, 2020, 8: 136434–136448. doi: [10.1109/ACCESS.2020.3011951](https://doi.org/10.1109/ACCESS.2020.3011951).
- [13] AKLAMATI D, ABDUS-SHAKUR B, and KACEM T. Security analysis of AWS-based video surveillance systems[C]. 2021 International Conference on Engineering and Emerging Technologies, Istanbul, Turkey, 2021: 1–6. doi: [10.1109/ICEET53442.2021.9659574](https://doi.org/10.1109/ICEET53442.2021.9659574).
- [14] LI Rongheng, ZHANG Jian, and SHEN Wenfeng. Replicas strategy and cache optimization of video surveillance systems based on cloud storage[J]. *Future Internet*, 2018, 10(4): 34. doi: [10.3390/fi10040034](https://doi.org/10.3390/fi10040034).
- [15] SUBUDHI B N, ROUT D K, and GHOSH A. Big data analytics for video surveillance[J]. *Multimedia Tools and*

- Applications*, 2019, 78(18): 26129–26162. doi: [10.1007/s11042-019-07793-w](https://doi.org/10.1007/s11042-019-07793-w).
- [16] DO T T T, HUYNH Q T, KIM K, *et al.* A survey on video big data analytics: Architecture, technologies, and open research challenges[J]. *Applied Sciences*, 2025, 15(14): 8089. doi: [10.3390/app15148089](https://doi.org/10.3390/app15148089).
- [17] 杨铮, 贺晓武, 吴家行, 等. 面向实时视频流分析的边缘计算技术[J]. *中国科学: 信息科学*, 2022, 52(1): 1–53. doi: [10.1360/SSI-2021-0133](https://doi.org/10.1360/SSI-2021-0133).  
YANG Zheng, HE Xiaowu, WU Jiahang, *et al.* Edge computing technologies for streaming video analytics[J]. *Scientia sinica Informationis*, 2022, 52(1): 1–53. doi: [10.1360/SSI-2021-0133](https://doi.org/10.1360/SSI-2021-0133).
- [18] AXIS Communications. The history of ARTPEC, the foundation of our product quality[EB/OL]. <https://newsroom.axis.com/en-sg/article/artpec-foundation-quality>, 2024.
- [19] 王秉路, 靳杨, 张磊, 等. 基于多传感器融合的协同感知方法[J]. *雷达学报(中英文)*, 2024, 13(1): 87–96. doi: [10.12000/JR23184](https://doi.org/10.12000/JR23184).  
WANG Binglu, JIN Yang, ZHANG Lei, *et al.* Collaborative perception method based on multisensor fusion[J]. *Journal of Radars*, 2024, 13(1): 87–96. doi: [10.12000/JR23184](https://doi.org/10.12000/JR23184).
- [20] JAIN S, ANANTHANARAYANAN G, JIANG Junchen, *et al.* Scaling video analytics systems to large camera deployments[C]. The 20th International Workshop on Mobile Computing Systems and Applications, Santa Cruz, USA, 2019: 9–14. doi: [10.1145/3301293.3302366](https://doi.org/10.1145/3301293.3302366).
- [21] NEFF C, MENDIETA M, MOHAN S, *et al.* REVAMP<sup>2</sup>T: Real-time edge video analytics for multicamera privacy-aware pedestrian tracking[J]. *IEEE Internet of Things Journal*, 2020, 7(4): 2591–2602. doi: [10.1109/JIOT.2019.2954804](https://doi.org/10.1109/JIOT.2019.2954804).
- [22] WANG Bingshu, LI Qiang, MAO Qianchen, *et al.* A survey on vision-based anti unmanned aerial vehicles methods[J]. *Drones*, 2024, 8(9): 518. doi: [10.3390/drones8090518](https://doi.org/10.3390/drones8090518).
- [23] ZHAO Jie, ZHANG Jingshu, LI Dongdong, *et al.* Vision-based anti-UAV detection and tracking[J]. *IEEE Transactions on Intelligent Transportation Systems*, 2022, 23(12): 25323–25334. doi: [10.1109/TITS.2022.3177627](https://doi.org/10.1109/TITS.2022.3177627).
- [24] 冯卫强, 李哲, 周强, 等. LSS-YOLO: 面向低慢小无人机的目标检测方法[J]. *火炮发射与控制学报*, 2025, 46(6): 17–25. doi: [10.19323/j.issn.1673-6524.202506001](https://doi.org/10.19323/j.issn.1673-6524.202506001).  
