

中图法分类号: TP18; TP391 文献标识码: A 文章编号: 1006-8961(2026)03-0797-14

论文引用格式: Zhang Y M, Bao W T, Xiao Q, Yang Y, Wan W G, Luo Y T, Zou X T and Zhang L. 2026. Selective attention-based for infrared small target detection. Journal of Image and Graphics, 31(3):0797-0810(张迎梅, 鲍王涛, 肖沁, 杨勇, 万伟国, 罗亦韬, 邹雪婷, 张磊. 2026. 基于选择性注意力的红外小目标检测. 中国图象图形学报, 31(3):0797-0810)[DOI:10.11834/jig.250313]

基于选择性注意力的红外小目标检测

张迎梅¹, 鲍王涛¹, 肖沁^{2*}, 杨勇³, 万伟国¹, 罗亦韬¹, 邹雪婷¹, 张磊¹

1. 江西财经大学软件与物联网工程学院, 南昌 330032; 2. 江西财经大学信息管理与数学学院, 南昌 330032;

3. 天津工业大学计算机科学与技术学院, 天津 300387

摘要: 目的 红外小目标检测旨在从复杂背景中准确识别和定位弱小红外目标, 在海上侦查、军事救援等任务中具有重要的应用前景。然而, 受限于红外图像中目标尺寸小、对比度低等因素, 当前的检测方法仍难以实现检测精度与误报率之间的平衡。针对上述问题, 提出一种基于选择性注意力的红外小目标检测网络(selective attention-based network for infrared small target detection, SANet)。方法 具体而言, 设计了双路径语义感知模块, 用于增强网络对弱小目标的感知能力。该模块融合了标准卷积与风车型卷积两种路径, 兼顾局部空间一致性与全局上下文感知能力, 并进一步引入空间/通道注意力机制以细化特征表达, 从而有效提升了目标与背景的可区分性。此外, 为克服U-Net中静态跳跃连接在特征融合中的局限性, 进一步提出了选择性注意力融合模块。该模块基于空间动态权重机制实现跨尺度特征的自适应融合, 能够根据空间显著性选择性增强关键区域, 从而提升对真实目标与误报的辨识能力。结果 在3个公开基准数据集上的实验结果验证了所提出的SANet在交并比(intersection over union, IoU)、nIoU、 P_d 和 F_d 4个指标上均优于现有先进方法, 其中, 本文方法的IoU指标在上述基准数据集上比次优方法分别提升1.93%、4.32%和2.21%。结论 SANet有效增强了网络对小目标的感知能力、关键特征的表达能力以及背景干扰的抑制能力。源代码可以在https://gitcode.com/m0_61988291/SANet上获取。

关键词: 红外小目标检测(IRSTD); 双路径语义感知模块(DSM); 风车型卷积; 选择性注意力融合模块(SAFM); 空间动态权重机制

Selective attention-based for infrared small target detection

Zhang Yingmei¹, Bao Wangtao¹, Xiao Qin^{2*}, Yang Yong³, Wan Weiguo¹, Luo Yitao¹,

Zou Xueting¹, Zhang Lei¹

1. School of Software and Internet of Things Engineering, Jiangxi University of Finance and Economics, Nanchang 330032, China;

2. School of Information Management and Mathematics, Jiangxi University of Finance and Economics, Nanchang 330032, China;

3. School of Computer Science and Technology, Tiangong University, Tianjin 300387, China

Abstract: **Objective** Infrared small target detection (IRSTD) is critical for numerous real-world applications, including maritime surveillance, military search and rescue, early warning systems, and precision-guided strikes. These systems

收稿日期: 2025-07-21; 修回日期: 2025-10-14; 预印本日期: 2025-10-21

* 通信作者: 肖沁 xiaoqin@jxufe.edu.cn

基金项目: 国家自然科学基金项目(62072218, 62261025); 江西省职业早期青年科技人才项目(20244BCE52075, 20244BCE52088); 江西省赣鄱俊才一急需紧缺海外人才项目(20242BCE50059); 天津市自然科学基金项目(24JCZDJC00130); 江西省研究生创新专项基金项目(YC2024-S375, YC2025-S472)

Supported by: National Natural Science Foundation of China(62072218, 62261025)

depend on the precise identification of dim, small targets within cluttered infrared backgrounds. Compared with conventional targets in natural scene images, infrared small targets present several significant challenges: 1) extremely limited spatial resolution, often covering only a few pixels; 2) a low signal-to-clutter ratio and weak contrast, particularly in dynamic or noisy backgrounds such as sea surfaces and clouds; and 3) structurally complex and diverse backgrounds that frequently induce false alarms. These factors may cause the target to be obscured by the complex background, leading to false positives or missed detections, thereby makingIRSTD a persistently difficult task. ExistingIRSTD methods are mainly divided into two categories: traditional methods and deep learning-based methods. Traditional methods are generally classified into three subcategories: filtering-based methods, low-rank representation-based methods, and local contrast-based methods. Although these methods yield satisfactory results in specific scenarios, they largely depend on hand-crafted feature extraction and manually tuned parameters, limiting their ability to capture the diversity and variability of targets in complex scenes. Consequently, their detection performance and robustness are often inadequate in practical applications. In recent years, deep learning-based methods have gained widespread adoption inIRSTD owing to their powerful feature learning capabilities and have contributed to notable advancements in the field. However, an information bottleneck remains in target feature extraction. Moreover, the static skip connection structures commonly used in mainstream deep neural networks struggle to adaptively distinguish real targets from pseudo-target regions in infrared images, further constraining model performance in complex backgrounds. To address these limitations, this paper proposes a novel selective attention-based network (SANet). **Method** The proposed network builds upon the classical U-Net encoder-decoder framework and incorporates two key modules: 1) a dual-path semantic-aware module (DSM) and 2) a selective attention fusion module (SAFM). These enhancements are intended to mitigate the information bottleneck inherent in previous U-Net-based approaches and to address issues such as semantic misalignment and background noise interference caused by static skip connections. The DSM improves the model's sensitivity to dim targets by improving its ability to preserve fine-grained spatial details while modeling contextual semantics. Specifically, this module integrates standard convolution (to extract local fine-scale features) with pinwheel-shaped convolution—a dilation-inspired structure that expands the receptive field in multiple directions without a significant computational burden. This combination enables the network to jointly capture local spatial continuity and global contextual information, which is essential for suppressing background clutter and enhancing target saliency. A classical spatial and channel attention mechanism is introduced following the DSM to further improve feature representation. This mechanism performs fine-grained recalibration by adaptively weighting spatial locations and feature channels, enabling the network to focus more effectively on regions with high target probability while suppressing irrelevant background signals. Additionally, SANet addresses a major limitation of the original U-Net architecture: the static and uniform nature of skip connections. Traditional skip connections uniformly transfer encoder features to the decoder without considering spatial-semantic correlations, often introducing background noise and degrading detection performance. The SAFM is proposed to overcome this issue. This module introduces a globally learnable scalar parameter to perform spatially aware, weighted fusion of multiscale features, facilitating more discriminative information transfer. Through this mechanism, the network adaptively filters the most relevant and representative features, thereby enhancing its response to real targets while reducing false alarms. **Result** Extensive experiments were conducted on three widely usedIRSTD datasets—NUAA-SIRST,IRSTD-1K, and NUDT-SIRST—to verify the effectiveness of the proposed SANet. SANet was comprehensively compared with a variety of current mainstream methods, including five traditional methods (FKRW, TLLCM, NWMTH, NOLC, and PSTNN) and nine deep learning-based methods (ACM, RDIAN, AGPCNet, DNANet, UIUNet, RPCANet, MSHNet, SCTransNet, and L²SKNet). Four standard evaluation metrics were employed to objectively assess detection performance: intersection over union (IoU), normalized IoU, probability of detection (P_d), and false alarm rate (F_a). The experimental results demonstrate that SANet consistently outperforms existing state-of-the-art methods across all evaluation metrics. Specifically, for the IoU metric, SANet achieves improvements of 1.93%, 4.32%, and 2.21% over the second-best methods on the NUAA-SIRST,IRSTD-1K, and NUDT-SIRST datasets, respectively, demonstrating strong generalization capability and robustness across diverse scenarios. **Conclusion** This paper introduces SANet, anIRSTD network grounded in a selective attention mechanism. Built upon the U-Net architecture, the proposed method incorporates a dual-path semantic-aware module and a selective attention fusion module,

which enhance the model's capacity for faint target perception, key feature representation, and background suppression. Experimental results on multiple publicIRSTD datasets demonstrate that SANet outperforms existing mainstream approaches in terms of detection accuracy, robustness, and generalization, highlighting its practical relevance and broad applicability. Future work will concentrate on model lightweighting and inference optimization to fulfill deployment requirements on edge devices. Additionally, we aim to investigate the model's adaptability under complex meteorological conditions (e. g. , rain, fog, thermal disturbances) to broaden its utility across diverse infrared imaging scenarios. The source code is available on https://gitcode.com/m0_61988291/SANet.

