基于多感受野的红外弱小目标检测算法

李奇峰，郝问裕，安羽翔，张 峰，郭继光

中国电子科学研究院 北京

【摘要】针对红外图像中弱小目标像素占比低、特征稀疏、背景干扰严重等问题，提出一种基于多感受野特征融合的红外弱小目标检测算法。该算法以 YOLOv5s 为基础检测框架，通过设计多分支感受野扩展模块，在颈部网络中构建多尺度上下文信息融合机制，增强了网络对弱小目标的特征表达能力；同时引入自适应特征加权策略，有效抑制复杂背景噪声干扰。在公开红外弱小目标数据集上的实验结果表明，改进后算法相较于基线 YOLOv5s，在检测精度、召回率及定位准确性等指标上均有显著提升，验证了多感受野融合策略对红外弱小目标检测任务的有效性。

【关键词】红外弱小目标检测；多感受野；特征融合；YOLOv5s；深度学习

【基金项目】国家自然科学基金企业创新发展联合基金（U23B2050）

【收稿日期】2026 年 6 月 12 日 【出刊日期】2026 年 7 月 15 日 【DOI】10.12208/j.jer.20260035

Infrared dim-small target detection algorithm based on multiple perception fields

Qifeng Li, Wenyu Hao, Yuxiang An, Feng Zhang, Jiguang Guo

China Academy of Electronic and Information Technology, Beijing

【**Abstract**】To address the challenges of extremely low pixel occupancy, sparse features, and severe background interference in infrared dim and small target detection, this paper proposes an infrared dim and small target detection algorithm based on multi-receptive field feature fusion. The algorithm adopts YOLOv5 as the baseline detection framework and constructs a multi-scale contextual information fusion mechanism within neck networks by designing a multi-branch receptive field expansion module, thereby enhancing the network's feature representation capability for dim and small targets. Meanwhile, an adaptive feature weighting strategy is introduced to effectively suppress interference from complex background clutter. Experimental results on public infrared dim and small target datasets demonstrate that, compared with the baseline YOLOv5, the improved algorithm achieves significant improvements in detection precision, recall, and localization accuracy, validating the effectiveness of the multi-receptive field fusion strategy for infrared dim and small target detection tasks.

【**Keywords**】Infrared dim and small target detection; Multi-receptive field; Feature fusion; YOLOv5; Deep learning

引言

当前，红外探测技术以其被动工作、隐蔽性强、可全天候作业等技术特点，在预警监视、精确制导、安防监控等军事和民用领域得到了广泛应用[1]。但是，对于红外弱小目标而言，由于其图像尺寸小、信噪比低且易受复杂背景杂波干扰，导致对其进行有效检测一直是红外探测领域的难点技术[2]。

红外目标检测因其重要的军民应用价值，一直受到国内外学者的重视和研究。叶斌、彭嘉雄提出的基于形态学 Top-Hat 算子的目标检测方法[3]，通过结构元素预先估计背景并利用差分运算实现背景抑制，算法简单、实时性好，但当目标尺寸超出预设结构元素范围时性能明显下降。周姣、辛云宏等引入尺度空间理论[4]，利用多尺度高斯核构建尺度空间进行目标增强，在一定程度上缓解了固定尺寸滤波核的适应性不足问题，但在非平稳背景下虚警率仍然较高。吴一全等提出基于 NSCT 的红外小目标检测方法[5]，利用非下采样 Contourlet 变换的多尺度、多方向分解能力，在变换域中分析目标和背景子带能量分布差异，对复杂纹理背景具有较好的抑制能力，但算法计算复杂度较高、实时性难以保证。

随着深度学习的发展，基于卷积神经网络的目标

检测框架逐渐成为主流。顾佼佼等通过多尺度特征拼接和 Anchor 重设计改进 Faster R-CNN[6]，在红外舰船检测中平均精确度提升至 83.98%。YOLO 系列因其良好的速度-精度平衡受到广泛关注[7-9]，DCS-YOLO 提出双域融合与类别感知自监督机制，TY-RIST 通过步长感知骨干和高分辨率检测头实现实时高精度检测。然而，上述通用检测框架直接应用于红外弱小目标时，仍面临深层特征小目标信息丢失、感受野与目标尺度不匹配等问题。