FENG Weiqiang, LI Zhe, ZHOU Qiang, *et al.* LSS-YOLO: A target detection method for low-slow-small UAVs[J]. *Journal of Gun Launch & Control*, 2025, 46(6): 17–25. doi: [10.19323/j.issn.1673-6524.202506001](https://doi.org/10.19323/j.issn.1673-6524.202506001).
- [25] 王迎龙, 孙备, 丁冰, 等. BG-YOLO: 复杂大视场下低慢小无人机目标检测方法[J]. *仪器仪表学报*, 2025, 46(2): 255–266. doi: [10.19650/j.cnki.cjsi.J2413551](https://doi.org/10.19650/j.cnki.cjsi.J2413551).  
WANG Yinglong, SUN Bei, DING Bing, *et al.* BG-YOLO: A low-altitude slow-moving small UAV targets detection method in complex large field of view[J]. *Chinese Journal of Scientific Instrument*, 2025, 46(2): 255–266. doi: [10.19650/j.cnki.cjsi.J2413551](https://doi.org/10.19650/j.cnki.cjsi.J2413551).
- [26] JIANG Nan, WANG Kuiran, PENG Xiaoke, *et al.* Anti-UAV: A large-scale benchmark for vision-based UAV tracking[J]. *IEEE Transactions on Multimedia*, 2023, 25: 486–500. doi: [10.1109/TMM.2021.3128047](https://doi.org/10.1109/TMM.2021.3128047).
- [27] ZHANG Tianlu, GUO Hongyuan, JIAO Qiang, *et al.* Efficient RGB-T tracking via cross-modality distillation[C]. The IEEE/CVF Conference on Computer Vision and Pattern Recognition, Vancouver, Canada, 2023: 5404–5413. doi: [10.1109/CVPR52729.2023.00523](https://doi.org/10.1109/CVPR52729.2023.00523).
- [28] HUI Tianrui, XUN Zizheng, PENG Fengguang, *et al.* Bridging search region interaction with template for RGB-T tracking[C]. The IEEE/CVF Conference on Computer Vision and Pattern Recognition, Vancouver, Canada, 2023: 13630–13639. doi: [10.1109/CVPR52729.2023.01310](https://doi.org/10.1109/CVPR52729.2023.01310).
- [29] 申炜豪, 马浩统, 樊红星, 等. 复杂场景下多目标高精度探测技术研究进展[J/OL]. *激光与光电子学进展*. <https://link.cnki.net/urlid/31.1690.TN.20251027.2254.116>, 2025.  
SHEN Weihao, MA Haotong, FAN Hongxing, *et al.* Research progress on multi-target high-accuracy detection technology in complex scenarios[J/OL]. *Laser & Optoelectronics Progress*. <https://link.cnki.net/urlid/31.1690.TN.20251027.2254.116>, 2025.
- [30] ROHLING H. Radar CFAR thresholding in clutter and multiple target situations[J]. *IEEE Transactions on Aerospace and Electronic Systems*, 1983, AES-19(4): 608–621. doi: [10.1109/TAES.1983.309350](https://doi.org/10.1109/TAES.1983.309350).
- [31] BAR-SHALOM Y, LI X R, and KIRUBARAJAN T. Estimation with Applications to Tracking and Navigation: Theory Algorithms and Software[M]. New York: John Wiley & Sons, Inc., 2001. doi: [10.1002/0471221279](https://doi.org/10.1002/0471221279).
- [32] ŚLESICK B, ŚLESICKA A, KAWALEC A, *et al.* Improving recognition of road users via Doppler radar data and deep learning convolutional networks[J]. *Electronics*, 2024, 13(20): 4070. doi: [10.3390/electronics13204070](https://doi.org/10.3390/electronics13204070).
- [33] LUO Xiaoyu and LI Qiusheng. Human motion recognition based on feature fusion and residual networks[J]. *Scientific Reports*, 2024, 14(1): 29097. doi: [10.1038/s41598-024-80783-7](https://doi.org/10.1038/s41598-024-80783-7).
- [34] 陈小龙, 陈唯实, 饶云华, 等. 飞鸟与无人机目标雷达探测与识别技术进展与展望[J]. *雷达学报*, 2020, 9(5): 803–827. doi: [10.12000/JR20068](https://doi.org/10.12000/JR20068).  
CHEN Xiaolong, CHEN Weishi, RAO Yunhua, *et al.* Progress and prospects of radar target detection and recognition technology for flying birds and unmanned aerial vehicles[J]. *Journal of Radars*, 2020, 9(5): 803–827.

