

多维融合注意力增强的红外小目标检测方法

李卫星^① 王 帅^② 陈怀宇^① 盛卫东^{*①}

^①(国防科技大学电子科学学院 长沙 410073)

^②(中国科学院西安光学精密机械研究所 西安 710061)

摘 要: 针对红外小目标在深层卷积神经网络中特征弥散、易受复杂背景干扰而导致检测性能下降的问题, 提出一种基于多维融合注意力机制的红外小目标检测方法。设计通道-空间注意力特征增强模块, 通过并行分支结构聚合跨维度显著性特征, 抑制目标特征在深层网络的弥散趋势; 通过自适应跨维度信息交互, 建立通道-空间的跨域依赖关系, 增强微弱目标在深层网络的特征响应。该方法仅需极低计算复杂度即可灵活嵌入基线模型, 实现目标深度特征聚拢。在NUDT-SIRST数据集上的实验表明, 所提方法在高杂波、低信噪比场景下表现出更强的鲁棒性和泛化能力。基于自研的轻量化智能处理单元进行边缘部署性能验证, 实测单帧推理时延为46.7 ms, 满足工程应用中实时性与高精度要求。

关键词: 红外小目标; 注意力机制; 特征融合; 边缘推理

中图分类号: TN216

文献标识码: A

文章编号: 1009-5896(2026)00-0001-12

DOI: 10.11999/JEIT260040

CSTR: 32379.14.JEIT260040

1 引言

红外成像通过捕获目标的热辐射信息实现目标探测, 具有作用距离远、覆盖范围广、隐蔽性强、全天候可视等优势, 在空天目标监视、海上应急救援、森林火灾监测、地球遥感等军用和民用领域具有广泛应用^[1-4]。相比于传统目标检测, 红外目标检测面临以下挑战: ①红外成像距离远, 目标尺寸小, 缺乏纹理与形态特征; ②受大气吸收和衰减影响, 目标热辐射信号弱易淹没在强杂波中。因此, 在远距离成像和复杂背景的约束下, 如何实现高精度、高可靠红外小目标检测是红外成像领域的研究热点。

传统红外弱小目标检测一般采用模型和专家知识驱动的方法, 基于先验知识进行特性建模, 通过目标和背景在时空域上的差异性实现目标提取。传统方法可分为基于图像滤波、基于局部对比度和基于稀疏矩阵低秩分解的方法。图像滤波方法常用的有高通、中值、均值滤波器、Top-Hat以及改进的Top-Hot等形态学滤波器等^[5-7]。基于局部对比度的方法利用目标区域和邻域的灰度值差异性构建局部显著图, 通过阈值分割检测小目标^[8-11]。文献^[12-15]将红外小目标检测问题转化为稀疏重构与低秩恢复问题, 并借助凸优化工具实现目标分割。传统方法

依赖目标和背景辐射、形态、运动等先验知识, 在时变复杂的场景下难以保证目标检测性能的鲁棒性。

深度学习技术的发展催生了基于数据驱动的红外小目标检测理论体系。通过构建深度神经网络, 直接从图像数据中学习和提取目标、背景深层语义特征, 无需预设先验模型, 可实现复杂多变场景下的宽适应检测^[16-19]。文献^[20]首次将卷积神经网络(Convolutional Neural Network, CNN)应用于红外小目标检测, 通过多层感知机实现目标特征学习。文献^[21]提出了由多个残差卷积块构成的深度神经网络ResNet, 通过残差学习避免了深层网络引起的梯度消失和梯度爆炸。ResNet在红外小目标检测领域获得广泛应用, ALCNet、ISNet、AGP-CNet等均采用ResNet为骨干网络^[22-26]。文献^[27]提出了一种基于编码器-解码器结构的U型卷积神经网络(UNet), 引入跳跃连接结构, 可融合深层语义特征与浅层细节特征。文献^[28-29]设计了非对称上下文调制模块(Asymmetric Context Modulation, ACM), 改进了UNet网络的跳跃连接结构。文献^[30]在UNet的基础上, 增加了密集嵌套特征提取模块, 在各神经元之间构建密集跳跃连接, 解决了点目标在深层网络中易丢失问题, 显著提升了红外小目标检测性能。

针对目标尺寸小、在网络深层易弥散的问题, 学者们将自注意力机制^[31-33]嵌入到目标检测网络, 以实现对重要区域的深度特征聚合。基于注意力机制的目标检测方法主要分为两类: 通道注意力^[34-36]和通道-空间注意力^[37-40]。文献^[34]开创了通道自注意的先河, 设计了挤压和激励网络模块(Squeeze

收稿日期: 2026-01-13; 改回日期: 2026-06-27; 网络出版: 2026-07-19

*通信作者: 盛卫东 shengweidong07@nudt.edu.cn

基金项目: 国家自然科学基金(62401589, 62403476)

Foundation Items: The National Natural Science Foundation of China (62401589, 62403476)

and Excitation, SENet), 在抑制噪声的同时突出高质量通道的作用。文献[35-36]提出了高效通道注意模块(ECANet), 通过一维卷积进行通道间信息交互, 从而降低模型复杂度。文献[37-43]提出了通道-空间注意力和坐标机制, 并引入全局最大池化和全局平均池化, 捕获跨域交互信息。然而, 现有通道-空间自注意力方法主要通过级联或位置编码进行信息融合, 未在通道和空间特征之间建立跨域协同机制, 检测性能提升有限。

本文针对红外小目标在深层网络响应中特征易丢失问题, 提出一种通道-空间多维融合注意力机制(Multi-dimension Fusion Attention Module, MFAM), 捕捉通道、高度、宽度多维特征及红外小目标在各维度的依赖关系, 通过特征融合和跨维度协同交互, 抑制小目标在深层网络的特征弥散, 提高红外小目标检测的鲁棒性。并行融合架构使得两维度显著性特征在融合前均保持完整信息, 有效消除串行结构中的误差累积, 抑制红外小目标在深层网络中的特征弥散, Sigmoid与乘法融合机制实现通道语义与空间位置的互补融合, 进一步增强了对微弱目标的敏感性。同时, 多维融合注意力模块具有轻量、灵活的特点, 可即插即用集成于多种目标检测骨干网络中。基于典型公开红外数据集实验验证了本文方法的有效性, 并在FPGA+GPU处理单元上验证了其边缘侧推理性能。

2 基于注意力机制的红外小目标检测网络

近年来, 基于注意力机制的深度学习网络在红外小目标检测领域发展迅速。注意力机制的核心思想是通过聚合输入特征信息, 实现对重要特征的筛选与增强。在红外小目标检测任务中, 注意力机制通常被应用于特征提取阶段, 如图1所示。注意力机制可以有效抑制小目标在深层神经网络中的弥散, 从而更有效地聚合重要特征。根据注意力作用机制作用域的不同, 可分为通道域注意力机制与混合域注意力机制等。

通道域注意力机制[34-36]的核心思想是在输入特征图的通道维度引入自注意力机制, 可抑制信息稀

疏通道的响应, 并增强判别性特征丰富通道的权重。由于红外小目标尺寸小, 在空间域往往呈现稀疏特性。因此, 进一步引入了空间域注意力机制[37-43], 如图2所示, 通过在空间维度突出潜在目标区域并抑制背景干扰, 进一步增强特征的判别能力。

然而, 现有方法多采用通道注意力与空间注意力串联连接(如CBAM), 后一级模块的输入已由前一级调制, 易导致小目标特征在逐级传递中逐渐稀释; 此外, 注意力机制伴随较高的计算复杂度与存储资源, 导致其在边缘设备上的实际应用受到限制。因此, 开发一种专用于红外小目标检测任务、兼顾性能与效率的轻量型注意力机制, 具有重要的工程应用价值。

