

基于改进卷积神经网络的红外图像 弱小目标检测方法研究

唐亮[†], 王海云, 陈晓范

(西安思源学院电子信息工程学院, 陕西 西安 710038)

摘要:针对现有红外小目标检测精度低、虚警率高的现实问题,提出了一种基于改进 U-Net 的红外小目标检测方法。利用级联通道和空间注意模块捕获红外目标特征,实现渐进式特征交互和自适应特征增强,并将其应用特征金字塔融合模块提取不同尺度的特征信息。通过重复特征融合和增强,从而有效融合和充分利用小目标的上下文信息,提高模型处理和提取深层特征的能力。实验结果表明,与 Transformer 检测方法相比,所提方法 IOU、POD 分别提升 1.16%、1.40%, FAR 降低 19.16%。实验结果验证了所提方法的有效性和实用性,该方法具有广阔的应用前景。

关键词:红外图像;小目标检测;卷积神经网络;特征提取;注意力;特征金字塔

中图分类号:TP393

文献标识码:A

Research on Infrared Image Weak Object Detection Method Based on Improved Convolutional Neural Network

TANG Liang[†], WANG Haiyun, CHEN Xiaofan

(School of Electronic Information Engineering, Xi'an Siyuan University, Xi'an, Shaanxi 710038, China)

Abstract: To address the issues of low accuracy and high false alarm rates in current infrared small target detection, a novel method based on an enhanced U-Net architecture is proposed. The method utilizes cascaded channels and spatial attention modules to capture infrared target features, enabling progressive feature interaction and adaptive feature enhancement. Additionally, it employs feature pyramid fusion modules to extract feature information across different scales. Through repeated feature fusion and enhancement, the contextual information of small targets is effectively integrated and fully leveraged, thereby enhancing the model's capability to process and extract deep features. Experimental results demonstrate that, compared to the Transformer detection method, the proposed approach improves IOU and POD by 1.16% and 1.40% respectively, while reducing FAR by 19.16%. These findings confirm the effectiveness and practicality of the proposed method, which suggests its potential for broad application.

Key words: infrared images; small object detection; convolutional neural networks; feature extraction; attention; feature pyramid

红外小目标检测在海上监视、预警探测、红外搜索跟踪及医学成像等领域^[1-3]展现出广泛应用价值,然而其检测面临诸多挑战。由于红外目标距

离传感器较远,图像中目标尺寸微小(仅数个至数十个像素),且易受特征清晰度、纹理干扰及低信号杂波比(Signal to Clutter Ratio, SCR)影响,导致

目标与背景难以有效分离,尤其是在复杂噪声和杂波背景下^[4,5]。

针对上述问题,学术界已探索多种解决方案,包括基于聚类的检测^[6]、双层局部能量因子方法^[7]及加权增强局部对比算法^[8]等。然而,这些方法高度依赖手工特征质量,场景适应性差,限制了其广泛应用。随着深度学习技术的兴起,卷积神经网络(Convolutional Neural Network, CNN)因其强大的拟合能力和鲁棒性,在图像处理等领域取得显著进展^[9]。尽管已有研究尝试将自适应注意力机制^[10]、YOLO 系列^[11,12]及 Transformer^[13]等 CNN 方法应用于红外小目标检测,但直接应用往往因目标尺寸过小而导致深层信息丢失,影响检测精度和虚警率。

受 U-Net 架构启发,本文提出了一种优化的 U-Net 结构,旨在提升红外小目标检测性能。该结

构包含特征提取、特征金字塔融合模块(Feature Pyramid Fusion Module, FPFM)、领域聚类模块(Domain Clustering Module, DCM)等架构,融合了分层梯度卷积核(Hierarchical Gradient Convolutional Kernel, HGCK)、级联通道和空间注意模块(Channels and Spatial Attention Modules, CSAM)、密集嵌套交互模块(Densely Nested Interaction Modules, DIM)等改进策略。这些改进有效增强了模型对红外小目标的检测能力,提升了检测精度和鲁棒性。