本文首先对红外弱小目标特性进行分析，对 YOLOv5s 检测框架的结构及其在弱小目标检测中的局限性进行深入研究；在此基础上，提出一种基于多感受野特征融合的红外弱小目标检测改进算法，通过设计多分支感受野扩展模块与自适应特征融合策略，增强网络对弱小目标的多尺度感知能力与背景抑制能力，以提升红外弱小目标的检测性能。

1 红外弱小目标特性分析

红外弱小目标检测之所以面临诸多困难，其根源在于目标自身所具有的固有属性，以及这些属性同成像环境之间形成的复杂耦合效应。唯有对这些特性加以深入剖析，方能为有效检测算法的设计提供可靠的前提与依据。基于此，本节将依次围绕目标定义与特性、成像模型简析、检测算法的设计要求三个方面展开讨论。

1.1 红外弱小目标特性

红外弱小目标一般是指在成像面上占据像素数不超过总像素 0.15%、且信噪比不足 3 的一类目标[10]。此类目标呈现出以下几个方面的典型特征：

（1）尺度极为有限。在常规红外成像系统下，远距离目标投射至焦平面后仅覆盖数个至数十个像素，可用的形状、轮廓及纹理信息均严重缺失。图像中的目标表现为近乎点状的局部亮斑，仅凭几何形态特征几乎无法将其与噪声点加以有效区分。

（2）辐射能量微弱。受大气衰减、光学系统损耗以及远距离传输等因素的综合作用，抵达探测器的目标辐射能量极为有限。从灰度分布来看，目标区域像素值仅略微超出局部背景均值，对比度极低，很容易被背景杂波与探测器噪声所遮蔽。

（3）形态缺乏规律性。大气湍流、光学散焦及探测器采样等效应的共同影响，使得目标在图像中的灰度分布常表现出无确定形态的点扩散特征，难以用结构化的方式加以描述，这也导致基于模板匹配或形状先验的检测思路难以有效施展。

（4）背景干扰突出。在实际场景中，目标极易被云层边缘、地形起伏、海面杂波等复杂背景所淹没。上述背景区域往往具有与目标相近的局部亮度特性或边缘结构，由此产生大量虚警源，使得检测难度进一步加大。

1.2 红外成像模型简析

红外弱小目标图像的成像过程可描述为目标信号、背景杂波与探测器噪声三者的线性叠加。设红外图像为$f(x，y)$，则其数学模型可表示为：

$$f(x，y)=f_T(x，y)+f_B(x，y)+n(x，y) \quad （1）$$

其中，$f_T(x，y)$为目标信号分量，通常可用二维高斯函数近似描述其在空间域的扩散特性；$f_B(x，y)$为背景杂波分量，在空间上通常表现为低频缓变的连续区域，但在云层边缘、地形交界等处具有较强的非平稳性；$n(x，y)$为探测器引入的随机噪声分量，通常服从高斯分布或泊松分布。

1.3 对检测算法的设计要求

红外弱小目标检测的核心任务在于从背景与噪声的联合分布中剥离出微弱的目标信号。当背景杂波呈现非平稳特性或噪声强度偏高时，目标信号同干扰分量在幅值域与空间域上的可分性将大幅下降，检测难度随之骤增。因此，一个有效的红外弱小目标检测算法应当满足如下设计要求：

（1）充分保留微弱空间响应。检测算法需要对微弱信号保持足够的敏感度，在网络前向传播环节中尽可能减少目标信息的损失。这对深度学习检测框架构成了重要约束——不宜直接沿用通用目标检测网络中频繁的下采样策略，而应采取针对性措施对低层特征图中的微弱空间响应信号加以保护。

（2）具备强背景抑制能力。算法应当有效压制云层、地物等复杂背景所产生的杂波干扰，从而降低虚警率。这意味着算法需在局部区域中提取可区分目标与背景的有效特征，例如局部对比度、方向梯度等，或将背景建模与目标增强进行有机结合。