3 基于多维融合注意力增强的红外小目标检测网络

本文提出通道-空间多维融合注意力模块(MFAM)。MFAM将通道与空间注意力并行作用于同一原始输入特征图, 两路输出经Sigmoid变换后逐元素相乘融合。并行融合设计能够确保两维度显著性特征在融合前均保持完整信息, 消除串行结构中的误差累积, 抑制红外小目标在深层网络中的特征弥散。同时, Sigmoid与乘法融合机制实现通道语义与空间位置的互补融合, 进一步增强了对微弱目标的敏感性, 并有效抑制了单维度造成误判的背景杂波。此外, MFAM以极低的参数量增量和极小的时延代价换取了显著的性能提升, 兼顾高精度与轻量化。

MFAM总体架构如图3所示, 其中, 左侧为嵌入Res-Block的示意图, 右侧为MFAM的详细结构。对于任何一个输入特征图 $F \in \mathbb{R}^{C \times H \times W}$, 输出为一个注意力增强的特征图 Y , 其尺寸与 F 保持一致。整个过程可以表示为: $Y = I_{MFAM}(F) \otimes F$ 。其中 $\otimes$ 表示逐元素相乘, C 、 H 、 W 分别表示输入特征图的通道数、高度和宽度。实现流程如下: 通道注意力和空间注意力以并联方式直接作用于输入特征图 F 上; 在通道域中, 对特征进行压缩并通过

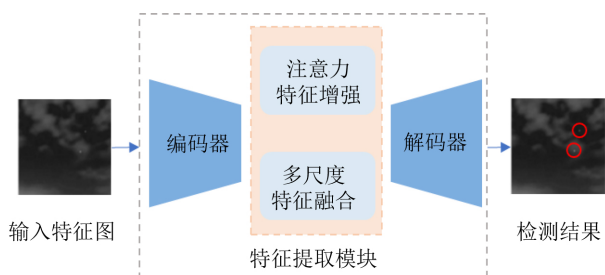


图1 基于注意力机制的红外小目标检测总体架构

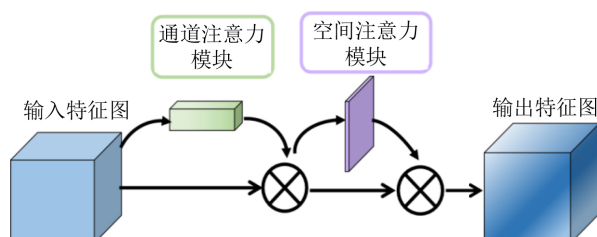


图2 通道-空间注意力机制

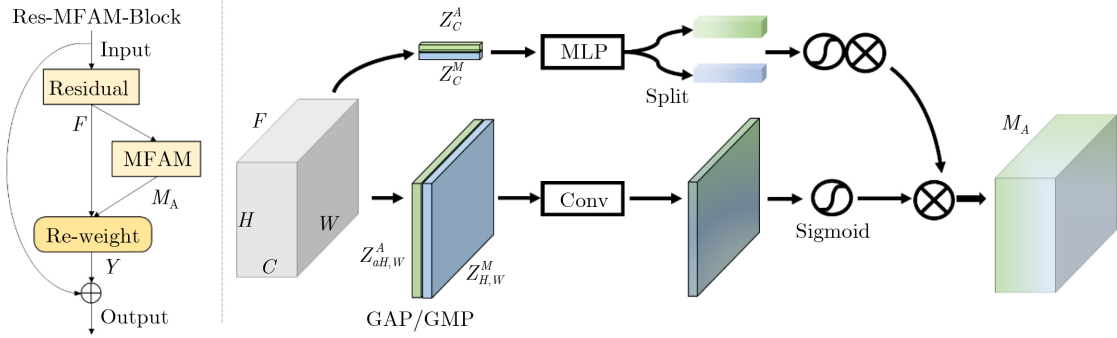


图 3 基于多维融合注意力机制的网络结构

多层感知机(Multi-layer perceptron, MLP)捕捉通道间关系；在空间域中，通过全局平均池化(Global average pooling, GAP)和全局最大池化(Global maximum pooling, GMP)进行宽度、高度编码，聚合空间信息；最后通过Sigmoid激活函数实现特征融合。

3.1 通道注意力模块

通道注意力机制通过挖掘特征通道之间的依赖关系来生成通道注意力图。现有方法通常采用全局平均池化对通道信息进行编码。然而，红外小目标通常表现为其所在区域中最亮的元素，而最大池化能够捕捉最显著区域的特征响应，从而更有效提取关键目标信息。因此，本文同时采用平均池化与最大池化两种特征，通过信息融合增强特征表示能力。

首先，通过平均池化和最大池化分别聚合特征图的空间信息，生成两个不同的通道上下文描述符，记为 Z_C^A 和 Z_C^M ，分别表示平均聚合特征和最大聚合特征；其次，将这两个描述符拼接后送入一个多层感知机MLP网络，生成通道注意力图 $Z_C \in \mathbb{R}^{C \times 1 \times 1}$ ；为了减少参数量，将MLP隐藏层激活尺寸设置为 $\mathbb{R}^{C/r \times 1 \times 1}$ ，其中 r 是缩减率；最后，对两个描述符的MLP输出通过逐元素相乘进行融合，得到通道注意力权重。计算过程如下：

$$\begin{aligned} Z_C &= \text{mul}(\sigma(\text{split}(\text{MLP}([P_{avg}^C(F), P_{max}^C(F)])))) \\ &= \text{mul}(\sigma(\text{split}(\text{MLP}([Z_C^A, Z_C^M]))) \end{aligned} \quad (1)$$

其中 σ 表示Sigmoid操作，mul为元素乘积， $[\cdot]$ 表示矩阵拼接操作。 $P_{avg}^C(F)$ 、 $P_{max}^C(F)$ 分别为通道域全局、最大平均池化函数， F 为输入特征图。

3.2 空间注意力模块

红外小目标尺寸极小，在空间域中表现出显著的稀疏特性。为有效提升目标的定位精度，本文设计了一种空间注意力模块，旨在提取空间维度上的显著性特征。

为了聚合输入特征图的空间信息，同时使用全

局平均池化与最大池化操作，生成了两个二维图： $Z_{H,W}^A \in \mathbb{R}^{1 \times H \times W}$ 与 $Z_{H,W}^M \in \mathbb{R}^{1 \times H \times W}$ ，分别表示平均池化特征与最大池化特征。将二者沿通道维度拼接之后，送入 3×3 标准卷积层生成二维空间注意力图。整个空间注意力的计算过程如下所示：

$$\begin{aligned} Z_{H,W} &= \sigma(f^{3 \times 3}([P_{avg}^S(F), P_{max}^S(F)])) \\ &= \sigma(f^{3 \times 3}([Z_{H,W}^A, Z_{H,W}^M])) \end{aligned} \quad (2)$$

其中 $f^{3 \times 3}$ 表示核大小为 3×3 的卷积操作。 $P_{avg}^S(F)$ 和 $P_{max}^S(F)$ 分别表示空间域全局平均池化函数和最大平均池化函数。

3.3 特征融合模块

传统的混合注意力机制将通道和空间注意力机制分别作用于输入特征图，池化操作将会丢失细节信息。本文将通道注意力模块与空间注意力模块并行直接作用于输入特征图，分别在全局感受野下对通道、高度和宽度三个维度进行精细化编码，能够同时捕获包含空间细节的浅层特征与富含上下文信息的深层语义特征，生成通道感知描述符 Z_C 与空间感知描述符 $Z_{H,W}$ 。将 Z_C 和 $Z_{H,W}$ 经由Sigmoid函数进行特征变换，实现跨维度信息交互，进而融合为统一的通道-空间混合域权重矩阵。最终，将该权重矩阵与输入特征图逐元素相乘，输出MFAM增强特征图：

$$Y = F \times Z_C \times Z_{H,W} \quad (3)$$

MFAM具备即插即用特性，无需引入复杂附加结构即可灵活嵌入ResNet、DNA-Net等多种基线模型，展现出良好的兼容性。以ResNet为例，该模块能够无缝集成至其网络结构中的残差块等基本单元中，并可与增强后的网络进行端到端联合训练，其集成方式如图4所示。