1 红外小目标检测架构

所提红外小目标检测架构如图 1 所示。该框架主要分为 3 个阶段:特征提取、FPFM、DCM。

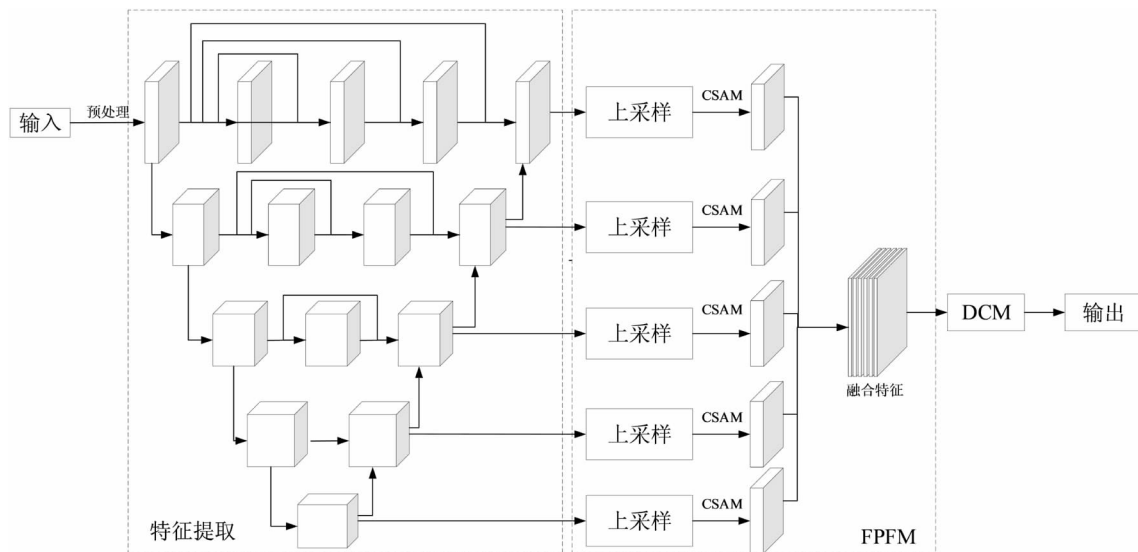


图 1 红外小目标检测架构

1.1 特征提取

特征提取网络如图 2 所示。参考 U-Net 结构,所提特征提取结构为由一个编码器、一个解码器和跳过连接层组成的改进 U-Net 结构。编码器用于扩大接收野并提取高级信息。解码器帮助恢复特征图的大小,最终达到与输入图像相同的大小,并实现渐进的多尺度特征融合。跳过连接层作为桥梁,将这些低级和高级特性从编码器传输到解码器。与传统 U-Net 不同,所提特征提取网络主要包含以下改进措施:HGCK、DIM、CSAM。

1.1.1 分层梯度核

根据红外图像的局部强度特性,目标在红外图像中的亮度值大于其周围像素的亮度值。因此红

外小目标的强度通常比周围背景更显著,故可基于高斯函数建模将目标集中在中心。红外小目标区域划分结构如图 3 所示。该结构包含 3 层,其中目标层(T)由目标在其中心的最大亮度强度值组成;背景层(B)包括邻近目标的杂波;保护层(G)包括目标层和背景层之间的过渡范围,其大小由红外图像的分辨率和衍射效应决定。

图像的梯度可以通过下式给出的方向导数计算:

$$\nabla f(x, y) = \left[\frac{\partial f(x, y)}{\partial x}, \frac{\partial f(x, y)}{\partial y} \right]$$

其中 $\nabla f(x, y)$ 为图像的梯度; $\partial f(x, y)/\partial x$ 和 $\partial f(x, y)/\partial y$ 表示对 x 轴和 y 轴的导数。