Key words: infrared small target detection (IRSTD); dual-path semantic-aware module (DSM); pinwheel-shaped convolution; selective attention fusion module (SAFM); spatial dynamic weight mechanism

0 引言

红外小目标检测(infrared small target detection,IRSTD)旨在从复杂背景中精确识别尺寸微小、对比度较低的目标(寇人可等,2024;潘晓英等,2023;刘忠林等,2019)。该任务在现实世界中具有广泛的应用价值(Yang等,2024a),已成功应用于海上监视、搜救以及早期预警系统等关键场景。在这些关键应用中,检测结果的准确性直接影响着系统对潜在威胁的响应速度与可靠性,进而关系到任务的成败与人员安全。因此,如何提升IRSTD模型在复杂场景下的检测精度与鲁棒性,使其具备更高的实用性与部署价值,多年来一直是人们关注的焦点。

传统的红外小目标检测方法通常将小目标检测视为滤波与目标增强的问题,主要依赖于人工设计的特征提取流程和精细调参的超参数设置。然而,这些方法在面对真实场景中目标尺度、形状及背景干扰高度多变的情况时,往往表现出灵活性不足和鲁棒性较差的缺点。近年来,卷积神经网络(convolutional neural network,CNN)等深度学习技术引入到IRSTD领域,在小目标检测精度方面取得了显著进展。例如,Dai等人(2021a)提出了第一种基于图像分割的检测方法,并设计了一种非对称调制特征融合结构以替代传统的U-Net框架(Ronneberger等,2015);Li等人(2023)开发了密集嵌套注意力网络,实现了红外小目标的多尺度特征融合;Zhang等人(2022)则提出了红外形状网络,在多个数据集上表现出优越的检测性能;Yang等人(2024b)提出了渐进式背景感知Transformer的方法,通过多阶段背景上下文细化参考机制,有效克服了特征退化与错误积累问题,显著增强了检测性能并抑制背景干扰;

Wu等人(2025)提出了可学习的局部显著性核模块,取代U-Net中传统的跨层特征融合,从而引导网络捕捉红外目标的显著特征。

尽管现有的方法推动了IRSTD的发展,但目标特征提取过程中的信息瓶颈依然是限制检测性能提升的关键因素之一,且当前的诸多研究更多是对局部问题的策略性补救,而未能从根本上解决红外小目标检测的核心挑战。这些方法在以下两个关键方面仍存在显著局限,亟需从结构设计与特征建模机制上进行系统性重构。1)现有网络架构中的初级卷积层在处理红外小目标时常表现出对细粒度信息感知不足的问题(Yu等,2021)。这种灵敏度缺失导致目标与复杂背景之间的特征对比度较低,限制了网络对微弱目标的区分能力,从而影响检测的精度和可靠性;2)当前主流网络中广泛采用的静态特征连接方式,缺乏针对显著性区域的动态调节能力。这种静态融合机制通常基于固定权重进行特征融合,难以自适应地区分红外图像中的真实目标与误报区域,进而限制了模型在复杂背景条件下的检测表现。

为克服现有方法在红外小目标检测中的局限性,本文提出了一种基于选择性注意力的红外小目标检测网络(selective attention-based network for infrared small target detection,SANet)。首先,为增强网络对微弱目标的感知能力,本文构建了双路径语义感知模块(dual-path semantic-aware module,DSM)作为主干网络的基础特征提取结构。该模块集成了标准卷积和风车型卷积的双路径分支,其中,标准卷积分支用于保留基础的判别能力,而风车型卷积(Yang等,2025)分支用于扩大感受野,从而增强对目标区域的响应能力。在双路径结构之后,进一步引入经典的空间注意力与通道注意力模块(convolutional block attention module,CBAM)(Woo等,2018),

用于细化关键信息的表达,并抑制背景噪声的干扰。其次,准确区分真伪目标需要对全局背景进行综合分析,而非仅依赖于局部区域的图像信息。为此,本文提出了选择性注意力融合模块(selective attention fusion module, SAFM),用于替代传统 U-Net 架构中静态的跳跃连接。SAFM 采用动态融合策略,结合选择性注意力机制与大感受野的特性,能够根据上下文信息自适应地整合浅层与深层特征。该机制显著提升了网络在复杂场景中对真实目标与背景干扰之间的判别能力,从而增强整体检测性能。

综上所述,本文的主要贡献如下:1)提出了一种基于选择性注意力机制的红外小目标检测方法,增强了模型在复杂背景中区分真实目标与误检的能力,有效提升了检测的鲁棒性与精度。2)构建了双路径语义感知模块,将标准卷积与风车型卷积相结合,借助更广的感受野实现对背景噪声的有效抑制,同时引入空间与通道注意力机制,进一步提升细节特征的表达能力。3)设计了选择性注意力融合模块,用于替代传统 U-Net 架构中的静态跳跃连接。该模块通过动态特征融合策略引导信息选择性传递,显著增强网络对误报的抑制能力。4)在 3 个公开数据集上的大量实验验证了所提出方法的有效性,其结果表明本文方法在多个性能指标上均优于现有最先进的检测方法,展示出良好的检测性能。

1 相关工作

1.1 基于传统方法的红外小目标检测

在红外小目标检测领域,传统方法长期以来构成了研究的基础与出发点。早期方法主要依赖图像处理 and 统计建模等人工设计策略,在低分辨率与低信噪比的背景下具有一定的鲁棒性。例如,基于滤波器的方法(Deng 等,2021)通过空间域或频率域滤波器抑制背景干扰,在背景较为均匀时表现良好,但在目标与背景频率成分重叠的复杂场景中往往效果有限。基于低秩表示的方法(Dai 和 Wu,2017;Sun 等,2021)将图像分解为低秩(背景)与稀疏(目标)矩阵,能够在低信噪比条件下有效增强目标,因此受到广泛关注。然而,此类方法难以有效区分稀疏噪声与真实目标,易导致较高的误报率。基于局部对比度的方法(Han 等,2020,2021)通过计算局部区域的强度差异以增强目标显著性,对小目标具有一定检

测能力,但其对固定窗口大小的依赖限制了对多尺度目标的适应性。

综上所述,传统方法大多依赖固定的人工特征提取与参数设定,难以充分建模复杂背景中目标的多样性与不确定性,导致其在实际应用中的检测性能与鲁棒性仍存在明显不足,也为后续基于深度学习的方法提供了广阔的发展空间。

1.2 基于深度学习的红外小目标检测

相比传统方法,基于数据驱动范式的深度学习技术在红外小目标检测领域取得了显著进展。MDvsFA (miss detection vs. false alarm)(Wang 等,2019)首次引入对抗学习机制,有效缓解了误检与漏检之间的矛盾。ACM (asymmetric contextual modulation)(Dai 等,2021a)基于 U-Net 架构,采用不对称上下文调制策略融合高层语义信息与底层精细节,从而显著提升了检测精度。ALCNet (attentional local contrast network)(Dai 等,2021b)以无参数形式将传统局部对比度测量模块嵌入网络结构,突破了小卷积核感受野有限的限制。UIU-Net (U-Net in U-Net)(Wu 等,2023)通过在 U-Net 主干网络中嵌入小型 U-Net 结构,实现了目标的多层次与多尺度表征学习。GCI-Net (Gaussian curvature inspired network)(Zhang 等,2024b)利用高斯曲率信息以增强目标区域的几何特征,进而提高检测性能。ABC (attention with bilinear correlation)(Pan 等,2023)融合卷积与线性结构,引入 Transformer 模块以增强特征提取与融合能力,不仅提升了目标表征效果,还有效抑制了背景噪声干扰。SCTransNet (spatial-channel cross Transformer network)(Yuan 等,2024)构建空间—通道交叉的 Transformer 模块,在不同语义层次上增强目标与背景杂波的可分性。CFD-Net (computational fluid dynamic network)(Zhang 等,2025b)借鉴计算流体力学思想,将红外图像特征的演化过程类比为像素粒子的流动,有效提升了目标的表征能力。HDNet (hybrid domain network)(Xu 等,2025)联合空间域的多尺度目标感知机制与频率域的低频信息抑制策略,进一步提高了检测的准确性与鲁棒性。