4 实验与结果分析

4.1 实验设置

采用红外小目标检测领域广泛使用的NUDT-SIRST数据集，包含地面、海洋、云层等背景，以

及无人机、飞机、舰船等目标,共1327帧图像,其中613帧用于训练,614帧用于测试。针对复杂背景下红外小目标检测任务,在训练前对图像进行剪裁、模糊、旋转等数据增强,提高模型的鲁棒性和泛化能力。将图像尺寸统一调整为 256×256 ,并输入至网络模型训练。

本文模型默认下采样为4,各卷积块的通道数分别为16、32、64、128、256。训练和推理均基于Pytorch深度学习框架实现,硬件环境为搭载1颗Intel(R) Xeno Gold 6 248 CPU@2.5GHz和1块NVIDIA V100 GPU的工作站。详细训练参数配置如表1所示。训练次数1500次,采用Adagrad优化器,损失函数采用Soft-IoU,默认学习率为0.05。

4.2 评价指标

选用在该领域表现良好且适应性强的DNA-Net^[30]作为基线模型,嵌入不同注意力机制,并统计各算

法在两类数据集上的评估指标,主要包括交并比(IoU)、检测率(Pd)和虚警率(Fa),其中IoU和Pd越高越好,Fa越低越好。

(1)交并比:通过度量预测区域与真实标注区域的重叠程度来评估分割精度,其值为两区域交集面积与并集面积的比值,计算公式为: $\text{IoU} = A_{\text{inter}}/A_{\text{all}}$,其中 A_{inter} 和 A_{all} 分别代表预测区域和真实区域间的交集区域和并集区域。

(2)检测率:定义为正确检出的目标数目与场景内全部目标数目的比值,主要反映算法对正确目标实例的定位精准程度,计算公式为: $P_d = T_{\text{correct}}/T_{\text{all}}$,其中 T_{correct} 和 T_{all} 分别表示正确预测的目标数量和所有目标的数量。

(3)虚警率:主要反映算法对错误目标实例的误报情况,计算公式为: $F_a = P_{\text{false}}/P_{\text{all}}$,其中, P_{false} 和 P_{all} 分别表示错误预测的像素数和图像总像素数。

4.3 结果分析

基于NUDT-SIRST数据集,对本文算法与现有红外小目标检测算法进行了性能对比分析,包括Top-Hat^[5]、IPI^[8]和RIPT^[11]等知识驱动传统算法,以及SENet^[34]、SKNet^[35]、ECA-Net^[36]、CBAM^[37]、BAM^[38]、GAM^[39]、CA^[42]、TA^[43]等深度学习算法。所有算法均在相同的软硬件环境下、使用相同数据集进行训练与测试,以确保实验的公平性。

4.3.1 定量分析

为了评估本文提出的MFAM模块在红外小目标检测任务上的性能,选用在该领域表现良好且适应性强的DNA-Net作为基线模型,嵌入不同注意力机制,并统计各算法在两类数据集上的评估指标。

在NUDT-SIRST数据集上的检测性能如表1所示,最优结果用红色加粗、次优结果用蓝色加粗显示。本文算法取得了最高的IoU、Pd和最低的Fa,整体性能最优。相较于Top-Hat、IPI和RIPT等传统算法,本文算法的性能提升显著。这主要源于NUDT-SIRST数据集涵盖了目标形态多变、大动态信杂比、杂波起伏、背景复杂等有挑战性的场景。传统算法依赖先验知识与特定假设,在复杂多变环境中泛化能力有限。相比之下,本文提出的MFAM算法能够提取跨场景中具有高判别力的语义特征,表现出更强的泛化性与鲁棒性。此外,深度学习算法在IoU指标上显著优于传统算法,原因在于传统方法更关注目标质心定位,对目标整体轮廓的建模能力较弱。

相较于DNA-Net基线(无注意力机制)的IoU为

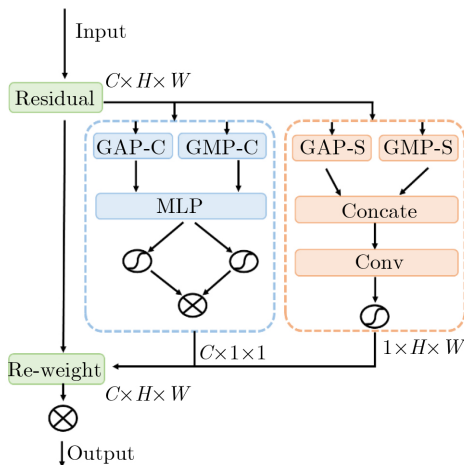


图4 MFAM计算流程

表1 不同算法目标检测性能(NUDT-SIRST)

类型	算法	IoU/(%)	Pd/(%)	Fa/(×1e-6)
传统算法	Top-Hat	20.7	78.41	166.7
	IPI	17.76	74.49	41.23
	RIPT	29.44	91.85	344.3
	DNA-Net(基线)	85.39	97.67	9.49
深度学习 算法	+SE	86.2	98.64	3.98
	+SK	85.3	98.41	12.8
	+ECA	85.50	98.41	6.78
	+CBAM	86.94	98.2	3.24
	+BAM	85.66	98.2	5.79
	+GAM	84.88	98.41	33.96
	+CA	85.25	97.57	9.77
	+TA	85.75	98.31	12.82
	+MFAM(本文)	87.34	98.72	3.22

85.39%，在引入MFAM模块后提升至87.34%，检测率 P_d 从97.67%提升到98.72%，虚警率 F_a 从 9.49×10^{-6} 降低到 3.22×10^{-6} 。性能的大幅提升得益于MFAM的并行跨维度融合设计。相较于纯通道注意力机制SENet、SKNet、ECA-Net, MFAM在IoU指标上分别提高1.14%、2.04%、1.84%。性能提升归因于MFAM在建模通道依赖的同时引入了空间注意力机制，使模型能够更准确地感知红外小目标的空间位置信息。相较于通道-空间注意力机制CBAM、BAM、GAM、CA、TA, MFAM仍然保持更优性能。CBAM将空间注意力模块串联于通道注意力之后，未直接作用于输入特征，可能导致信息损失，其IoU相比于MFAM下降0.4%。CA和BAM仅使用全局平均池化进行信息编码，忽略了最大池化在深层特征提取中的作用；TA通过旋转操作建模跨维度依赖，但未实现三维同时交互且计算复杂较高。MFAM同时融合通道与空间信息，实现充分的跨域交互，并结合全局平均与最大池化增强上下文与细节特征提取能力，从而实现更精细的红外小目标检测。

MFAM模块具有良好的即插即用灵活性，可嵌入多种主流检测架构。除DNA-Net外，本文还将MFAM嵌入ALC-Net、AMFU-Net等基线模型中，测试结果如表2所示。实验表明，MFAM与

CBAM均能嵌入不同模型，但MFAM带来的整体性能提升更为显著。

4.3.2 定性分析

(a)MFAM的特征增强作用

为验证通道-空间注意力特征增强模块的作用及微弱目标在深层网络的特征响应，本文采用同一组高杂波场景测试图像，对比无注意力基线、CBAM、MFAM三种方式在编码器深层(Encoder4-5)与解码器浅层(Decoder1-2)的特征图，结果如图5所示。