（3）多尺度感知与自适应处理能力。由于目标尺寸的先验信息往往未知且存在一定变化范围，算法需具备在不同感受野尺度下并行感知目标的能力，防止因固定感受野与目标尺度不匹配而引发漏检或虚警。与此同时，对不同场景类型的自适应处理能力也是增强算法鲁棒性的关键所在。

2 融合多感受野的 YOLOv5s 改进算法

YOLOv5s 作为单阶段目标检测算法的代表性框架

之一，以其良好的检测精度与推理速度平衡而受到工业界和学术界的广泛关注。本文选择 YOLOv5s 作为基线检测框架，在其基础上针对红外弱小目标检测任务进行改进设计。

2.1 YOLOv5s 检测框架

YOLOv5 网络包括 YOLOv5s、YOLOv5m、YOLOv5l、YOLOv5x 等多种网络模型。其中 YOLOv5s 的参数量少、推理速度快，且网络深度与宽度适中，多用于红外弱小目标检测，因此本文主要在该网络模型基础上进行优化，并将该模型作为基线模型进行结果对比，其网络模型如下图所示。

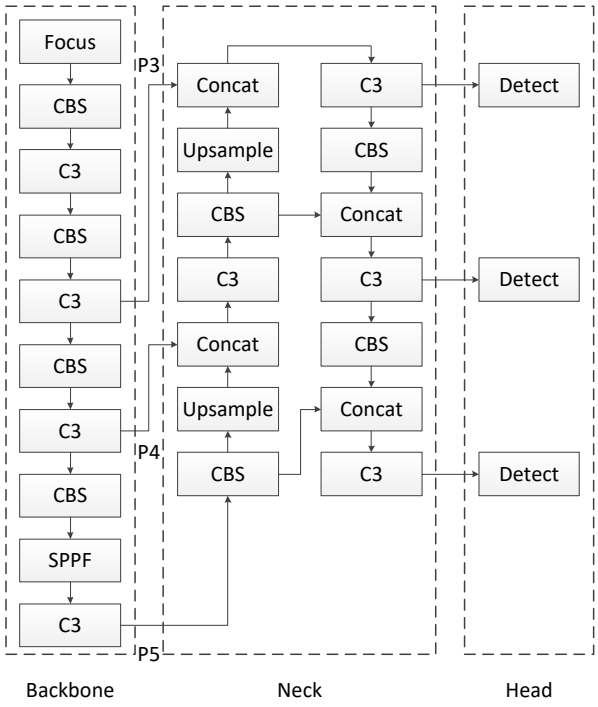


图 1　YOLOv5s 网络模型结构图

YOLOv5s 由主干网络（Backbone）、颈部网络（Neck）和检测头（Head）三个主要部分组成。

其中 Backbone 采用 CSPDarknet 结构，其核心组成包括：Focus 模块，通过切片操作将输入特征图的空间分辨率减半而通道数扩增为四倍，实现无信息损失的下采样；CBS 模块，由卷积层、批归一化层与 SiLU 激活函数串行构成的标准卷积单元；C3 模块，基于跨阶段局部网络设计思想，通过部分通道的跨越连接减少计算瓶颈，同时保持丰富的梯度流动；SPPF（Spatial Pyramid Pooling - Fast）模块，通过串联多个相同尺寸的最大池化层，以较小的计算代价实现多尺度感受野的融合。

Neck 采用 PANet 结构的特征金字塔，在 FPN 自顶向下的语义信息传递通道基础上，额外增加自底向上的空间信息增强通道。通过双向跨层特征融合，颈部网络将骨干网络不同层级的特征图进行聚合，使浅层特征获得更丰富的语义上下文、深层特征保留更精确的空间定位信息。

Head 在颈部网络输出的多尺度特征图上施加卷积预测层，输出目标的类别概率、置信度分数及边界框回归参数。各尺度检测头独立完成预测，最终通过非极大值抑制融合多尺度检测结果。

