

## 基于物理感知频空注意力与运动学约束的半监督视频SAR阴影跟踪

高森浩<sup>①</sup> 杨夏青\*<sup>①</sup> 师君<sup>②</sup> 崔国龙<sup>②</sup> 武俊杰<sup>②</sup>

<sup>①</sup>(电子科技大学资源与环境学院 成都 611731)

<sup>②</sup>(电子科技大学信息与通信工程学院 成都 611731)

**摘要:** 针对视频合成孔径雷达(Video SAR)动目标阴影跟踪任务中面临的强相干斑噪声、显著非刚性形变及有监督训练样本稀缺等挑战, 该文提出一种融合物理感知频空注意力与运动学约束的半监督跟踪方法。在检测阶段, 构建基于UniMatchV2架构改进的半监督特征增强网络, 通过在解码器高层特征嵌入所设计的物理感知混合频空注意力(PA-HSSA)模块, 利用其频域分支实现对特征域宽带高频干扰的全局抑制, 并借助空域分支对阴影几何结构进行局部锚定, 同时引入动态权重生成器以自适应融合双域特征, 从而在极低标注率条件下生成高质量的预测掩码。在跟踪阶段, 提出一种面向半监督不确定性的时空关联框架, 首先基于线性高斯状态空间模型(LGSSM)构建运动学先验波门以平滑检测结果的边缘抖动, 进而设计融合运动学残差与几何一致性约束的多维代价矩阵, 有效缓解因目标机动与形变引起的关联模糊问题。基于桑迪亚国家实验室(SNL)实测数据的实验结果表明, 在仅使用1/32有标签数据的情况下, 所提方法的多目标跟踪准确率(MOTA)达到64.19%, 较基准方法提升6.73%, 显著提升了强杂波背景下弱小阴影目标的稳健跟踪性能。

**关键词:** 视频合成孔径雷达; 阴影跟踪; 半监督学习; 频空融合注意力; 相干斑抑制; 多目标跟踪

**中图分类号:** TN957

**文献标识码:** A

**文章编号:** 2095-283X(2026)02-0479-24

**DOI:** 10.12000/JR25265

**CSTR:** 32380.14.JR25265

**引用格式:** 高森浩, 杨夏青, 师君, 等. 基于物理感知频空注意力与运动学约束的半监督视频SAR阴影跟踪[J]. 雷达学报(中英文), 2026, 15(2): 479–502. doi: 10.12000/JR25265.

**Reference format:** GAO Senhao, YANG Xiaqing, SHI Jun, *et al.* Semi-supervised video synthetic aperture radar shadow tracking based on physics-aware spectral spatial attention and kinematic constraints[J]. *Journal of Radars*, 2026, 15(2): 479–502. doi: 10.12000/JR25265.

## Semi-supervised Video Synthetic Aperture Radar Shadow Tracking Based on Physics-aware Spectral Spatial Attention and Kinematic Constraints

GAO Senhao<sup>①</sup> YANG Xiaqing\*<sup>①</sup> SHI Jun<sup>②</sup> CUI Guolong<sup>②</sup> WU Junjie<sup>②</sup>

<sup>①</sup>(School of Resources and Environment, University of Electronic Science and Technology of China,  
Chengdu 611731, China)

<sup>②</sup>(School of Information and Communication Engineering, University of Electronic Science and  
Technology of China, Chengdu 611731, China)

**Abstract:** Moving target shadow tracking in Video Synthetic Aperture Radar (Video SAR) faces significant

收稿日期: 2025-12-08; 改回日期: 2026-03-05; 网络出版: 2026-03-23

\*通信作者: 杨夏青 yangxiaqing@uestc.edu.cn \*Corresponding Author: YANG Xiaqing, yangxiaqing@uestc.edu.cn

基金项目: 国家自然科学基金(62301118, 62371104), 四川省自然科学基金(2025ZNSFSC0466), 电子科技大学科研启动基金(Y030232059002019)

Foundation Items: The National Natural Science Foundation of China (62301118, 62371104), Natural Science Foundation of Sichuan Province (2025ZNSFSC0466), Starting Foundation of the University of Electronic Science and Technology of China (Y030232059002019)

责任编辑: 丁金闪 Corresponding Editor: DING Jinshan

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

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

challenges, including strong speckle noise interference, substantial nonrigid target deformation, and a scarcity of supervised training samples. To address these issues, we propose a semi-supervised tracking method that integrates physics-aware spectral spatial attention and kinematic constraints. For the detection stage, we construct a semi-supervised feature enhancement network based on an improved UniMatchV2 architecture. Specifically, to account for the spectral spatial characteristics of SAR images, a Physics-Aware Hybrid Spectral Spatial Attention (PA-HSSA) module is designed and embedded into the high-level feature space of the decoder. This module utilizes its spectral branch to globally suppress wideband high-frequency interference and its spatial branch to locally anchor the shadow's geometric structure. A dynamic weight generator is introduced to adaptively fuse the dual-domain features, thereby generating high-quality prediction masks under extremely low annotation ratios. For the tracking stage, we propose a spatiotemporal association framework tailored for semi-supervised uncertainty. The framework includes a kinematic prior gate based on a Linear Gaussian State Space Model (LGSSM) to smooth and correct jittery detection edges. Subsequently, a multidimensional cost matrix integrating kinematic residuals and geometric consistency is built to resolve association ambiguities caused by target maneuverability and deformation. Experimental results on measured data from Sandia National Laboratories (SNL) demonstrate that the proposed method achieves a Multiple Object Tracking Accuracy (MOTA) of 64.19% using only 1/32 of the labeled data, outperforming baseline methods by 6.73%. This method effectively addresses the challenge of robustly tracking weak and small shadows in heavily cluttered backgrounds.

**Key words:** Video Synthetic Aperture Radar (Video SAR); Shadow tracking; Semi-Supervised Learning (SSL); Frequency-spatial fusion attention; Speckle reduction; Multi-target tracking

## 1 引言

合成孔径雷达(Synthetic Aperture Radar, SAR)作为一种主动式微波成像传感器,具备全天候作业能力和较强的穿透性,在军事侦察、灾害监测与地形测绘等领域具有重要应用价值。近年来,随着雷达硬件性能的提升、宽带发射体制的发展与实时成像算法的进步,SAR系统正逐步从传统静态成像模式向高分辨率、高帧率的视频SAR (Video SAR)模式演进。视频SAR通常工作于太赫兹或高频微波频段(如Ku, Ka波段),能够以每秒数帧至数十帧的速率对地面感兴趣区域进行连续凝视观测,生成类似光学视频的SAR图像序列,从而为地面目标的持续监视与行为分析提供有效数据支撑<sup>[1,2]</sup>。

在视频SAR图像序列中,运动目标在成像积分时间内对背景散射形成遮挡,产生稳定且能量较低的阴影区域(Moving Target Shadow)<sup>[3]</sup>。相较于基于多普勒能量的传统SAR地面动目标显示(SAR-GMTI)方法<sup>[4]</sup>,阴影位置更贴近目标真实物理坐标,不受多普勒频移引起的定位偏差影响,因此更适用于高精度定位任务<sup>[5]</sup>。此外,阴影特征不依赖于目标的雷达散射截面(Radar Cross Section, RCS),对低RCS或隐身目标具有更好的探测能力;其几何形态呈面状分布,相较于散射点更为稳定,有利于后续的目标分类、识别与轨迹关联。因此,基于阴影的检测与跟踪已成为视频SAR动目标监视的重要研究方向。

尽管视频SAR阴影跟踪具有显著潜力,但在实际复杂场景中实现稳健、持续的阴影跟踪仍面临多

重挑战,主要体现在以下3方面:其一,SAR图像固有的相干斑噪声具有乘性、宽带高频特性<sup>[6]</sup>,在阴影等低灰度区域中噪声干扰尤为显著,易破坏区域均匀性与边缘结构,导致阴影与背景难以区分。传统基于空域卷积的深度学习模型通常难以充分利用频域噪声可分性,在低信杂比条件下分割性能明显下降。其二,视频SAR成像视角持续变化,导致背景杂波统计特性呈现显著非平稳性。在凝视成像模式下,波束入射角与方位角随平台运动不断变化,引起背景散射剧烈波动及阴影非刚性形变,传统单一特征提取方法难以适应此类动态变化,易导致目标丢失或身份切换。其三,SAR图像专业标注成本高昂,有监督训练样本稀缺。SAR图像解译门槛高,视频数据规模大,全监督学习方法容易过拟合。半监督学习(Semi-Supervised Learning, SSL)通过联合利用少量标注样本与大量无标注数据,为解决该问题提供了可行路径<sup>[7,8]</sup>。

为应对上述挑战,国内外学者在视频SAR动目标检测与阴影跟踪方向开展了大量研究。在检测方面,早期研究主要致力于提升弱目标的信杂比,其中,基于联合能量-阴影的双域融合策略旨在利用互补特征提升弱目标的检测概率<sup>[2]</sup>;通道和空间注意力机制的引入则增强了特征提取能力,在一定程度上缓解了由阴影对比度不足引发的分割难题<sup>[9,10]</sup>;部分研究探索了利用U-Net<sup>[11]</sup>, DeepLab<sup>[12]</sup>, Transformer<sup>[13]</sup>等模型对SAR阴影进行端到端分割,但这些模型大多沿用光学图像的网络结构,未显式考虑SAR图像的频域结构与非平稳噪声特性,在高杂波

条件下性能受限。在跟踪方面，近年来针对视频SAR阴影特性的专用算法取得了显著进展。例如，针对阴影特征微弱导致的目标表征难题，文献[14]提出了专用的ShadowDeNet网络，通过显著性引导机制大幅增强了动目标的特征提取能力，为稳健跟踪提供了可靠的观测基础；文献[15]则提出Siam-Sort框架，将改进的孪生网络(Siamese Network)引入多目标关联阶段，利用深度外观特征显著提升了在复杂机动场景下的身份保持能力；此外，针对相干斑噪声干扰，文献[16]通过构建联合核相关滤波器(Joint Kernelized Correlation Filter, JKCF)，利用多特征融合与联合更新策略，在一定程度上改善了低信噪比环境下的跟踪稳定性。

然而，现有研究在实际应用中仍面临双重困境：上述专用方法依赖于全监督学习范式，需要大量昂贵的逐帧标注数据。且基于匈牙利算法<sup>[17]</sup>和卡尔曼滤波<sup>[18]</sup>的传统多目标跟踪框架(如DeepSORT<sup>[19]</sup>, ByteTrack<sup>[20]</sup>)在光学场景表现良好，但直接迁移至视频SAR领域时，往往难以处理阴影形变和外观特征微弱带来的关联模糊问题。另外，以UniMatchV2<sup>[21]</sup>为代表的半监督学习框架在自然图像领域取得了突破性进展。尽管该类方法凭借强大的特征表征能力在视频SAR数据集上展现了良好的泛化性能与应用潜力，但其本质上仍是面向光学视觉特征设计的通用模型。由于其原生架构缺乏对SAR图像频域噪声可分性及阴影低频结构稳定性的显式建模，在面对强斑噪干扰与非平稳背景时，难以充分利用物理先验信息，限制了跟踪精度的进一步跃升。

针对上述问题，本文采用“检测—跟踪”一体化的技术路线，提出一种基于自适应频空融合注意力的半监督视频SAR阴影跟踪方法，通过频域低频增强、空域结构建模与时空关联优化，协同提升弱阴影的检测与跟踪性能。本文的主要贡献包括：

(1) 提出一种物理感知驱动的混合频空注意力机制(Physics-Aware Hybrid Spectral Spatial Attention, PA-HSSA)，实现对弱阴影特征的稳健建模。针对SAR图像中宽带高频分量与阴影(低频)在深层特征域表现出的频域可分性，在检测网络解码器中引入频域分析分支，利用快速傅里叶变换全局抑制高频噪声并保留表征阴影结构的低频成分。同时，构建与之并行的空域感知分支，借助全局池化提取空间上下文描述符并进行通道重标定。该双路结构在频域去噪的基础上引入显式空间几何约束，有效保障了复杂背景下弱阴影目标的形态完整性。

(2) 设计一种环境-架构协同的自适应感知策略，赋予模型应对非平稳杂波的动态调节能力。针

对视频SAR背景杂波的非平稳特性，设计动态权重生成器，根据输入场景的频谱熵与杂波统计特性自适应调节频空分支融合权重：在强噪声场景中增强频域去噪，在复杂纹理背景下聚焦空域细节。此外，提出层级选择性注入策略，规避高频噪声密集的浅层特征，将PA-HSSA模块精准嵌入解码器高层语义空间，以最小计算代价实现特征鲁棒性最大化。

(3) 构建一种面向弱监督不确定性的运动感知关联与轨迹校正机制，解决半监督检测漂移导致的跟踪难题。针对半监督检测器在强杂波干扰下的判别不确定性问题，建立独立于特征学习的运动学纠偏机制，利用线性高斯状态空间模型对阴影运动进行建模，并通过递归贝叶斯估计预测目标时空状态。在此基础上，构建融合运动学残差与几何一致性的多维耦合代价矩阵，在检测器因弱监督信号失效时，借助连续运动先验实现目标的精准重捕获。

## 2 视频SAR阴影成像建模与特性分析

### 2.1 视频SAR阴影成像模型

视频SAR通常工作于太赫兹或高频微波频段(如Ku, Ka波段)，通过对感兴趣区域(Region of Interest, ROI)进行聚束或圆周扫描模式的高帧率连续观测，生成反映场景动态变化的时间序列图像。对于场景中的像素点 $(x, y)$ ，第 $k$ 帧的复数图像 $I_k(x, y)$ 可建模为目标与背景的散射回波经相干斑调制后与系统噪声的叠加。其成像几何与信号构成如图1所示。

在视频SAR成像模式下，当运动目标进入场景时，其径向速度分量会引入两类显著的成像畸变：一方面，额外的多普勒频移导致目标回波在方位向发生位置偏移；另一方面，径向运动累积产生的跨距离单元走动导致能量无法有效聚焦，从而造成图像散焦<sup>[1,19]</sup>。与此形成鲜明对比的是，位于真实位置的目

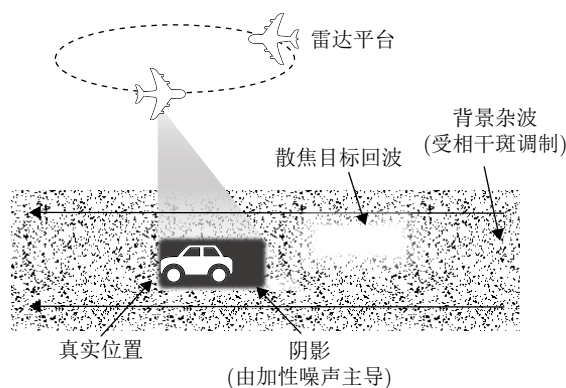


图1 视频SAR动目标多普勒散焦与阴影成像几何示意图  
Fig. 1 Geometric schematic of Doppler defocusing and shadow imaging for moving targets in Video SAR

标实体遮挡了该处地物背景的雷达回波,从而在充满相干斑噪声的强杂波背景中形成显著的低能量阴影区域。这一物理特性表明,阴影区域能够比散焦的目标回波更精准地表征运动目标的真实地理位置<sup>[3,5]</sup>。定义二值掩膜 $M_k(x,y)$ 表示目标存在状态(背景区域为0,阴影区域为1),则第 $k$ 帧的成像模型可表示为

$$I_k(x,y) = [(1 - M_k(x,y)) \cdot C_k(x,y)] \cdot S(x,y) + N_k(x,y) \quad (1)$$

其中, $C_k(x,y)$ 为背景杂波,代表了场景中地物的真实后向散射特性(RCS),是包含低频物理结构信息的物理量; $S(x,y)$ 为SAR成像系统固有的相干斑噪声,是一种由相干叠加原理引起的乘性随机扰动<sup>[6,22]</sup>,满足 $E[S] = 1$ ;  $N_k(x,y)$ 为接收机系统热噪声,通常建模为服从零均值复高斯分布的加性干扰。