近年来,通用分割模型被逐步引入红外小目标检测领域,并展现出良好性能。IRSAM (advancing segment anything model for infrared small target detection)(Zhang 等,2024a)首次将“任意分割模型”(Kirillov 等,2023)应用于红外小目标检测任务,并

结合基于小波的 Perona-Malik 扩散模块,以增强边缘特征的提取效果。状态空间序列模型(state space model,SSM)(Gu 等,2021)在长序列建模方面表现出了较大潜力,为解决视觉任务中的长期依赖问题提供了有效替代方案。作为SSM架构的代表性方法,Mamba(Gu 和 Dao,2023)受到了广泛关注。MiM-ISTD(Mamba-in-Mamba for efficient infrared small-target detection)(Chen 等,2024)是首个将Mamba架构应用于红外小目标检测的工作。在此基础上,IRMamba(pixel difference Mamba with layer restoration for infrared small target detection)(Zhang 等,2025a)进一步通过建模扫描位置及其中心邻域之间的像素差异(包括强度与方向信息),将这些差异信息融入状态空间的状态更新方程中,从而有效增强了目标细节的表征能力。

2 本文方法

2.1 总体网络结构

本文提出的SANet整体架构如图1所示,它基

于经典的U-Net结构构建,并通过集成两个关键模块——双路径语义感知模块和选择性注意力融合模块以增强其检测性能。

在编码器路径中,网络由5个编码单元构成。每个编码单元由一个双路径语义感知模块和一个最大池化操作组成,逐层减小特征图的空间分辨率。双路径语义感知模块用于融合更大感受野下的上下文信息,同时通过CBAM(Woo 等,2018),有效增强目标特征表达并抑制背景干扰,从而提升网络对微弱小目标的感知能力。

在解码器路径中,网络包含4个选择性注意力融合模块与对应的上采样卷积层。其中,选择性注意力融合模块用于动态融合来自编码器与解码器不同尺度的特征信息,自适应加强显著区域的响应,提升对真实小目标与误报区域的辨别能力。上采样卷积部分则保持U-Net中原有的标准卷积设计,用于逐步恢复空间分辨率。

最终,解码器输出的特征图送入到一个由 3×3 卷积和sigmoid激活函数组成的检测头,用于生成最终的目标预测结果。

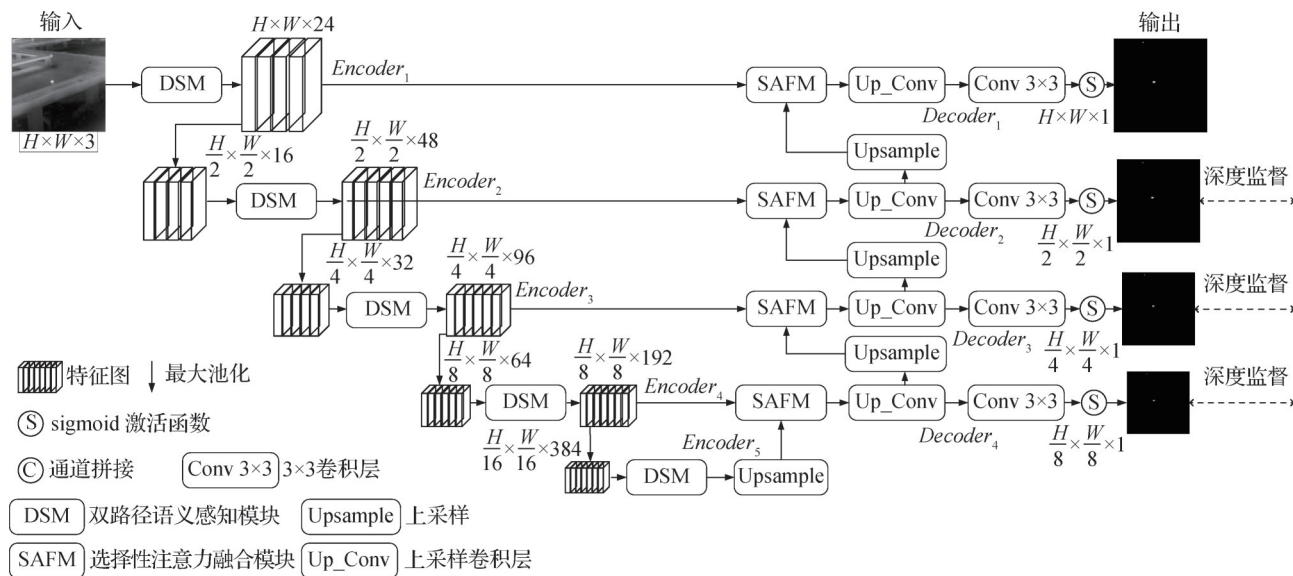


图1 SANet网络整体架构

Fig. 1 Overall framework of the SANet network

2.2 双路径语义感知模块

卷积神经网络中的传统卷积层受限于其固定且有限的感受野,难以有效建模长距离的特征依赖关系,进而限制了对红外小目标等微弱目标的感知能力。针对这一问题,本文提出了一种双路径语义感知模块,旨在增强网络的基础特征提取能力。

如图2所示,DSM由两个并行分支构成,协同提升特征表达能力。第1分支采用标准卷积(卷积核大小均为 3×3),以保持局部空间结构的一致性,确保基础的空间细节信息得以保留;第2分支则引入风车型卷积结构,依次采用 3×3 、 5×5 和 3×3 的卷积核组合,扩展感受野,从而更全面地捕捉上下文信息。

风车型卷积的结构如图3所示。与传统卷积不同,该分支通过不对称填充构造方向敏感的水平与垂直卷积核,使网络能够感知不同方向上的目标形态,提升对目标形变的适应能力。此外,与标准的膨胀卷积等其他感受野扩展技术相比,风车型卷积拥有更少的参数量和计算开销。该多尺度扩展机制使得DSM可在不同空间尺度下建模上下文语义,显著丰富了特征表达。随后,两个分支提取的特征图通过一个 3×3 卷积进行融合,生成统一的聚合特征表示。

在此基础上,DSM利用CBAM来细化融合后的特征。如图4所示,CBAM由级联的空间和通道注意力两个模块组成。空间注意力模块通过建模任意空间位置间的长程依赖关系,增强空间结构信息的表

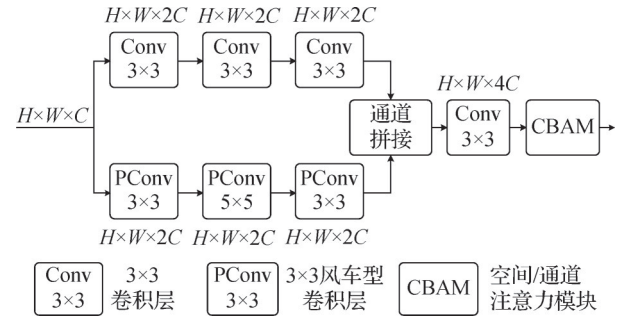


图2 双路径语义感知模块结构

Fig. 2 Dual-path semantic perception module structure

达;而通道注意力模块则从全局角度建模通道之间的相关性,突出与小目标检测任务高度相关的语义通道,从而提升特征的判别能力。

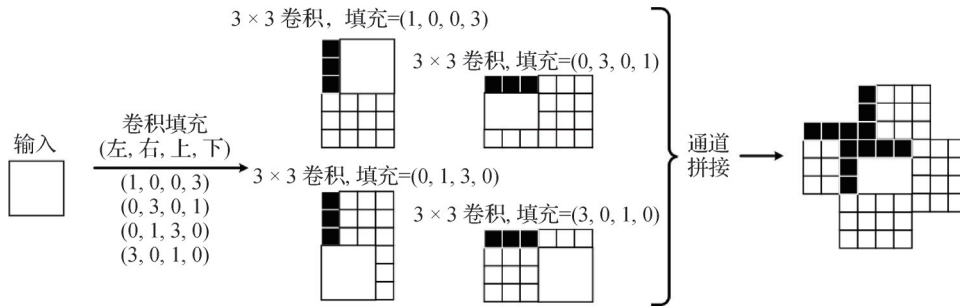


图3 风车型卷积结构

Fig. 3 Pinwheel-shaped convolution structure

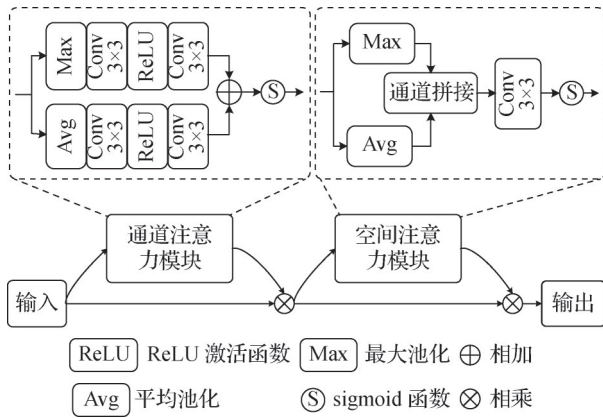


图4 卷积块注意力模块结构

Fig. 4 Convolutional block attention module structure

2.3 选择性注意力融合模块

为提升网络在小目标检测中对真伪目标的判别能力,本文提出了一种选择性注意力融合模块。该模块通过在U-Net结构中引入动态空间选择机制,有效利用多尺度特征,使网络能够更加关注空间中信息密度高、与目标相关性强的区域,如图5