从编码器深层特征来看，无注意力基线(绿色行)杂波背景开始响应且目标逐渐丢失，淹没于热斑干扰中；CBAM(蓝色行)虽能延缓衰减，但Encoder5层目标响应已接近背景水平，难以与杂波区分；MFAM(红色行)则凭借并行结构的完整信息保持能力，使微弱目标在Encoder4层仍保持突出的亮

表 2 注意力模块嵌入不同基线模型的性能

算法	IoU/(%)	Pd/(%)	Fa/($\times 10^{-6}$)
ALC-Net (基线)	71.9	97.14	29
+CBAM	72.15	95.24	4.16
+MFAM (本文)	72.58	97.46	8.92
AMFU-Net (基线)	81.17	97.04	32.24
+CBAM	83.04	98.41	15.49
+MFAM (本文)	83.32	97.35	11.4

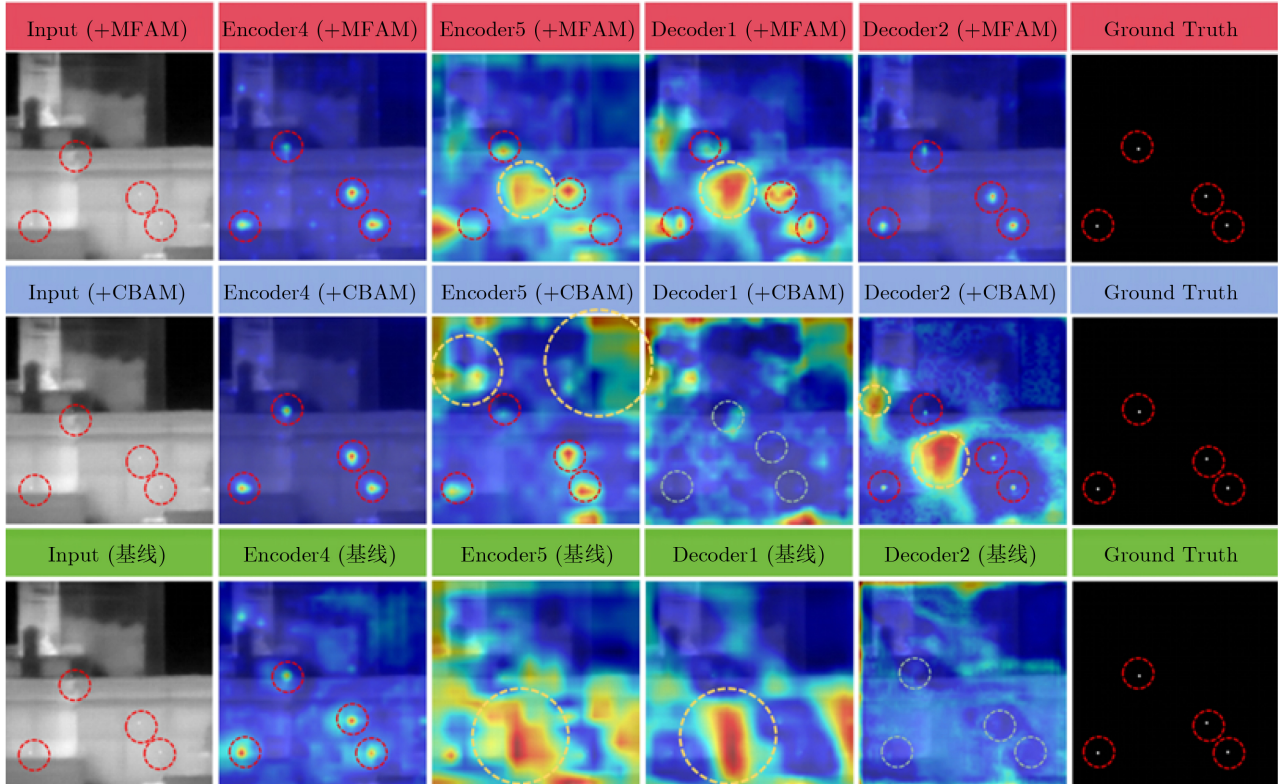


图 5 不同注意力机制的特征图可视化对比

斑响应, Encoder5层目标特征未出现明显衰减。

从解码器恢复特征来看, 无注意力基线因编码阶段信息完全丢失, Decoder1-2层呈现大面积背景噪声; 而CBAM因编码器深层特征质量劣化, 出现严重的背景虚警, 大面积热斑散布于非目标区域; MFAM则凭借编码阶段保持的显著特征, 从Decoder1层即精准聚焦真实目标, Decoder2层目标轮廓清晰、背景干净, 无需过度补偿即可精准恢复。

上述可视化结果证明了MFAM通过并行跨维度聚合机制, 有效解决了红外小目标检测中的特征弥散与微弱目标丢失问题: 并行双分支直接作用于原始特征, 缓解串联结构的误差累积, 使目标特征在网络深层仍能较好的保留; Sigmoid与乘法融合实现通道-空间交叉验证, 抑制单一维度误判的不确定性。

(b)多场景下的目标检测效果

图6直观对比了多种场景下不同注意力模块嵌入DNA-Net后的目标检测效果。为便于观察, 图6左上角提供了目标区域的放大示意图。图中红色、黄色和蓝色虚线框分别标识出正确检测的目标、虚警和漏检区域。选取了六组典型复杂场景进行可视

化对比, 涵盖云层起伏、建筑反射、海面波纹、植被纹理、多目标密集及强杂波城市环境。例如第2、第4行这类高杂波场景, 其建筑产生强热辐射形成高亮扩展区域以及云层边缘灰度突变剧烈, 均与弱小目标在灰度分布上高度相似。因此现有方法例如SE、ECA等缺乏空间位置约束, 将云层局部高亮、建筑顶部强反射误判为目标, 出现虚警。CBAM、BAM虽有空间注意力约束但由于其串联结构中通道注意力先对高亮区域进行筛选, 后续空间注意力基于已混淆的特征进一步产生虚警; 对于第1、第5行等这类海面波纹形成亮度渐变, 目标极弱小以及云层覆盖小目标等低信噪比场景, SE、ECA、TA等通道注意力方法对均匀渐变杂波抑制不足, 部分目标因响应接近背景而被过滤, 从而出现蓝色漏检; 而本文提出的MFAM算法通过通道+空间并行结构减少了现有方法的串联逐级筛选而导致的信息损失, 使微弱目标原始响应完整传递至深层; 同时还运用了GMP(全局平均池化)增强了对局部最亮像素的敏感性, 确保低信噪比目标在浅层即被激活并保留; 最后还通过sigmoid乘法对两个并联注意力部分进行充分融合互补, 极大的增

