

基于深度学习的红外小目标检测方法综述

阿勒尔呷

西南民族大学电子信息学院, 四川 成都

收稿日期: 2025年5月24日; 录用日期: 2025年6月23日; 发布日期: 2025年6月30日

摘要

红外小目标检测作为智能安防的核心技术, 针对低信噪比、弱表征等难题, 系统构建了红外小目标检测的理论框架与方法体系。首先, 通过公开数据集开展多维度特性分析, 结合背景杂波建模与目标辐射特性揭示成像机理。继而聚焦深度学习范式, 建立分类研究体系。单阶段检测方向聚焦YOLO与SSD系列算法的效率与精度平衡机制改进。在CNN方法层面, 关注多尺度特征融合、注意力协同、轻量化编解码、密集嵌套特征交互关键技术以及Transformer与小样本学习在红外小目标检测中的应用。最后指出, 未来应倾向于物理模型驱动的新型网络架构设计, 通过辐射特性建模与深度学习特征融合提升算法鲁棒性。发展边缘计算友好的轻量化模型, 采用神经架构搜索技术实现模型参数量压缩。探索红外、可见光、雷达多模态协同检测机制。开发自监督小样本学习范式, 利用生成对抗网络缓解数据稀缺问题。这些突破将检测技术向全天候、高实时、强泛化方向持续演进。

关键词

红外小目标, 目标检测, 深度学习, 小样本学习, 多模态

A Review of Deep Learning-Based Infrared Weak Target Detection Methods

Aleerxia

College of Electronics and Information, Southwest Minzu University, Chengdu Sichuan

Received: May 24th, 2025; accepted: Jun. 23rd, 2025; published: Jun. 30th, 2025

Abstract

Infrared Small Target Detection (ISTD), a pivotal technology in intelligent security systems, addresses challenges such as low signal-to-noise ratio (SNR) and weak feature representation through a systematic framework. The methodology begins with multi-dimensional characteristic analysis using public datasets, integrating background clutter suppression and target radiation modeling to

elucidate imaging mechanisms. Under a deep learning paradigm, the research establishes a dual-axis classification: Single-stage detection optimizes efficiency and accuracy trade-offs in YOLO and SSD architectures via anchor-free designs and feature pyramid enhancements. CNN-based innovations focus on multi-scale fusion, attention mechanisms (spatial-channel and coordination), lightweight encoders (depthwise separable convolutions), dense cross-layer interaction for enhanced feature extraction, as well as cutting-edge applications of Transformer and few-shot learning in infrared small target detection. Future priorities emphasize: Physics-guided networks merging radiation physics with deep features for robustness. Edge-optimized lightweight models via neural architecture search (NAS) and pruning. Explore the multimodal cooperative detection mechanism of infrared, visible light and radar. Self-supervised few-shot learning using GANs to overcome data scarcity. These breakthroughs will continue to evolve the detection technology towards all-weather, high real-time, and strong generalization.

Keywords

Infrared Weak Targets, Target Detection, Deep Learning, Few-Shot Learning, Multimodal

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

准确且高效地定位出图像中目标所在位置，是红外目标检测的核心任务[1]。作为计算机视觉领域的研究重点和热点，红外小目标检测一直是具有挑战性和广泛研究的课题。该技术在军事、安防、无人驾驶等领域具有广泛的应用价值[2]。

传统图像检测技术依赖于其高分辨率和丰富可见光图像。然而，可见光成像在夜间或低照度场景中的性能急剧下降，且受限于穿透能力与成像距离。相比之下，红外检测技术依托长波红外辐射特性，能够突破可见光依赖环境光照的瓶颈，具备全天候工作能力。穿透力强、可识别隐蔽目标、适用于远距离探测，这些特性使其在安防监控、军事侦察及工业无损检测领域具有不可替代性。

传统的检测算法，如基于滤波的红外小目标检测算法[3]，主要分为基于空域滤波和频域滤波的算法。此类算法的思路简单、计算量小、整体效果差；基于人类视觉系统的检测方法[4]，主要分为基于频谱残差的算法和基于局部对比度的算法两类。此类算法本身局限性太大；但基于深度学习的目标检测算法在检测中取得越来越好的结果，采用深度学习对红外目标进行检测，显示出较强的泛化能力和可行性，这对于红外目标检测算法的发展有很好的前景。

本文首先分析红外微小目标特性，继而综述基于 One-stage 检测和 CNN 的深度学习方法在红外小目标检测中的应用，对比其优缺点及适用场景。针对当前技术瓶颈，重点探讨数据匮乏、目标尺度变化及复杂背景干扰等挑战。最后从网络架构优化、数据增强及多模态协同检测机制策略三方面展望未来发展方向。

2. 红外微小目标特性分析

2.1. 红外小目标公开数据集

红外图像的获取是利用物体发射的红外辐射进行图像获取，通过红外相机或红外传感器捕捉并转换红外辐射成可视化的灰度图像[5]。红外微小目标图像由目标、背景和噪声三部分组成[6]。

红外微小目标检测面临多重挑战：其一，目标辐射强度弱、与背景温差小，导致检测定位困难；其二，目标常像素尺寸较小，易被噪声、背景干扰或模糊效应掩盖；其三，热成像系统的低空间分辨率使目标呈现边缘模糊、对比度低、形状特征不显著且缺乏纹理细节的特性[7]。为了进一步分析红外小目标的特征，对公开数据集进行统计分析。