由式(1)可知,视频SAR图像中背景与阴影的统计特性存在本质差异。

**背景区域( $M_k = 0$ ):**受相干成像机理的影响,地物回波 $C_k$ 会被相干斑噪声 $S$ 进行乘性调制,即 $I_k \approx C_k \cdot S$ 。这导致原本均匀的地物区域在图像上表现出高频的随机颗粒状起伏,从而掩盖了地物的真实纹理细节。

**阴影区域( $M_k = 1$ ):**背景散射项被抑制,图像强度主要由加性热噪声 $N_k$ 决定,即 $I_k \approx N_k$ 。理想情况下,阴影应为低灰度区域,但实际上受加性噪声底电平和旁瓣泄露的影响,阴影内部仍存在随机起伏,表现为低能量特征。

需要特别说明的是,尽管阴影区域本身受乘性相干斑影响较小,但抑制相干斑对于阴影检测依然至关重要。这是因为背景区域剧烈的相干斑波动会严重模糊背景与阴影的交界线,导致边缘定位困难及虚警率上升。因此,本文提出的方法旨在通过抑制背景区域的相干斑调制以平滑背景,同时保留阴影区域的低能量结构,从而最大化“背景-阴影”的对比度(即信杂比),实现鲁棒的目标分割。

## 2.2 频空域特征的互补性分析

为了在强相干斑噪声和非平稳背景下稳健地提取弱阴影特征,单纯依靠空域或频域信息均存在物理局限。视频SAR场景中充斥着宽带相干斑噪声(绿色标记)和具有频谱相似但结构不同特性的静态背景(紫色标记)如图2所示。为了揭示二者的本质差异并构建互补特征,本节从频域可分性与空域几何约束两个维度进行机理分析。

### 2.2.1 全局频谱感知与特征解耦

尽管相干斑在原始SAR图像域常被建模为乘性

噪声,其对纹理的扰动通常呈现多尺度、宽带特性。因此,仅依赖预设截止频率的固定频域滤波,往往难以在抑制噪声的同时稳定保留阴影等关键结构信息。为了在深度特征空间中构建更具理论依据的结构-斑噪分离机制,本文首先阐明频域分析的理论基础。Qin等人<sup>[10]</sup>指出在卷积网络中广泛使用的全局平均池化(Global Average Pooling, GAP)本质上是对特征图最低频全局分量的提取(原文以二维DCT形式严格推导)。为便于与本文后续FFT表达体系统一,现给出基于DFT的等价性推导。设某一通道特征图为 $\mathbf{X} \in \mathbb{R}^{H \times W}$ ,其二维离散傅里叶变换的零频分量 $F(0,0)$ 与GAP的关系可推导为

$$\begin{aligned} F(0,0) &= \sum_{h=0}^{H-1} \sum_{w=0}^{W-1} \mathbf{X}(h,w) \cdot e^{-j2\pi(0)} \\ &= \sum_{h=0}^{H-1} \sum_{w=0}^{W-1} \mathbf{X}(h,w) \end{aligned} \quad (2)$$

$$\text{GAP}(\mathbf{X}) = \frac{1}{HW} \sum_{h=0}^{H-1} \sum_{w=0}^{W-1} \mathbf{X}(h,w) \quad (3)$$

$$\text{GAP}(\mathbf{X}) \equiv \frac{1}{HW} F(0,0) \quad (4)$$

式(4)表明:GAP本质上是对通道响应的全局能量汇聚,与频域最低频分量具有一致的数学含义(该等式在非归一化DFT定义下成立且成比例)。基于这一数学事实,Qin等人<sup>[10]</sup>指出仅利用GAP会丢失其他频率分量中包含的丰富信息,并引入频域扩展以增强通道描述能力。需要指出的是,与自然图像中中高频分量往往承载精细纹理细节不同,在SAR成像机理<sup>[23]</sup>下,中高频分量更容易受到相干斑

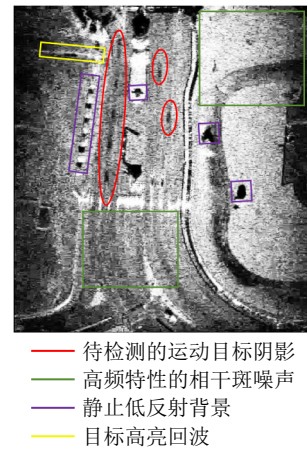


图2 典型视频SAR场景下弱阴影目标与频空域干扰要素的视觉特性分析

Fig. 2 Visual characteristics analysis of weak shadow targets and spatio-frequency interference elements in typical Video SAR scenes

随机干涉所引入的颗粒起伏影响，并呈现更宽带的谱分布。因此，本文利用FFT获取全频谱统计特征，并以通道频谱统计生成软权重，在统计意义上对特征通道进行重标定：当某通道更体现为宽带随机起伏扰动主导时，降低其权重以抑制干扰；当某通道呈现更强的结构连续性(低频集中性更明显)时，提高其权重以增强结构表征。需要强调的是，这种区分体现为统计倾向而非严格的频带互斥：阴影边界及散射体细节同样会贡献中高频成分，而相干斑也可能影响多个尺度频段。因此，本文摒弃了硬截止的固定滤波策略，转而采用基于频谱统计的自适应重标定机制，以在不同场景下自适应地平衡结构保真与扰动抑制。

进一步地，本文方法工作于深度特征空间而非像素域。为了获得对深度特征空间中相干斑噪声演变规律的直观理解，我们尝试提供一种基于局部线性近似的启发式视角。设输入信号满足SAR乘性扰动模型  $\mathbf{X} = \mathbf{S}(1 + \varepsilon)$ ，其中  $\mathbf{S}$  为结构成分， $\varepsilon$  为随机扰动 ( $|\varepsilon| \ll 1$ )。对于已训练的可微编码器  $f(\cdot)$ ，在语义成分  $\mathbf{S}$  的局部邻域内，根据一阶泰勒展开理论，其特征映射可近似表示为

$$f(\mathbf{X}) = f(\mathbf{S} + \mathbf{S}\varepsilon) \approx f(\mathbf{S}) + \mathbf{J}_f(\mathbf{S})[\mathbf{S}\varepsilon] \quad (5)$$

其中， $\mathbf{J}_f(\mathbf{S})$  表示编码器在  $\mathbf{S}$  处的雅可比算子。该式暗示：在局部线性假设下，输入端的乘性扰动经过非线性映射后，倾向于在特征域转化为叠加在语义表征  $f(\mathbf{S})$  上的近似加性残差项  $\mathbf{J}_f(\mathbf{S})[\mathbf{S}\varepsilon]$ 。需要指出的是，虽然泰勒展开假设扰动项  $\varepsilon$  较小，但这并不意味着其对检测任务的影响可以忽略。在视频SAR图像中，阴影目标通常表现为弱纹理和低对比度，即使是幅值较小的相干斑扰动，由于其显著的宽带高频特性，在深层特征中仍可能模糊阴影边缘或导致虚警。这正是本文引入频域注意力机制进行抑制的物理动机。这一数学推导也与Goodfellow等人<sup>[24]</sup>在对抗扰动研究中关于深度网络局部近似线性响应的观点一致。结合SAR相干斑的多尺度、宽带随机起伏特性，该残差在频谱统计上更易呈现宽带能量占比上升的趋势。本文所用编码器采用DINOv2，其多层非线性表征可对输入进行重参数化，使上述结构成分和扰动残差的区分在特征空间中更便于被频谱统计量刻画与利用。需要强调的是，本文不主张“低频=结构、高频=噪声”的硬性划分，而是利用全频谱统计进行通道自适应重标定：自适应抑制宽带扰动主导通道、增强结构信息主导通道，从而提升特征层面的判别相关有效信噪比。

为客观验证上述统计倾向并降低单帧示例可能

带来的偶然性，我们随机挑选一个视频序列：一方面，从该序列中选取一帧作为可视化示例，用于直观展示深层特征在空域与频域上的差异；另一方面，在序列尺度上对该序列内全部有效帧进行统计检验，以检验该差异是否具有稳定的一致性。基于真实标注，我们对阴影区域与背景区域的深层特征进行定量分析结果如图3所示。图3(a)给出了深层特征图的可视化结果，其中红色轮廓标示真实阴影区域，用以建立后续统计对比的物理参照。图3(b)展示了阴影区域与背景区域的空域特征强度分布直方图：两类分布存在明显重叠，表明仅依赖空域幅值差异难以实现稳定分离。相比之下，频域分析揭示了更具区分性的统计特征。如图3(c)的功率谱密度(Power Spectral Density, PSD)曲线所示，阴影区域的能量主要集中于低频段，体现出更强的结构连续性；而背景区域呈现更接近宽带的能量分布，并在较高频段保持相对更高的能量占比，符合随机扰动在频域上的典型表现。进一步地，图3(d)对低频能量占比进行了量化比较：阴影区域的低频能量占比显著高于背景区域，两者在该指标上存在32%的统计差异(PSD采用Welch方法估计并做总能量归一化；本文将低频带定义为PSD频率采样点的前30%，低频能量占比为该频带内归一化谱能量的累积和)。

为了进一步验证深度网络非线性映射对特征解耦的有效性，并探究原始信号在特征提取过程中的演变规律，我们选取了与图3完全一致的视频SAR图像帧，并对其原始信号进行了与深层特征相同的空域及频域定量分析，结果如图4和图5所示。

首先，在空域维度，原始SAR图像受背景地物后向散射多样性、乘性相干斑噪声及加性热噪声的共同调制，背景杂波与阴影的像素强度分布呈现出严重的重叠如图4所示。相比之下，在深层特征中，图3(b)的分布形态更为平滑和集中。虽然两者仍存在交叠，但这进一步印证了单纯依赖空域特征进行分割的局限性。

其次，在频域维度，这种转变更为本质。如图5所示，在原始图像中，背景杂波的低频能量占比高达0.98。这主要是因为原始像素空间，大尺度的地物结构(如连续的路面等)主导了能量分布，其强大的低频能量在频谱上掩盖了宽带高频相干斑噪声。同时，阴影区域(低频占比0.82)受热噪声影响也表现为低频主导。这导致两者在原始频域上呈现混叠状态，缺乏有效区分度。

然而，对比图3(c)可知，深度网络表现出了卓越的特征解耦能力：经过非线性映射后，背景区域

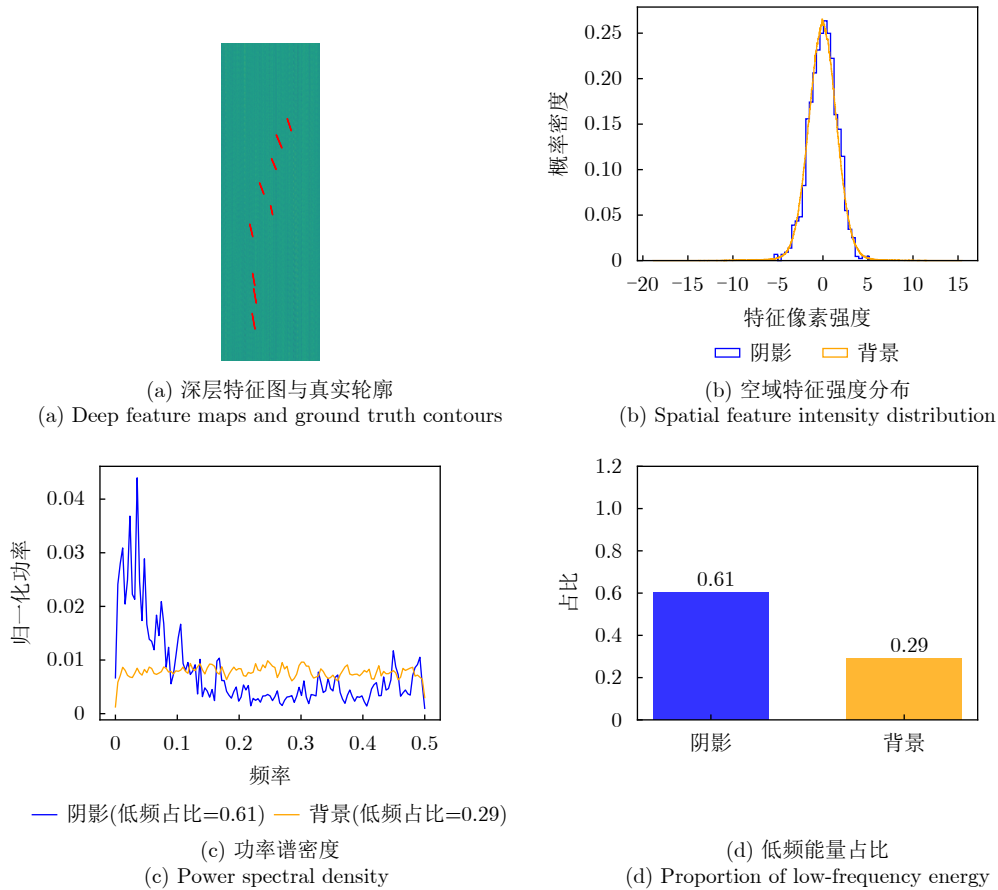


图3 基于真实标签的阴影与杂波特征定量频谱分析

Fig. 3 Quantitative spectral analysis of shadow and clutter features based on ground truth

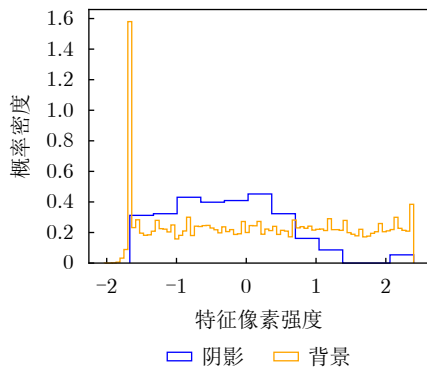


Fig. 4 Histogram of spatial feature intensity distribution in original Video SAR images

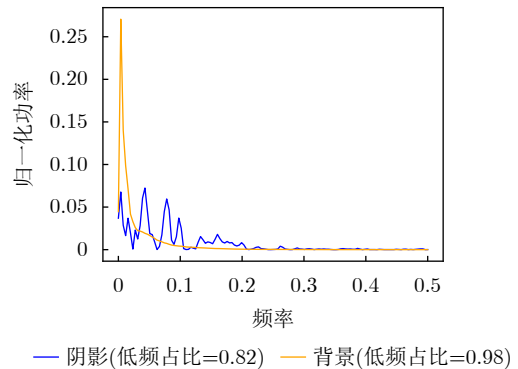


Fig. 5 Power spectral density analysis of original Video SAR images

原本主导的低频地物结构被有效抑制，使得原本被结构信息掩盖的宽带高频分布特性得以显现(低频占比骤降至0.29)；而阴影目标则依然保持了低频能量的主导地位(占比0.61)。这种从原始空间的频谱混叠到特征空间的频谱差异化转变，直观地展示了网络如何将背景与阴影在频域上拉开距离，从而为后续的频域抑制提供了坚实的可分性基础。