- doi: [10.12000/JR20068](https://doi.org/10.12000/JR20068).
- [35] 陈小龙, 袁旺, 杜晓林, 等. 多波段多角度FMCW雷达低慢小探测数据集(LSS-FMCWR-2.0)及特征融合分类方法[J]. 雷达学报(中英文), 2025, 14(5): 1276–1293. doi: [10.12000/JR25004](https://doi.org/10.12000/JR25004).  
CHEN Xiaolong, YUAN Wang, DU Xiaolin, *et al.* Multi-band multi-angle FMCW radar low-slow-small target detection dataset (LSS-FMCWR-2.0) and feature fusion classification methods[J]. *Journal of Radars*, 2025, 14(5): 1276–1293. doi: [10.12000/JR25004](https://doi.org/10.12000/JR25004).
- [36] 陈小龙, 饶桂林, 关键, 等. 被动雷达低慢小探测数据集(LSS-PR-1.0)及多域特征提取和分析方法[J]. 雷达学报(中英文), 2025, 14(2): 249–268. doi: [10.12000/JR24145](https://doi.org/10.12000/JR24145).  
CHEN Xiaolong, RAO Guilin, GUAN Jian, *et al.* Passive radar low slow small detection dataset (LSS-PR-1.0) and multi-domain feature extraction and analysis methods[J]. *Journal of Radars*, 2025, 14(2): 249–268. doi: [10.12000/JR24145](https://doi.org/10.12000/JR24145).
- [37] 肖振, 谷延锋, 蒋彦泽, 等. 时空耦合模型驱动的激光雷达多目标回波轻量化检测算法[J]. 雷达学报(中英文), 2025, 14(3): 548–561. doi: [10.12000/JR24245](https://doi.org/10.12000/JR24245).  
XIAO Zhen, GU Yanfeng, JIANG Yanze, *et al.* Full-waveform small-footprint LiDAR multi-target echo waveform lightweight detection by spatio-temporal coupling models[J]. *Journal of Radars*, 2025, 14(3): 548–561. doi: [10.12000/JR24245](https://doi.org/10.12000/JR24245).
- [38] 刘斌越, 杨建强, 徐波, 等. 5G-A通感一体基站组网低空感知关键技术[J]. 信号处理, 2025, 41(5): 787–806. doi: [10.12466/xhcl.2025.05.002](https://doi.org/10.12466/xhcl.2025.05.002).  
LIU Binyue, YANG Jianqiang, XU Bo, *et al.* Key technologies for low-altitude sensing in 5G-A integrated communication and sensing networks[J]. *Journal of Signal Processing*, 2025, 41(5): 787–806. doi: [10.12466/xhcl.2025.05.002](https://doi.org/10.12466/xhcl.2025.05.002).
- [39] 唐爱民, 王书涵, 曲文泽. 面向远距离高速无人机检测的OFDM通信感知一体化参考信号设计[J]. 雷达学报(中英文), 2025, 14(4): 842–853. doi: [10.12000/JR24240](https://doi.org/10.12000/JR24240).  
TANG Aimin, WANG Shuhan, and QU Wenze. Reference signal design in OFDM ISAC for long-range and high-speed UAV detection[J]. *Journal of Radars*, 2025, 14(4): 842–853. doi: [10.12000/JR24240](https://doi.org/10.12000/JR24240).
- [40] 陈辉, 杜双燕, 连峰, 等. Track-MT3: 一种基于Transformer的新型多目标跟踪算法[J]. 雷达学报(中英文), 2024, 13(6): 1202–1219. doi: [10.12000/JR24164](https://doi.org/10.12000/JR24164).  
CHEN Hui, DU Shuangyan, LIAN Feng, *et al.* Track-MT3: A novel multitarget tracking algorithm based on transformer network[J]. *Journal of Radars*, 2024, 13(6): 1202–1219. doi: [10.12000/JR24164](https://doi.org/10.12000/JR24164).
- [41] NIE Wei, HAN Zhichao, LI Yi, *et al.* UAV detection and localization based on multi-dimensional signal features[J]. *IEEE Sensors Journal*, 2022, 22(6): 5150–5162. doi: [10.1109/JSEN.2021.3105229](https://doi.org/10.1109/JSEN.2021.3105229).