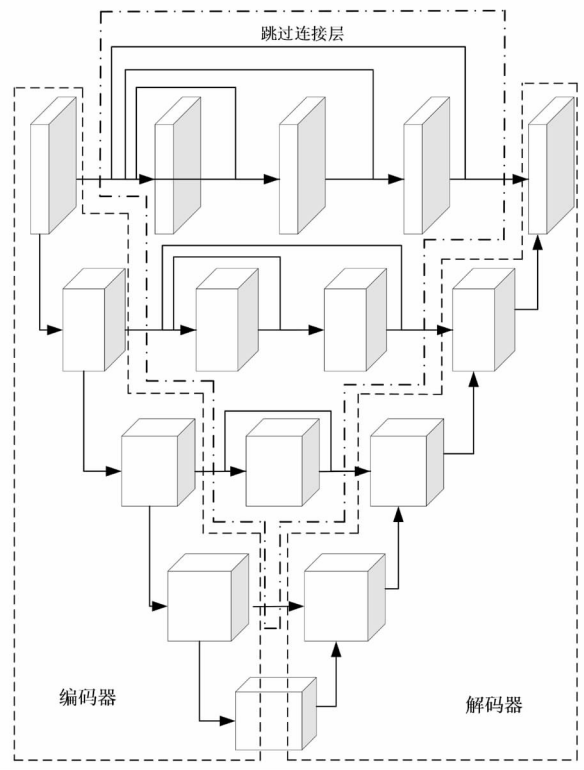


图 2 特征提取网络结构

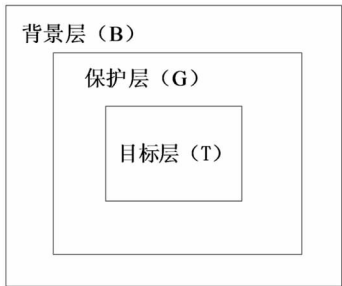


图 3 红外小目标区域划分结构

接着,图像的二阶导数计算如下:

$$\nabla^2 f(x,y) = \frac{\partial^2 f(x,y)}{\partial x^2} + \frac{\partial^2 f(x,y)}{\partial y^2}$$

其中 $\nabla^2 f(x,y)$ 为图像的二阶导数。

一般情况下,图像中边缘的位置为 $\partial f(x,y)/\partial x$ 和 $\partial f(x,y)/\partial y$ 的极值位置,即 $\nabla^2 f(x,y)$ 中强度为零的阈值点。考虑到图像局部特性,计算二阶导数的零点比找到一阶导数的极值更容易。为此,可通过图像之间的卷积近似计算二阶导数。

图像中的卷积可计算如下:

$$g(x,y) = \sum_{u=-a}^a \sum_{v=-b}^b f(x-u,y-v)K(u,v)$$

其中 $g(x,y)$ 为卷积后的图像; $f(x,y)$ 为原始图像; $K(u,v)$ 为卷积核,其尺寸为 (u,v) ,且

$$-a \leq u \leq a, -b \leq v \leq b.$$

一般来说,梯度图的值越大,表示局部梯度的和越大,因此该位置为目标的可能性越大。为此,本研究使用分层梯度核与输入的红外图像进行卷积,得到其梯度映射:

$$I_{GM}(x,y) = \sum_{u=-a}^a \sum_{v=-b}^b I(x-u,y-v)K_G(u,v)$$

其中 (x,y) 为原始红外图像的尺寸; (u,v) 为分层梯度核的尺寸; I 为原始图像; K_G 为分层梯度核,且核的每一个元素都为 $-a \leq u \leq a, -b \leq v \leq b$ 。

1.1.2 密集嵌套交互模块

令 $L_{i,j}$ 表示密集嵌套交互模块中编码器下采样层 i 和卷积层 j 的输出特征。当 $j = 0$ 时,每个节点只接收来自密集嵌套交互模块连接层的特征。因此,由 $L_{i,j}$ 表示的特征映射堆栈计算如下:

$$L_{i,j} = P_{\max}[F(L_{i-1,j})]$$

其中 $F(\cdot)$ 为同一卷积块的多个级联卷积层; $L_{i-1,j}$ 为密集嵌套交互模块中编码器下采样层 $i-1$ 和卷积层 j 的输出特征; $P_{\max}(\cdot)$ 为最大池化操作。当 $j > 0$ 时,每个节点接收到密集跳跃连接、嵌套双向交互跳跃连接等方向的输出组成,则生成的特征映射可更新为:

$$L_{i,j} = \{F[(L_{i,k})], P_{\max}[F(L_{i+1,j-1})], U[F(L_{i-1,j})]\}$$

其中 $U(\cdot)$ 为上采样层。

1.1.3 通道和空间注意模块

本研究将通道和空间注意模块放在密集嵌套交互模块的每一层特征融合后,以自适应提升特征提取能力。

通道和空间注意模块结构如图 4 所示。该模块由两个级联的通道注意单元和空间注意单元组成。通道注意单元执行过程可以总结如下:

$$\begin{cases} L_c = M_c(L) \otimes L \\ M_c(L) = \sigma[M_{LP}(P_{\max}(L)) + M_{LP}(P_{\text{avg}}(L))] \end{cases}$$

其中 L_c 为经通道注意单元后的通道输出特征; M_c 为经通道注意单元执行后的通道注意特征图,且 $M_c \in \mathbb{R}^{C_i \times H_i \times W_i}$; C_i, W_i, H_i 分别为输入特征 $L_{i,j}$ 的通道数、高度和宽度; $\otimes$ 为元素乘法操作符; σ 为 sigmoid 函数; $P_{\text{avg}}(\cdot)$ 为平均池化操作; $M_{LP}(\cdot)$ 为共享网络,由一个多层感知器(Multi-layer Perceptron, MLP)和一个隐藏层组成。需注意,在执行元素乘法操作之前, $M_c(L)$ 的尺寸将被拉伸为 $C_i \times H_i \times W_i$ 。