YOLOv5s 在进行红外弱小目标检测时，存在以下几个方面的问题。

（1）下采样导致目标信息严重丢失

Backbone 经过多次下采样操作后，特征图的空间分辨率已降至原始图像的几分之一甚至十几分之一。由于红外弱小目标本身可能仅有几个像素，因此目标的空间响应信号极易被压缩至不可分辨的程度。

（2）标准卷积的感受野与目标尺度不匹配

Backbone 和 Neck 以标准 3×3 卷积为主要特征提取算子，其感受野范围相对固定，难以全面覆盖不同尺度目标的需求，导致小感受野缺乏对周围背景关联的有效感知，易产生虚警；大感受野则可能引入过多无关背景信息，稀释目标的微弱信号。

（3）缺乏针对弱小目标的专项优化设计

YOLOv5s 对通用场景的设计，将导致针对红外弱小目标正样本匹配困难、锚框与目标尺度严重偏离、损失函数中对弱小目标的监督信号被大目标淹没等问题。

本文以 YOLOv5s 作为基础网络框架，在保持其原有端到端单阶段检测范式的基础上，针对红外弱小目标嵌入多感受野模块，提升对红外弱小目标的检测能力。

2.2 红外目标检测流程

红外弱小目标检测流程如图 2 所示。多感受野模块（MRF）主要用于 Neck 部分进行融合增强（图中红框部分）。

红外图像输入网络后首先进行马赛克拼接、多尺度缩放、归一化等预处理操作；然后 Backbone 逐层进行特征提取，输出 P3、P4、P5 特征图；Neck 通过 PANet 双向特征金字塔对三层特征图进行融合，其中 P3 在输出前融合多感受野模块并进行注意力增强；最后，Head 三个检测头分别进行密集预测。

2.3 多感受野模块

在 Neck 中设计在 P3/8 小目标特征层的 C3 输出之后插入一个轻量化多感受野模块，并进行注意力增

强（Attention）。Neck 中 P3 分支对红外小目标最为敏感，该设计可在提升对红外弱小目标检测能力的同时，避免因大范围替换 C3 模块造成的模型训练不稳定和背景噪声放大。

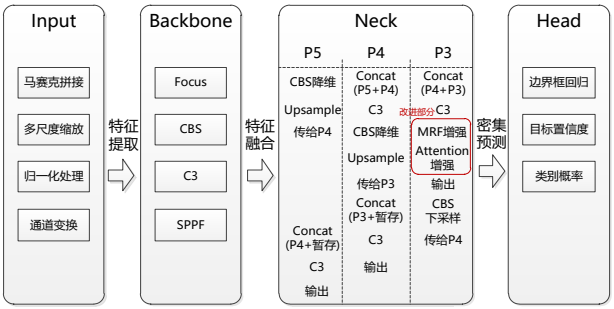


图 2　红外弱小目标检测流程图

轻量化多感受野模块计算流程如下：

$$F1 = Conv1 \times 1\ （Fp3）\tag{2}$$

$$Flocal = DWConv3 \times 3\ （F1）\tag{3}$$

$$Fc = ECA\ （Flocal）\tag{4}$$

$$Fs = SA\ （Fc）\tag{5}$$

$$Fout = Fp3 + Fs\tag{6}$$

其中$DWConv3 \times 3$表示3×3深度可分离卷积，ECA表示高效通道注意力，SA表示空间注意力。残差连接用于保留原始 P3 特征，降低额外增加的多感受野模块对主干特征分布的影响。

ECA通道注意力可表示为：

$$ECA(F) = F \cdot sigmoid\ （Conv1D\ （GAP\ （F）））\tag{7}$$

SA空间注意力可表示为：

$$SA = F \cdot sigmoid\ （Conv7 \times 7\ （[\text{AvgPool_c}（F），\text{MaxPool_c}（F）]））\tag{8}$$