此外，我们在序列尺度上对该序列内全部有效

帧重复上述计算，并对每帧得到的低频能量占比进行汇总与检验，结果如图6所示。图6给出了阴影与背景在序列维度上的低频能量占比统计。序列统计结果与图3的单帧现象一致，说明该频谱差异具有稳定的序列一致性，而非由个别帧的偶然波动导致。上述结果表明：在深度特征空间中，阴影相关信息更倾向表现为结构化、低频集中的成分，而相干斑等背景扰动更倾向表现为宽带随机起伏，从而

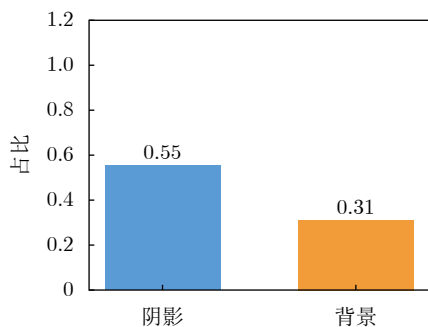


图6 序列尺度低频能量占比统计

Fig. 6 Sequence-level statistics of low-frequency energy proportion

支持本文采用全频谱统计进行通道自适应加权的合理性。

因此，本文将深层特征映射到频域并依据频谱统计自适应生成权重，可在全局范围内有效抑制宽带高频分量，同时保留并增强阴影相关的结构化语义成分。需要指出的是，该机制并不等价于对特征施加固定频率滤波，而是基于通道频谱统计进行自适应抑制与自适应增强，从而提升特征层面的有效信噪比并增强后续阴影跟踪的稳定性。

### 2.2.2 空域几何先验与语义锚定

在前述频域分析中，我们有意仅利用幅度谱来提取通道级的全局频谱统计。其目的并非在频域完成定位或重建，而是从能量分布角度判断：某一通道的响应更偏向于结构连续成分，还是更偏向于随机起伏扰动。由于幅度谱对目标在图像中的具体位置不敏感，因此适合用于生成通道权重并进行全局判别；相位谱主要携带空间对齐与位置关系信息，若用于计算通道的全局权重，容易引入位置相关的不稳定因素。基于此，本文频域分支仅输出通道级权重，并以乘法方式作用于原始空域特征图，实现按通道的软加权重标定；频域分支不替代空域特征，也不承担空间定位任务。

需要指出的是，仅依赖幅度谱进行全局判别存在天然局限：不同空间结构可能出现相似的频谱能量分布。例如，某些低频占比更高的静止背景暗区（地形阴影、水体等）在频谱统计上可能与目标阴影接近，从而带来虚警风险。为降低这类混淆，我们引入空域特征作为补充约束，利用局部几何与上下文信息进行语义锚定。尽管运动目标阴影的形态会随观测几何与径向速度变化而产生拉伸或形变，但其通常仍具有较强的连通性与形态紧致性；空域分支通过聚合上下文信息，可有效抑制两类典型干扰：其一是连通性差、形态破碎的相干斑孔洞；其二是大尺度或形态不规则的自然背景暗区。

对于与运动目标具有相似几何结构的静止地物

（如停放车辆），本阶段仅利用空域特征将其保留为候选对象，以避免漏检；最终的动静判别将由后续的运动学关联模块依据其轨迹行为完成。综上所述，频域与空域特征呈现出高度互补性：频域分支利用全局频谱差异性抑制噪声，空域分支利用局部几何上下文锚定目标。这种频空协同感知机制能够在深层特征空间中实现降噪与保形的最优平衡。

### 2.3 阴影运动特性与关联不确定性分析

为了实现对视频SAR序列中弱小阴影目标的连续跟踪，不仅需要检测器输出的高质量预测掩码，还需对目标的运动规律进行物理建模与分析。然而，受SAR成像视角变化及检测器定位抖动的影响，数据关联过程仍面临着状态估计与观测匹配的双重不确定性。在短时间间隔的视频SAR帧序列中，地面运动目标通常遵循平滑的运动轨迹，其局部状态演化可近似为线性高斯马尔可夫过程。然而，由于弱监督检测器输出的预测掩码边缘不可避免地存在定位抖动，观测值（即掩码质心）包含测量误差。值得注意的是，尽管原始视频SAR图像包含非平稳的相干斑噪声，但经过前级深度检测网络的非线性映射与背景抑制，输出的检测结果已具备较高的信噪比。因此，我们将检测器的输出坐标建模为观测向量 $z_k$ ，并假设在局部跟踪窗口内，其测量噪声 $v_k$ 近似服从高斯分布。基于此假设，本文引入递归平滑算法（如卡尔曼滤波），利用其对高斯噪声的最优估计特性，卡尔曼滤波能够有效平滑检测器的量化误差与随机抖动，从而基于运动学规律实现鲁棒的轨迹预测。

在获得状态估计的基础上，复杂场景下的数据关联仍面临严峻挑战，单纯使用欧氏距离极易失效。首先是高相似度背景干扰与目标机动滞后导致的关联歧义。如图7所示，以预测位置 $\hat{z}_{k|k-1}$ （蓝色

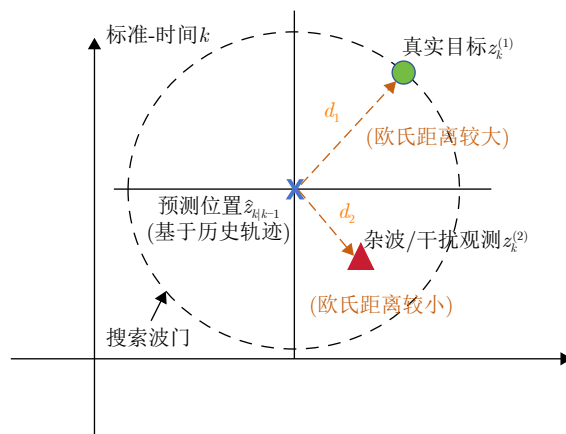


图7 关联模糊挑战

Fig. 7 Association ambiguity challenge

叉号)为中心构建搜索波门。干扰观测 $z_k^{(2)}$ (红色三角形)虽然与预测点的欧氏距离 $d_2$ 较小,但其本质往往是场景中的静止地物(如小型房屋、停放车辆)或运动趋势不一致的邻近目标。这类观测虽然在空间位置上落入波门,但缺乏历史速度矢量的连续性支持,未遵循目标的惯性运动规律。相反,真实目标观测 $z_k^{(1)}$ (绿色圆形)由于发生了非线性机动(如转弯),其真实位置偏离了基于线性运动假设的预测点,使得欧氏距离 $d_1$ 较大(即产生动态滞后),但从全局轨迹看,其运动方向符合惯性延伸,且几何形态保持时空连续性。若仅依据基于线性预测的距离度量,算法极易误将空间邻近的杂波关联为目标。因此,必须构建包含惯性距离补偿与几何一致性的多维关联约束:利用上一帧实测位置进行非线性自适应修正,并结合阴影形态的演化规律,从而有效剔除虽在空间上邻近但物理特征(如运动趋势与几何结构)不匹配的干扰观测。

其次是视角变化引起的非刚性形变挑战。如图8所示,随着雷达平台从时刻 $t_1$ 运动至 $t_2$ ,波束入射角发生改变,同一目标的阴影形状(如长宽比、投影面积)会发生显著的非刚性形变<sup>[3]</sup>。这种外观的不稳定性使得传统基于模板匹配的方法失效。因此,必须引入几何一致性代价,通过衡量前后帧阴影面积的相对变化率来辅助关联决策,从而解决视频SAR弱目标跟踪的不连续问题。

### 3 方法论

#### 3.1 方法整体框架

针对视频SAR动目标监视任务中面临的有监督训练样本匮乏与相干斑噪声干扰强这一核心矛盾,本文提出了一种物理感知驱动的级联式半监督检测与跟踪框架。如图9所示,该框架由两个紧密衔接的核心阶段构成:前端是基于UniMatchV2改进的半监督特征增强检测网络,旨在通过物理可解释的

频空注意力机制嵌入解码器,对有标签与无标签数据进行协同特征增强,从而在强噪声干扰下挖掘稳健的阴影特征;后端是运动学约束下的多目标关联跟踪,利用运动状态先验与几何形态约束对检测结果进行时空维度的纠偏与平滑。

在检测阶段,为了充分利用海量无标注视频SAR数据并克服标注稀缺难题,本文以先进的UniMatchV2半监督范式为基础架构,构建了适合SAR图像特性的半监督特征增强检测网络结构。该框架利用弱增强驱动强增强的一致性正则化机制,通过教师网络生成的无标签图像高置信度伪标签,作为监督信号指导学生网络在强扰动下的预测,从而引导模型挖掘对外观变化具有鲁棒性的阴影语义不变表征。然而,尽管UniMatchV2提供了抗噪的一致性监督信号,但其通用解码器缺乏对SAR特有频空域物理属性的显式建模能力,难以在强相干斑干扰下实现精准分割。为此,本文设计了物理感知解码器(Physics Aware Decoder),我们将物理感知混合频空注意力机制(PA-HSSA)精准嵌入到高层特征重构过程中。作为物理去噪的关键组件,该模块使得网络在响应半监督一致性约束的同时,能够协同利用频域信息抑制全局斑噪并借助空域上下文保持几何结构,从而显著提升了预测掩码的质量。

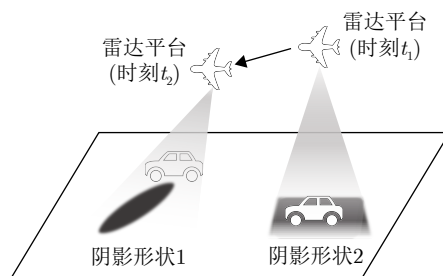


图8 雷达视角变化下的阴影形状变化

Fig. 8 Shadow shape variation under changing radar viewing angles

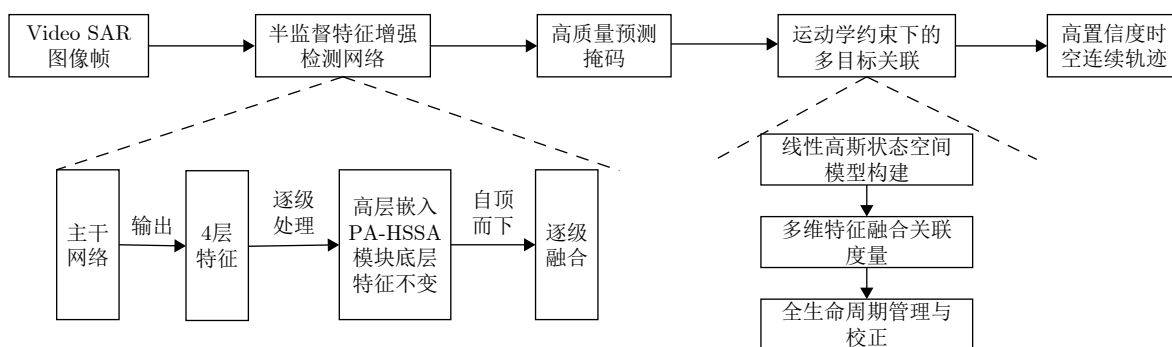


图9 物理感知驱动的半监督检测与跟踪框架

Fig. 9 Physics-aware semi-supervised detection and tracking framework

在数据关联阶段，针对半监督检测环境下观测噪声非平稳及虚警干扰严重的难题，构建了基于线性高斯状态空间模型(Linear Gaussian State Space Model, LGSSM)的运动学纠偏框架。该框架利用递归贝叶斯估计对目标状态进行最优预测<sup>[18]</sup>，并通过融合运动学残差与几何形态一致性的多维关联度量，对检测结果进行时空维度的规整与平滑，从而有效克服单帧检测的不确定性，输出高置信度的时空连续轨迹。

### 3.2 物理感知混合频空注意力机制(PA-HSSA)

为了有效解决2.2节分析中宽带高频分量干扰与阴影几何形变带来的检测难题，本文在检测网络的解码器(Decoder)高层语义特征处嵌入了PA-HSSA

模块。如图10所示的模块内部结构，该模块突破了传统注意力机制仅关注空间或通道维度的局限<sup>[9]</sup>，深度结合SAR图像的物理成像特性，创新性地在频率域和空间域同时对特征进行解耦与重标定。

通过构建频域感知、空域感知与动态融合3条并行支路，PA-HSSA能够在抑制宽带高频斑噪的同时，精准锚定阴影的几何拓扑结构。

#### 3.2.1 频域感知分支：全局噪声抑制

在深层特征域中背景扰动在频域表现为宽带高频分量，而阴影能量集中于低频，频域分支旨在从全局频谱的角度，自适应地识别并抑制噪声主导的特征通道。

设输入特征图为 $\mathbf{X} \in \mathbb{R}^{C \times H \times W}$ ，其中 $C, H, W$ 分

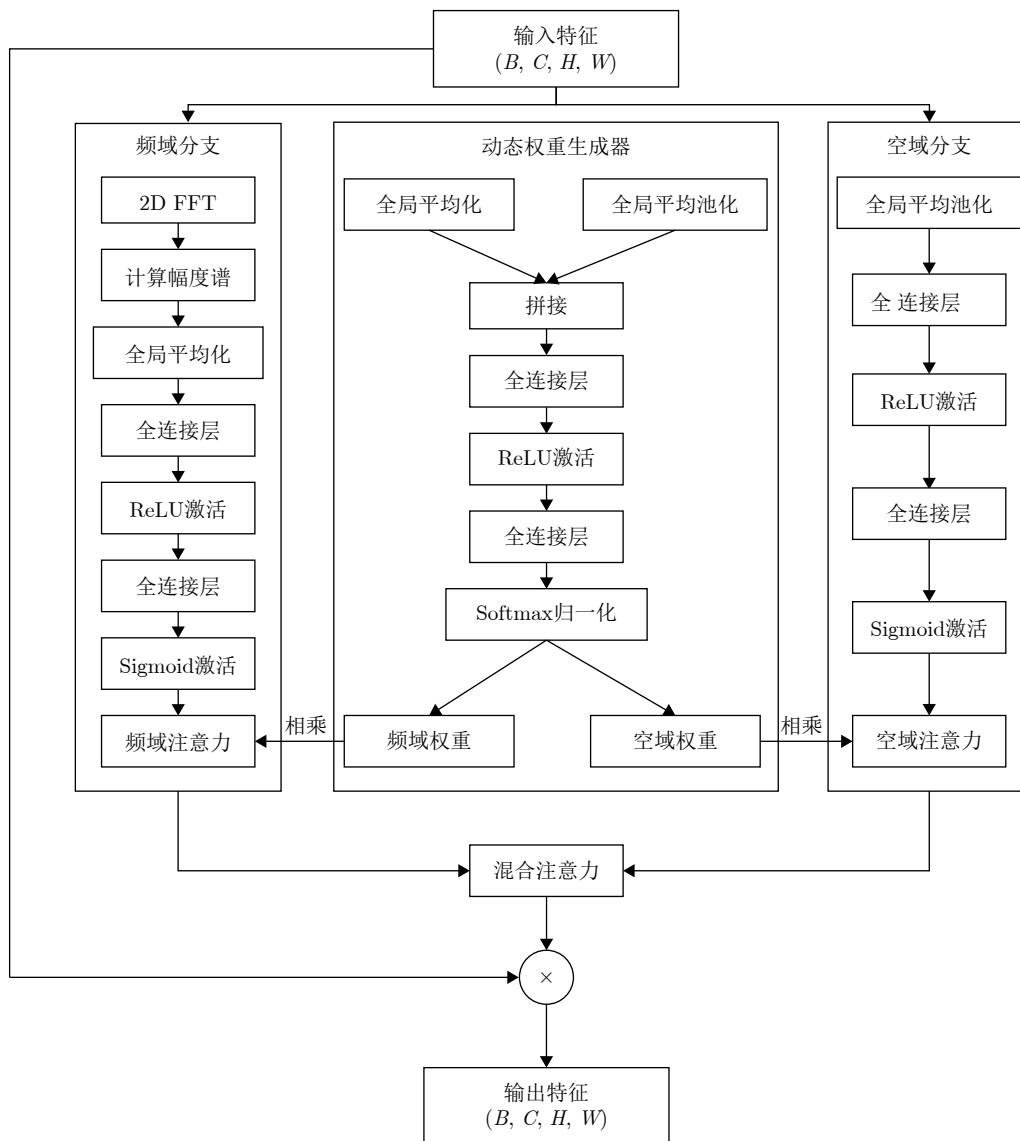


图 10 PA-HSSA模块内部结构细节

Fig. 10 Detailed internal structure of the PA-HSSA module

别为通道数、高度和宽度。首先,利用二维快速傅里叶变换(Two-Dimensional Fast Fourier Transform, 2DFFT)将特征映射至频域,并计算其幅度谱以获得平移不变的频谱特征:

$$\mathbf{X}_{\text{freq}} = |\mathcal{F}(\mathbf{X})| \in \mathbb{R}^{C \times H \times (\frac{W}{2} + 1)} \quad (6)$$