在深度学习中，数据集的规模大小和多样性对提高模型的性能和鲁棒性至关重要。通过对现有文献的深入分析，表 1 中，总结了五个公共红外小目标数据集。

Table 1. Infrared small targets expose datasets
表 1. 红外小目标公开数据集

数据集	训练集	测试集
MDvsFA	9978	100
SIRST-Aug	8525	545
SIRST	341	86
NUDT-SIRST	1061	265
IRSTD-1k	800	201

数据来源：MDvsFA (<https://mdvsfadataset.org>)；SIRST-Aug (<https://github.com/SIRST-Benchmark/SIRST-Aug>)；SIRST (<https://ieee-dataport.org/documents/sirst-dataset>)；NUDT-SIRST (<https://github.com/YeRen123455/Infrared-Small-Target-Detection>)；IRSTD-1k (<https://github.com/udasl/IRSTD-1K>)。

2.2. 红外小目标特征分析

2.2.1. 红外目标的背景特性

红外小目标图像特征分析中的红外背景特性是指在红外图像中，背景区域的特征和性质，对于红外小目标的检测和识别具有重要的影响[8]。

红外背景特性主要体现为三方面，辐射强度分布不均，地面、建筑等物体热辐射差异形成复杂纹理；动态温度变化，自然环境和物体热特性变化引发背景温度时空动态波动；噪声干扰显著，热噪声、大气湍流等干扰源易掩盖弱小目标，需通过滤波和去噪技术抑制。

2.2.2. 弱小目标特性

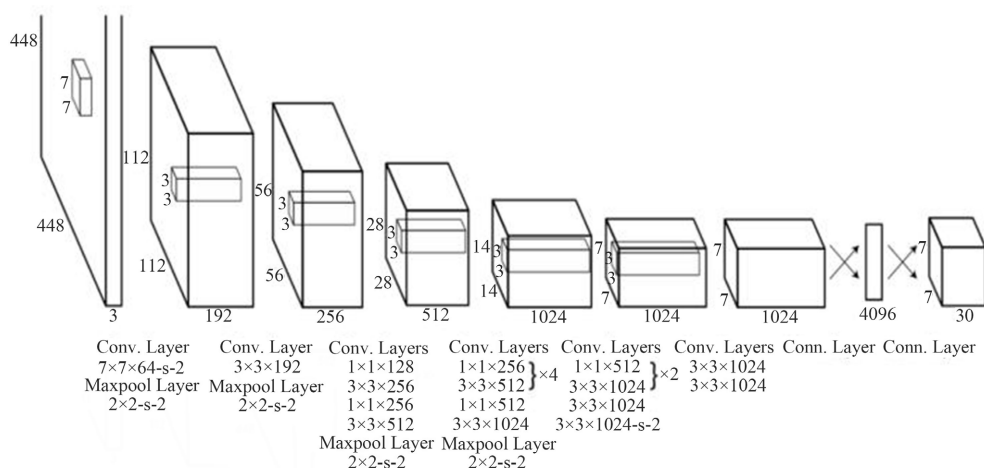
红外小目标特性是指在红外图像中，目标区域相对于背景区域的特征和性质[9]。红外小目标的特性分析对于目标检测、识别和跟踪等应用具有重要的意义。

红外小目标检测面临三重核心挑战，一是远距成像致目标能量衰减严重，信噪比低且边界模糊；二是目标像素稀疏，呈无纹理点状特征；最后是背景占主导形成极端样本不平衡[10]。相较可见光，红外图像缺乏色彩与结构信息，尤其在复杂背景下目标边缘解析困难。这些特性既构成技术瓶颈，也催生出算法创新机遇——通过开发抗噪增强、多尺度融合及弱特征强化技术，可突破现有检测精度极限，推动红外感知向隐蔽目标识别等战略领域延伸。

3. 基于 One-Stage 检测的红外小目标检测方法

3.1. 基于 YOLO 系列算法及其改进算法检测方法

在 2016 年，Joseph Redmon 等人提出了一种著名的 YOLO 的算法结构见图 1 [11]，通过取消候选框生成步骤实现端到端检测。其单阶段架构虽大幅提升实时性，却因网格粗粒度划分导致小目标漏检率较高，定位精度受限。这一核心矛盾推动了后续 YOLOv3 引入多尺度预测等优化策略。



来源于 Redmon, Joseph, *et al.* "You only look once: Unified, real-time object detection." Proceedings of the IEEE conference on computer vision and pattern recognition. 2016。

Figure 1. YOLO network structure

图 1. YOLO 网络结构

YOLOv2 在第一代算法模型的基础上引入了 Batch Normalization、Anchor Box 结构来提升网络检测精度。Lin Xupeng [12]等人在此基础上，将红外图像中的移动目标作为检测对象采用维度聚类的方法来确定目标所在的边界框位置。通过使检测框边界更加贴合目标边界，该方法在不同程度上减少了误检和漏检的情况，从而提高了目标检测的准确度。

更为经典的 YOLOv3 [13]通过引入残差连接与特征金字塔网络架构实现多尺度特征融合底层特征保留小目标细节，高层特征提取语义信息。其多尺度预测机制使小目标检测精度较 YOLOv2 提升 30%，成为工业级实时检测的标杆模型。