所示。

具体而言,SAFM在第 i 个阶段对编码器与解码器输出的特征映射进行融合处理,分别记为 $E_i, D_i \in \mathbb{R}^{C_i \times \frac{H}{2^i} \times \frac{W}{2^i}}$,其中 C_i 表示通道数, H, W 为原始红外图像的高与宽。首先对编码器和解码器的特征进行通道维度的拼接,得到融合输入,具体为

$$X_i = \text{Concat}(E_i, D_i) \quad (1)$$

为了减少计算复杂度,本文对融合特征 X_i 分别施加通道平均池化 $P_{\text{avg}}(\cdot)$ 与最大池化操作 $P_{\text{max}}(\cdot)$,表示为

$$\begin{cases} SF_i^{\text{avg}} = P_{\text{avg}}(X_i) \\ SF_i^{\text{max}} = P_{\text{max}}(X_i) \end{cases} \quad (2)$$

式中, SF_i^{avg} 和 SF_i^{max} 分别表示池化后的空间特征表示。随后将两者连接(Concat),形成融合的空间表示矩阵,以促进不同空间注意信息的互补与交互,具体为

$$SAF_i = f_{7 \times 7}(\text{Concat}(SF_i^{\text{avg}}, SF_i^{\text{max}})) \quad (3)$$

接着,通过将生成的空间注意力特征图 SAF_i 分

别与编码器特征 E_i 和解码器特征 D_i 进行逐元素加权融合, 得到融合表示 Y_i , 具体为

$$Y_i = E_i \otimes \sigma(SAF_i) + D_i \otimes \sigma(1 - SAF_i) \quad (4)$$

式中, σ 表示 sigmoid 激活函数, $\otimes$ 表示逐元素乘法操作。

最后, 将 Y_i 送入一个 1×1 卷积层以扩展其通道表达能力, 再与原始融合特征 X_i 逐元素相乘, 并引入

一个可学习的标量权重 λ , 形成最终输出, 具体为

$$F_{SAFM} = \lambda \otimes (X_i \times f_{1 \times 1}(Y_i)) + X_i \quad (5)$$

式中, λ 是一个全局单一的可学习标量参数, 在整个 SAFM 模块中统一使用, 并可在网络训练中实现自动更新与优化。该参数用于控制注意力增强特征与残差信息之间的融合强度, 既能保留原始特征结构, 又能有效强化关键区域的判别性表达。

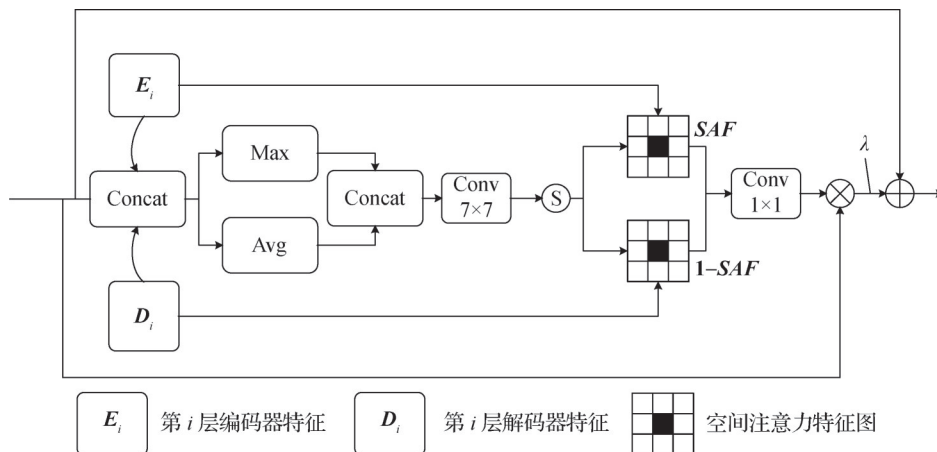


图5 选择性注意力融合模块结构

Fig. 5 Selective attention fusion module structure

3 实验结果与分析

3.1 数据集与实验设置

为了全面评估所提出方法的有效性和优越性, 在3个公开的基准数据集 NUAA-SIRST (Nanjing University of Aeronautics and Astronautics—single-frame infrared small target) (Dai 等, 2021a)、IRSTD-1K (infrared small target detection) (Zhang 等, 2022) 和 NUDT-SIRST (National University of Defense Technology) (Li 等, 2023) 上进行实验。这3个数据集分别包含427幅、1 001幅和1 327幅标注完备的红外图像, 涵盖多种典型红外场景, 确保了评估环境的严谨性和代表性。在实验设置方面, 3个数据集的训练集与测试集的比例分别为: NUAA-SIRST 为 1:1、IRSTD-1K 为 4:1、NUDT-SIRST 为 1:1。此外, 在 NUAA-SIRST 和 IRSTD-1K 数据集上的训练图像尺寸统一为 320×320 像素, 而在 NUDT-SIRST 数据集上设置为 256×256 像素。对应的数据集训练批次大小 (batch size) 分别为 4、12 和 16。

本文采用 Soft-IoU 损失函数 (Rahman 和 Wang,

2016) 训练网络, 并选用 AdamW 优化器, 训练轮数设为 600 轮 (epoch)。初始学习率为 0.001, 权重衰减为 1×10^{-2} 。为确保模型稳定收敛, 训练过程中引入余弦退火策略, 将学习率逐步衰减为 1×10^{-5} 。同时, 本文中显著图的分割阈值统一设定为 0.5。

3.2 评估指标

为了全面评估检测性能, 在实验中采用了4个广泛使用的指标: 交并比 (intersection over union, IoU)、归一化交并比 (normalized intersection over union, nIoU)、检测概率 (P_d) 和误报率 (F_a), 分别定义分别为

$$f_{IoU} = \frac{TP}{T + P - TP} \quad (6)$$

$$f_{nIoU} = \frac{1}{N} \sum_i \frac{TP(i)}{T(i) + P(i) - TP(i)} \quad (7)$$

$$P_d = \frac{TP}{TP + FN} \quad (8)$$

$$F_a = \frac{FP}{FP + TN} \quad (9)$$

式中, N 为样本总数, T 和 P 分别表示真实值和预测正像素的数量。 TP 、 FP 和 FN 分别表示真正像素、

假正像素和假负像素的数量。

3.3 定量评估

本文将所提出的方法与14种现有的红外小目标检测方法进行了比较。其中,包括5种传统方法FKRW (facet kernel and random walker) (Qin 等, 2019)、TLLCM (tri-layer local contrast method) (Han 等, 2020)、NWMTH (new method top-hat) (Bai 和 Zhou, 2010)、NOLC (non-convex optimization with L_{p_1} norm constraint) (Zhang 等, 2019b)、PSTNN (partial sum of the tensor nuclear norm) (Zhang 和 Peng, 2019a) 和9种基于深度学习的方法ACM (asymmetric contextual modulation) (Dai 等, 2021a)、RDIAN (receptive-field and direction induced attention network) (Sun 等,

2023)、AGPCNet (attention-guided pyramid context) (Zhang 等, 2023)、DNANet (dense nested attention) (Li 等, 2023)、UIUNet (U-Net in U-Net) (Wu 等, 2023)、RPCANet (robust principal component analysis) (Wu 等, 2024)、MSHNet (multi-scale head) (Liu 等, 2024)、SCTransNet (spatial-channel cross transformer network) (Yuan 等, 2024) 和 L^2 SKNet (learnable local saliency kernel network) (Wu 等, 2025)。另外,在ACM和 L^2 SKNet中用于比较的网络主干是U-Net。

为了确保公平比较,在训练过程中严格遵循原始论文中描述的设置,以相同的训练基准重新训练所有深度学习模型。实验结果如表1所示。

表1 NUAA-SIRST、IRSTD-1K和NUDT-SIRST数据集上不同方法客观指标

Table 1 The objective metrics for different methods on the NUAA-SIRST, IRSTD-1K, and NUDT-SIRST datasets