随后,为了提取各通道的全局频谱能量分布(即频谱熵特征),对幅度谱执行全局平均池化(GAP),得到频域统计特征 $\mathbf{v}_{\text{freq}}$ ,并通过多层感知机(Multi-Layer Perceptron, MLP)生成频域注意力权重 $\mathbf{A}_{\text{freq}}$ :

$$\mathbf{v}_{\text{freq}} = \text{GAP}(\mathbf{X}_{\text{freq}}) \in \mathbb{R}^C \quad (7)$$

$$\mathbf{A}_{\text{freq}} = \sigma(\mathbf{W}_2 \delta(\mathbf{W}_1 \mathbf{v}_{\text{freq}})) \in \mathbb{R}^{C \times 1 \times 1} \quad (8)$$

其中, $\sigma(\cdot)$ 为Sigmoid激活函数, $\delta(\cdot)$ 为ReLU激活函数, $\mathbf{W}_1, \mathbf{W}_2$ 为可学习的权重矩阵。 $\mathbf{A}_{\text{freq}}$ 本质上反映了网络对不同特征通道信噪比的判断,用于抑制噪声主导的特征通道。

需要特别指出的是,尽管SAR相干斑在成像机理上表现为乘性噪声,但本模块的处理对象为深层语义特征。基于2.2.1节的理论分析与实验验证,原始图像中的乘性干扰在经过骨干网络的非线性特征映射后,已在特征空间隐式解耦为与语义信息分离的近似加性高频残差。因此,PA-HSSA的频域分支无需显式执行对数变换,而是通过感知特征通道的频谱分布,自适应地抑制高频能量主导的噪声通道,即可在特征层面实现对乘性相干斑影响的有效削减。

### 3.2.2 空域感知分支:几何结构保持

尽管频域分支在特征域能有效抑制宽带高频分量,但傅里叶变换过程中相位信息的丢失导致特征图缺乏显式的空间定位能力,易导致阴影边缘细节的丢失与几何轮廓的定位模糊。为了弥补这一缺陷,本文并行构建了空域感知分支。该分支利用全局平均池化(GAP)作为空间描述符,将二维特征图压缩为一维通道统计量,从而在保留全局感受野的同时聚合空间上下文信息:

$$\mathbf{v}_{\text{spa}} = \text{GAP}(\mathbf{X}) \in \mathbb{R}^C \quad (9)$$

$$\mathbf{A}_{\text{spa}} = \sigma(\mathbf{W}_4 \delta(\mathbf{W}_3 \mathbf{v}_{\text{spa}})) \in \mathbb{R}^{C \times 1 \times 1} \quad (10)$$

这一设计使得网络能够感知目标的整体拓扑形态而非局部纹理。空域注意力权重 $\mathbf{A}_{\text{spa}}$ 本质上充当了语义锚定的角色,它强化了那些对阴影连通区域(如矩形、椭圆形等规则结构)响应强烈的特征通道。这种机制在频域去噪的基础上,进一步引入了几何约束,确保了弱小阴影目标在空间域上的结构完整性,并有效剔除形态不规则的静止背景干扰。

### 3.2.3 动态权重生成与层级特征重构

视频SAR场景具有显著的非平稳性,某些帧可能背景干扰较强(需主导频域去噪),而某些帧背景复杂(需主导空域特征提取)。为了实现环境自适应的特征增强,并确保模块在网络中的高效部署,本文提出了一套包含动态权重生成与层级选择性注入的协同重构策略。

首先,针对非平稳背景下的特征自适应融合问题,本文设计了一个轻量级的动态权重生成器(Dynamic Weight Generator),在频域降噪和空域保形之间寻找最优平衡点。该生成器将频域统计特征 $\mathbf{v}_{\text{freq}}$ 与空域统计特征 $\mathbf{v}_{\text{spa}}$ 进行通道维度的拼接,输入到一个预测网络(由全连接层构成)中,利用Softmax函数自适应生成两个互补的标量权重 $\alpha$ 和 $\beta$ :

$$[\alpha, \beta] = \text{Softmax}(\mathbf{W}_{\text{dyn}}[\mathbf{v}_{\text{freq}}; \mathbf{v}_{\text{spa}}]) \quad (11)$$

最终,PA-HSSA模块的输出特征 $\mathbf{Y}$ 定义为频域注意力与空域注意力的加权融合对原始特征的重标定:

$$\mathbf{A}_{\text{hybrid}} = \alpha \cdot \mathbf{A}_{\text{freq}} + \beta \cdot \mathbf{A}_{\text{spa}} \quad (12)$$

$$\mathbf{Y} = \mathbf{X} \odot \mathbf{A}_{\text{hybrid}} \quad (13)$$

通过这一机制,网络能够实时感知输入图像的频谱熵(噪声水平)与空间复杂度(纹理细节),动态调节双域分支的贡献比例,从而在多变环境下保持特征提取的鲁棒性。

其次,特征重构的有效性不仅取决于融合方式,还取决于其在网络架构中的部署位置。考虑到解码器的浅层特征主要由高频纹理和微观噪声细节构成,在此处引入频域滤波操作不仅计算冗余,且极易放大斑噪干扰;相反,深层特征包含了丰富的全局语义信息与对象级拓扑结构,是进行几何保形与去噪的最佳阶段。基于此物理认知,本文采取层级选择性注入策略,仅在解码器的最高两层语义特征(即Layer3和Layer4)中嵌入PA-HSSA模块,如图11中物理感知解码器(Physics Aware Decoder, PA-Decoder)结构所示。这一策略引导注意力机制避开底层噪点,聚焦于深层特征中的全局对象结构,在显著降低计算开销的同时,实现了对弱阴影目标的精准特征重构。

## 3.3 运动学约束下的轨迹关联与校正

在通过前端检测网络获取高质量的阴影预测掩码后,如何将离散的帧间观测转化为连续稳定的时空轨迹成为关键。尽管检测器引入了PA-HSSA机制以抑制斑噪,但受限于半监督学习的弱监督属

性，输出的掩码质心仍不可避免地存在高频定位抖动与虚警干扰。传统的全局最近邻(Global Nearest Neighbor, GNN)或联合概率数据关联(Joint Probabilistic Data Association, JPDA)算法多基于欧氏距离度量，忽视了视频SAR阴影目标的物理运动惯性与几何拓扑连续性，难以应对此类非平稳观测噪声。

鉴于此，本文提出了一种面向半监督增强的改进型时空关联框架。如图12所示，该框架由3个核心模块构成：首先，利用基于递归贝叶斯推断的线性高斯状态空间模型(LGSSM)构建运动学先验波门，实现对观测状态的平滑与纠偏；其次，重构了

包含运动学残差与几何一致性的多维关联代价矩阵，通过全局最优匹配解决关联模糊；最后，建立基于有限状态机(FSM)的全生命周期管理与时空高斯校正机制，确保输出轨迹的连续性与平滑度。

### 3.3.1 线性高斯状态空间模型构建

针对半监督检测器输出掩码存在的边缘定位抖动与非平稳噪声，如图12左侧所示本文利用线性高斯状态空间模型(LGSSM)引入显式运动学约束。鉴于视频SAR成像的高帧率特性，目标在极短的帧间间隔内机动幅度有限，其局部状态演化可近似建模为匀速直线运动。基于此假设，我们定义包含质

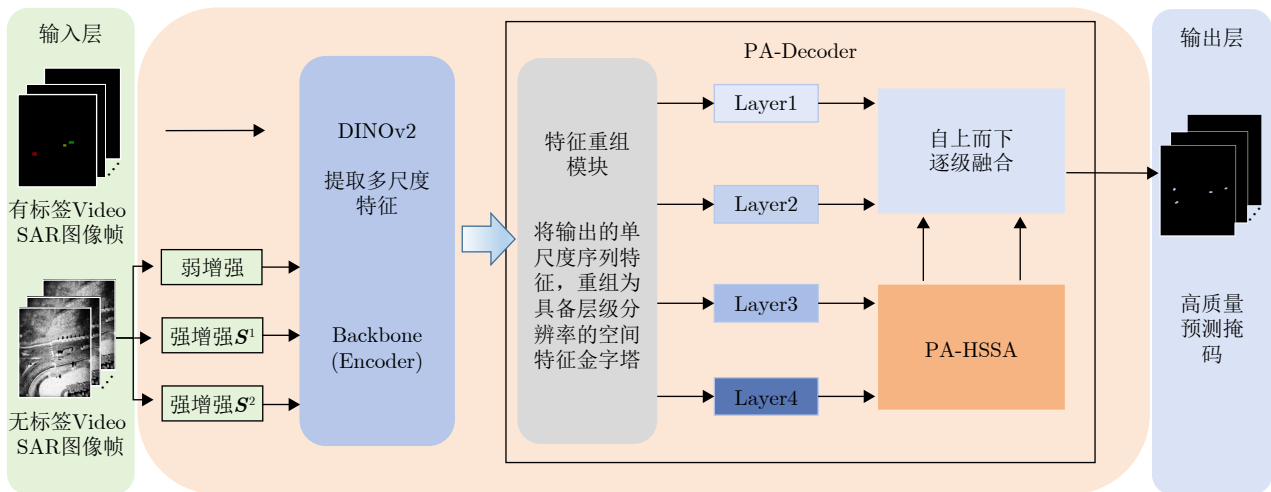


图 11 半监督特征增强检测网络结构与PA-Decoder细节

Fig. 11 Architecture of semi-supervised feature-enhanced detection network and details of PA-Decoder

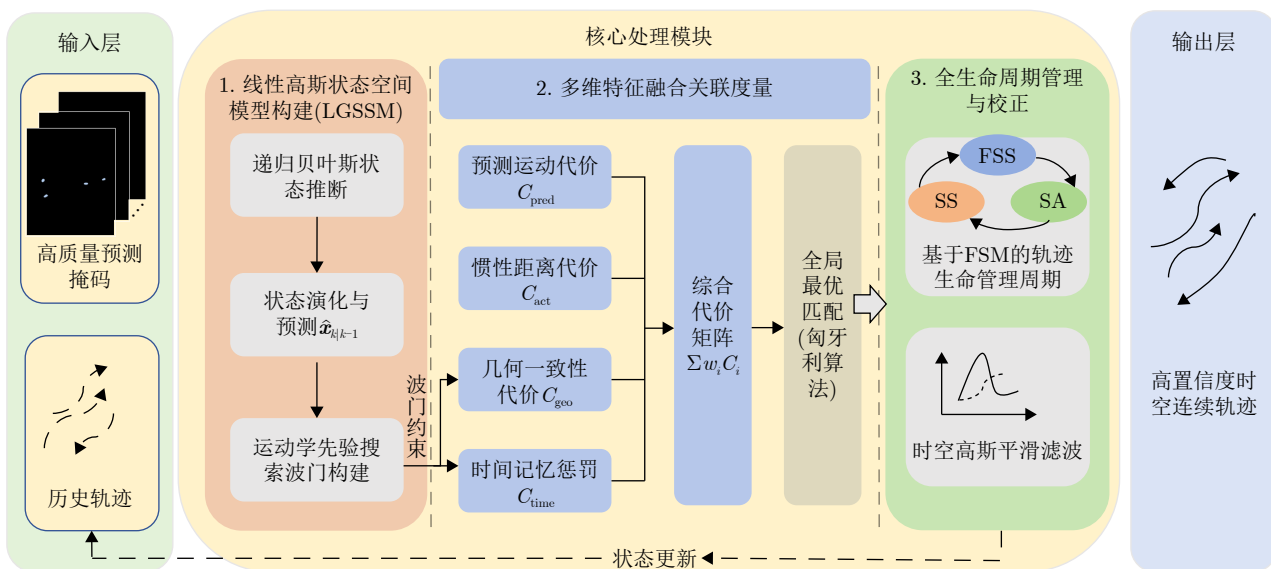


图 12 面向半监督增强的改进型时空关联框架

Fig. 12 Improved spatiotemporal association framework oriented toward semi-supervised enhancement

心坐标与轴向的状态向量  $\mathbf{x}_k = [x, y, v_x, v_y]^T$ , 并利用递归贝叶斯估计对目标状态进行更新。

该步骤并未止步于标准滤波, 其核心价值在于构建运动学搜索波门。利用LGSSM推断的先验状态预测  $\hat{\mathbf{x}}_{k|k-1}$  及其协方差矩阵, 我们在时域上平滑了检测器的量化误差, 并为当前帧的数据关联划定了高置信度的候选搜索区域。这一物理约束有效地将搜索空间限制在符合惯性运动规律的邻域内, 从而在关联阶段前置剔除了大量不符合运动学逻辑的背景杂波干扰。

### 3.3.2 多维特征融合的关联代价度量

为了应对目标非线性机动导致的预测偏差及半监督推断中的虚警杂波, 本文构建了包含线性预测与非线性补偿的多维特征融合关联代价矩阵, 如图12中间模块所示。设第  $i$  条轨迹的预测位置为  $\mathbf{p}_{\text{pred}}$ , 上一帧实际位置为  $\mathbf{p}_{\text{last}}$ , 第  $j$  个预测掩码的质心为  $\mathbf{p}_{\text{det}}$ , 其匹配代价  $C_{i,j}$  定义为

$$C_{i,j} = w_1 \cdot C_{\text{pred}}(i, j) + w_2 \cdot C_{\text{act}}(i, j) + w_3 \cdot C_{\text{geo}}(i, j) + C_{\text{time}}(i) \quad (14)$$

各分项的定义及物理含义如下。

(1) 预测运动代价  $C_{\text{pred}}$ : 计算LGSSM先验预测位置与当前检测的欧氏距离。该项衡量观测值符合线性运动模型的程度, 是剔除运动学异常杂波的主导约束(即运动学残差)。

(2) 惯性距离代价  $C_{\text{act}}$  (非线性补偿项): 计算上一帧实测位置与当前检测的欧氏距离。当目标发生急转弯或机动导致线性预测失效时, 该项利用位置邻近性原理进行自适应修正, 弥补线性滤波器的动态滞后性。

(3) 几何一致性代价  $C_{\text{geo}}$ : 针对阴影随视角变化的非刚性形变, 利用面积相对变化率作为柔性几何约束, 抑制与目标尺度差异显著的虚警。