Li Mukai 等人在此网络基础上进行优化，提出了一种满足实时性要求的红外行人检测算法[14]。由于 YOLOv3 中分类准确率还可以优化，借鉴 SENet 中对特征进行权重重标定的思路，将 SE block 模块引入 YOLOv3 网络结构中，提升了网络的检测精度。Qiwei Xu [15]等人在 YOLOv3 骨干网络中引入扩展卷积，增强网络图像特征提取，提高网络对红外小目标的识别精度与实时检测。Liu [16]等人基于 YOLOv3 结构，连接两个具有相同宽度和高度的 ResNet 单元来优化 darknet 中的 Resblock；并在前一层增加卷积运算来丰富空间信息，从而改进整个 darknet 结构，从而提高了小目标检测的性能。

YOLOv5 [17]作为性能更强的检测网络，仍可在压缩通道、优化参数等方面进一步提升检测性能。S. Li [18]等人在 YOLOv5 基础上，在残差块中引入注意模块，以聚焦对象和抑制背景，并增加了多尺度检测，提高了对小目标检测的精度。Mou X [19]等人在提出了基于特征重组采样的检测模 YOLO-FR，并将骨干网络提取的特征与颈部网络的顶级语义信息融合，Wang Y [20]等人提出 YOLOv5-AFAM 算法，将 Maxpool 和卷积相结合，改进了下采样层，提升微小目标的检测精度与实时性。Fan X [21]等人在 YOLOv5s 的基础上，将 Bottleneck Transformer 结构和 CoordConv 添加到网络中，极大改善检测精度。Z. Wei [22]等人在 YOLOv5 网络结构中嵌入注意力机制再增加多尺度检测策，将红外微小目标检测精度进一步提升。Cai Renhao [23]等人提出的 JH-YOLOv5-RDAB 算法，运用密集卷积层来充分提取局部特征，引入注意力机制获取更有效的通道与像素信息，获得更好的检测效果。Dang L [24]等人在 YOLOv7 的基础上引入 C3-Faster 模块增强特征提取，结合 NWD-GIoU 损失优化小目标回归，集成 CA 注意力机制进一步提升了检测精度。YUE Mingkai [25]等人在 YOLOv8 的基础上新增小目标检测层，引入 SA (Shuffle Attention)

注意力机制，采用 WIoU_v3 动态损失函数将精度提升 4.83%。

Table 2. Performance comparison of different algorithm detection frameworks on infrared data sets

表 2. 不同算法检测框架在红外数据集上的性能比较

Algorithm	Data	Backbone	Precision (%)
YOLO	PascalVOC	GoogleNet	63.4%
YOLOv2	trainval135k	Darknet-19	77.8%
YOLOv3	PascalVOC	Darknet-53	33.0%
YOLOv5	PascalVOC	FocusCSPDarknet-53	49.0%
YOLOv5s	Infrared dim-small airplane target dataset	YOLOv5 improvement	93.7%
YOLO-FR	Infrared dim-small airplane target dataset	YOLOv5 improvement	97.0%
JH-YOLOv5-RDAB	5023 images of infrared vehicle targets	YOLOv5 improvemen	95.1%。
YOLO-FNC	DIOR	YOLOv7	84.4%
YOLOv8_SG	VisDrone2019	YOLOv8	89.97%

资料来源：表中所有数据来源于公开数据集 PascalVOC 以及文献[17] [21] [23] [24]；YOLOv8_SG 数据来自 Vis-Drone2019 数据集。

如表 2 所示，分别列出了文中所介绍的检测算法基于固定骨干网络在公开数据集检测精度对比，从表中数据可以看出，YOLO 算法及其改进算法在红外检测领域检测精度稳步提升。

3.2. 基于 SSD 系列算法及其改进算法检测方法

2016 年 Liu W 等人提出了性能比 YOLO 更好得算法检测算法 SSD 算法[26]，算法从不同尺度进行目标特征提取并结合数据增强，有效提升对小目标的检测精度，SSD 在精度和速度方面都比当时最先进的对象检测器更优。实时 SSD300 模型以 59 帧/秒运行，比当时的实时 YOLO 替代方案更快。

在 SSD 模型算法基础上 Fu 等人提出 DSSD 算法，算法中用 ResNet 替换模型[27]，加入了反卷积层将图像进一步提升检测精度。Singh 等人[28]从多尺度的训练方向切入，有效利用训练数据，提升检测效果。Lin 等人[29]利用特征金字塔网络结构(FPN)融合模型中高低层语义信息，并给出焦点损失函数，提升对小目标的特征提取能力。

Xin Xu 等人[30]改进了 Single Shot MultiBox 的网络架构(SSD-ST)，增强高分辨率层，并设计了自适应流水线滤波器(APF)进一步提升检测精度。Z. Shen 等人[31]按照单次检测(SSD)框架开发 DSOD，从头开始学习对象检测器的框架，通过密集的逐层连接实现的深度监督在学习提升检测精度。

Table 3. Detection results on the MS COCO test-dev set

表 3. MS COCO 数据集的检测结果

Method	Data	Backbone	Avg.Precision		
			AP_S	AP_M	AP_L
YOLOv2	trainval135k	Darknet-19	22.1	44.8	53.2
SSD300	trainval135k	VGG16	5.6	25.9	41.2
SSD321	trainval135k	ResNet-101	8.7	30.1	44.5
DSSD321	trainval135k	ResNet-101	9.2	30.5	45.8
DSOD300	trainval	DS/64-192-48-1	6.9	28.3	42.1