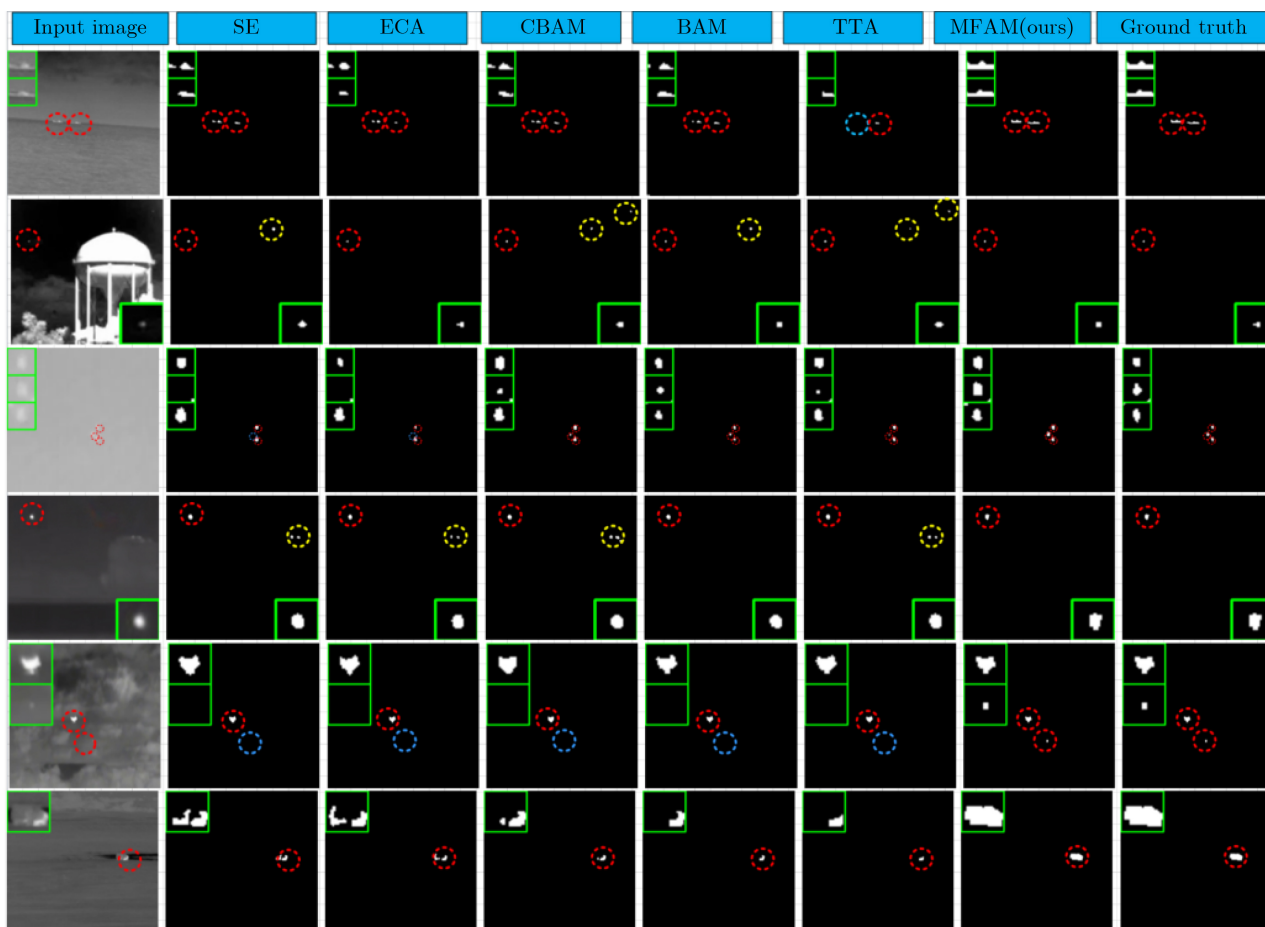


图6 不同注意力机制的红外小目标检测二维可视化图

强了对不同环境，例如高杂波、低信噪比等环境的鲁棒性。

4.4 消融实验

本节通过消融实验验证所提算法的有效性。首先，通过对比不同注意力融合方式，评估通道与空间维度并联融合策略的贡献；然后，分析不同信息编码方法对目标检测性能的影响。

4.4.1 多维融合方式的影响

为了评估通道-空间多维融合机制的有效性，本文比较了不同注意力组合方式下的检测性能，包括仅通道注意力、仅空间注意力、通道-空间级联融合以及通道-空间并联融合(即本文方法)，结果如表3所示。实验表明，在基线模型基础上引入通道注意力(+C)或空间注意力(+S)均能改善IoU、检测率(Pd)和虚警率(Fa)，说明注意力机制有助于增强红外目标特征表示。进一步引入通道-空间注意力(+CS)可进一步提升性能。与串联融合方式相比，本文所提出的并联结构在IoU和Pd上分别提升0.37%和0.39%，Fa降低 1.41×10^{-6} 。其优势源于并联结构将通道与空间注意力直接作用于输入特征，缓解了目标特征在深层网络中的弥散。

4.4.2 信息编码方式的影响

为比较不同信息编码策略的效果，本文分别测试了全局平均池化(GAP)、全局最大池化(GMP)以及GAP与GMP联合使用三种方式，结果如表4所示。相比基线模型，单独使用GAP或GMP均带来一定性能提升，其中GMP因其能捕捉显著区域特征，在红外小目标场景中表现更优。联合使用

GAP与GMP既可提取局部显著性特征以敏感于潜在目标，又能融合全局上下文以提高虚警抑制能力，因此获得最优的综合性能。

5 边缘部署与实现

面向无人机载、星载等边缘侧部署应用，设计研制了基于FPGA+GPU异构处理器的轻量型智能信息处理单元，并部署了本文提出的多维融合注意力增强的红外小目标检测算法，以验证算法在边缘端的部署效果。

该智能信息处理单元以高性能FPGA、嵌入式GPU、DDR4高速缓存及大容量eMMC嵌入式存储为核心，构建了存算一体的边缘侧处理平台，具备开放式体系架构、高吞吐全互联路由和通用化计算模块，支持功能软件定义与算法在线重构。核心处理模块采用FPGA与GPU协同架构，如图7所示。其中，FPGA选用Xilinx XC7K325T，主要负责高速图像收发与预处理；GPU采用NVIDIA Jetson AGX Xavier，集成512核Volta架构GPU、64个张量核心、8核Arm® CPU(主频为2.265 GHz)、16 GB LPDDR4x内存和32 GB eMMC 5.1存储。

本文搭建“图像模拟—边缘侧推理—结果解析”验证系统，步骤如下：(1)在智能处理单元AGX Xavier边缘侧GPU上部署DNA-Net+MFAM红外小目标检测网络模型；(2)通过图像模拟器发送测试图像至智能信息处理单元，由FPGA进行接收、预处理后送入AGX Xavier GPU进行推理，并将处理结果回传后分析。

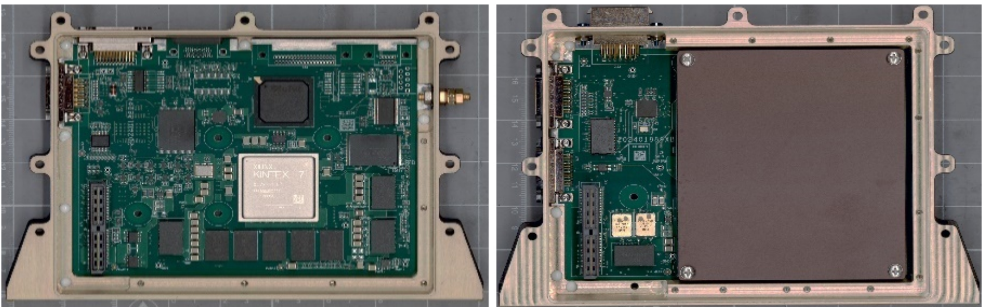


图 7 自研FPGA和GPU处理模块

表 3 不同注意力融合方式的性能对比

算法	IoU/(%)	Pd/(%)	Fa/($\times 1e-6$)
基线模型	85.39	97.67	9.49
+C	86.78	98.18	8.95
+S	85.27	97.43	13.24
+CS, 串联	86.97	98.33	4.63
+CS, 并联(本文)	87.34	98.72	3.22

表 4 不同信息编码方式的性能对比

算法	IoU/(%)	Pd/(%)	Fa/($\times 1e-6$)
基线模型	85.39	97.67	9.49
GAP	85.86	98.12	9.43
GMP	86.98	98.43	6.75
GAP+GMP(本文)	87.34	98.72	3.22

经实测验证, 图像发送、边缘侧推理、结果回传全流程工作正常, 对典型场景的推理结果及AGX Xavier GPU的运行时间如图8所示, 可正确检测出目标。对于 256×256 的输入图像, 单