方法	NUAA-SIRST				IRSTD-1K				NUDT-SIRST			
	IoU/%	nIoU/%	P_d /%	$F_d/10^{-6}$	IoU/%	nIoU/%	P_d /%	$F_d/10^{-6}$	IoU/%	nIoU/%	P_d /%	$F_d/10^{-6}$
FKRW	22.06	28.08	81.77	16.32	10.39	16.25	69.54	24.37	12.67	21.73	79.51	67.13
TLLCM	18.66	27.71	82.24	16.95	10.26	18.24	68.55	24.48	11.08	23.82	77.97	67.26
NWMTH	15.77	17.28	71.12	55.61	18.94	16.91	49.26	21.72	11.72	10.26	52.71	46.81
NOLC	26.64	35.68	81.62	17.44	12.39	22.18	75.38	21.94	23.87	34.90	85.47	58.20
PSTNN	22.40	29.59	77.95	29.11	24.57	28.71	71.99	35.26	27.72	39.80	66.13	44.17
ACM	65.28	65.67	90.49	47.13	58.84	58.23	91.92	27.59	60.63	60.94	91.75	20.06
RDIAN	70.66	74.04	93.92	43.08	63.37	62.63	88.55	30.18	77.77	79.46	95.66	26.36
AGPCNet	75.69	76.60	96.48	14.99	66.29	65.23	92.83	13.12	88.87	90.64	97.20	10.02
DNANet	76.34	79.19	95.82	15.23	66.50	66.13	91.58	18.31	91.96	92.88	98.94	9.28
UIUNet	77.17	78.86	95.44	14.34	65.34	65.55	90.91	15.01	90.43	89.87	98.94	7.72
RPCANet	53.93	61.01	95.44	149.20	62.09	61.48	89.35	48.61	87.79	89.80	95.98	32.03
MSHNet	73.88	72.56	95.82	19.25	67.17	60.38	92.52	12.60	76.95	79.72	95.77	20.52
SCTransNet	75.46	78.06	95.82	18.80	66.54	66.36	93.27	13.68	93.01	93.50	98.73	4.80
L^2 SKNet	72.90	73.89	93.92	40.34	67.38	66.36	91.58	14.78	93.52	92.86	97.57	5.29
SANet	79.10	80.01	96.58	11.52	71.70	66.59	93.27	11.88	95.73	95.43	99.26	2.80

注:加粗字体表示各列最优结果。

表1揭示了以下观察结果:1)本文方法在3个数据集的所有指标上始终保持优异性能,超越其他检测方法,这凸显了双路径语义感知模块和选择性注意力融合模块能够在复杂背景中区分真伪目标,最终有效提升了检测精度。2)基于深度学习的方法显著优于传统方法。这种差异凸显了传统方法的局限

性,它们依赖于人工设计的特征和先验知识,难以对图像的语义内容进行建模。相比之下,深度学习方法擅长提取语义信息,同时最大限度地减少超参数依赖性,从而对场景变化具有更强的鲁棒性。3)提出的SANet在3个数据集上均始终优于最新的检测方法 L^2 SKNet。 L^2 SKNet通过其可学习的局部显著性

核模块有效地捕捉红外目标显著性,而本文方法集成了目前最新的基础特征提取模块和选择性注意力融合模块,从而实现更准确的小目标检测和更优越的整体性能。

3.4 定性评估

图6展示了不同IRSTD方法在3个基准数据集上的定性比较结果。包括 NUAA-SIRST(第1、2行)、IRSTD-1K(第3、4行)和 NUDT-SIRST(第5、6行)。其中,红色和黄色框分别代表正确检测的目标和误检。

与传统方法相比,如第2、3、4行所示,基于深度学习的检测方法以其数据驱动的特征提取能力,在大多数情况下可以显著降低误检。虽然其他主流方法在检测性能上具有竞争力,但在大多数情况下它们的表现仍然不如 SANet。在复杂场景中,如第3行所示,除 UIUNet和本文方法外,其他深度学习方法均存在误检现象;但同样是复杂场景(第4行)的图中,UIUNet出现了大量误检。而本文方法在不同场景下始终保持较高的检测精度和鲁棒性,充分验证了其优越性。

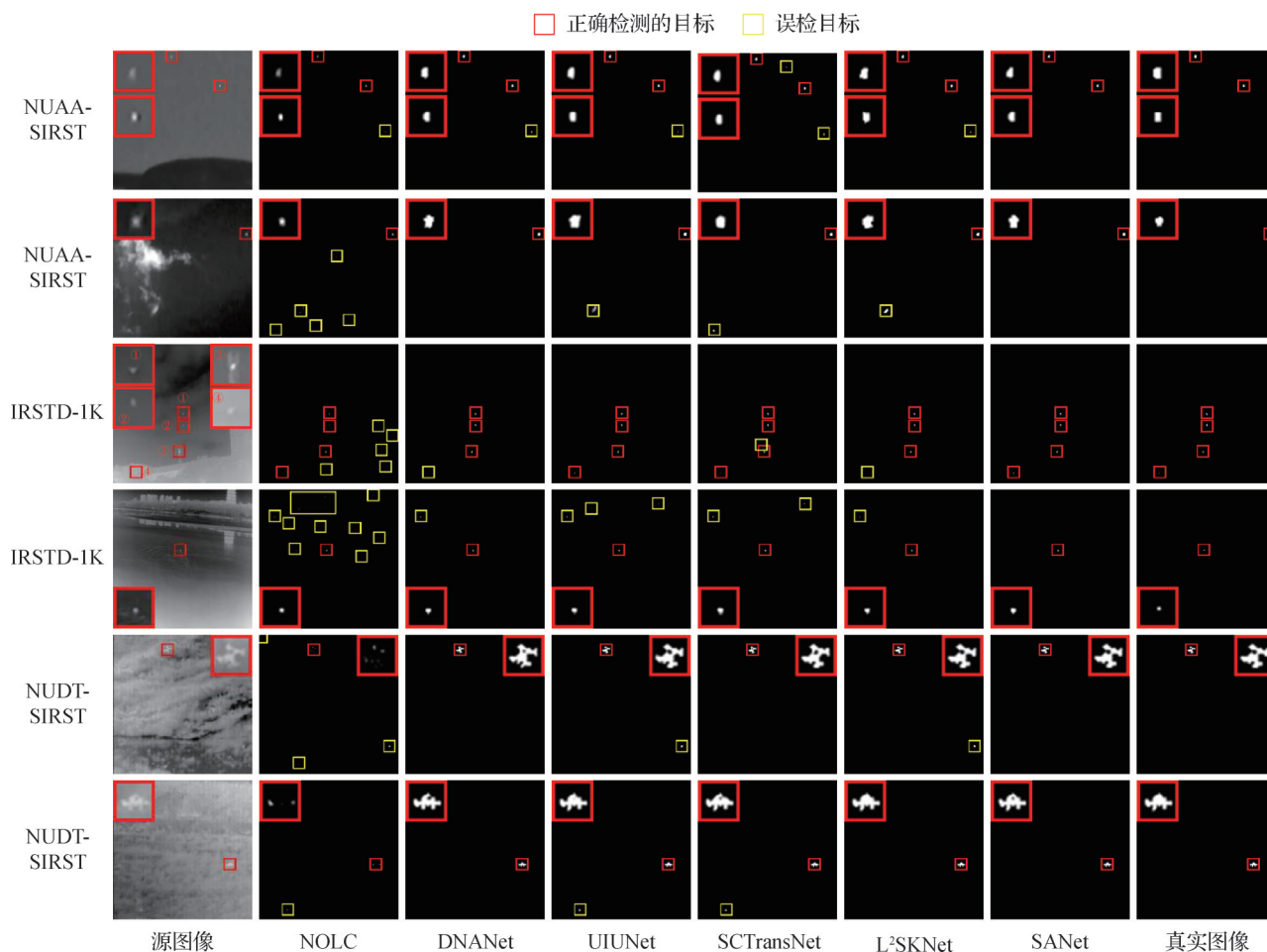


图6 在 NUAA-SIRST、IRSTD-1K 和 NUDT-SIRST 数据集上对不同检测方法的可视化

Fig. 6 Visualization of different detection methods on the NUAA-SIRST, IRSTD-1K, and NUDT-SIRST datasets