资料来源：表中所有数据来源于公开数据集 trainval135k、trainval 以及文献[26]-[28] [31]。

针对弱小目标检测的挑战，SSD 系列算法通过多尺度检测框架与特征提取技术优化网络架构，显著提升了小目标检测能力。表 3 中详细列举了 YOLOv2、SSD300、SSD321、DSSD321 以及 DSOD300 在 MS COCO 数据集上的检测结果。从表中结果可知，SSD 的改进版本 DSSD321 与 DSOD300 通过引入残差结构(ResNet-101 Backbone)、反卷积层及多尺度特征融合策略，有效解决了传统 SSD 网络对小目标特征表达能力不足的问题。对比传统模型，YOLOv2 (DarkNet-19 Backbone)因单尺度预测限制，其小目标检测能力较弱，而 SSD 系列通过多层次特征图联合预测实现了更优的尺度适应性。

4. 基于 CNN 方法的红外微小目标检测方法

4.1. 多尺度特征与全局和局部区块注意力融合方法

多尺度特征与全局 - 局部注意力融合方法，针对红外小目标检测中目标尺度多变、像素占比低等挑战。多尺度特征融合，通过跨层级特征金字塔整合不同分辨率的上下文与细节信息，结合混合注意力机制，增强网络对微小目标的跨尺度感知能力；全局 - 局部块注意力融合，利用全局上下文建模定位潜在目标区域，通过局部块注意力强化目标边缘与纹理特征，实现细粒度特征增强。

K. Wang 等人[32]引入了一种特征注意机制增强有用的特征，实现了特征通道的自适应校准。基于多尺度特征融合的小目标检测算法 MSFYOLO 算法[33]提升检测精度。H. Fang 等人将检测结果转换为预测残差图像的问题，运用到 U-Net 模型中[34]，并结合上下文结构融合不同尺度的特征，提升检测效果。YOLOv5-AFAM [35]算法，结合多尺度、多特征提取机制更好地适应目标尺寸的变化和复杂背景条件，从而提高对微小目标的检测精度。

F. Chen 等人[36]结合全局和局部特征提出具有全局关注性的局部贴片网络 LPNet。AFFPN 算法[37]结合了注意力与特征金字塔、FTC-Net [38]引入全局注意力特征提取对小目标的检测性能进一步提升。Z.Wang 等人[39]结合强化学习及全局上下文边界注意，提出了 RLPGB-Net 算法，该算法在公开的 SAITD 和 SIRST 数据集上表现优异。多尺度特征融合、全局与局部块注意力融合方法结合，在红外微小目标检测中取得了一定的研究进展，为解决小目标检测问题提供了更加有效的思路 and 手段。

Table 4. Detection results of infrared datasets
表 4. 红外数据集检测结果

Model	Data	Method	MAP
MSFT-YOLO	NEU-DET	YOLOV5 + Transform	75.2%
AFFPN	SIRST/IDSAT	Attention Fusion Feature Pyramid Network	96.4%/98.1%
YOLOv5-AFAM	路面病害数据集	YOLOv5 + Multiscale Feature Fusion and Anchor Adaptive	94.7%
RLPGB-Net	SAITD/SIRST	Reinforcement Learning and Global Boundary Attention	超越现有 SOTA

资料来源：表中所有数据来源于公开数据集 SIRST、SAITD 以及文献[33] [35] [39]。

从表 4 给出的各算法对比可知，多尺度特征融合与注意力机制在复杂背景下表现最佳，但需权衡计算资源。与 YOLO 和 SSD 系列对比，传统模型适用于通用实时检测，改进型 YOLO 系列在特定任务中优势明显，而多尺度融合、注意力模型是复杂场景的最优解，但需牺牲部分实时性。

4.2. 基于全卷积网络及其改进方法

全卷积网络(Fully Convolutional Network, FCN)通过卷积层替代全连接层，保留空间信息并输出密集特征图。基于深度卷积神经网络的小目标检测方法[40]，通过全卷积网络对红外小目标进行增强和初步筛

选，分类网络对目标点后续筛选，网络结构中引入注意力机制对特征图进行选择，提升小目标检测精度。基于全卷积网络的空间和图匹配检测方法[41]、FCST [42]、DBGNet [43]算法结合时间特征，设计高精度小目标检测的深度全卷积回归网络，运用多尺度全卷积网络来提取不同层次的特征，提高了对红外小目标的检测率。

4.3. 密集嵌套的交互式特征融合与轻量化设计方法

Li 等人[44]设计了 The Dense Nested Attention Network (DNANet)，分别采用 SPSCNet 网络和图像增强网，通过重复的融合和增强，将弱小目标的上下文信息结合起来并得到充分的利用，提升检测率。其中 APANet [45]构建了一个整合双注意和紧密连接的金字塔结构，细化高级网络的细节特征，突出小目标特征，抑制背景干扰，使目标检测效率进一步提升。