帧平均推理时延为46.7 ms。实验表明, 本文算法可成功部署至AGX Xavier GPU, 实现边缘侧高效推理, 充分验证了本文算法的灵活性和轻量化。



图8 边缘侧推理结果及时间

6 结论

针对红外小目标检测中面临的目标尺寸小、特征易弥散、背景复杂等技术难点, 本文构建了多维融合注意力模块(MFAM), 实现了通道与空间显著性特征的有效聚合与跨维度自适应交互, 增强了目标在深层网络中的特征保持与鲁棒输出。该模块在维持较低计算复杂度的同时具有良好的即插即用特性。测试结果表明, 相比传统知识驱动算法与现有基于注意力的深度学习方法, 本文算法在交并比、检测率和虚警率等指标上均表现更优。面向边缘计算应用场景, 研制了基于FPGA+GPU的轻量型智能处理单元, 成功实现算法在边缘设备上的部署与高实时推理, 展现出良好的工程适用性与应用前景。

参 考 文 献

- [1] 周海, 李保权, 王怀超, 等. 低空复杂场景红外弱小目标快速精准检测[J]. 国防科技大学学报, 2023, 45(1): 74–85. doi: [10.11887/j.cn.202301008](https://doi.org/10.11887/j.cn.202301008).
ZHOU Hai, LI Baoquan, WANG Huaichao, *et al.* Fast and accurate detection of infrared dim small target in low altitude complex scenes[J]. *Journal of National University of Defense Technology*, 2023, 45(1): 74–85. doi: [10.11887/j.cn.202301008](https://doi.org/10.11887/j.cn.202301008).
- [2] LIU J, LIU Y G, ZHANG Y Q, *et al.* Robust infrared small target detection using a novel four-leaf model[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2023, 61: 1–15. doi: [10.1109/TGRS.2023.3241234](https://doi.org/10.1109/TGRS.2023.3241234).
- [3] 王密, 郭贝贝, 皮英冬, 等. 珞珈三号01星在轨处理技术及验证[J]. 测绘学报, 2024, 53(4): 599–609. doi: [10.11947/j.AGCS.2024.20230412](https://doi.org/10.11947/j.AGCS.2024.20230412).
WANG Mi, GUO Beibei, PI Yingdong, *et al.* On-orbit processing technology and verification of LuoJia-3 01 satellite[J]. *Acta Geodaetica et Cartographica Sinica*, 2024, 53(4): 599–609. doi: [10.11947/j.AGCS.2024.20230412](https://doi.org/10.11947/j.AGCS.2024.20230412).
- [4] 李召良, 唐伯惠, 吴骅, 等. 热红外遥感发展历程与展望[J]. 遥感学报, 2025, 29(6): 1529–1550. doi: [10.11834/jrs.20252301](https://doi.org/10.11834/jrs.20252301).
LI Zhaoliang, TANG Bohui, WU Hua, *et al.* Development and prospects of thermal infrared remote sensing[J]. *National Remote Sensing Bulletin*, 2025, 29(6): 1529–1550. doi: [10.11834/jrs.20252301](https://doi.org/10.11834/jrs.20252301).
- [5] HARALICK R M, STERNBERG S R, and ZHUANG Xinhua. Image analysis using mathematical morphology[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1987, PAMI-9(4): 532–550. doi: [10.1109/TPAMI.1987.4767941](https://doi.org/10.1109/TPAMI.1987.4767941).
- [6] HAO Congyu, LI Zhengzhou, ZHANG Yuting, *et al.* Infrared small target detection based on adaptive size estimation by multidirectional gradient filter[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2024, 62: 5007915. doi: [10.1109/TGRS.2024.3502421](https://doi.org/10.1109/TGRS.2024.3502421).
- [7] 张晶晶, 曹思华, 崔文楠, 等. 基于改进顶帽变换的红外弱小目标检测[J]. 电子与信息学报, 2024, 46(1): 267–276. doi: [10.11999/JEIT230892](https://doi.org/10.11999/JEIT230892).
ZHANG Jingjing, CAO Sihua, CUI Wennan, *et al.* Improved top-hat transform-based algorithm for infrared dim and small target detection[J]. *Journal of Electronics & Information Technology*, 2024, 46(1): 267–276. doi: [10.11999/JEIT230892](https://doi.org/10.11999/JEIT230892).
- [8] GAO Chenqiang, MENG Deyu, YANG Yi, *et al.* Infrared patch-image model for small target detection in a single image[J]. *IEEE Transactions on Image Processing*, 2013, 22(12): 4996–5009. doi: [10.1109/TIP.2013.2281420](https://doi.org/10.1109/TIP.2013.2281420).
- [9] 李守昌, 吴滢跃. 复杂背景下双邻域局部权重对比度红外小目标检测算法[J]. 半导体光电, 2024, 45(5): 853–860. doi: [10.11999/JEIT230892](https://doi.org/10.11999/JEIT230892).