- [42] 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. doi: [10.1016/j.future.2019.05.007](https://doi.org/10.1016/j.future.2019.05.007).
- [43] XU Jingren, LI Zhen, ZHANG Kai, *et al.* The principle, methods and recent progress in RFID positioning techniques: A review[J]. *IEEE Journal of Radio Frequency Identification*, 2023, 7: 50–63. doi: [10.1109/JRFID.2022.3233855](https://doi.org/10.1109/JRFID.2022.3233855).
- [44] CHEN Huawei and ZHAO Junwei. On locating low altitude moving targets using a planar acoustic sensor array[J]. *Applied Acoustics*, 2003, 64(11): 1087–1101. doi: [10.1016/S0003-682X\(03\)00073-2](https://doi.org/10.1016/S0003-682X(03)00073-2).
- [45] TONG Jianfei, XIE Wei, HU Yuheng, *et al.* Estimation of low-altitude moving target trajectory using single acoustic array[J]. *The Journal of the Acoustical Society of America*, 2016, 139(4): 1848–1858. doi: [10.1121/1.4944567](https://doi.org/10.1121/1.4944567).
- [46] SONG Chunhuan and LI Hanshan. An acoustic array sensor signal recognition algorithm for low-altitude targets using multiple five-element acoustic positioning systems with VMD[J]. *Applied Sciences*, 2024, 14(3): 1075. doi: [10.3390/app14031075](https://doi.org/10.3390/app14031075).
- [47] LIU Miao, YU Jiyan, and YANG Zhengpeng. Research on the improvement of the signal time delay estimation method of acoustic positioning for anti-low altitude UAVs[J]. *Sensors*, 2025, 25(9): 2735. doi: [10.3390/s25092735](https://doi.org/10.3390/s25092735).
- [48] YANG Bowon, MATSON E T, SMITH A H, *et al.* UAV detection system with multiple acoustic nodes using machine learning models[C]. The Third IEEE International Conference on Robotic Computing, Naples, Italy, 2019: 493–498. doi: [10.1109/IRC.2019.00103](https://doi.org/10.1109/IRC.2019.00103).
- [49] MIĘSIKOWSKA M. Classification of unmanned aerial vehicles based on acoustic signals obtained in external environmental conditions[J]. *Sensors*, 2024, 24(17): 5663. doi: [10.3390/s24175663](https://doi.org/10.3390/s24175663).
- [50] PASZKOWSKI W, GOLA A, and SWIC A. Acoustic-based drone detection using neural networks—a comprehensive analysis[J]. *Advances in Science and Technology Research Journal*, 2024, 18(1): 36–47. doi: [10.12913/22998624/175863](https://doi.org/10.12913/22998624/175863).
- [51] NAJAFI J, MIRZAKUCHAKI S, and SHAMAGHDARI S. Autonomous drone detection and classification using computer vision and Prony algorithm-based frequency feature extraction[J]. *Journal of Intelligent & Robotic Systems*, 2025, 111(1): 8. doi: [10.1007/s10846-024-02216-x](https://doi.org/10.1007/s10846-024-02216-x).
- [52] DING Siyi, GUO Xiao, PENG Ti, *et al.* Drone detection and tracking system based on fused acoustical and optical approaches[J]. *Advanced Intelligent Systems*, 2023, 5(10):