此外,图7中的3D可视化结果进一步直观地展示了不同深度学习方法检测性能。其中,红色和黄色框分别代表正确检测的目标和误检。从图7中可以看出,当目标数量较多且背景复杂时,其他检测方法难以稳定地检测目标,尤其是 UIUNet,在第2种场景下产生了大量的误检。这一现象进一步证明了本文方法的鲁棒性和在复杂环境下的卓

越性能。

3.5 消融实验

3.5.1 SANet中核心组件的有效性

本文以经典的 U-Net 架构作为基准模型,逐步引入提出的两个关键模块。通过这种逐层集成的策略,能够系统性地评估各个模块对整体检测性能的独立贡献及其协同效应。实验结果如表2所示,随

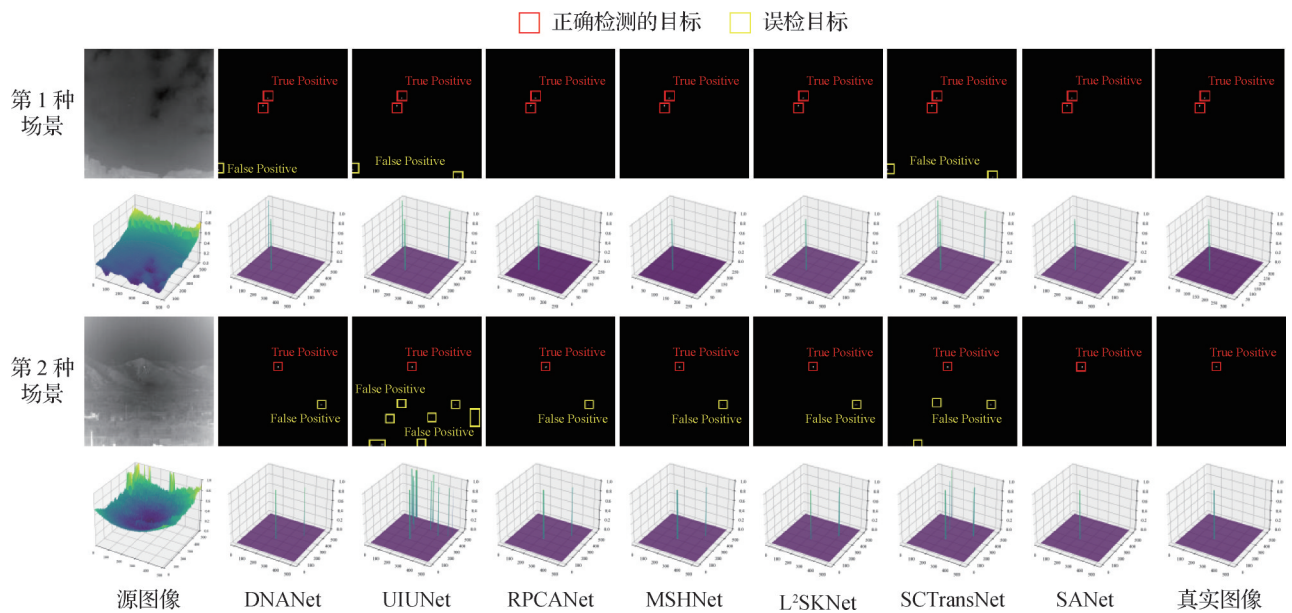


图7 IRSTD-1K数据集上不同IRSTD方法的3D可视化
Fig. 7 3D visualization of differentIRSTD methods on theIRSTD-1K dataset

着DSM和SAFM模块逐步集成至基线模型,网络的检测性能持续提升,且每个模块在独立工作时亦表现出显著的性能增益,进一步验证了本文的设计动机与方法有效性。具体而言,在仅集成DSM的情况下,3个公开数据集上的IoU指标分别提升了3.32%、6.71%和9.63%,充分说明DSM在基础特征

提取阶段增强了网络对红外小目标细节的捕捉与建模能力。进一步地,随着SAFM模块的引入,模型在全部评估指标上均达到了最优表现。SAFM通过引入动态空间选择机制,实现了对多尺度特征的自适应融合,有效提升了网络对真实目标与误报之间的判别能力,显著增强了整体检测性能。

表2 定量消融实验评估SANet中核心组件DSM和SAFM的影响													
Table 2 Quantitative ablation experiments to evaluate the impact of the core components DSM and SAFM in SANet													
DSM	SAFM	NUAA-SIRST数据集				IRSTD-1K数据集				NUDT-SIRST数据集			
		IoU/%	nIoU/%	P_d /%	$F_a/\times 10^{-6}$	IoU/%	nIoU/%	P_d /%	$F_a/\times 10^{-6}$	IoU/%	nIoU/%	P_d /%	$F_a/\times 10^{-6}$
×	×	74.91	79.01	95.44	28.33	64.75	63.24	91.58	18.77	85.42	85.58	98.10	6.99
√	×	78.23	79.51	95.06	13.86	71.46	67.08	90.57	16.25	95.05	94.94	98.84	2.85
√	√	79.10	80.01	96.58	11.52	71.70	66.59	93.27	11.88	95.73	95.43	99.26	2.80

注:加粗字体表示各列最优结果。√表示含有该模块,×表示不含有该模块。

3.5.2 双路径语义感知模块的有效性

DSM是在传统特征提取结构基础上的改进,融合了风车型卷积和空间/通道注意力机制,旨在生成更鲁棒、更具判别性的特征表示。为全面评估DSM各组成部分的有效性,本文对其中各子模块的独立贡献进行了消融实验,实验结果如表3所示。该结果表明,风车型卷积的引入在一定程度上提升了检测性能,表明其通过扩大感受野有效增强了对微弱目标的感知能力,尤其在降低漏检率方面具有积极

作用。然而,感受野的扩大也带来了额外的背景干扰,导致一定程度上的误检增加。这说明,虽然大感受野有助于捕捉远距离上下文信息,但若缺乏精细提取,容易引入冗余或噪声信息。相比之下,空间和通道注意力机制的集成对检测性能的提升更为显著。注意力机制通过自适应地调整不同空间位置和通道维度的响应强度,使模型更聚焦于潜在目标区域,从而显著提高了特征的判别性与目标响应强度,有效抑制了背景噪声带来的干扰。

表 3 定量消融实验评估 DSM 中核心组件的影响

Table 3 Quantitative ablation experiments to evaluate the impact of core components in DSM

Conv	PConv	CBAM	NUAA-SIRST 数据集				IRSTD-1K 数据集				NUDT-SIRST 数据集			
			IoU/%	nIoU/%	P_d /%	$F_a \times 10^{-6}$	IoU/%	nIoU/%	P_d /%	$F_a \times 10^{-6}$	IoU/%	nIoU/%	P_d /%	$F_a \times 10^{-6}$
√	×	×	74.90	77.53	93.16	35.74	66.99	64.03	91.25	34.56	91.28	91.09	98.41	4.71
√	√	×	75.26	78.84	93.92	25.57	67.84	64.41	93.27	28.45	92.95	92.56	99.05	6.96
√	√	√	79.10	80.01	96.58	11.52	71.70	66.59	93.27	11.88	95.73	95.43	99.26	2.80

注:加粗字体表示各列最优结果。√表示含有该模块,×表示不含有该模块。

3.5.3 选择性注意力融合模块的有效性

SAFM 基于空间动态权重机制实现跨尺度特征的自适应融合,能够根据空间显著性选择性增强关键区域。为全面评估 SAFM 各组成部分的有效性,本文对其中各子模块的独立贡献进行了消融实验,实验结果如表 4 所示。从表中可以看出,去除残差连接(residual)或可学习标量参数(λ)都会导致性能下降。在 3 个数据集的性能表现上,完整模型在 3 个数据集上的 IoU 分别达到 79.10%、71.70% 和

95.73%,均优于去除任一模块的情况,说明两个子模块在提升检测精度方面具有互补性。此外,引入可学习标量参数时普遍带来更高的 IoU 和更低的误检率,如在 NUAA-SIRST 数据集上 F_a 从 13.10×10^{-6} 降至 11.52×10^{-6} ,验证了选择性注意力融合的有效性。而残差连接的引入则有助于保持目标信息的完整性,并在多个数据集上提升了检测准确率,尤其是在 IRSTD-1K 数据集上提升明显(P_d 从 89.56% 提升至 93.27%)。

表 4 定量消融实验评估 SAFM 中核心组件的影响

Table 4 Quantitative ablation experiments to evaluate the impact of core components in SAFM