轻量化设计方法旨在减少网络模型的参数量和计算量，提高实时检测效率。Iandola, F. N 等人提出的 SqueezeNet [46]是公认的轻量级模型设计最早期的工作之一。网络剪枝与深度可分离卷积是轻量化设计常用方法，通过减少网络的参数量来提高模型的计算效率同时保持检测性能。Zang, Y 等人 LMANet [47]将网络扩展到两个阶段获取上下文信息，使网络轻量化。RDLCM [48]提出一个基于轻量级卷积神经网络，提升小目标的检测率。

Table 5. Comparison table of infrared small target detection algorithms

表 5. 红外小目标检测算法对比表

Method	Data	Backbone	MAP
DNANet	NUDT-SIRST、SIRST	U-Net	Pd = 0.98
APANet	自建红外小目标数据集	Asymmetric pyramid aggregation network	IoU = 87.6%
SqueezeNet	海面红外数据集	SqueezeNet 轻量化	虚警率降低 40%
RDLCM	海上近海天线红外序列数据集	区域划分局部对比度测量	检测率提升 25%

资料来源：表中所有数据来源于公开数据集 NUDT-SIRST、SIRST 以及文献[45] [46] [48]。

表 5 中模型都为轻量化设计，改进传统局部对比度方法，检测率提升 25%。DNANet 和 APANet 在复杂背景下表现优异，但需权衡计算资源。SqueezeNet 和 Full-Scale U-Net 通过结构优化实现高效检测，适合嵌入式部署。与 YOLO 与 SSD 系列算法对比，YOLO 系列在通用检测中综合优势显著且需结合注意力机制或专用架构。轻量化设计虽是边缘计算的关键，但需权衡精度与效率。

4.4. Transformer 与小样本学习红外小目标检测方法

Transformer 架构与小样本学习(Few-shot Learning, FSL)技术已成为红外小目标检测领域的研究热点。Transformer 凭借其全局注意力机制在长程依赖建模上的优势，能够有效捕捉红外小目标与复杂背景的细微差异；而小样本学习则通过少量标注数据实现模型快速适应，解决了红外数据稀缺的核心瓶颈。

Transformer 通过自注意力机制和位置编码实现了对红外图像全局上下文的建模。基于视觉 Transformer (ViT)的 LPformer 算法通过拉普拉斯金字塔多级 Transformer 结构[49]，从不同尺度提取高频边界信息，并与主干网络特征融合，补偿小目标丢失的纹理细节，提升在 SIRST 数据集上的检测精度。Yang H 等人[50]提出 PBT (Progressive Background-aware Transformer)模型采用非对称编码器 - 解码器架构，通过任务自适应的门控 Transformer 模块分别提取目标响应和背景上下文，在 NUDT-SIRST 数据集上虚警率降低 44.4%。

小样本学习通过元学习和迁移学习策略，将基类(base class)学到的知识快速迁移到新类检测任务中。YOLO 结合 a contrario 范式和自监督学习(SSL) [51]，通过统计背景分布建模，将小目标视为“异常事件”，从而抑制背景噪声；而 SSL 通过实例判别预训练 backbone，显著提升了模型在少样本条件下的泛化能力，在标注数据不足的情况下仍能保持高检测率与低误报率。Han G 等人提取的 FM-FSOD [52]，融合视觉基础模型如 DINOv2 和大语言模型如 Vicuna，通过多模态提示学习实现零样本到少样本的过渡。在 SIRST 数据集上，5-shot 设置下 AP50 达 78.6%。Zhang J 等人提出的 SuperYOLO [53]，通过超分辨率辅助分支和像素级多模态融合，参数减少 18 倍的同时保持高分辨率检测能力，适用于机载红外平台实时处理。在保持高性能实时性检测的同时模型更加轻量化。

Table 6. Performance comparison between Transformer and few-shot learning in infrared small target detection
表 6. Transformer 与小样本学习在红外小目标检测中的性能对比

方法	优势	局限性	适用场景
Transformer	全局特征融合、多模态兼容性强	计算资源需求高，训练数据量大	高精度需求、复杂背景场景
小样本学习	数据效率高，适应新类别快	依赖生成模型质量，泛化性受限于支持集多样性	数据稀缺、快速部署场景

表 6 全面总结了 Transformer 与小样本学习在红外小目标检测中的技术特性与适用场景。该表表明，Transformer 与小样本学习在红外小目标检测中呈现互补特性。前者通过全局建模提升精度，后者通过数据高效利用解决标注瓶颈，二者协同创新将是未来研究的重要方向。

5. 结论

本文系统构建了红外小目标检测的理论框架与深度学习方法体系，在单阶段检测范式优化与 CNN 特征融合机制上取得突破性进展。基于 YOLO 系列算法的改进策略，SE 注意力模块、密集跨层交互、动态损失函数等使检测精度提升，SSD 系列通过残差反卷积架构与多尺度训练突破小目标特征衰减瓶颈。总结多尺度全局 - 局部注意力融合、密集嵌套特征交互、Transformer 架构与小样本学习等 CNN 优化方法，结合轻量化设计，深度可分离卷积、NAS 架构探索实现模型参数量降低。

然而，当前红外小目标检测面临数据稀缺、目标特征弱、算法鲁棒性不足等瓶颈，尤其在实时性与泛化性方面难以适应复杂场景，制约了该领域发展，使得需要进一步去研究更好得方法。因此，对于红外小目标检测算法研究趋势应是：