2300251. doi: [10.1002/aisy.202300251](https://doi.org/10.1002/aisy.202300251).
- [53] MARTINEZ-GARCÍA F P, CONTRERAS-DE-VILLAR A, and MUÑOZ-PEREZ J J. Review of wind models at a local scale: Advantages and disadvantages[J]. *Journal of Marine Science and Engineering*, 2021, 9(3): 318. doi: [10.3390/jmse9030318](https://doi.org/10.3390/jmse9030318).
- [54] WU Yuankang and HONG Jingshan. A literature review of wind forecasting technology in the world[C]. The IEEE Lausanne Power Tech, Lausanne, Switzerland, 2007: 504–509. doi: [10.1109/PCT.2007.4538368](https://doi.org/10.1109/PCT.2007.4538368).
- [55] 刘夏, 韩雁飞, 李海, 等. 基于数值天气预报模式的机载气象雷达降雨目标极化特性仿真[J]. 雷达学报, 2016, 5(2): 190–199. doi: [10.12000/JR16048](https://doi.org/10.12000/JR16048).
- LIU Xia, HAN Yanfei, LI Hai, *et al.* Polarization characteristics simulation of airborne weather radar rainfall target based on numerical weather prediction[J]. *Journal of Radars*, 2016, 5(2): 190–199. doi: [10.12000/JR16048](https://doi.org/10.12000/JR16048).
- [56] ZHANG Zongwei, LIN Lianlei, GAO Sheng, *et al.* A machine learning model for hub-height short-term wind speed prediction[J]. *Nature Communications*, 2025, 16(1): 3195. doi: [10.1038/s41467-025-58456-4](https://doi.org/10.1038/s41467-025-58456-4).
- [57] HAN Tao, GUO Song, LING Fenghua, *et al.* FengWu-GHR: Learning the kilometer-scale medium-range global weather forecasting[EB/OL]. <https://arxiv.org/abs/2402.00059>, 2024.
- [58] BODNAR C, BRUINSMA W P, LUCIC A, *et al.* A foundation model for the Earth system[J]. *Nature*, 2025, 641(8065): 1180–1187. doi: [10.1038/s41586-025-09005-y](https://doi.org/10.1038/s41586-025-09005-y).
- [59] ESPEHOLT L, AGRAWAL S, SØNDERBY C, *et al.* Deep learning for twelve hour precipitation forecasts[J]. *Nature Communications*, 2022, 13(1): 5145. doi: [10.1038/s41467-022-32483-x](https://doi.org/10.1038/s41467-022-32483-x).
- [60] ANDRYCHOWICZ M, ESPEHOLT L, LI Di, *et al.* Deep learning for day forecasts from sparse observations[EB/OL]. <https://arxiv.org/abs/2306.06079>, 2023.
- [61] PATHAK J, COHEN Y, and GARG P. Kilometer-scale convection-allowing model emulation using generative diffusion modeling[J]. *Science Advances*, 2026, 12(5): eadv0423. doi: [10.1126/sciadv.adv0423](https://doi.org/10.1126/sciadv.adv0423).
- [62] XU Pengbo, ZHENG Xiaogu, GAO Tianyan, *et al.* An artificial intelligence-based limited area model for forecasting of surface meteorological variables[J]. *Communications Earth & Environment*, 2025, 6(1): 372. doi: [10.1038/s43247-025-02347-5](https://doi.org/10.1038/s43247-025-02347-5).
- [63] 王俊, 郑彤, 雷鹏, 等. 深度学习在雷达中的研究综述[J]. 雷达学报, 2018, 7(4): 395–411. doi: [10.12000/JR18040](https://doi.org/10.12000/JR18040).
- WANG Jun, ZHENG Tong, LEI Peng, *et al.* Study on deep learning in radar[J]. *Journal of Radars*, 2018, 7(4): 395–411. doi: [10.12000/JR18040](https://doi.org/10.12000/JR18040).