- 16818/j.issn1001-5868.2024040301.
- LI Shouchang and WU Yingyue. Dual-neighborhood local weighted contrast algorithm for infrared small target detection within complex backgrounds[J]. *Semiconductor Optoelectronics*, 2024, 45(5): 853–860. doi: [10.16818/j.issn1001-5868.2024040301](https://doi.org/10.16818/j.issn1001-5868.2024040301).
- [10] 苟士淼, 刘兆瑜, 马鹏阁, 等. 基于IHBF的加权局部对比度红外小目标检测[J]. *电光与控制*, 2025, 32(8): 86–91. doi: [10.3969/j.issn.1671-637X.2025.08.014](https://doi.org/10.3969/j.issn.1671-637X.2025.08.014).
- GOU Shimiao, LIU Zhaoyu, MA Pengge, *et al.* Weighted local contrast infrared small target detection based on IHBF[J]. *Electronics Optics & Control*, 2025, 32(8): 86–91. doi: [10.3969/j.issn.1671-637X.2025.08.014](https://doi.org/10.3969/j.issn.1671-637X.2025.08.014).
- [11] DAI Yimian and WU Yiquan. Reweighted infrared patch-tensor model with both nonlocal and local priors for single-frame small target detection[J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2017, 10(8): 3752–3767. doi: [10.1109/JSTARS.2017.2700023](https://doi.org/10.1109/JSTARS.2017.2700023).
- [12] 蹇渊, 黄自力, 王询. 基于随机化张量算法的红外弱小目标检测[J]. *激光技术*, 2024, 48(1): 127–134. doi: [10.7510/jgjs.issn.1001-3806.2024.01.020](https://doi.org/10.7510/jgjs.issn.1001-3806.2024.01.020).
- JIAN Yuan, HUANG Zili, and WANG Xun. Infrared small target detection based on randomized tensor algorithm[J]. *Laser Technology*, 2024, 48(1): 127–134. doi: [10.7510/jgjs.issn.1001-3806.2024.01.020](https://doi.org/10.7510/jgjs.issn.1001-3806.2024.01.020).
- [13] SUN Yang, YANG Jungang, LONG Yunli, *et al.* Infrared small target detection via spatial-temporal total variation regularization and weighted tensor nuclear norm[J]. *IEEE Access*, 2019, 7: 56667–56682. doi: [10.1109/ACCESS.2019.2914281](https://doi.org/10.1109/ACCESS.2019.2914281).
- [14] LIU Ting, YANG Jungang, LI Boyang, *et al.* Nonconvex tensor low-rank approximation for infrared small target detection[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2021, 60: 5614718. doi: [10.1109/TGRS.2021.3130310](https://doi.org/10.1109/TGRS.2021.3130310).
- [15] 胡亮, 杨德贵, 赵党军, 等. 改进非凸估计与非对称时空正则化的红外小目标检测方法[J]. *国防科技大学学报*, 2024, 46(3): 180–194. doi: [10.11887/j.cn.202403018](https://doi.org/10.11887/j.cn.202403018).
- HU Liang, YANG Degui, ZHAO Dangjun, *et al.* Infrared small target detection method based on improved non-convex estimation and asymmetric spatial-temporal regularization[J]. *Journal of National University of Defense Technology*, 2024, 46(3): 180–194. doi: [10.11887/j.cn.202403018](https://doi.org/10.11887/j.cn.202403018).
- [16] LI Fenghong, RAO Peng, SUN Wen, *et al.* A new motion feature-enhanced multiframe spatial-temporal infrared target detection network[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2025, 63: 5006819. doi: [10.1109/TGRS.2025.3603784](https://doi.org/10.1109/TGRS.2025.3603784).
- [17] XU Jielei, HAN Xinheng, WANG Jiacheng, *et al.* PCLC-Net: Parallel connected lateral chain networks for infrared small target detection[J]. *Remote Sensing*, 2025, 17(12): 2072. doi: [10.3390/rs17122072](https://doi.org/10.3390/rs17122072).
- [18] 龙畅, 张弦, 刘艳阳, 等. 天基红外小目标探测技术新进展、挑战与对策[J]. *光电工程*, 2025, 52(12): 250300. doi: [10.12086/oe.2025.250300](https://doi.org/10.12086/oe.2025.250300).
- LONG Chang, ZHANG Xian, LIU Yanyang, *et al.* Recent progress, challenges, and countermeasures in the detection of dim small targets using space-based infrared systems[J]. *Opto-Electronic Engineering*, 2025, 52(12): 250300. doi: [10.12086/oe.2025.250300](https://doi.org/10.12086/oe.2025.250300).
- [19] LI Fenghong, RAO Peng, SUN Wen, *et al.* A low signal-to-noise ratio infrared small-target detection network[J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2025, 18: 8643–8658. doi: [10.1109/JSTARS.2025.3550581](https://doi.org/10.1109/JSTARS.2025.3550581).
- [20] LIU Ming, DU Haoyuan, ZHAO Yuejin, *et al.* Image small target detection based on deep learning with SNR controlled sample generation[J]. *Current Trends in Computer Science and Mechanical Automation*, 2017, 1(1): 211–220. doi: [10.1515/9783110584974-025](https://doi.org/10.1515/9783110584974-025).
- [21] HE Kaiming, ZHANG Xiangyu, REN Shaoqing, *et al.* Deep residual learning for image recognition[C]. Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, USA, 2016: 770–778. doi: [10.1109/CVPR.2016.90](https://doi.org/10.1109/CVPR.2016.90).
- [22] 盛卫东, 吴双林, 肖超, 等. 可微稀疏掩模引导的红外小目标快速检测网络[J]. *电子与信息学报*, 2025, 47(12): 4779–4789. doi: [10.11999/JEIT250989](https://doi.org/10.11999/JEIT250989).
- SHENG Weidong, WU Shuanglin, XIAO Chao, *et al.* Differentiable sparse mask guided infrared small target fast detection network[J]. *Journal of Electronics & Information Technology*, 2025, 47(12): 4779–4789. doi: [10.11999/JEIT250989](https://doi.org/10.11999/JEIT250989).
- [23] ZHANG Mingjin, ZHANG Rui, YANG Yuxiang, *et al.* ISNet: Shape matters for infrared small target detection[C]. Proceedings of the 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), New Orleans, USA, 2022: 867–876. doi: [10.1109/CVPR52688.2022.00095](https://doi.org/10.1109/CVPR52688.2022.00095).
- [24] ZHANG Tianfang, LI Lei, CAO Siying, *et al.* Attention-guided pyramid context networks for detecting infrared small target under complex background[J]. *IEEE Transactions on Aerospace and Electronic Systems*, 2023, 59(4): 4250–4261. doi: [10.1109/TAES.2023.3238703](https://doi.org/10.1109/TAES.2023.3238703).
- [25] LI Chengyu, ZHANG Yan, SHI Zhiguang, *et al.* Moderately dense adaptive feature fusion network for infrared small target detection[J]. *IEEE Transactions on Geoscience and*

- Remote Sensing*, 2024, 62: 5616712. doi: [10.1109/TGRS.2024.3381006](https://doi.org/10.1109/TGRS.2024.3381006).
- [26] LIN Fanzhao, BAO Kexin, LI Yong, *et al.* Learning contrast-enhanced shape-biased representations for infrared small target detection[J]. *IEEE Transactions on Image Processing*, 2024, 33: 3047–3058. doi: [10.1109/TIP.2024.3391011](https://doi.org/10.1109/TIP.2024.3391011).
- [27] ZHOU Zongwei, RAHMAN SIDDIQUEE M M, TAJBAKHS N, *et al.* Unet++: A nested u-net architecture for medical image segmentation[C]. Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support 4th International Workshop, Granada, Spain, 2018: 3–11. doi: [10.1007/978-3-030-00889-5_1](https://doi.org/10.1007/978-3-030-00889-5_1).
- [28] DAI Yimian, WU Yiquan, ZHOU Fei, *et al.* Asymmetric contextual modulation for infrared small target detection[C]. Proceedings of the 2021 IEEE Winter Conference on Applications of Computer Vision (WACV), Waikoloa, USA, 2021: 949–958. doi: [10.1109/WACV48630.2021.00099](https://doi.org/10.1109/WACV48630.2021.00099).
- [29] DAI Yimian, WU Yiquan, ZHOU Fei, *et al.* Attentional local contrast networks for infrared small target detection[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2021, 59(11): 9813–9824. doi: [10.1109/TGRS.2020.3044958](https://doi.org/10.1109/TGRS.2020.3044958).
- [30] LI Boyang, XIAO Chao, WANG Longguang, *et al.* Dense nested attention network for infrared small target detection[J]. *IEEE Transactions on Image Processing*, 2023, 32: 1745–1758. doi: [10.1109/TIP.2022.3199107](https://doi.org/10.1109/TIP.2022.3199107).
- [31] LIN T Y, DOLLÁR P, GIRSHICK R, *et al.* Feature pyramid networks for object detection[C]. Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, USA, 2017: 2117–2125. doi: [10.1109/CVPR.2017.106](https://doi.org/10.1109/CVPR.2017.106).
- [32] 高凯, 王晟宇, 付强, 等. 基于多层级交互式特征融合的三维目标检测算法[J]. 吉林大学学报: 理学版, 2026, 64(3): 591–602. doi: [10.13413/j.cnki.jdxblxb.2025021](https://doi.org/10.13413/j.cnki.jdxblxb.2025021).
- GAO Kai, WANG Shengyu, FU Qiang, *et al.* Three-dimensional object detection algorithm based on multi-level interactive feature fusion[J]. *Journal of Jilin University: Science Edition*, 2026, 64(3): 591–602. doi: [10.13413/j.cnki.jdxblxb.2025021](https://doi.org/10.13413/j.cnki.jdxblxb.2025021).
- [33] GUO Huinan, ZHANG Nengshuang, ZHANG Jing, *et al.* Location-guided dense nested attention network for infrared small target detection[J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2024, 17: 18535–18548. doi: [10.1109/JSTARS.2024.3472041](https://doi.org/10.1109/JSTARS.2024.3472041).
- [34] HU Jie, SHEN Li, and SUN Gang. Squeeze-and-excitation networks[C]. Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, USA, 2018: 7132–7141. doi: [10.1109/CVPR.2018.00745](https://doi.org/10.1109/CVPR.2018.00745).
- [35] LI Xiang, WANG Wenhui, HU Xiaolin, *et al.* Selective kernel networks[C]. Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, USA, 2019: 510–519. doi: [10.1109/CVPR.2019.00060](https://doi.org/10.1109/CVPR.2019.00060).
- [36] WANG Qilong, WU Banggu, ZHU Pengfei, *et al.* ECA-Net: Efficient channel attention for deep convolutional neural networks[C]. Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Seattle, USA, 2020: 11531–11539. doi: [10.1109/CVPR42600.2020.01155](https://doi.org/10.1109/CVPR42600.2020.01155).
- [37] WOO S H, 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).
- [38] PARK J, WOO S, LEE J Y, *et al.* BAM: Bottleneck attention module[C]. Proceedings of the 2018 British Machine Vision Conference (BMVC), Newcastle, UK, 2018: 1–20. doi: [10.5244/C.32.1](https://doi.org/10.5244/C.32.1). (查阅网上资料,未找到本条文献doi信息,请确认).
- [39] LIU Yichao, SHAO Zongru, and HOFFMANN N. Global attention mechanism: Retain information to enhance channel-spatial interactions[EB/OL]. <https://arxiv.org/abs/2112.05561>, 2021.
- [40] ZHONG Yuhua, SHI Zhiguang, ZHANG Yan, *et al.* CSAN-UNET: Channel spatial attention nested UNet for infrared small target detection[J]. *Remote Sensing*, 2024, 16(11): 1894. doi: [10.3390/rs16111894](https://doi.org/10.3390/rs16111894).
- [41] ZUO Zhen, TONG Xiaozhong, WEI Junyu, *et al.* AFFPN: Attention fusion feature pyramid network for small infrared target detection[J]. *Remote Sensing*, 2022, 14(14): 3412. doi: [10.3390/rs14143412](https://doi.org/10.3390/rs14143412).
- [42] HOU Qibin, ZHOU Daquan, and FENG Jiashi. Coordinate attention for efficient mobile network design[C]. Proceedings of 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Nashville, USA, 2021: 13708–13717. doi: [10.1109/CVPR46437.2021.01350](https://doi.org/10.1109/CVPR46437.2021.01350).
- [43] MISRA D, NALAMADA T, ARASANIPALAI A U, *et al.* Rotate to attend: Convolutional triplet attention module[C]. Proceedings of 2021 IEEE Winter Conference on Applications of Computer Vision (WACV), Waikoloa, USA, 2021: 3138–3147. doi: [10.1109/WACV48630.2021.00318](https://doi.org/10.1109/WACV48630.2021.00318).
- 李卫星: 男, 副教授, 研究方向为红外图像处理及轻量化。
王 帅: 男, 副教授, 研究方向为边缘计算。
陈怀宇: 男, 助理研究员, 研究方向为红外图像处理。
盛卫东: 男, 副研究员, 研究方向为光电探测。