残差连接	λ	NUAA-SIRST 数据集				IRSTD-1K 数据集				NUDT-SIRST 数据集			
		IoU/%	nIoU/%	P_d /%	$F_a \times 10^{-6}$	IoU/%	nIoU/%	P_d /%	$F_a \times 10^{-6}$	IoU/%	nIoU/%	P_d /%	$F_a \times 10^{-6}$
×	√	78.80	80.70	96.58	13.93	70.99	65.58	89.56	13.49	95.23	94.94	99.05	2.71
√	×	78.13	79.00	95.82	13.10	70.04	67.13	91.92	26.85	95.11	94.72	98.94	1.54
√	√	79.10	80.01	96.58	11.52	71.70	66.59	93.27	11.88	95.73	95.43	99.26	2.80

注:加粗字体表示各列最优结果。√表示含有该模块,×表示不含有该模块。

3.6 模型复杂度分析

本文对 SANet 与近年来具有代表性的红外小目标检测模型(包括 DNANet、UIUNet、SCTransNet 和 L²SKNet)在模型复杂度与检测性能方面进行了系统对比分析,结果如图 8 所示,其中 IoU 指标为在 3 个数据集的平均结果。从对比结果可见,SANet 在 3 个基准数据集上取得了最高的平均 IoU 值(82.18%),显著优于其余对比方法,性能提升幅度达到 3.84%~4.56%,充分展现了其出色的检测能力。在模型复杂度方面,SANet 的参数量为 12.72 M,计算量为 19.75 G。尽管在 5 种方法中并非最轻量级模型,但与 UIUNet 和 SCTransNet 相比,其计算成本处于相对适中的水平,仍具有一定的效率优势。

综上所述,SANet 在检测精度和模型复杂度之

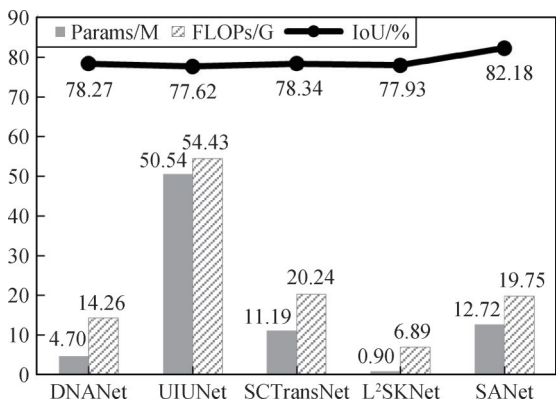


图 8 5 种算法在 3 个指标上的性能评估

Fig. 8 Performance evaluation of five algorithms across three metrics

间实现了良好的权衡。尽管模型结构相对略为复杂,但凭借其出色的检测性能,仍在实际应用中具备

较高的实用性和推广潜力。

4 结 论

针对红外小目标检测中存在的目标弱小、背景复杂和误报率高等挑战,本文提出了一种基于选择性注意力机制的红外小目标检测网络——SANet。通过引入双路径语义感知模块和选择性注意力融合模块,有效增强了网络对小目标的感知能力、关键特征的表达能力以及背景干扰的抑制能力。

具体而言,DSM模块利用标准卷积与风车型卷积的互补特性,在保留基础判别能力的同时扩大感受野,实现对复杂背景下小目标的精准响应;SAFM模块则通过动态融合策略,替代传统U-Net中的静态跳跃连接,实现更为智能的特征融合,增强网络对真实目标与误报的区分能力。

在多个公开数据集上的实验结果表明,所提出的SANet在检测精度、鲁棒性以及泛化能力方面均优于现有主流方法,展现出良好的实用价值与应用前景。未来工作将聚焦于模型轻量化及复杂场景下的检测性能提升,进一步拓展其在红外图像不同领域的应用范围。

参考文献(References)

- Bai X Z and Zhou F G. 2010. Analysis of new top-hat transformation and the application for infrared dim small target detection. *Pattern Recognition*, 43 (6) : 2145-2156 [DOI: 10.1016/j.patcog.2009.12.023]
- Chen T X, Ye Z, Tan Z T, Gong T, Wu Y, Chu Q, et al. 2024. MiM-ISTD: Mamba-in-Mamba for efficient infrared small-target detection. *IEEE Transactions on Geoscience and Remote Sensing*, 62: #5007613 [DOI: 10.1109/TGRS.2024.3485721]
- Dai Y M and Wu Y Q. 2017. Reweighted infrared patch-tensor model with both nonlocal and local priors for single-frame small target detection. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 10(8): 3752-3767 [DOI: 10.1109/JSTARS.2017.2700023]
- Dai Y M, Wu Y Q, Zhou F and Barnard K. 2021a. Asymmetric contextual modulation for infrared small target detection//*Proceedings of 2021 IEEE Winter Conference on Applications of Computer Vision*. Waikoloa, USA: IEEE: 949-958 [DOI: 10.1109/WACV48630.2021.00099]
- Dai Y M, Wu Y Q, Zhou F and Barnard K. 2021b. Attentional local contrast networks for infrared small target detection. *IEEE Transactions*

- on *Geoscience and Remote Sensing*, 59(11): 9813-9824 [DOI: 10.1109/TGRS.2020.3044958]
- Deng L Z, Zhang J K, Xu G X and Zhu H. 2021. Infrared small target detection via adaptive M-estimator ring top-hat transformation. *Pattern Recognition*, 112: #107729 [DOI: 10.1016/j.patcog.2020.107729]
- Gu A and Dao T. 2023. Mamba: linear-time sequence modeling with selective state spaces [EB/OL]. [2025-07-21]. <https://arxiv.org/pdf/2312.00752.pdf>
- Gu A, Goel K and Ré C. 2021. Efficiently modeling long sequences with structured state spaces [EB/OL]. [2025-07-21]. <https://arxiv.org/pdf/2111.00396.pdf>
- Han J H, Moradi S, Faramarzi I, Liu C Y, Zhang H H and Zhao Q. 2020. A local contrast method for infrared small-target detection utilizing a tri-layer window. *IEEE Geoscience and Remote Sensing Letters*, 17(10): 1822-1826 [DOI: 10.1109/LGRS.2019.2954578]
- Han J H, Moradi S, Faramarzi I, Zhang H H, Zhao Q and Zhao X J. 2021. Infrared small target detection based on the weighted strengthened local contrast measure. *IEEE Geoscience and Remote Sensing Letters*, 18(9): 1670-1674 [DOI: 10.1109/LGRS.2020.3004978]
- Kirillov A, Mintun E, Ravi N, Mao H Z, Rolland C, Gustafson L, et al. 2023. Segment anything//*Proceedings of 2023 IEEE/CVF International Conference on Computer Vision*. Paris, France: IEEE: 3992-4003 [DOI: 10.1109/ICCV51070.2023.00371]
- Kou R K, Wang C P, Luo Y, Zhang Y, Xu Z L, Peng Z M, et al. 2024. Multiscale small-target detection techniques in single-frame infrared images: a review. *Journal of Image and Graphics*, 29(9): 2625-2649 (寇人可, 王春平, 罗迎, 张勇, 徐泽龙, 彭真明, 等. 2024. 单帧红外图像多尺度小目标检测技术综述. *中国图象图形学报*, 29(9): 2625-2649) [DOI: 10.11834/jig.230788]
- Li B Y, Xiao C, Wang L G, Wang Y Q, Liu Z P and Li M. 2023. Dense nested attention network for infrared small target detection. *IEEE Transactions on Image Processing*, 32: 1745-1758 [DOI: 10.1109/TIP.2022.3199107]
- Liu Q K, Liu R, Zheng B L, Wang H K and Fu Y. 2024. Infrared small target detection with scale and location sensitivity//*Proceedings of 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Seattle, USA: IEEE: 17490-17499 [DOI: 10.1109/CVPR52733.2024.01656]
- Liu Z L, Wu Y Q and Zou Y. 2019. Multiscale infrared superpixel-image model for small-target detection. *Journal of Image and Graphics*, 24(12): 2159-2173 (刘忠林, 吴一全, 邹宇. 2019. 多尺度红外超像素图像模型的小目标检测. *中国图象图形学报*, 24(12): 2159-2173) [DOI: 10.11834/jig.190068]
- Pan P W, Wang H, Wang C Y and Nie C. 2023. ABC: attention with bilinear correlation for infrared small target detection//*Proceedings of 2023 IEEE International Conference on Multimedia and Expo*. Brisbane, Australia: IEEE: 2381-2386 [DOI: 10.1109/ICME55011.