- [64] 黄钟冷, 姚西文, 韩军伟. 面向SAR图像解译的物理可解释深度学习技术进展与探讨[J]. 雷达学报, 2022, 11(1): 107–125. doi: [10.12000/JR21165](https://doi.org/10.12000/JR21165).
- HUANG Zhongling, YAO Xiwen, and HAN Junwei. Progress and perspective on physically explainable deep learning for synthetic aperture radar image interpretation[J]. *Journal of Radars*, 2022, 11(1): 107–125. doi: [10.12000/JR21165](https://doi.org/10.12000/JR21165).
- [65] 闫文君, 刘康晟, 凌青, 等. 跨场景辐射源个体识别技术综述[J]. 雷达学报(中英文), 待出版. doi: [10.12000/JR25166](https://doi.org/10.12000/JR25166).
- YAN Wenjun, LIU Kangsheng, LING Qing, *et al.* Survey of cross-scenario specific emitter identification technology[J]. *Journal of Radars*. in press, doi: [10.12000/JR25166](https://doi.org/10.12000/JR25166).
- [66] 万显荣, 易建新, 占伟杰, 等. 基于多照射源的被动雷达研究进展与发展趋势[J]. 雷达学报, 2020, 9(6): 939–958. doi: [10.12000/JR20143](https://doi.org/10.12000/JR20143).
- WAN Xianrong, YI Jianxin, ZHAN Weijie, *et al.* Research progress and development trend of the multi-illuminator-based passive radar[J]. *Journal of Radars*, 2020, 9(6): 939–958. doi: [10.12000/JR20143](https://doi.org/10.12000/JR20143).
- [67] LEVIE R, YAPAR Ç, KUTYNIOK G, *et al.* RadioUNet: Fast radio map estimation with convolutional neural networks[J]. *IEEE Transactions on Wireless Communications*, 2021, 20(6): 4001–4015. doi: [10.1109/TWC.2021.3054977](https://doi.org/10.1109/TWC.2021.3054977).
- [68] LEE J H, SERBETCI O G, SELVAM D P, *et al.* PMNet: Robust pathloss map prediction via supervised learning[C]. Global Communications Conference, Kuala Lumpur, Malaysia, 2023: 4601–4606. doi: [10.1109/GLOBECOM54140.2023.10437562](https://doi.org/10.1109/GLOBECOM54140.2023.10437562).
- [69] CHEN Guokai, LIU Yongxiang, ZHANG Tao, *et al.* A graph neural network based radio map construction method for urban environment[J]. *IEEE Communications Letters*, 2023, 27(5): 1327–1331. doi: [10.1109/LCOMM.2023.3260272](https://doi.org/10.1109/LCOMM.2023.3260272).
- [70] CHAVES-VILLOTA A and VITERI-MERA C A. DeepREM: Deep-learning-based radio environment map estimation from sparse measurements[J]. *IEEE Access*, 2023, 11: 48697–48714. doi: [10.1109/ACCESS.2023.3277248](https://doi.org/10.1109/ACCESS.2023.3277248).
- [71] LUO Xuanhao, ZHIZHEN L, PENG Zhiyuan, *et al.* RM-Gen: Conditional diffusion model-based radio map generation for wireless networks[C]. 2024 IFIP Networking Conference, Thessaloniki, Greece, 2024: 543–548. doi: [10.23919/IFIPNetworking62109.2024.10619829](https://doi.org/10.23919/IFIPNetworking62109.2024.10619829).
- [72] WANG Xiucheng, TAO Keda, CHENG Nan, *et al.* RadioDiff: An effective generative diffusion model for sampling-free dynamic radio map construction[J]. *IEEE Transactions on Cognitive Communications and Networking*, 2025, 11(2): 738–750. doi: [10.1109/TCCN.2024.3504489](https://doi.org/10.1109/TCCN.2024.3504489).