1) 算法架构创新

物理辐射模型与 Transformer 的异构融合，构建可解释性网络架构。多尺度特征融合与注意力协同机制，通过跨层密集连接与注意力模块的协同优化，提升小目标特征表达能力。轻量化编解码设计，引入深度可分离卷积与神经架构搜索，减少参数量的同时提升检测精度，实现边缘设备的高效部署。

2) 数据驱动与物理模型融合

创建基于元学习的自监督增量训练体系，利用生成式量子对抗网络解决跨场景数据迁移难题。结合生成对抗网络(GAN)生成高保真红外小目标数据，利用自监督预训练的对比学习预训练多尺度特征提取器，缓解标注数据不足问题。

3) 多模态协同检测机制

探索红外、可见光、雷达多模态协同检测机制，借鉴 CNN-Transformer 混合架构，如 SOAR 模型，融合红外与可见光模态的几何、语义特征，提升在遮挡场景下检测精度。融合红外热辐射与可见光纹理特征，设计双流注意力网络提升复杂背景下的跨模态注意力。

参考文献

- [1] Ren, S., He, K., Girshick, R. and Sun, J. (2017) Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **39**, 1137-1149. <https://doi.org/10.1109/tpami.2016.2577031>
- [2] Liu, T., Yin, Q., Yang, J., Wang, Y. and An, W. (2023) Combining Deep Denoiser and Low-Rank Priors for Infrared Small Target Detection. *Pattern Recognition*, **135**, Article ID: 109184. <https://doi.org/10.1016/j.patcog.2022.109184>
- [3] Deng, H., Liu, J. and Chen, Z. (2010) Infrared Small Target Detection Based on Modified Local Entropy and EMD. *Chinese Optics Letters*, **8**, 24-28.
- [4] Chen, J., Zhu, Z., Hu, H., Qiu, L., Zheng, Z. and Dong, L. (2023) Multi-Scale Local Contrast Fusion Based on LOG in Infrared Small Target Detection. *Aerospace*, **10**, Article 449. <https://doi.org/10.3390/aerospace10050449>
- [5] Shao, X., Fan, H., Lu, G. and Xu, J. (2012) An Improved Infrared Dim and Small Target Detection Algorithm Based on the Contrast Mechanism of Human Visual System. *Infrared Physics & Technology*, **55**, 403-408. <https://doi.org/10.1016/j.infrared.2012.06.001>
- [6] Liu, C., Xie, F., Zhang, H., Jiang, Z. and Zheng, Y. (2023) Infrared Small Target Detection Based on Multi-Perception of Target Features. *Infrared Physics & Technology*, **135**, Article ID: 104927. <https://doi.org/10.1016/j.infrared.2023.104927>
- [7] Gao, R. and Cai, Z. (2023) Object Detection in Thermal Infrared Image Based on Improved YOLOX. *IEEE Geoscience and Remote Sensing Letters*, **20**, 1-4. <https://doi.org/10.1109/lgrs.2023.3312148>
- [8] Zhang, L., Lin, W., Shen, Z., Zhang, D., Xu, B., Wang, K., et al. (2023) Infrared Dim and Small Target Sequence Dataset Generation Method Based on Generative Adversarial Networks. *Electronics*, **12**, Article 3625. <https://doi.org/10.3390/electronics12173625>
- [9] He, Y., Li, M., Zhang, J. and An, Q. (2015) Small Infrared Target Detection Based on Low-Rank and Sparse Representation. *Infrared Physics & Technology*, **68**, 98-109. <https://doi.org/10.1016/j.infrared.2014.10.022>
- [10] Gao, C.Q., Zhang, T.Q. and Li, Q. (2012) Small Infrared Target Detection Using Sparse Ring Representation. *IEEE Aerospace and Electronic Systems Magazine*, **27**, 21-30. <https://doi.org/10.1109/maes.2012.6196254>
- [11] Redmon, J., Divvala, S., Girshick, R. and Farhadi, A. (2016) You Only Look Once: Unified, Real-Time Object Detection. 2016 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, 27-30 June 2016, 779-788. <https://doi.org/10.1109/cvpr.2016.91>
- [12] Huang, X., Wang, X., Lv, W., et al. (2021) PP-YOLOv2: A Practical Object Detector. arXiv: 2104.10419.
- [13] Lin, X.P. (2019) Object Detection System Based on Vehiclebrone Infrared Image. Beijing Jiaotong University.
- [14] Redmon, J. and Farhadi, A. (2018) Yolov3: An Incremental Improvement. arXiv: 1804.02767.
- [15] Mukai, L., Tao, Z. and Wennan, C. (2020) Research of Infrared Small Pedestrian Target Detection Based on YOLOv3. *Infrared Technology*, **42**, 176-181.
- [16] Xu, Q., Huang, H., Zhou, C. and Zhang, X. (2021) Research on Real-Time Infrared Image Fault Detection of Substation High-Voltage Lead Connectors Based on Improved YOLOv3 Network. *Electronics*, **10**, Article 544. <https://doi.org/10.3390/electronics10050544>
- [17] Liu, M., Wang, X., Zhou, A., Fu, X., Ma, Y. and Piao, C. (2020) UAV-YOLO: Small Object Detection on Unmanned Aerial Vehicle Perspective. *Sensors*, **20**, Article 2238. <https://doi.org/10.3390/s20082238>
- [18] Jocher, G., Stoken, A., Borovec, J., et al. (2021) ultralytics/yolov5: v4. 0-nn. SiLU () Activations, Weights & Biases logging, PyTorch Hub Integration. Zenodo.
- [19] Li, S., Li, Y., Li, Y., Li, M. and Xu, X. (2021) YOLO-FIRI: Improved YOLOv5 for Infrared Image Object Detection. *IEEE Access*, **9**, 141861-141875. <https://doi.org/10.1109/access.2021.3120870>
- [20] Mou, X., Lei, S. and Zhou, X. (2023) YOLO-FR: A YOLOv5 Infrared Small Target Detection Algorithm Based on Feature Reassembly Sampling Method. *Sensors*, **23**, Article 2710. <https://doi.org/10.3390/s23052710>
- [21] Fan, X., Ding, W., Qin, W., Xiao, D., Min, L. and Yuan, H. (2023) Fusing Self-Attention and CoordConv to Improve the YOLOv5s Algorithm for Infrared Weak Target Detection. *Sensors*, **23**, Article 6755. <https://doi.org/10.3390/s23156755>
- [22] Wei, Z., Zhao, J., Chen, X., Wang, A., Li, F. and Gu, Y. (2023) Infrared Target Detection Based on the Fusion of Attention Mechanism and YOLOv5. 2023 *IEEE 2nd International Conference on Electrical Engineering, Big Data and Algorithms (EEBDA)*, Changchun, 24-26 February 2023, 933-936. <https://doi.org/10.1109/eebda56825.2023.10090650>
- [23] 蔡仁昊, 程宁, 彭志勇, 等. 基于深度学习的轻量化红外弱小车辆目标检测算法研究[J]. 红外与激光工程, 2022, 51(12): 357-367.