3 对比实验与结果分析

3.1 实验环境与数据集选择

本文基于 NCHU-SIRST 与 IRST640 合并数据集对基线模型 YOLOv5s 以及融合多感受野的改进模型进行算法性能的评估实验。其中训练集图像共 638 幅，包含 677 个红外弱小目标；测试集图像共 409 幅，包含 459 个红外弱小目标。实验环境为 Python3.13.9，Pytorch2.5.1，两个模型分别训练 100 轮。

3.2 评价指标

实验选择准确率（Precision）、召回率（Recall）、mAP50 和 mAP50-95 作为主要指标。

准确率P的含义是模型所有检测结果中正确预测结果所占的比例，其表示如下所示：

$$P = \frac{TP}{TP+FP}\tag{9}$$

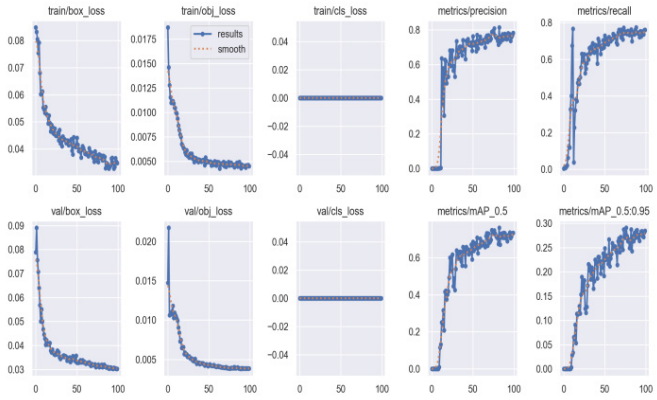
召回率R的含义是模型对红外弱小目标检测中被正确检测的红外弱小目标所占的比例，其表示如下所示：

$$R = \frac{TP}{TP+FN}\tag{10}$$

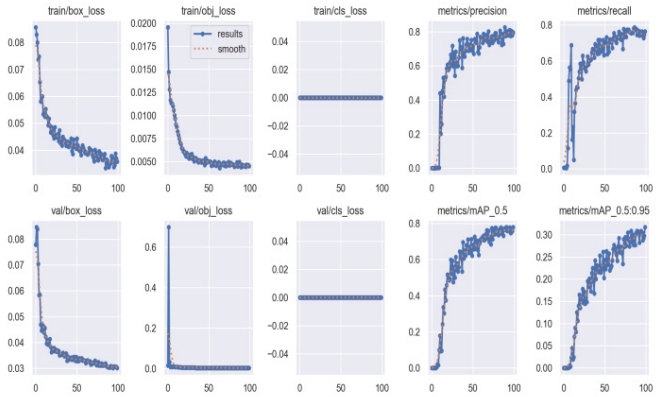
$mAP50$的含义是交对比（IoU）阈值为 0.5 时的平均准确率；$mAP50-90$的含义是 IoU 阈值从 0.5 到 0.95、步长为 0.05 时的平均准确率。

3.3 对比实验

本节将本文算法与 YOLOv5s 模型进行实验对比，模型训练过程如下图 3 所示。



（a）YOLOv5s 模型训练过程曲线



（b）融合多感受野模型训练过程曲线

图 3　模型训练过程对比图

从训练过程曲线上可以发现，YOLOv5s 模型的最佳$mAP50-90$出现在第 76 轮，融合多感受野模型的最佳$mAP50$和$mAP50-90$均出现在第 99 轮。实验结果表明融合多感受野后模型在后期仍能继续获得检测性能收益。

两个模型在数据集上的验证结果如下表 1 所示。

结果显示，融合多感受野模型在合并测试集上$mAP50$提升了 3.1%，$mAP50-90$提升了 2.5%；在 NCHU 子集上准确率、召回率、$mAP50$和 $mAP50-90$均提升；在 IRST640 子集上准确率下降，但召回率、$mAP50$和 $mAP50-90$提升，说明该结构更倾向于增强弱小目标召回能力。

两个模型的误差分析结果如下表 2 所示。

结果显示，在合并验证集上，融合多感受野模型相比 Baseline 增加 10 个 TP，减少 30 个 FP，减少 10 个 FN，错误图像数从 151 降至 134。该结果说明融合多感受野模块后，模型不只是提高平均准确率，同时在固定阈值下也改善了实际检测错误分布。部分检测结果如下图 4 所示。