- [73] JIANG TAO, GELLER J, NI Daiheng, *et al.* Unmanned aircraft system traffic management: concept of operation and system architecture[J]. *International Journal of Transportation Science and Technology*, 2016, 5(3): 123–135. doi: [10.1016/j.ijtst.2017.01.004](https://doi.org/10.1016/j.ijtst.2017.01.004).
- [74] 广东省通信学会, 中国信息通信研究院, 中国联合网络通信有限公司广东省分公司. 2024低空智联网发展研究报告[R]. 2024.  
Guangdong Communication Society, China Academy of Information and Communications Technology, and China United Network Communications Group Co., Ltd. Guangdong Branch. 2024 research report on the development of low-altitude intelligent network[R]. 2024.
- [75] 张学军, 刘法旺, 张祖耀, 等. 低空智能网联体系[J]. *北京航空航天大学学报*, 2025, 51(6): 1793–1815. doi: [10.13700/j.bh.1001-5965.2025.0060](https://doi.org/10.13700/j.bh.1001-5965.2025.0060).  
ZHANG Xuejun, LIU Fawang, ZHANG Zuyao, *et al.* Overview of low-altitude intelligent networked system[J]. *Journal of Beijing University of Aeronautics and Astronautics*, 2025, 51(6): 1793–1815. doi: [10.13700/j.bh.1001-5965.2025.0060](https://doi.org/10.13700/j.bh.1001-5965.2025.0060).
- [76] LIU Yue, TIAN Yunjie, ZHAO Yuzhong, *et al.* VMamba: Visual state space model[C]. The 38th International Conference on Neural Information Processing Systems, Vancouver, Canada, 2024: 3273. doi: [10.52202/079017-3273](https://doi.org/10.52202/079017-3273).
- [77] WANG Zhaozhi, LIU Yue, TIAN Yunjie, *et al.* Building vision models upon heat conduction[C]. The IEEE/CVF Conference on Computer Vision and Pattern Recognition, Nashville, USA, 2025: 9707–9717.
- [78] 陈翔, 汪连栋, 许雄, 等. 基于Raw I/Q和深度学习的射频指纹识别方法综述[J]. *雷达学报*, 2023, 12(1): 214–234. doi: [10.12000/JR22140](https://doi.org/10.12000/JR22140).  
CHEN Xiang, WANG Liandong, XU Xiong, *et al.* A review of radio frequency fingerprinting methods based on Raw I/Q and deep learning[J]. *Journal of Radars*, 2023, 12(1): 214–234. doi: [10.12000/JR22140](https://doi.org/10.12000/JR22140).
- [79] FANG Zheng, LIU Kangjun, CHEN Ke, *et al.* RadioFormer: A multiple-granularity radio map estimation transformer with 1% spatial sampling[EB/OL]. <https://arxiv.org/abs/2504.19161>, 2025.
- [80] FENG Wenzhi, LI Xutao, WU Zhe, *et al.* Perceptually constrained precipitation nowcasting model[C]. The Forty-Second International Conference on Machine Learning, Vancouver, Canada, 2025.
- [81] GAO Zhangyang, TAN Cheng, WU Lirong, *et al.* SimVP: Simpler yet better video prediction[C]. The IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans, Louisiana, USA, 2022: 3170–3180. doi: [10.1109/CVPR52688.2022.00317](https://doi.org/10.1109/CVPR52688.2022.00317).
- [82] CAO Guiping, LUO Shengda, HUANG Wenjian, *et al.* Strip-MLP: Efficient token interaction for vision MLP[C]. The IEEE/CVF International Conference on Computer Vision, Paris, France, 2023: 1494–1504. doi: [10.1109/ICCV51070.2023.00144](https://doi.org/10.1109/ICCV51070.2023.00144).
- [83] 张寅, 张平, 庾兴宇, 等. 扫描雷达未知天线方向图误差下的稀疏目标角超分辨重建方法[J]. *雷达学报(中英文)*, 2024, 13(3): 646–666. doi: [10.12000/JR23208](https://doi.org/10.12000/JR23208).  
ZHANG Yin, ZHANG Ping, TUO Xingyu, *et al.* Sparse targets angular super-resolution reconstruction method under unknown antenna pattern errors for scanning radar[J]. *Journal of Radars*, 2024, 13(3): 646–666. doi: [10.12000/JR23208](https://doi.org/10.12000/JR23208).
- [84] CAO Guiping, HUANG Wenjian, LAN Xiangyuan, *et al.* Cross-DINO: Cross the deep MLP and transformer for small object detection[J]. *IEEE Transactions on Multimedia*, 2025, 27: 7369–7379. doi: [10.1109/TMM.2025.3599074](https://doi.org/10.1109/TMM.2025.3599074).
- [85] 万昊, 梁菁. 基于多重对比损失的雷达传感器网络HRRP无监督目标特征提取方法[J]. *雷达学报(中英文)*, 2025, 14(5): 1294–1305. doi: [10.12000/JR24200](https://doi.org/10.12000/JR24200).  
WAN Hao and LIANG Jing. HRRP unsupervised target feature extraction method based on multiple contrastive loss in radar sensor networks[J]. *Journal of Radars*, 2025, 14(5): 1294–1305. doi: [10.12000/JR24200](https://doi.org/10.12000/JR24200).
- [86] CAO Guiping, WANG Tao, HUANG Wenjian, *et al.* Open-Det: An efficient learning framework for open-ended detection[C]. The 42nd International Conference on Machine Learning, Vancouver, Canada, 2025: 6654–6674.
- [87] LIU Shilong, ZENG Zhaoyang, REN Tianhe, *et al.* Grounding DINO: Marrying DINO with grounded pre-training for open-set object detection[C]. The 18th European Conference on Computer Vision, Milan, Italy, 2024: 38–55. doi: [10.1007/978-3-031-72970-6\\_3](https://doi.org/10.1007/978-3-031-72970-6_3).
- [88] LIN Chuang, JIANG Yi, QU Lizhen, *et al.* Generative region-language pretraining for open-ended object detection[C]. The IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, USA, 2024: 13958–13968. doi: [10.1109/CVPR52733.2024.01324](https://doi.org/10.1109/CVPR52733.2024.01324).
- [89] AGRIM G, PIOTR D, and ROSS G. LVIS: A dataset for large vocabulary instance segmentation[C]. The IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, USA, 2019: 5351–5359. doi: [10.1109/CVPR.2019.00550](https://doi.org/10.1109/CVPR.2019.00550).
- [90] XIAO Linhui, YANG Xiaoshan, PENG Fang, *et al.* HiVG: Hierarchical multimodal fine-grained modulation for visual grounding[C]. The 32nd ACM International Conference on Multimedia, Melbourne, Australia, 2024: 5460–5469. doi: [10.1145/3664647.3681071](https://doi.org/10.1145/3664647.3681071).
- [91] XIAO Linhui, YANG Xiaoshan, PENG Fang, *et al.* OneRef: Unified one-tower expression grounding and segmentation with mask referring modeling[C]. The 38th