(4) 时间记忆惩罚  $C_{\text{time}}$ : 对长时间未匹配的轨迹施加线性惩罚, 抑制陈旧状态对当前关联的干扰。

综上所述, 作为面向半监督增强的改进型时空关联框架的空域决策核心, 该多维特征耦合策略通过实验确定的权重配置 ( $w_1 = 0.6, w_2 = 0.3, w_3 = 0.1$ ), 在抑制半监督检测器非平稳观测噪声的同时, 有效补偿了线性模型的动态滞后。该权重设置基于短时物理可置信度的考虑: 视频SAR帧间间隔较短, 目标在短时间尺度上的惯性运动假设相对稳定, 因此对运动学预测项赋予相对更大的权重; 用于补偿目标在转弯、变速等非线性机动下线性运动模型可能产生的预测偏差, 从而在不破坏整体运动先验的前提下提供必要的关联纠偏; 相比之下, 阴影外观会

随视角与散射条件变化而产生形变, 导致几何特征的可靠性低于运动特征。因此, 我们将其作为辅助约束并赋予较小权重, 旨在利用阴影与杂波在形态学上在形态统计上的差异, 辅助抑制显著虚警。上述权重在可行域内通过验证集网格搜索确定(步长0.1), 结果表明, 当权重在最优值附近的区间内变化时, MOTA的波动小于1%, 说明该配置位于性能平坦区, 对超参数扰动不敏感, 从而降低了对特定数值选择的依赖。需要特别说明的是, 式(14)中的时间记忆惩罚项  $C_{\text{time}}$  被设计为一个独立的加性偏置, 不参与上述权重的归一化分配。它根据轨迹的未匹配时长直接对总代价施加线性惩罚, 从而在全局关联中抑制陈旧轨迹的优先级。最终, 结合匈牙利算法的全局最优匹配, 该策略完成了从时域运动学波门约束到空域精细化关联的逻辑递进, 实现了在复杂杂波背景下对目标质心的精准锁定。

### 3.3.3 全生命周期管理的轨迹时空校正

作为时空级联框架的最终环节, 本节聚焦于解决半监督检测中普遍存在的轨迹断裂与定位高频抖动问题, 通过建立如图12右侧所示的轨迹全生命周期管理机制与时空平滑策略, 实现对目标运动状态的闭环校正。

首先, 针对遮挡或漏检导致的轨迹不连续性, 本文设计了基于有限状态机(Finite State Machine, FSM)的生命周期管理策略。我们将轨迹定义为3种状态: 活跃状态(Stable Active, SA)、滑行补盲状态(Stable Switching, SS)和终止状态(Fully Stopped State, FSS)。设置最大丢失容忍阈值  $T_{\text{coast}}$  (本文取5帧)。当关联阶段判定目标丢失时(轨迹由SA状态转移至SS状态), 算法并不立即终止轨迹, 而是利用LGSSM生成的先验预测状态  $\hat{\mathbf{x}}_{k|k-1}$  直接填补缺失帧。这一机制实现了轨迹的惯性传播, 有效修复了因检测器瞬时失效造成的断裂。若连续丢失帧数超过  $T_{\text{del}}$  (本文设为8帧), 则判定目标离开视场并终止跟踪(即轨迹由SS状态转移至FSS状态)。此外, 为抑制背景杂波产生的瞬态虚警, 引入最小轨迹长度约束  $L_{\text{min}}$  (本文设为5帧), 剔除生命周期不足的碎片化噪声, 防止将其误判为真实目标。本框架采用级联阈值策略进行轨迹生命周期管理: 设置最小轨迹长度  $L_{\text{min}} = 5$  帧以抑制由偶发虚警触发的短寿命轨迹; 设置丢失容忍阈值  $T_{\text{coast}} = 5$  帧以应对短时漏检下的惯性滑行; 设置最大丢失阈值  $T_{\text{del}} = 8$  帧作为重捕获缓冲区, 用在避免过早删除导致轨迹断裂与抑制不确定性累积引发的误关联之间进行权衡。上述阈值在验证集上进行逐项扰动与对比分析确定, 实验表明, 在不同视频序列中对这些阈值进

行小幅调整( $\pm 1$ 帧)时,跟踪指标变化不显著,说明该生命周期管理策略对参数选择不敏感,具有较好的稳定性与鲁棒性。

其次,针对半监督预测掩码边缘抖动残留的高频噪声,在输出最终轨迹前,对全生命周期的空间坐标序列应用一维高斯滤波(Gaussian Smoothing)。设原始轨迹位置序列为 $P = \{p_1, p_2, \dots, p_T\}$ ,平滑后的位置 $\tilde{p}_t$ 计算如下:

$$\tilde{p}_t = \frac{1}{Z} \sum_{k=-w}^w p_{t-k} \cdot \exp\left(-\frac{k^2}{2\sigma^2}\right) \quad (15)$$

其中, $Z$ 为归一化系数,确保滤波器权重之和为1; $\sigma$ 为高斯标准差,控制平滑程度。该后处理步骤在保留目标宏观运动趋势的同时,有效平滑了高频定位抖动,确保了最终生成的轨迹在时空维度上具有高度的平滑性与连续性。

## 4 实验结果与分析

### 4.1 实验设置

#### 4.1.1 数据集

为了验证本文方法在真实场景下的有效性,选用美国桑迪亚国家实验室(Sandia National Laboratories, SNL)公开发布的视频SAR实测数据集作为实验基准。该数据集由工作在Ku波段(中心频率16.7 GHz)的雷达系统采集,包含900帧连续的高分辨率SAR图像。图像分辨率为0.1 m,原始尺寸为660像素 $\times$ 720像素,帧率为30 FPS,涵盖了车辆在复杂背景下的直线行驶、转弯及相互遮挡等多种运动状态。为了构建适用于半监督学习的训练与评估体系,本文将原始视频序列划分为9个子序列,每个子序列包含100帧连续图像。其中,前6个子序列(共600帧)被划分为训练集,用于模型的半监督训练;后3个子序列(共300帧)作为测试集,用于评估框架的跟踪性能。在训练集中,我们设计了不同比例的半监督抽样实验,分别随机抽取1/2, 1/4, 1/8, 1/16及1/32的图像帧作为有标签数据,其余作为无标签数据,以此模拟真实场景中从数据丰富到极度匮乏的各类标注条件,验证方法的鲁棒性。

#### 4.1.2 训练设置

本文所有实验是在CUDA 12.4上进行的,硬件平台为Intel Xeon Gold 5320 CPU和NVIDIA RTX A6000 GPU。软件实验环境为Windows 11, Python 3.10和PyTorch 2.5.1。

在半监督检测网络的训练阶段,采用UniMatch V2作为基础框架。输入图像统一裁剪为518像素 $\times$

518像素。特征编码器选用基于Vision Transformer (ViT)架构的DINOv2-Base预训练模型<sup>[25]</sup>。解码器选用本文提出的PA-Decoder,通过嵌入PA-HSSA模块在解码器高层(Layer3与Layer4)协同进行频域去噪与空域保形。优化器采用AdamW,初始学习率设为 $5 \times 10^{-6}$ ,解码器部分的学习率倍率设为40.0,以加速收敛。训练总轮次为70, Batch Size设为6。为保证半监督伪标签的质量,置信度阈值(Confidence Threshold)设为0.95。在后端的运动学数据关联环节,数据关联的最大欧氏距离阈值设为80像素,为了抑制宽带高频分量产生的碎片化虚警,设置最小目标面积阈值为15像素。此外,为了应对遮挡或检测漏检情况,轨迹的最大丢失容忍阈值设为8帧。

### 4.2 实验结果

理想的多目标跟踪结果应具备高准确性和高精度,能够实时、稳定地跟踪所有目标,并在整个过程中保持ID的连续性和一致性。为了全面量化算法在处理目标遮挡、非线性机动及非平稳噪声等复杂情况下的性能,本文采用多目标跟踪领域的各项核心指标进行评估<sup>[26,27]</sup>。具体而言,我们重点关注多目标跟踪准确度(Multiple Object Tracking Accuracy, MOTA)以评估综合检测与跟踪性能,采用识别F1分数(Identification F1-score, IDF1)衡量轨迹的长期一致性,并统计漏检(False Negative, FN)、虚警(False Positive, FP)及身份切换(ID Switch, IDSW)次数以细粒度分析失效原因。

#### 4.2.1 物理感知解码器的积极影响

为了精准评估我们提出的物理感知解码器(PA-Decoder,集成了PA-HSSA模块)相较于原始解码器的性能增益,我们在SNL数据集上进行了一组严格的定量对比实验。为保证比较的公平性,所有实验共享相同的半监督训练设定(UniMatch V2框架)、主干网络(DINOv2-Base)及后端跟踪关联策略,仅改变解码器结构这一单一变量。实验不仅对比了本文方法(PA-Decoder)与基准方法(Baseline)在不同有标签数据比例(从1/2到1/32)下的性能差异,还将本文方法与当前先进的半监督分割算法AllSpark<sup>[28]</sup>和SemiVL<sup>[29]</sup>进行了横向可视化对比。

实验结果如表1所示,进一步结合图13中不同标注比例下的性能变化趋势,以及图14和图15的可视化结果,可以得出以下核心结论。

首先,物理感知解码器在所有半监督设置下均取得了显著的性能优势,且在极低标注比例下表现出卓越的鲁棒性。数据表明,PA-Decoder的MOTA

指标在所有组别中均优于基准方法。尤为值得注意的是，随着有标签样本的减少，性能差距被显著拉大，如图13所示，在1/2比例下MOTA提升约1.02%，而在1/16和1/32比例下，提升幅度分别激增至4.77%和6.73%。这有力地证明了，当监督信号极度匮乏时，PA-HSSA提供的频空物理先验能

够有效防止网络对背景噪声的过拟合，成为保障模型性能的关键支撑。

其次，该模块显著增强了对弱小阴影的感知灵敏度，并在抑制复杂背景虚警方面优于现有先进算法。图14展示了在1/2有标签数据训练下，本文方法与基准方法以及AllSpark, SemiVL的弱阴影检测

表 1 不同标注比例下物理感知解码器与基准方法的跟踪性能对比  
Tab. 1 Tracking performance comparison between PA-Decoder and baseline under different labeled ratios

解码器	标签	MOTA (%) ↑	IDF1 (%) ↑	IDSW ↓	FP ↓	FN ↓
PA-Decoder	1/2	70.93	75.55	12	102	227
	1/4	69.48	74.82	10	101	247
	1/8	68.20	68.57	11	109	253
	1/16	65.30	65.36	16	98	293
	1/32	64.19	68.13	10	149	261
基准方法	1/2	69.91	75.91	6	115	232
	1/4	69.05	70.56	11	83	269
	1/8	67.43	69.80	14	84	284
	1/16	60.53	67.09	8	83	372
	1/32	57.46	69.18	22	261	216

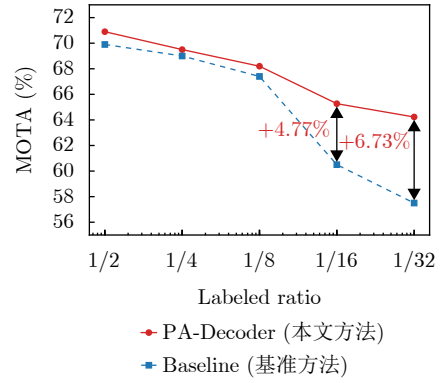
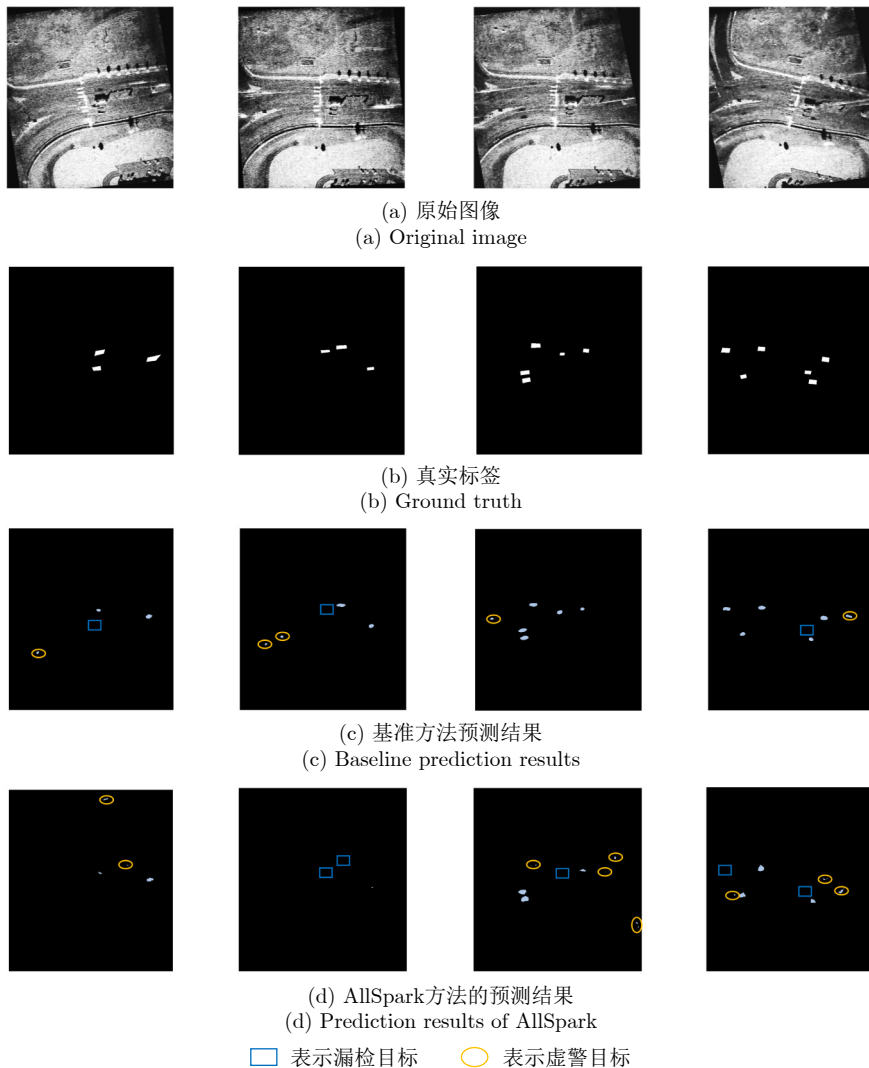


图 13 低资源半监督设定下的跟踪鲁棒性评估

Fig. 13 Robustness evaluation of tracking under low-resource semi-supervised settings



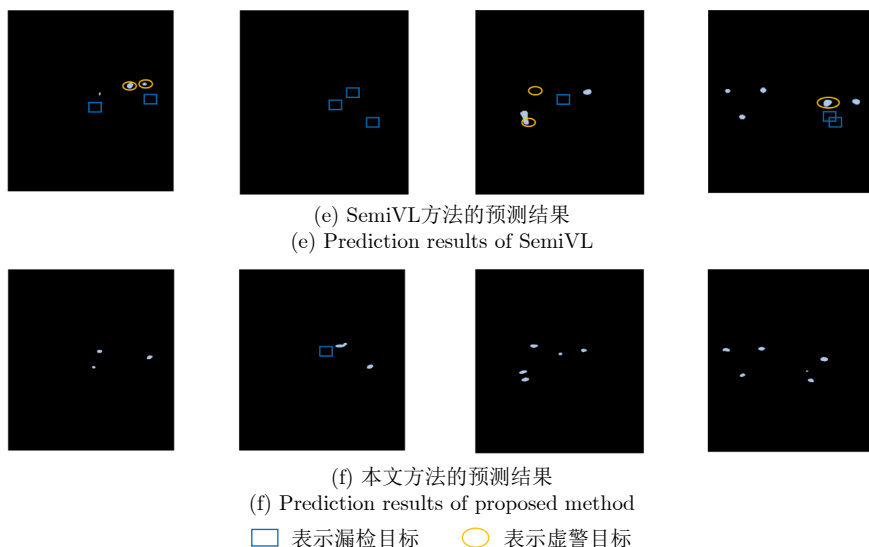


图 14 基于1/2有标签数据训练下的弱阴影检测可视化对比

Fig. 14 Visual comparison of weak shadow detection trained with 1/2 labeled data

可视化对比。从图14可以看出，基准方法(图14(c))缺乏物理约束，难以区分背景杂波与真实目标，导致预测结果破碎且虚警频发。而同类先进算法也面临严峻挑战：AllSpark(图14(d))在处理低对比度暗区时容易发生误判，产生了明显的虚警(如图中黄色椭圆所示)；SemiVL(图14(e))则因过度平滑策略导致对弱小目标敏感度不足，出现了严重的漏检现象(如图中蓝色方框所示)。相比之下，本文方法(图14(f))得益于频空协同感知机制，展现出了双重优势：既能精准捕捉被其他算法遗漏的弱小目标，又能有效识别并抑制被误判为目标背景噪声。