表 1　验证结果对比

验证集	模型	P	R	mAP50	mAP50-95
合并 val	YOLOv5s	0.808	0.749	0.747	0.291
合并 val	多感受野	0.795	0.765	0.778	0.316
NCHU val	YOLOv5s	0.833	0.771	0.810	0.317
NCHU val	多感受野	0.850	0.779	0.837	0.342
IRST640 val	YOLOv5s	0.708	0.663	0.569	0.225
IRST640 val	多感受野	0.653	0.707	0.592	0.241

表 2　误差分析

验证集	模型	TP	FP	FN	P	R	error images
合并 val	YOLOv5s	359	154	100	0.6998	0.7821	151
合并 val	多感受野	369	124	90	0.7485	0.8039	134
NCHU val	YOLOv5s	295	110	72	0.7284	0.8038	116
NCHU val	多感受野	303	87	64	0.7769	0.8256	101
IRST640 val	YOLOv5s	63	43	29	0.5943	0.6848	35
IRST640 val	多感受野	66	35	26	0.6535	0.7174	32



图 4　融合多感受野模型部分检测结果

4 结语

本文对比了原始 YOLOv5s 模型与融合多感受野模型在红外弱小目标上的检测能力。试验结果表明，在 P3 小目标分支加入轻量化多感受野注意力模块即可有效提升红外弱小目标的检测能力。融合多感受野模型在合并验证集$mAP50$从 74.7%提升至 77.8%，$mAP50-90$从 29.1%提升至 31.6%；除 IRST640 子集准确率外，其余指标均优于 YOLOv5s 模型。FP/FN 分析进一步表明，融合多感受野模型能同时降低误检和漏检，具有更好的实际检测可靠性。后续针对当前实验结果展现的模型在 IRST640 子集准确率下降的问题，继续开展目标误检和漏检分析，解决当前 IRST640 子集背景干扰导致的目标误检问题。

参考文献

[1] 凡遵林, 王浩, 管乃洋,等. 单帧红外图像弱小目标检测研究综述[J]. 红外技术, 2023, 45(11): 1133-1140.

[2] 罗俊海, 余杭. 基于低秩稀疏分解的红外弱小目标检测算法研究进展[J]. 激光与光电子学进展, 2023, 60(16): 1600004.

[3] 叶斌,彭嘉雄. 基于形态学 Top_Hat 算子的小目标检测方法[J]. 中国图像图形学报,2022,9(7)：638-642.

[4] 周姣,辛云宏. 基于显著性与尺度空间的红外弱小目标检测[J]. 激光与红外,2015,45(4)：452-456.

[5] 吴一全,罗子娟,吴文怡. 基于 NSCT 的红外图像小目标检测技术[J]. 中国图像图形学报,2009,14(3)：477-481.

[6] 顾佼佼,李炳臻,刘克,等. 基于改进 Faster_R_CNN 的红外舰船目标检测算法[J]. 红外技术,2021,43(2)：170-178.

[7] Redmon J, Farhadi A.YOLO9000 : Better , faster , stronger[C]// Proc of 2017 IEEE Conference on Computer Visionand Pattern Recognition,2017:6517-6525

[8] Redmon J, Farhadi J R A.YOLOv3:An incremental improvement[J]. arXiv: 1804.02767, 2018.

[9] Bochkovskiy A,Wang C Y,Liao H Y M.YOLOv4: Optimalspeed and accuracy of object detection[J]. arXiv:2004.10934,2020.

[10] 杨昳,徐长彬,马玉莹,等. 低信噪比下的红外弱小目标检测算法研究综述[J]. 激光与红外,2019,49(6)：643-649.

版权声明：©2026 作者与开放获取期刊研究中心(OAJRC)所有。本文章按照知识共享署名许可条款发表。
https://creativecommons.org/licenses/by/4.0/
OPEN ACCESS