- [24] Dang, L., Liu, G., Hou, Y., *et al.* (2024) YOLO-FNC: An Improved Method for Small Object Detection in Remote Sensing Images Based on YOLOv7. *IAENG International Journal of Computer Science*, **51**, 1281-1290.
- [25] Mingkai, Y.U.E., Kangnan, Q., Cong, Z., *et al.* (2024) Research on Infrared Small Target Detection Algorithm Based on Improved YOLOv8. *Infrared Technology*, **46**, 1286-1292.
- [26] Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C., *et al.* (2016) SSD: Single Shot MultiBox Detector. In: Leibe, B., Matas, J., Sebe, N. and Welling, M., Eds., *Computer Vision—ECCV 2016*, Springer, 21-37. https://doi.org/10.1007/978-3-319-46448-0_2
- [27] Fu, C.Y., Liu, W., Ranga, A., *et al.* (2017) DSSD: Deconvolutional Single Shot Detector. arXiv: 1701.06659.
- [28] Singh, B., Najibi, M. and Davis, L.S. (2018) Sniper: Efficient Multi-Scale Training. arXiv: 1805.09300.
- [29] Lin, T., Dollar, P., Girshick, R., He, K., Hariharan, B. and Belongie, S. (2017) Feature Pyramid Networks for Object Detection. 2017 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, 21-26 July 2017, 936-944. <https://doi.org/10.1109/cvpr.2017.106>
- [30] Ding, L., Xu, X., Cao, Y., Zhai, G., Yang, F. and Qian, L. (2021) Detection and Tracking of Infrared Small Target by Jointly Using SSD and Pipeline Filter. *Digital Signal Processing*, **110**, Article ID: 102949. <https://doi.org/10.1016/j.dsp.2020.102949>
- [31] Shen, Z., Liu, Z., Li, J., Jiang, Y., Chen, Y. and Xue, X. (2017) DSOD: Learning Deeply Supervised Object Detectors from Scratch. 2017 *IEEE International Conference on Computer Vision (ICCV)*, Venice, 22-29 October 2017, 1937-1945. <https://doi.org/10.1109/iccv.2017.212>
- [32] Wang, K., Li, S., Niu, S. and Zhang, K. (2019) Detection of Infrared Small Targets Using Feature Fusion Convolutional Network. *IEEE Access*, **7**, 146081-146092. <https://doi.org/10.1109/access.2019.2944661>
- [33] Song, Z., Zhang, Y., Liu, Y., Yang, K. and Sun, M. (2022) MSFYOLO: Feature Fusion-Based Detection for Small Objects. *IEEE Latin America Transactions*, **20**, 823-830. <https://doi.org/10.1109/tla.2022.9693567>
- [34] Fang, H., Xia, M., Zhou, G., Chang, Y. and Yan, L. (2022) Infrared Small UAV Target Detection Based on Residual Image Prediction via Global and Local Dilated Residual Networks. *IEEE Geoscience and Remote Sensing Letters*, **19**, 1-5. <https://doi.org/10.1109/lgrs.2021.3085495>
- [35] Wang, Y., Zhao, L., Ma, Y., Shi, Y. and Tian, J. (2023) Multiscale YOLOv5-AFAM-Based Infrared Dim-Small-Target Detection. *Applied Sciences*, **13**, Article 7779. <https://doi.org/10.3390/app13137779>
- [36] Chen, F., Gao, C., Liu, F., Zhao, Y., Zhou, Y., Meng, D., *et al.* (2022) Local Patch Network with Global Attention for Infrared Small Target Detection. *IEEE Transactions on Aerospace and Electronic Systems*, **58**, 3979-3991. <https://doi.org/10.1109/taes.2022.3159308>
- [37] Zuo, Z., Tong, X., Wei, J., Su, S., Wu, P., Guo, R., *et al.* (2022) AFFPN: Attention Fusion Feature Pyramid Network for Small Infrared Target Detection. *Remote Sensing*, **14**, Article 3412. <https://doi.org/10.3390/rs14143412>
- [38] Qi, M., Liu, L., Zhuang, S., Liu, Y., Li, K., Yang, Y., *et al.* (2022) FTC-Net: Fusion of Transformer and CNN Features for Infrared Small Target Detection. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, **15**, 8613-8623. <https://doi.org/10.1109/jstars.2022.3210707>
- [39] Wang, Z., Zang, T., Fu, Z., Yang, H. and Du, W. (2023) RLPGB-Net: Reinforcement Learning of Feature Fusion and Global Context Boundary Attention for Infrared Dim Small Target Detection. *IEEE Transactions on Geoscience and Remote Sensing*, **61**, 1-15. <https://doi.org/10.1109/tgrs.2023.3304755>
- [40] Wu, S.C. and Zuo, Z.R. (2019) Small Target Detection in Infrared Images Using Deep Convolutional Neural Networks. *Journal of Infrared, Millimeter, and Terahertz Waves*, **38**, 371-380.
- [41] Wang, H., Li, H., Zhou, H. and Chen, X. (2021) Low-altitude Infrared Small Target Detection Based on Fully Convolutional Regression Network and Graph Matching. *Infrared Physics & Technology*, **115**, Article ID: 103738. <https://doi.org/10.1016/j.infrared.2021.103738>
- [42] Zhang, L., Han, P., Xi, J. and Zuo, Z. (2023) Infrared Small Target Detection Based on a Temporally-Aware Fully Convolutional Neural Network. *Remote Sensing*, **15**, Article 4198. <https://doi.org/10.3390/rs15174198>
- [43] Chi, W., Liu, J., Wang, X., Feng, R. and Cui, J. (2023) DBGNet: Dual-Branch Gate-Aware Network for Infrared Small Target Detection. *IEEE Transactions on Geoscience and Remote Sensing*, **61**, 1-14. <https://doi.org/10.1109/tgrs.2023.3309952>
- [44] Li, B., Xiao, C., Wang, L., Wang, Y., Lin, Z., Li, M., *et al.* (2023) Dense Nested Attention Network for Infrared Small Target Detection. *IEEE Transactions on Image Processing*, **32**, 1745-1758. <https://doi.org/10.1109/tip.2022.3199107>
- [45] Lv, G., Dong, L., Liang, J. and Xu, W. (2022) Novel Asymmetric Pyramid Aggregation Network for Infrared Dim and Small Target Detection. *Remote Sensing*, **14**, Article 5643. <https://doi.org/10.3390/rs14225643>
- [46] Iandola, F.N., Han, S., Moskewicz, M.W., *et al.* (2016) SqueezeNet: AlexNet-Level Accuracy with 50x Fewer Parameters and < 0.5 MB Model Size. arXiv: 1602.07360.