最后，PA-HSSA引入的几何约束在低资源场景下显著保障了长时序跟踪的轨迹连续性。图15直观展示了在1/2有标签数据下的轨迹可视化对比：基准方法(图15(b))生成的轨迹呈现出明显的碎片化特征，容易出现断裂和遗漏(如图中红色虚线所示)，且在多目标交互区域极易产生偏离真实的错误轨迹(如图中红色虚线椭圆所示)。相比之下，本文方法生成的轨迹(图15(c))更加平滑且连续。尽管在部分高标签比例下，由于PA-Decoder能够从背景中挖掘出更多高难度的边缘目标(参见表1中大幅降低的FN)导致身份切换数略有增加，但在极低标注(1/32)的关键挑战下，基准方法因缺乏约束导致IDSW激增至22次，而本文方法依然能够将其稳定控制在10次，降幅超过50%。这充分证明了物理先验知识能够有效抑制身份频繁跳变，确保核心轨迹的长期稳定。

这种在目标形态完整性恢复与背景噪声抑制上的优势，在定量数据中得到了进一步印证，尤其是在低资源场景下：如表1所示，在1/32的极端条件

下，基准方法的虚警数飙升至261，导致其MOTA跌至57.46%；而PA-Decoder不仅凭借频域分支将FP稳定控制在149，更在综合指标MOTA上保持了64.19%的高水准。这一显著的性能差距表明，本文方法不仅构建了比现有SOTA更有效的噪声过滤机制，更通过精准的几何特征恢复保障了核心跟踪性能，从而在极端低资源条件下展现出了更优异的鲁棒性。

#### 4.2.2 与其他先进半监督方法的性能对比

为了全面验证本文提出的半监督特征增强检测网络在视频SAR阴影跟踪任务上的先进性，本节将其与当前计算机视觉领域内最具代表性的主流半监督学习框架进行了横向对比。为了确保对比的严谨性与公平性，所有实验共享相同的半监督训练设定，采用相同的训练数据划分(1/2)、相同的主干网络(DINOv2-Base)以及相同的后端运动学轨迹关联引擎。本次对比实验选取的先进框架具体如下。

**UniMatchV2:** 作为当前半监督学习领域的强基准(Strong Baseline)，该框架也是本文方法的基础架构。它通过多级强弱数据增强和一致性正则化策略，在自然图像的分类与分割任务上表现卓越。

**AllSpark:** 一种针对密集预测任务设计的高效半监督学习方法，通过引入通道语义重组与空间正则化策略，在减少计算开销的同时保持了具有竞争力的特征提取能力。

**SemiVL:** 代表了结合视觉-语言(Vision-Language)预训练知识的前沿半监督范式。该类方法试图利用大规模跨模态数据的语义先验来提升模型在少样本条件下的理解能力。

**Ours:** 本文提出的物理感知驱动半监督特征

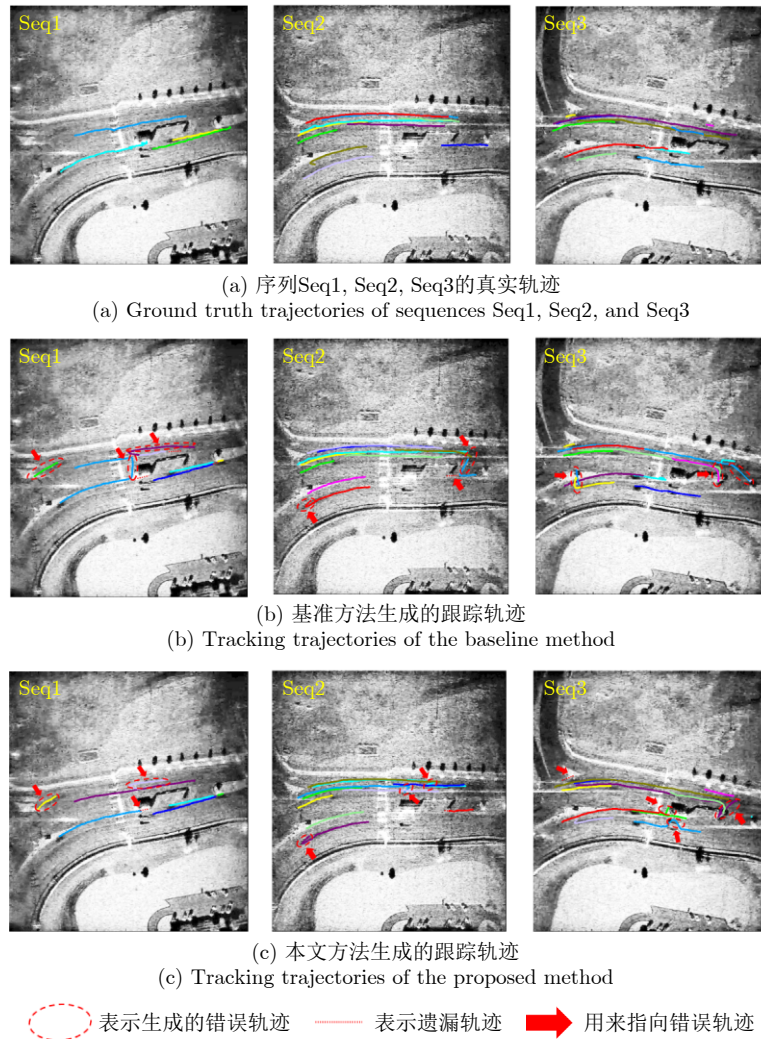


图 15 基于1/2有标签数据训练下的复杂运动场景轨迹连续性可视化对比

Fig. 15 Visual comparison of trajectory continuity in complex motion scenarios trained with 1/2 labeled data

增强检测与跟踪框架。与上述通用模型不同，其核心在于集成了物理感知解码器，通过嵌入PA-HSSA模块协同利用频域去噪与空域几何约束，专门针对SAR图像在深层特征域中的阴影特性进行了深度优化。

定量结果如表2所示。数据揭示了通用视觉模型向视频SAR领域迁移时面临的严峻挑战，以及本文物理感知框架的显著优势。

首先，通用密集预测方法在强相干斑噪声下表现出严重的泛化性能下降。以AllSpark为例，尽管其在自然图像分割中表现高效，但在SAR场景下性能显著退化，MOTA仅为25.83%。最致命的问题在于其高达371的虚警数，远超其他所有方法。这表明，缺乏物理约束的通道重组与扰动策略无法区分高频斑噪与真实纹理，导致模型将大量背景噪点误判为目标，陷入了严重的过拟合，基本丧失了视频SAR领域的可用性。

其次，基于自然图像的跨模态先验知识难以克

表 2 不同半监督学习框架在SNL数据集上的跟踪性能  
Tab. 2 Tracking performance of different semi-supervised learning frameworks on the SNL dataset

方法	MOTA (%) ↑	IDF1 (%) ↑	IDSW ↓	FP ↓	FN ↓
AllSpark	25.83	53.43	28	371	471
SemiVL	41.77	54.33	27	149	507
UniMatchV2	69.91	<b>75.91</b>	<b>6</b>	115	232
Ours	<b>70.93</b>	75.55	12	<b>102</b>	<b>227</b>

注：加粗数值表示最优。

服SAR图像显著的域差异。SemiVL虽然引入了视觉-语言预训练知识，但性能提升有限，MOTA仅为41.77%。与AllSpark产生的大量无序虚警不同，SemiVL的主要瓶颈在于其漏检数高达507，是所有方法中最高的。根本原因在于：虽然SAR图像具备纹理信息，但其本质是高频相干斑噪声与低能量热噪声的随机叠加，这与SemiVL预训练所依赖的、富含判别性语义特征(如物体材质、光影细节)的自

然光学纹理存在本质差异。这种巨大的模态域差异导致模型难以将车辆阴影的文本描述与图像中的噪声空洞特征进行对齐，从而对模糊、微弱的阴影目标无法有效感知，产生严重的漏检现象。最后，本文方法通过物理感知机制超越了强基准，在目标捕获与杂波抑制能力上实现了双重突破。相比于强基准UniMatchV2，本文所提出的方法将MOTA刷新至70.93%。更具说服力的是，我们在虚警数和漏检数上实现了双向优化。这有力地证明了，通过PA-HSSA模块引入频空物理约束，不仅有效抑制了类似AllSpark面临的背景斑噪，又显著增强了SemiVL难以捕捉的低能量弱目标特征。尽管在IDF1指标上基本与Baseline持平，但这主要源于本文方法成功召回了更多处于临界状态的困难目标，证明了PA-Decoder在处理复杂、非平稳SAR场景时具备更优的综合判别力。

#### 4.2.3 与全监督视频SAR阴影跟踪算法对比

为了进一步评估本文所提半监督框架在视频SAR领域的优越性，特别是验证物理感知机制在解决SAR图像特有难点上的有效性，本节将该方法与当前主流的先进全监督跟踪算法进行了横向对比。选取的对比基准分为两类：一类是通用计算机视觉跟踪框架，包括联合检测与跟踪框架FairMOT<sup>[30]</sup>、基于检测关联的ByteTrack<sup>[20]</sup>以及专注于非线性运动建模的OC-SORT<sup>[31]</sup>。另一类是针对视频SAR特性设计的最新专用算法，包括基于图神经网络联合特征学习的GNN-JFL<sup>[32]</sup>和基于图特征表示的Graph-based Tracker<sup>[33]</sup>。为了全面考察性能差异，上述对比算法均在100%全量有标签数据下进行训练，而本文方法仅使用50%的比例有标签数据。详细的定量对比结果如表3所示。

实验结果表明，相比于通用算法的领域适应性不足和全监督专用算法的高昂标注成本，本文方法展现出了卓越的领域适应性与数据高效性。

首先，物理感知机制有效克服了视频SAR领域

的域偏移难题。通用光学跟踪算法由于过度依赖清晰纹理，在强斑点噪声干扰下性能显著退化。如表3所示，基于外观特征的FairMOT难以适应强噪声干扰，MOTA仅为30.19%，且身份切换(IDSW)高达202次，表明其特征提取器在处理SAR图像时面临严重的泛化瓶颈；然而ByteTrack和OC-SORT虽然依靠关联策略将MOTA维持在53%~55%区间，但仍远低于本文方法的70.93%。本文方法凭借PA-HSSA模块的频空域去噪能力，在仅使用一半数据的条件下实现了性能的跨越式提升，证明了在SAR图像解译中针对物理特性建模的必要性。

其次，与领域专用算法的对比凸显了本文方法极高的数据利用效率。GNN-JFL和Graph-based Tracker方法通过引入图结构来挖掘阴影间的拓扑关系，相比通用算法取得了明显的性能提升(MOTA提升至60%以上区间)。然而，本文方法在仅使用一半标注数据的条件下，依然取得了超越这些全监督专用算法的SOTA性能。这一优势归因于PA-HSSA模块的频空域去噪能力：不同于上述方法侧重于后端关联优化，本文方法致力于在前端特征提取阶段直接抑制宽带高频分量并恢复阴影几何结构。这种底层特征增强策略为后端关联提供了更高质量的输入，从而在大幅降低数据依赖的同时实现了精度的跃升。

最后，运动学约束保障了优异的检测完整性与跟踪鲁棒性。针对弱小阴影常被过滤的问题，本文方法将漏检数(FN)大幅降低至227，约为通用算法的一半，证明了其捕捉低信噪比微弱目标的能力。同时，在衡量轨迹连续性的IDF1指标上，结合LGSSM对非线性运动的纠偏，本文方法达到了75.55%。这表明，相比于单纯依赖图拓扑结构的关联方法，引入显式的运动学状态空间模型能更有效地应对阴影形变与遮挡，从而在复杂干扰下维持目标身份的全局一致性。

#### 4.3 消融实验

为了深入探究本文所提框架中各核心组件的贡

表3 本文方法与主流全监督跟踪算法在SNL数据集上的性能对比

Tab. 3 Performance comparison between the proposed method and mainstream fully-supervised tracking algorithms on the SNL dataset

方法	标签(%)	MOTA (%) ↑	IDF1 (%) ↑	IDSW ↓	FP ↓	FN ↓
FairMOT	100	30.19	33.60	202	<b>55</b>	460
ByteTrack	100	54.94	48.78	<b>2</b>	96	431
OC-SORT	100	53.24	48.98	4	103	442
GNN-JFL	100	60.70	63.60	43	237	385
Graph-based Tracker	100	63.50	64.20	46	292	<b>220</b>
Ours	50	<b>70.93</b>	<b>75.55</b>	12	102	227

注：加粗数值表示最优。

献度与必要性,我们在1/2有标签数据比例的半监督设定下,分别针对PA-HSSA注意力机制和多维关联代价函数进行了细粒度的逐项消融分析。

#### 4.3.1 频空双域协同机制的消融分析

PA-HSSA模块的核心思想在于利用频域去噪与空域保形的物理互补性来解决单一维度特征提取的局限性。

为了验证这一协同机制的必要性,我们在相同的半监督设定下,构建了以下3种变体模型与基准模型(Baseline,无注意力机制)进行对比。仅频域分支(Freq-Only):仅保留PA-HSSA中的频域感知路径,移除空域分支与动态权重生成器。仅空域分支(Spatial-Only):仅保留空域感知路径,移除频域分支与动态权重生成器。完整PA-HSSA(Full):同时包含双域分支及自适应权重生成机制。

实验结果如表4所示,通过对频域与空域分支的解耦分析,我们可以得到以下结论:

首先,频域分支通过全局滤波有效抑制了虚警,但存在牺牲细节的风险。对比数据可知,引入频域分支后,虚警数从115降至107,验证了FFT对特征域中的宽带高频分量的抑制作用。然而,值得注意的是,该配置下的漏检数却由232上升至252,导致整体MOTA不升反降。这表明,在缺乏空间几何约束的情况下,过度依赖频域低通滤波会导致弱小阴影的边缘细节被过度平滑,从而丢失部分低信噪比目标。其次,空域分支极大地提升了目标召回率,但易受背景杂波干扰。该配置取得了全场最低的漏检数,说明全局池化提供的空间上下文信息能够极好地锚定阴影形态,防止目标丢失。但与此

同时,其虚警数飙升至149。这证实了仅依靠空域难以区分真实阴影与具有频谱相似但结构不同特性的静止低反射背景,导致大量背景杂波被误检。

最后,频空动态协同实现了性能的最优平衡。完整的PA-HSSA模型通过动态权重生成器实现了双域特征的自适应融合,最终取得了最高的MOTA(70.93%)和最低的虚警数(FP=102)。这一结果具有重要的物理意义:动态融合机制不仅保留了频域分支的宽带噪声抑制特性,同时利用空域分支补偿了位置信息,将漏检数控制在较低水平。这证明了该模块成功在抑制宽带高频分量与保持阴影几何结构之间找到了最佳平衡点,最大化了特征提取的鲁棒性。