- International Conference on Neural Information Processing Systems, Vancouver, Canada, 2024: 4438. doi: [10.52202/079017-4438](https://doi.org/10.52202/079017-4438).
- [92] LI Xin, HUANG Yuqing, HE Zhenyu, *et al.* CiteTracker: Correlating image and text for visual tracking[C]. The IEEE/CVF International Conference on Computer Vision, Paris, France, 2023: 9974–9983. doi: [10.1109/ICCV51070.2023.00915](https://doi.org/10.1109/ICCV51070.2023.00915).
- [93] 徐开明, 王佰录, 李溯琪, 等. 低空监视雷达“走-停-走”目标跟踪技术[J]. 雷达学报, 2022, 11(3): 443–458. doi: [10.12000/JR21211](https://doi.org/10.12000/JR21211).  
XU Kaiming, WANG Bailu, LI Suqi, *et al.* Move-stop-move target tracking with low-altitude surveillance radars[J]. *Journal of Radars*, 2022, 11(3): 443–458. doi: [10.12000/JR21211](https://doi.org/10.12000/JR21211).
- [94] HUANG Yuqing, LI Xin, ZHOU Zikun, *et al.* RTracker: Recoverable tracking via pn tree structured memory[C]. The IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, USA, 2024: 19038–19047. doi: [10.1109/cvpr52733.2024.01801](https://doi.org/10.1109/cvpr52733.2024.01801).
- [95] XU Guangning, NG M K, YE Yunming, *et al.* TLS-MWP: A tensor-based long- and short-range convolution for multiple weather prediction[J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2024, 34(9): 8382–8397. doi: [10.1109/TCSVT.2024.3379291](https://doi.org/10.1109/TCSVT.2024.3379291).
- [96] TAN Cheng, GAO Zhangyang, WU Lirong, *et al.* Temporal attention unit: Towards efficient spatiotemporal predictive learning[C]. The IEEE/CVF Conference on Computer Vision and Pattern Recognition, Vancouver, Canada, 2023: 18770–18782. doi: [10.1109/CVPR52729.2023.01800](https://doi.org/10.1109/CVPR52729.2023.01800).
- [97] CHEN Taiqin, ZHOU Zikun, FANG Zheng, *et al.* RadioDUN: A physics-inspired deep unfolding network for radio map estimation[EB/OL]. <https://arxiv.org/abs/2506.08418>, 2025.
- [98] LIU Kangjun, QIU Chunyan, CHEN Ke, *et al.* Paying deformable attention to sparse spatial observations for deep radio map estimation[J]. *IEEE Transactions on Cognitive Communications and Networking*, 2026, 12: 1436–1450. doi: [10.1109/TCCN.2025.3613520](https://doi.org/10.1109/TCCN.2025.3613520).
- [99] ZHANG Shuhang, LIU Qingyu, CHEN Ke, *et al.* Large models for aerial edges: An edge-cloud model evolution and communication paradigm[J]. *IEEE Journal on Selected Areas in Communications*, 2025, 43(1): 21–35. doi: [10.1109/JSAC.2024.3460078](https://doi.org/10.1109/JSAC.2024.3460078).
- [100] DU Dawei, ZHU Pengfei, WEN Longyin, *et al.* VisDrone-DET2019: The vision meets drone object detection in image challenge results[C]. The IEEE/CVF International Conference on Computer Vision Workshop, Seoul, Korea, 2019: 213–226. doi: [10.1109/ICCVW.2019.00030](https://doi.org/10.1109/ICCVW.2019.00030).
- [101] TEKIN B, SINHA S N, and FUA P. Real-time seamless single shot 6d object pose prediction[C]. The 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, USA, 2018: 292–301. doi: [10.1109/CVPR.2018.00038](https://doi.org/10.1109/CVPR.2018.00038).
- [102] ZHOU Xingyi, WANG Dequan, and KRÄHENBÜHL P. Objects as points[EB/OL]. <https://arxiv.org/abs/1904.07850>, 2019.

## 作者简介

陈 轲, 副研究员, 主要研究方向为低空立体智能感知方法与应用。

吴少聪, 助理研究员, 主要研究方向为低空气象环境建模、多维时序数据分析。

吴 海, 助理研究员, 主要研究方向为三维目标感知、雷达数据处理、计算机视觉。

张舒航, 研究员, 主要研究方向为智能无线通信, 低空网络。

吴 哲, 助理研究员, 主要研究方向为智能交通感知、多元时空分析。

邹龙坤, 博士, 主要研究方向为通信感知一体化技术, 人工智能。

刘康俊, 助理研究员, 主要研究方向为电磁频谱感知、模式识别、图像生成。

曹桂平, 助理研究员, 主要研究方向为计算机视觉、多模态大模型。

肖麟慧, 助理研究员, 主要研究方向为多模态表征学习、视觉语言理解、多模态预训练及其迁移研究。

李 鑫, 副研究员, 主要研究方向为视觉目标感知、视频理解。

王耀威, 教授, 主要研究方向为大规模视觉智能感知、人工智能。

(责任编辑: 于青)