-
- [47] Zang, Y., Fan, C., Zheng, Z. and Yang, D. (2021) Pose Estimation at Night in Infrared Images Using a Lightweight Multi-Stage Attention Network. *Signal, Image and Video Processing*, **15**, 1757-1765.
<https://doi.org/10.1007/s11760-021-01916-3>
- [48] Fu, J., Li, F., Zhao, J., Tong, J. and Zhang, H. (2023) Infrared Small Dim Target Detection under Maritime near Sea-Sky Line Based on Regional-Division Local Contrast Measure. *IEEE Geoscience and Remote Sensing Letters*, **20**, 1-5.
<https://doi.org/10.1109/lgrs.2023.3316272>
- [49] 杜妮妮, 单凯东, 卫莎莎. LPformer: 基于拉普拉斯金字塔多级 Transformer 的红外小目标检测[J]. 红外技术, 2023, 45(6): 630-638.
- [50] Yang, H., Mu, T., Dong, Z., Zhang, Z., Wang, B., Ke, W., *et al.* (2024) PBT: Progressive Background-Aware Transformer for Infrared Small Target Detection. *IEEE Transactions on Geoscience and Remote Sensing*, **62**, 1-13.
<https://doi.org/10.1109/tgrs.2024.3415080>
- [51] Ciocarlan, A., Hégarat-Masclé, S.L., Lefebvre, S., *et al.* (2024) Robust Infrared Small Target Detection Using Self-Supervised and a Contrario Paradigms. arXiv: 2410.07437.
- [52] Han, G. and Lim, S. (2024) Few-Shot Object Detection with Foundation Models. 2024 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle, 16-22 June 2024, 28608-28618.
<https://doi.org/10.1109/cvpr52733.2024.02703>
- [53] Zhang, J., Lei, J., Xie, W., Fang, Z., Li, Y. and Du, Q. (2023) SuperYOLO: Super Resolution Assisted Object Detection in Multimodal Remote Sensing Imagery. *IEEE Transactions on Geoscience and Remote Sensing*, **61**, 1-15.
<https://doi.org/10.1109/tgrs.2023.3258666>