#### 4.3.2 多维关联机制的有效性分析

在后端跟踪阶段,为了验证本文提出的多维特征融合关联策略相较于传统方法的优越性,我们在保持前端检测器一致的前提下,对数据关联模块进行了对比消融。实验设置了两种典型的关联策略进行对比。

传统单一度量策略(Baseline):沿用经典多目标跟踪算法(如SORT)的设置,仅计算预测位置与观测位置的欧氏距离,并利用匈牙利算法进行最近邻匹配。

本文多维耦合策略(Ours):引入基于LGSSM的运动学预测,并构建包含运动学残差、惯性距离补偿、几何一致性以及时间记忆惩罚的综合代价矩阵。

实验结果如表5所示,需要指出的是,由于两组实验使用了完全相同的检测器,其漏检数保持一

表4 频空域特征融合的消融实验

Tab. 4 Ablation study on spatial frequency feature fusion

方法	MOTA (%) ↑	IDF1 (%) ↑	IDSW ↓	FP ↓	FN ↓
基准实验	69.91	<b>75.91</b>	<b>6</b>	115	232
频域注意力	68.46	72.60	11	107	252
空域注意力	68.63	69.90	13	149	<b>206</b>
频空动态融合	<b>70.93</b>	75.55	12	<b>102</b>	227

注:加粗数值表示最优。

表5 传统单一度量与本文多维关联策略的跟踪性能对比

Tab. 5 Tracking performance comparison between the traditional single metric and the proposed multi-dimensional association strategy

方法	MOTA (%) ↑	IDF1 (%) ↑	IDSW ↓	FP ↓	FN ↓
单一度量关联	69.39	52.45	28	104	227
多维关联策略	<b>70.93</b>	<b>75.55</b>	<b>12</b>	<b>102</b>	227

注:加粗数值表示最优。

致(均为227),这使得我们可以排除检测精度的干扰,纯粹评估关联策略的有效性。对比数据揭示了以下关键结论:

首先,传统单一欧氏距离度量在轨迹连续性上表现出严重缺陷。Baseline方法的IDF1仅为52.45%,且伴随高达28次的身份切换。这是因为在强相干斑噪声干扰下,高亮的杂波斑点往往随机出现且距离目标极近。仅依赖欧氏距离极易导致算法陷入最近邻陷阱,误将空间邻近的杂波关联为目标,或者在目标形变导致质心偏移时发生链路断裂。这说明单一的空间度量无法维持长时稳定的轨迹跟踪。

相比之下,本文提出的多维耦合策略实现了跟踪稳定性的质的飞跃。首先,通过引入运动学残差与惯性距离,算法构建了符合物理运动规律的概率波门。即便杂波在空间上距离预测点较近,但由于其缺乏连续的运动趋势,会被高额的运动学代价所抑制。这一机制将IDSW从28大幅降至12,有效遏制了虚假轨迹的产生。其次,最显著的提升体现在IDF1指标上,从52.45%激增至75.55%,这主要归功于几何一致性代价利用阴影面积变化的连续性作为辅助约束。在目标发生转弯或非线性机动导致位置预测出现偏差时,该约束充当了身份锚点,防止轨迹错误切换至周围位置相近但形状差异较大的干扰物上。

综上所述,本文提出的多维关联机制在几乎不增加计算代价的前提下,通过物理与几何的双重约束,极大地提升了视频SAR弱小目标的轨迹完整度与身份一致性。

### 4.3.3 与空域去噪策略的对比

首先,需要阐明的是处理域的差异。传统的相干斑抑制方法(如多视法、偏微分方程)通常作用于复图像域或原始像素域,旨在处理乘性噪声。然而,本文提出的PA-HSSA模块构建于深度特征空间。经过骨干网络的非线性映射,原始的乘性相干斑已被转化为特征通道上的加性高频扰动,这使得在特征域进行频域处理成为可能。

为了进一步验证文中提出的混合频空注意力机制的必要性,并探究其与传统单一空域去噪策略的本质差异,我们将自适应的PA-HSSA模块替换为

标准的静态高斯平滑层进行对比实验。该实验旨在评估在特征空间中,不同去噪机制对弱小阴影目标表征能力的影响。

表6的定量分析深刻揭示了不同去噪机制在深层特征处理上的物理差异。引入静态空域高斯平滑导致模型性能出现显著衰退,MOTA从70.93%降至63.77%,且漏检数从227攀升至269。这一数据的剧烈变化强有力地证明:对于视频SAR图像序列中占地像素极少且对比度微弱的阴影目标,高斯平滑作为一种静态低通滤波器,在抑制宽带高频分量的同时,无差别地平滑了构成目标轮廓的关键边缘梯度,直接导致弱小目标因特征模糊而丢失。相比之下,PA-HSSA凭借其构建的频空协同感知机制,在维持极低漏检水平(FN=227)的同时实现了最佳跟踪精度。这得益于其频域分支通过自适应门控显著抑制噪声主导的频率分量,并利用空域分支并行捕获全局语义上下文,从而有效弥补了单一频域处理可能带来的空间细节损失。实验结果证实,单纯的空域平滑无法解决视频SAR中的信噪比矛盾,只有频空双流互补机制才能在强噪声背景下有效兼顾噪声抑制与特征保持,这正是本文频域分支存在的根本必要性。

### 4.4 不同源实测数据的泛化性验证

为了进一步验证本文方法在不同雷达系统参数及成像环境下的泛化性能与实用性,除上述美国SNL公开数据集外,我们引入了来自中国航天科工集团二十三研究所的另外一组机载视频SAR实测数据(CASIC数据集)进行补充实验。

该数据集由飞行高度为3000 m的机载平台通过聚束式SAR (Spotlight SAR)模式采集获得。观测场景为包含环形道路的复杂地面区域,主要运动目标为道路上行驶的车辆。与SNL数据集(Ku波段)不同,该系统工作于Ka频段,载机飞行速度为80 m/s。数据集共包含369帧连续的SAR图像,图像分辨率为0.15 m,尺寸为1000像素×1000像素。

为了构建适用于半监督学习的实验基准,我们将完整数据划分为4个子序列。其中3个子序列作为训练集,剩余1个子序列作为测试集。在训练过程中,为了检验算法在低资源条件下的适应能力,我们保持了和SNL数据集一样的半监督设定。

表6 在特征空间中不同去噪机制的对比

Tab. 6 Comparison of different denoising mechanisms in the feature space

方法	MOTA (%) ↑	IDF1 (%) ↑	IDSW ↓	FP ↓	FN ↓
高斯平滑滤波	63.77	69.86	17	139	269
PA-HSSA	<b>70.93</b>	<b>75.55</b>	<b>12</b>	<b>102</b>	<b>227</b>

注:加粗数值表示最优。

实验结果如表7所示。在CASIC数据集上,尽管面临Ka波段成像机理差异(如对路面纹理细节更敏感)及场景结构变化的挑战,本文提出的物理感知解码器(PA-Decoder)在所有半监督标注比例下均展现出优于基准方法(Baseline)的性能。特别是在极低标签资源(1/32)的严苛条件下,本文方法的MOTA达到了39.33%,相较于基准方法的30.59%提升了约8.74%。尤为关键的是,得益于频域分支对宽带高频分量的有效抑制,本文方法的误检数(FP)显著降低(从103下降至49),这表明PA-HSSA模块能够有效区分背景杂波与真实阴影,在不同信杂比环境下具有良好的鲁棒性。

为了更直观地评估算法的跨域泛化能力,图16给出了在1/2有标签数据训练下的可视化检测对比结果。如图16(a)所示,CASIC数据的Ka波段成像呈现出与SNL数据不同的背景纹理特征,且环形道路周边的强反射点极易形成类似阴影的虚假暗区。从图16(c)可以看出,由于缺乏对SAR物理特性的建模,基准方法(Baseline)在复杂背景处产生了较多虚警(如黄色椭圆所示),且容易漏检微弱的阴影目标(如蓝色方框所示)。相比之下,如图16(d)所示,得益于频域分支对噪声的宽带抑制与空域分支对阴影形态的精细建模,本文方法不仅精准捕获了车辆阴影的完整轮廓,还成功抑制了道路边缘的静态背景杂波,实现了对虚警的有效控制。

综上所述,这一跨数据源(从Ku波段SNL到Ka波段CASIC)的实验结果有力地证明了:本文提出的物理感知频空注意力机制不依赖于特定的雷达系统参数或场景分布,具备处理多样化视频SAR场景的优越泛化性能与实用价值。

#### 4.5 模型复杂度与运行效率分析

为了全面评估本文所提PA-HSSA框架的实际

应用潜力,并确定性能提升是否源于计算代价的过度堆砌,本节从外部横向对比与内部消融分析两个维度,详细探讨了算法的复杂度与运行效率。

##### 4.5.1 外部先进框架横向效能对比

首先,我们将本文方法分别与主流的全监督跟踪算法及先进的半监督检测方法进行了横向对比,详细结果如表8所示。

如表8所示,以ByteTrack和OC-SORT为代表的SOTA全监督方法虽然参数量(98.94 M)与本文方法(97.80 M)处于同一水平,但受限于其高计算复杂度的检测器所带来的推理瓶颈,其推理速度仅为4.5 FPS左右。相比之下,本文方法在同等参数规模下,推理速度提升了近一倍(8.81 FPS),且MOTA指标高出16%~17%。这证明了本文方法并非盲目堆砌参数,而是通过物理感知设计实现了比通用大模型更优的精度与效率平衡。尽管轻量级算法FairMOT速度较快(14.31 FPS),但其30.19%的MOTA表明其难以应对视频SAR的强噪声挑战。

为了排除后端关联策略差异对实验结论的干扰,我们在实验中将表8中所有对比的半监督检测方法均统一集成了本文提出的运动学关联跟踪模块。在此统一基准下,不同方法的效能差异被清晰地揭示:虽然AllSpark和SemiVL凭借架构差异实现了较高的推理速度(14~15 FPS),但它们在视频SAR任务中表现欠佳,特别是参数量高达118.18M的AllSpark,在该检测方法下MOTA仅为25.83%,有力证明了若缺乏针对SAR物理特性的特征增强机制,单纯增加模型容量难以应对强噪声干扰。相比于基准模型 UniMatchV2 (97.53 M),本文框架虽然通过引入频空双流注意力机制实现了MOTA指标1.02%的跃升,但其总参数量仅微幅增加至97.80 M,且推理速度(8.81 FPS)仍保持在同一量级。这种极

表7 在CASIC数据集中不同标注比例下物理感知解码器与基准方法的跟踪性能对比

Tab. 7 Tracking performance comparison between PA-Decoder and Baseline under different labeled ratios on the CASIC dataset

解码器	标签	MOTA (%) ↑	IDF1 (%) ↑	IDSW ↓	FP ↓	FN ↓
PA-Decoder	1/2	51.41	71.74	0	36	153
	1/4	47.04	69.06	1	61	144
	1/8	43.96	68.31	0	64	154
	1/16	42.42	67.72	1	71	152
	1/32	39.33	63.12	0	49	187
基准方法	1/2	48.84	70.25	2	54	145
	1/4	44.73	58.08	4	66	145
	1/8	40.62	60.97	4	105	122
	1/16	39.07	65.01	1	72	164
	1/32	30.59	62.38	1	103	166

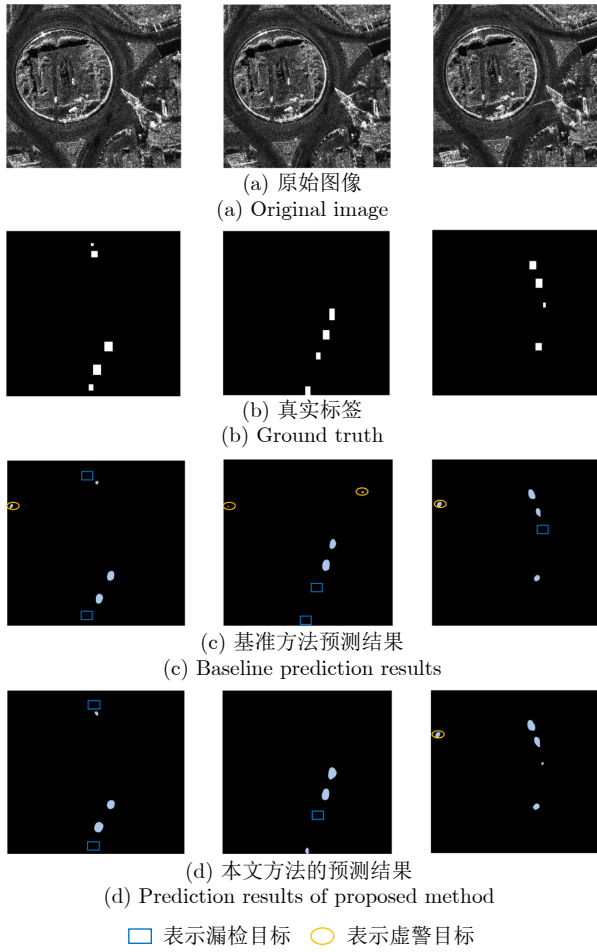


图 16 在CASIC数据集中基于1/2有标签数据训练下的弱阴影检测可视化对比

Fig. 16 Visual comparison of weak shadow detection trained with 1/2 labeled data on the CASIC dataset

低计算代价带来的性能跃升归因于PA-HSSA模块的主动降噪能力，相比于UniMatchV2易受背景干扰导致的特征模糊，PA-HSSA通过在特征域引入物理先验约束实现了特征级的主动重构，显著提升了特征信噪比，从而在几乎不牺牲推理速度的前提下，换取了更优的跟踪精度及对复杂环境的强大鲁棒性。

#### 4.5.2 内部核心组件纵向开销对比

为了进一步解析本文方法的参数效率，我们以单分支模型为基准，对PA-HSSA模块的内部开销进行了详细拆解，结果如表9所示。相比于频域或空域单分支基准模型(97.62 M)，完整的PA-HSSA模块虽然集成了双流分支与动态权重生成器，但其总参数量仅微幅增加至97.80 M。值得注意的是，频域与空域单分支模型表现出完全一致的参数量。这是由于两者在网络构建上均采用了同构的瓶颈MLP结构(降维比 $r = 16$ )，其核心区别仅在于输入时无参数变换操作(FFT与GAP)。此外，得益于这种高压缩比的轻量化设计及在深层特征的选择性嵌入，完整的PA-HSSA模块虽然集成了双流分支与动态权重生成器，但其相比单分支模型仅增加了0.18 M的参数量(在97.80 M的总容量中占比不到0.2%)。这种设计确保了模型在显著增强特征表达能力的同时，维持了与基准模型几乎一致的计算复杂度。在时间开销方面，尽管双域特征的并行处理与动态融合使得端到端推理速度从单分支方法的9.91 FPS略降至8.81 FPS，但系统的主要计算开销集中在特征提取与检测阶段(约9.25 FPS)，而跟

表 8 不同训练范式下跟踪框架的模型复杂度与运行效率对比

Tab. 8 Comparison of model complexity and runtime efficiency of tracking frameworks under different training paradigms

方法	训练范式	总参数量(M)	推理速度(FPS)	MOTA (%) ↑
FairMOT	Fully-Sup.	20.35	14.31	30.19
ByteTrack	Fully-Sup.	98.94	4.58	54.94
OC-SORT	Fully-Sup.	98.94	4.43	53.24
AllSpark	Semi-Sup.	118.18	15.70	25.83
SemiVL	Semi-Sup.	89.32	14.04	41.77
UniMatchV2	Semi-Sup.	97.53	9.91	69.91
Ours	Semi-Sup.	97.80	8.81	70.93