Infrared Small Target Detection Enhanced by Multi-dimensional Fusion Attention

LI Weixing^① WANG Shuai^② CHEN Huaiyu^① SHENG Weidong^①

^①(College of Electronic Science and Technology, National University of Defense Technology,
Changsha 410073, China)

^②(Xi'an Institute of Optics and Precision Mechanics, Chinese Academy of Sciences, Xi'an 710061, China)

Abstract:

Objective Infrared imaging boasts advantages such as long operating range, wide coverage, strong concealment, and all-weather visibility, making it widely applicable in aerospace target surveillance, maritime emergency rescue, forest fire monitoring, and earth remote sensing. Infrared payloads are typically mounted on platforms like satellites and aircraft, capturing images over long distances where targets appear small in the imagery and lack distinct texture and morphological features. Due to weak thermal radiation signals, targets are prone to being submerged in strong clutter. Current infrared small target detection networks face several key challenges. First, target segmentation networks expand the receptive field through continuous down-sampling, which can cause the features of infrared small targets to be easily lost in the deeper layers of neural networks. Second, features in the deeper neural network layers tend to diffuse. Therefore, under the constraints of long-range imaging and complex backgrounds, achieving high-precision and highly reliable infrared small target detection remains a research hotspot in the field of infrared imaging.

Methods A channel-spatial Multi-Dimensional Fusion Attention Mechanism (MFAM) is proposed to address the challenges of small target size and feature diffusion in deep networks. Multi-dimensional features across channels, height, and width are captured, as well as the dependencies across these dimensions. By integrating feature fusion and cross-dimensional synergistic interaction, feature diffusion of small targets in deep networks is mitigated and the robustness of infrared small-target detection is enhanced. The channel attention and spatial attention modules are applied in parallel directly to the input feature maps, performing feature extraction and local information interaction along the channel and spatial dimensions, respectively. In the channel domain, features are compressed and processed through a Multi-layer Perceptron (MLP) to capture inter-channel relationships. In the spatial domain, Global Average Pooling (GAP) and Global Maximum Pooling (GMP) are employed to encode width and height dimensions. Finally, feature fusion is achieved via a Sigmoid activation function. Compared with traditional hybrid attention mechanisms, the channel attention module and spatial attention module are applied in parallel directly to the input feature maps. This approach enables refined encoding of the channel, height, and width dimensions under a global receptive field. Meanwhile, shallow features with spatial details and deep semantic features rich in contextual information are captured. The proposed MFAM features a plug-and-play design, allowing flexible integration into various baseline models such as ResNet and DNA-Net without introducing complex additional structures, which demonstrates excellent compatibility and generalization capability.

Results and Discussions The publicly available dataset NUDT-SIRST is utilized to conduct a performance analysis of the proposed algorithm. By incorporating the proposed MFAM into the baseline DNA-Net, the Intersection over Union (IoU), detection rate (Pd), and false alarm rate (Fa) are 87.34%, 98.72% and 3.22×10^{-6} , respectively. Compared to CBAM, BAM, GAM, CA, and TA, MFAM improves IoU by 0.4%, 1.68%, 2.46%, 2.11%, and 1.61%, respectively (Table.1). CBAM places the spatial attention module in series of the channel attention, which may lead to information loss during feature propagation. CA and BAM only employ GAP for information encoding, overlooking the role of max pooling in deep feature extraction. TA realizes cross-dimensional dependencies through rotation operations but fails to achieve simultaneous three-dimensional interaction. MFAM simultaneously integrates channel and spatial information, enabling sufficient cross-domain interaction, and combines global average and max pooling to enhance context and detailed feature extraction capabilities, thereby achieving more refined infrared small target detection. MFAM is also embedded into ALC-

Net and AMFU-Net, resulting in IoU improvements of 0.43% and 0.28% compared with CBAM([Table.2](#)). Through ablation experiments, detection performance is compared under different attention combination strategies. Compared to the serial fusion approach, the proposed parallel structure achieves improvements of 0.37% and 0.39% in IoU and Pd, respectively, while reducing Fa by 1.41×10^{-6} . The advantage stems from the parallel structure applying channel and spatial attention directly to the input features, mitigating the diffusion of target features in deep networks. To validate the inference performance of the proposed algorithm on edge-side processors, a verification system is designed by using FPGA and NVIDIA Jetson AGX Xavier. Practical testing confirms that the proposed algorithm can be deployed and inferred on edge-side GPUs, with an average single-frame inference latency of 46.7 ms ([Fig. 8](#)) for 256×256 input images.

Conclusions A Multi-dimensional Fusion Attention Module (MFAM) is constructed in this paper. The MFAM module achieves effective aggregation of channel and spatial salient features and enables cross-dimensional adaptive interaction, enhancing the preservation of target features in deep networks and ensuring robust output. The MFAM module exhibits favorable plug-and-play characteristics. Experiments demonstrate that the proposed algorithm performs better in metrics of IoU, Pd, and Fa. A lightweight intelligent processing unit based on FPGA+GPU is developed, which successfully achieves deployment of the algorithm on edge devices and enables high real-time inference, demonstrating promising engineering applicability and future application prospects.

Key words: Infrared small target; Attention mechanism; Feature fusion; Edge intelligence