- 2023.00406]
- Pan X Y, Jia N X, Mu Y Z and Gao X R. 2023. Survey of small object detection. *Journal of Image and Graphics*, 28(9): 2587-2615 (潘晓英, 贾凝心, 穆元震, 高炫蓉. 2023. 小目标检测研究综述. *中国图象图形学报*, 28(9): 2587-2615) [DOI: 10.11834/jig.220455]
- Qin Y, Bruzzone L, Gao C Q and Li B. 2019. Infrared small target detection based on facet kernel and random walker. *IEEE Transactions on Geoscience and Remote Sensing*, 57(9): 7104-7118 [DOI: 10.1109/TGRS.2019.2911513]
- Rahman A and Wang Y. 2016. Optimizing intersection-over-union in deep neural networks for image segmentation//*Proceedings of the 12th International Symposium on Visual Computing*. Las Vegas, USA: Springer: 234-244 [DOI: 10.1007/978-3-319-50835-1_22]
- Ronneberger O, Fischer P and Brox T. 2015. U-Net: convolutional networks for biomedical image segmentation//*Proceedings of the 18th International Conference on Medical Image Computing and Computer-Assisted Intervention*. Munich, Germany: Springer: 234-241 [DOI: 10.1007/978-3-319-24574-4_28]
- Sun H, Bai J X, Yang F and Bai X Z. 2023. Receptive-field and direction induced attention network for infrared dim small target detection with a large-scale dataset IRDST. *IEEE Transactions on Geoscience and Remote Sensing*, 61: #5000513 [DOI: 10.1109/TGRS.2023.3235150]
- Sun Y, Yang J G and An W. 2021. Infrared dim and small target detection via multiple subspace learning and spatial-temporal patch-tensor model. *IEEE Transactions on Geoscience and Remote Sensing*, 59(5): 3737-3752 [DOI: 10.1109/TGRS.2020.3022069]
- Wang H, Zhou L P and Wang L. 2019. Miss detection vs. false alarm: adversarial learning for small object segmentation in infrared images//*Proceedings of 2019 IEEE/CVF International Conference on Computer Vision*. Seoul, Korea (South): IEEE: 8508-8517 [DOI: 10.1109/ICCV.2019.00860]
- Woo S, Park J, Lee J Y and Kweon I S. 2018. CBAM: convolutional block attention module//*Proceedings of the 15th European Conference on Computer Vision*. Munich, Germany: Springer: 3-19 [DOI: 10.1007/978-3-030-01234-2_1]
- Wu F Y, Liu A R, Zhang T F, Zhang L P, Luo J H and Peng Z M. 2025. Saliency at the helm: steering infrared small target detection with learnable kernels. *IEEE Transactions on Geoscience and Remote Sensing*, 63: #5000514 [DOI: 10.1109/TGRS. 2024. 3521947]
- Wu F Y, Zhang T F, Li L, Huang Y and Peng Z M. 2024. RPCANet: deep unfolding RPCA based infrared small target detection//*Proceedings of 2024 IEEE/CVF Winter Conference on Applications of Computer Vision*. Waikoloa, USA: IEEE: 4797-4806 [DOI: 10.1109/WACV57701.2024.00474]
- Wu X, Hong D F and Chanussot J. 2023. UIU-Net: U-Net in U-Net for infrared small object detection. *IEEE Transactions on Image Processing*, 32: 364-376 [DOI: 10.1109/TIP.2022.3228497]
- Xu M Z, Yu C L, Li Z X, Tang H Y, Hu Y P and Nie L Q. 2025. HDNet: a hybrid domain network with multiscale high-frequency information enhancement for infrared small-target detection. *IEEE Transactions on Geoscience and Remote Sensing*, 63: #5004115 [DOI: 10.1109/TGRS.2025.3574962]
- Yang D G, Han T H, Hu L and Bai Z Y. 2024a. Research status and prospect of single frame infrared dim small target detection technology. *Journal of Signal Processing*, 40(5): 887-906 (杨德贵, 韩同欢, 胡亮, 白正阳. 2024. 单帧红外弱小目标检测技术研究现状与展望. *信号处理*, 40(5): 887-906) [DOI: 10.16798/j.issn.1003-0530.2024.05.008]
- Yang H R, Mu T K, Dong Z Y, Zhang Z C, Wang B and Ke W. 2024b. PBT: progressive background-aware transformer for infrared small target detection. *IEEE Transactions on Geoscience and Remote Sensing*, 62: #5004513 [DOI: 10.1109/TGRS.2024.3415080]
- Yang J N, Liu S L, Wu J J, Su X Y, Hai N and Huang X L. 2025. Pinwheel-shaped convolution and scale-based dynamic loss for infrared small target detection//*Proceedings of the 39th AAAI Conference on Artificial Intelligence*. Vancouver, USA: AAAI: 9202-9210 [DOI: 10.1609/aaai.v39i9.32996]
- Yu Z T, Wan J, Qin Y X, Li X B, Li S Z and Zhao G Y. 2021. NAS-FAS: static-dynamic central difference network search for face anti-spoofing. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 43(9): 3005-3023 [DOI: 10.1109/TPAMI.2020.3036338]
- Yuan S, Qin H L, Yan X, Akhtar N and Mian A. 2024. SCTransNet: spatial-channel cross transformer network for infrared small target detection. *IEEE Transactions on Geoscience and Remote Sensing*, 62: #5002615 [DOI: 10.1109/TGRS.2024.3383649]
- Zhang L D and Peng Z M. 2019a. Infrared small target detection based on partial sum of the tensor nuclear norm. *Remote Sensing*, 11(4): #382 [DOI: 10.3390/rs11040382]
- Zhang M J, Li X L, Gao F and Guo J. 2025a. IRMamba: pixel difference Mamba with layer restoration for infrared small target detection//*Proceedings of the 39th AAAI Conference on Artificial Intelligence*. Philadelphia, USA: AAAI: 10003-10011 [DOI: 10.1609/aaai.v39i9.33085]
- Zhang M J, Wang Y C, Guo J, Li Y S, Gao X B and Zhang J. 2024a. IRSAM: advancing segment anything model for infrared small target detection//*Proceedings of the 18th European Conference on Computer Vision*. Milan, Italy: Springer: 233-249 [DOI: 10.1007/978-3-031-72855-6_14]
- Zhang M J, Yue K, Guo J, Zhang Q M, Zhang J and Gao X B. 2025b. Computational fluid dynamic network for infrared small target detection. *IEEE Transactions on Neural Networks and Learning Systems*, 36(8): 14777-14789 [DOI: 10.1109/TNNLS.2025.3548984]
- Zhang M J, Yue K, Li B Y, Guo J, Li Y S and Gao X B. 2024b. Single-frame infrared small target detection via Gaussian curvature inspired network. *IEEE Transactions on Geoscience and Remote*

Sensing, 62: #5005013 [DOI: 10.1109/TGRS.2024.3423492]

Zhang M J, Zhang R, Yang Y X, Bai H C, Zhang J and Guo J. 2022. ISNet: shape matters for infrared small target detection//Proceedings of 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New Orleans, USA: IEEE: 867-876 [DOI: 10.1109/CVPR52688.2022.00095]

Zhang T F, Li L, Cao S Y, Pu T and Peng Z M. 2023. Attention-guided pyramid context networks for detecting infrared small target under complex background. IEEE Transactions on Aerospace and Electronic Systems, 59 (4) : 4250-4261 [DOI: 10.1109/TAES.2023.3238703]

Zhang T F, Wu H, Liu Y H, Peng L B, Yang C P and Peng Z M. 2019b. Infrared small target detection based on non-convex optimization with Lp-Norm constraint. Remote Sensing, 11 (5) : #559 [DOI: 10.3390/rs11050559]

作者简介

张迎梅,女,讲师,硕士生导师,主要研究方向为红外小目标

检测、图像融合与计算机视觉。
E-mail: zhangyingmei@jxufe.edu.cn

肖沁,通信作者,女,讲师,硕士生导师,主要研究方向为数据驱动决策、计算机视觉和机器学习。
E-mail: xiaoqin@jxufe.edu.cn

鲍王涛,男,硕士研究生,主要研究方向为红外小目标检测与图像处理。E-mail: baowangtao29@gmail.com

杨勇,男,教授,博士生导师,主要研究方向为遥感图像处理、计算机视觉与机器学习。E-mail: yangyong@tiangong.edu.cn

万伟国,男,讲师,硕士生导师,主要研究方向为图像处理、人脸识别与计算机视觉。E-mail: wanweiguo@jxufe.edu.cn

罗亦韬,男,本科生,主要研究方向为红外小目标检测与图像处理。E-mail: 3014815803@qq.com

邹雪婷,女,硕士研究生,主要研究方向为图像融合与深度学习。E-mail: zouxueting0113@126.com

张磊,男,硕士研究生,主要研究方向为人脸识别与图像处理。E-mail: 13647037937@163.com