表 9 不同模块配置下的模型复杂度、运行效率与检测性能对比

Tab. 9 Comparison of model complexity, runtime efficiency and Detection Performance under different module configurations

方法	编码器(M)	解码器(M)	总参数(M)	检测(FPS)	数据关联(FPS)	端到端(FPS)	MOTA (%)
频域注意力	86.58	11.04	97.62	10.24	401.73	9.91	68.46
空域注意力	86.58	11.04	97.62	10.32	379.56	9.86	68.63
PA-HSSA	86.58	11.22	97.80	9.25	340.67	8.81	70.93

踪关联模块的运行速率高达340.67 FPS,并不构成计算瓶颈。

综上所述,本文方法在保持可接受的运行效率与模型规模的前提下,实现了跟踪精度的显著跃升,证明了算法在性能增益与模型复杂度之间取得了良好的平衡。

## 5 结语

针对视频SAR动目标监视任务中数据标注成本高昂且成像环境复杂的问题,本文提出一种物理感知驱动的半监督阴影检测与跟踪框架。该方法通过深度融合SAR物理成像机理,降低对大规模标注数据的依赖,显著提升了模型在强噪声与非平稳背景下的鲁棒性能。

本文的主要工作与贡献如下:(1)构建物理感知特征增强机制。针对在深层特征域中SAR图像宽带高频分量与阴影在频空域分布的差异性,提出物理感知混合频空注意力模块(PA-HSSA),通过频域抑制宽带高频分量与空域保持几何结构的协同机制,有效缓解通用半监督模型在强斑噪声干扰下的特征退化问题,提升网络对低能量弱小目标的感知能力。(2)提出运动学约束的时空关联策略。针对半监督检测器输出中存在的非平稳观测噪声,建立基于线性高斯状态空间模型(LGSSM)的运动学纠偏框架,引入运动学残差与几何一致性约束,显著缓解由阴影非刚性形变与背景杂波引起的轨迹漂移与身份切换问题。(3)验证极低标注资源下的算法鲁棒性。基于桑迪亚国家实验室(SNL)实测数据的实验表明,在仅使用1/32有标签数据的极端条件下,本文方法仍可保持较高的跟踪精度,多目标跟踪准确率(MOTA)达到64.19%,较基准模型提升6.73%,展现出在数据受限场景下的良好应用潜力。

此外,为了增强研究的完整性,本文进一步探讨了方法的适用范围及成像场景变化带来的潜在影响。首先,关于成像模式的普适性:得益于数据驱动的端到端架构,本框架直接作用于SAR图像域,不依赖特定的雷达系统参数(如载频、带宽)或观测几何先验。因此,理论上该方法不仅适用于SNL数据集,也可迁移至其他聚束式视频SAR成像系统,具备广泛的适应潜力。其次,关于成像几何与场景复杂度的影响:SAR阴影的显著性受成像俯仰角及背景杂波影响较大。例如,在高俯仰角观测下,阴影区域可能因几何压缩而变短,导致特征显著性下降;在复杂城市背景下,强散射干扰可能破坏阴影边缘。针对上述挑战,虽然极端物理遮挡仍受成像规律制约,但本文提出的PA-HSSA模块通过频空

域协同机制强化了对微弱特征的提取能力,同时结合运动学约束有效利用时空上下文信息对抗环境干扰。因此,相比于缺乏物理先验的通用视觉模型,本框架在应对成像场景变化时展现出了更优越的鲁棒性基础。

未来的研究工作将聚焦于两个重要方向:一方面,探索面向视频SAR的自监督学习范式,利用海量无标签数据中蕴含的时空一致性设计代理任务,开展大规模无监督预训练,以引导模型自主学习SAR成像的物理表征规律,进一步降低对人工标注的依赖,增强算法在开放环境中的泛化能力。另一方面,针对实际工程应用(如星载或弹载平台)中资源受限环境对推理速度与存储开销的要求,本文后续将开展模型轻量化与部署优化研究。未来工作计划引入结构化剪枝与低比特量化等技术以降低冗余计算与存储成本,并探索基于知识蒸馏的轻量化策略,利用当前高精度模型对小规模网络进行指导训练,从而在尽可能保持关键物理感知性能指标的前提下提升推理帧率,推动算法在嵌入式边缘计算平台上的实时部署。

**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] DING Jinshan, WEN Liwu, ZHONG Chao, *et al.* Video SAR moving target indication using deep neural network[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2020, 58(10): 7194–7204. doi: [10.1109/TGRS.2020.2980419](https://doi.org/10.1109/TGRS.2020.2980419).
- [2] 丁金闪, 仲超, 温利武, 等. 视频合成孔径雷达双域联合运动目标检测方法[J]. *雷达学报*, 2022, 11(3): 313–323. doi: [10.12000/JR22036](https://doi.org/10.12000/JR22036).  
DING Jinshan, ZHONG Chao, WEN Liwu, *et al.* Joint detection of moving target in video synthetic aperture radar[J]. *Journal of Radars*, 2022, 11(3): 313–323. doi: [10.12000/JR22036](https://doi.org/10.12000/JR22036).
- [3] RAYNAL A M, BICKEL D L, and DOERRY A W. Stationary and moving target shadow characteristics in synthetic aperture radar[C]. *Radar Sensor Technology XVIII*, Baltimore, USA, 2014: 90771B. doi: [10.1117/12.2049729](https://doi.org/10.1117/12.2049729).
- [4] ENDER J H G, GIERULL C H, and CERUTTI-MAORI D. Improved space-based moving target indication via alternate transmission and receiver switching[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2008, 46(12): 3960–3974. doi: [10.1109/TGRS.2008.2002266](https://doi.org/10.1109/TGRS.2008.2002266).
- [5] TIAN Xiaoqing, LIU Jing, MALLICK M, *et al.*

- Simultaneous detection and tracking of moving-target shadows in ViSAR imagery[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2021, 59(2): 1182–1199. doi: [10.1109/TGRS.2020.2998782](https://doi.org/10.1109/TGRS.2020.2998782).
- [6] ARGENTI F, LAPINI A, BIANCHI T, *et al.* A tutorial on speckle reduction in synthetic aperture radar images[J]. *IEEE Geoscience and Remote Sensing Magazine*, 2013, 1(3): 6–35. doi: [10.1109/MGRS.2013.2277512](https://doi.org/10.1109/MGRS.2013.2277512).
- [7] SOHN K, ZHANG Zizhao, LI Chunliang, *et al.* A simple semi-supervised learning framework for object detection[OL]. arXiv preprint arXiv: 2005.04757, 2020. doi: [10.48550/arXiv.2005.04757](https://doi.org/10.48550/arXiv.2005.04757).
- [8] LIU Yencheng, MA C Y, HE Zijian, *et al.* Unbiased teacher for semi-supervised object detection[C]. 9th International Conference on Learning Representations (ICLR), Virtual, Austria, May 2021. doi: [10.48550/arXiv.2102.09480](https://doi.org/10.48550/arXiv.2102.09480).
- [9] WOO S, PARK J, LEE J Y, *et al.* CBAM: Convolutional block attention module[C]. 15th European Conference on Computer Vision – ECCV 2018, Munich, Germany, 2018: 3–19. doi: [10.1007/978-3-030-01234-2\\_1](https://doi.org/10.1007/978-3-030-01234-2_1).
- [10] QIN Zequn, ZHANG Pengyi, WU Fei, *et al.* FcaNet: Frequency channel attention networks[C]. 2021 IEEE/CVF International Conference on Computer Vision (ICCV), Montreal, Canada, 2021: 763–772. doi: [10.1109/ICCV48922.2021.00082](https://doi.org/10.1109/ICCV48922.2021.00082).
- [11] ZHANG Peng, CHEN Lifu, LI Zhenhong, *et al.* Automatic extraction of water and shadow from SAR images based on a multi-resolution dense encoder and decoder network[J]. *Sensors*, 2019, 19(16): 3576. doi: [10.3390/s19163576](https://doi.org/10.3390/s19163576).
- [12] LI Qiupeng and KONG Yingying. An improved SAR image semantic segmentation Deeplabv3+ network based on the feature post-processing module[J]. *Remote Sensing*, 2023, 15(8): 2153. doi: [10.3390/rs15082153](https://doi.org/10.3390/rs15082153).
- [13] WANG Wei, ZHOU Yuanyuan, XIE Zhikun, *et al.* Moving target shadow detection using transformer in video sar[C]. IGARSS 2022 - 2022 IEEE International Geoscience and Remote Sensing Symposium, Kuala Lumpur, Malaysia, 2022: 2614–2617. doi: [10.1109/IGARSS46834.2022.9884510](https://doi.org/10.1109/IGARSS46834.2022.9884510).
- [14] BAO Jinyu, ZHANG Xiaoling, ZHANG Tianwen, *et al.* ShadowDeNet: A moving target shadow detection network for video SAR[J]. *Remote Sensing*, 2022, 14(2): 320. doi: [10.3390/rs14020320](https://doi.org/10.3390/rs14020320).
- [15] FANG Hui, LIAO Guisheng, LIU Yongjun, *et al.* Siam-Sort: Multi-target tracking in video SAR based on tracking by detection and Siamese network[J]. *Remote Sensing*, 2023, 15(1): 146. doi: [10.3390/rs15010146](https://doi.org/10.3390/rs15010146).
- [16] ZHONG Chao, DING Jinshan, and ZHANG Yuhong. Video SAR moving target tracking using joint kernelized correlation filter[J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2022, 15: 1481–1493. doi: [10.1109/JSTARS.2022.3146035](https://doi.org/10.1109/JSTARS.2022.3146035).
- [17] KUHN H W. The Hungarian method for the assignment problem[J]. *Naval Research Logistics Quarterly*, 1955, 2(1/2): 83–97. doi: [10.1002/nav.3800020109](https://doi.org/10.1002/nav.3800020109).
- [18] KALMAN R E. A new approach to linear filtering and prediction problems[J]. *Journal of Basic Engineering*, 1960, 82(1): 35–45. doi: [10.1115/1.3662552](https://doi.org/10.1115/1.3662552).
- [19] BEWLEY A, GE Zongyuan, OTT L, *et al.* Simple online and realtime tracking[C]. 2016 IEEE International Conference on Image Processing (ICIP), Phoenix, USA, 2016: 3464–3468. doi: [10.1109/ICIP.2016.7533003](https://doi.org/10.1109/ICIP.2016.7533003).
- [20] ZHANG Yifu, SUN Peize, JIANG Yi, *et al.* ByteTrack: Multi-object tracking by associating every detection box[C]. 17th European Conference on Computer Vision – ECCV 2022, Tel Aviv, Israel, 2022: 1–21. doi: [10.1007/978-3-031-20047-2\\_1](https://doi.org/10.1007/978-3-031-20047-2_1).
- [21] YANG Lihe, ZHAO Zhen, and ZHAO Hengshuang. UniMatch V2: Pushing the limit of semi-supervised semantic segmentation[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2025, 47(4): 3031–3048. doi: [10.1109/TPAMI.2025.3528453](https://doi.org/10.1109/TPAMI.2025.3528453).
- [22] GOODMAN J W. Some fundamental properties of speckle[J]. *Journal of the Optical Society of America*, 1976, 66(11): 1145–1150. doi: [10.1364/JOSA.66.001145](https://doi.org/10.1364/JOSA.66.001145).
- [23] GOODMAN J W. Speckle Phenomena in Optics: Theory and Applications[M]. Englewood, Roberts & Company Publishers, 2007 : 73-82.
- [24] GOODFELLOW I J, SHLENS J, and SZEGEDY C. Explaining and harnessing adversarial examples[C]. International Conference on Learning Representations (ICLR), San Diego, USA, 2015. doi: [10.48550/arXiv.1412.6572](https://doi.org/10.48550/arXiv.1412.6572).
- [25] OQUAB M, DARCET T, MOUTAKANNI T, *et al.* DINOv2: Learning robust visual features without supervision[OL]. arXiv preprint arXiv: 2304.07193, 2023. doi: [10.48550/arXiv.2304.07193](https://doi.org/10.48550/arXiv.2304.07193).
- [26] BERNARDIN K and STIEFELHAGEN R. Evaluating multiple object tracking performance: The CLEAR MOT metrics[J]. *EURASIP Journal on Image and Video Processing*, 2008, 2008(1): 246309. doi: [10.1155/2008/246309](https://doi.org/10.1155/2008/246309).
- [27] RISTANI E, SOLERA F, ZOU R, *et al.* Performance measures and a data set for multi-target, multi-camera tracking[C]. Computer Vision – ECCV 2016 Workshops, Amsterdam, The Netherlands, 2016: 17–35. doi: [10.1007/978-3-319-48881-3\\_2](https://doi.org/10.1007/978-3-319-48881-3_2).
- [28] WANG Haonan, ZHANG Qixiang, LI Yi, *et al.* AllSpark: Reborn labeled features from unlabeled in transformer for semi-supervised semantic segmentation[C]. 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition

- (CVPR), Seattle, USA, 2024: 3627–3636. doi: [10.1109/CVPR52733.2024.00348](https://doi.org/10.1109/CVPR52733.2024.00348).
- [29] HOYER L, TAN D J, NAEEM M F, *et al.* SemiVL: Semi-supervised semantic segmentation with vision-language guidance[C]. 18th European Conference on Computer Vision – ECCV 2024, Milan, Italy, 2025: 257–275. doi: [10.1007/978-3-031-72933-1\\_15](https://doi.org/10.1007/978-3-031-72933-1_15).
- [30] ZHANG Yifu, WANG Chunyu, WANG Xinggang, *et al.* FairMOT: On the fairness of detection and re-identification in multiple object tracking[J]. *International Journal of Computer Vision*, 2021, 129(11): 3069–3087. doi: [10.1007/s11263-021-01513-4](https://doi.org/10.1007/s11263-021-01513-4).
- [31] CAO Jinkun, PANG Jiangmiao, WENG Xinshuo, *et al.* Observation-centric SORT: Rethinking SORT for robust multi-object tracking[C]. 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Vancouver, Canada, 2023: 9686–9696. doi: [10.1109/CVPR52729.2023.00934](https://doi.org/10.1109/CVPR52729.2023.00934).
- [32] ZHANG Wensi, ZHANG Xiaoling, XU Xiaowo, *et al.* GNN-JFL: Graph neural network for video SAR shadow tracking with joint motion-appearance feature learning[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2024, 62: 5209117. doi: [10.1109/TGRS.2024.3383870](https://doi.org/10.1109/TGRS.2024.3383870).
- [33] SU Mingjie, NI Peishuang, PEI Hao, *et al.* Graph feature representation for shadow-assisted moving target tracking in video SAR[J]. *IEEE Geoscience and Remote Sensing Letters*, 2025, 22: 4004905. doi: [10.1109/LGRS.2025.3539748](https://doi.org/10.1109/LGRS.2025.3539748).

### 作者简介

高森浩, 硕士生, 主要研究方向为SAR图像智能解译、雷达信号处理等。

杨夏青, 副教授, 主要研究方向为雷达信号处理、合成孔径雷达成像、SAR图像智能解译等。

师君, 副教授, 主要研究方向为雷达信号处理、雷达系统、雷达图像处理、深度学习等。

崔国龙, 教授, 主要研究方向为最优化理论和算法、雷达目标检测理论、波形多样性以及城市环境目标探测等。

武俊杰, 教授, 主要研究方向为双/多基SAR、智能化雷达成像等。

(责任编辑: 于青)