与通道注意单元执行过程类似,空间注意单元

执行过程可以总结如下:

$$\begin{cases} L_s = M_s(L) \otimes L \\ M_s(L) = \sigma[f'(P_{\max}(L)), P_{\text{avg}}(L)] \end{cases}$$

其中 L_s 为经空间注意单元后的空间输出特征; M_s 为经空间注意单元执行后的空间注意特征

图, $M_s \in \mathbb{R}^{C_i \times H_i \times W_i}$; C_i 、 H_i 和 W_i 分别为输入特征 $L_{i,j}$ 的通道数、高度和宽度; f' 为分层梯度核卷积运算,且卷积核的尺寸为 7×7 。需注意,空间注意特征图 $M_s(L)$ 的尺寸在元素乘法之前将被拉伸到 $C_i \times H_i \times W_i$ 。

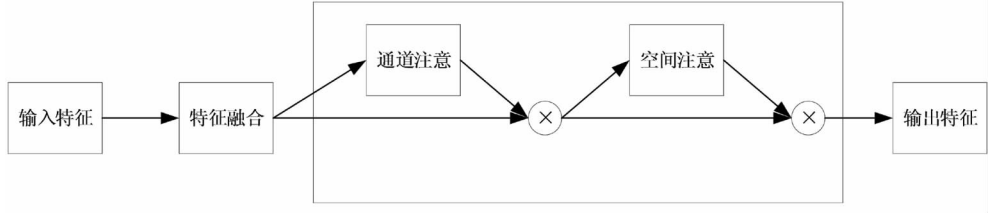


图 4 通道和空间注意模块结构

1.2 特征金字塔融合模块

在特征提取模块之后,本文提出了一种特征金字塔融合模块以聚合特征提取模块的多层特征。特征金字塔融合模块结构如图 5 所示。

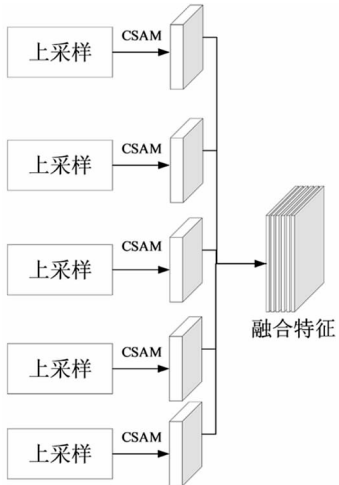


图 5 特征金字塔融合模块结构

首先将多层特征转化为相同尺寸。然后,将具有空间和通道信息的注意特征(CSAM)与具有丰富语义信息的深层特征进行连接,以生成全局鲁棒特征映射:

$$G = \{L_e^0, L_e^1, \dots, L_e^j, \dots, L_e^l\}$$

其中 G 为全局鲁棒特征图; L_e^j 为多层特征,且有 $L_e^j \in \mathbb{R}^{C_i \times H_0 \times W_0}$, C_i 、 H_0 和 W_0 为转化后特征图的通道数、高度和宽度。

1.3 邻域聚类模块

在特征金字塔融合模块之后,本文引入一个邻域聚类模块将属于同一目标的像素聚在一起,以计算每个目标的质心。如果全局鲁棒特征图 G 中任意两个像素 (m_0, n_0) 和 (m_1, n_1) 在邻域中有相交

区域,即:

$$N_{(m_0, n_0)} \cap N_{(m_1, n_1)} \neq \emptyset$$

其中 $N_{(m_0, n_0)}$ 和 $N_{(m_1, n_1)}$ 为像素 (m_0, n_0) 和 (m_1, n_1) 的邻域。上式表明判断像素 (m_0, n_0) 和 (m_1, n_1) 为相邻像素。

接着,如果这两个像素的值相同,即:

$$g_{(m_0, n_0)} = g_{(m_1, n_1)}, \forall g_{(m_0, n_0)}, g_{(m_1, n_1)} \in G$$

其中 $g_{(m_0, n_0)}$ 和 $g_{(m_1, n_1)}$ 为像素 (m_0, n_0) 和 (m_1, n_1) 的灰度值。上式表明像素 (m_0, n_0) 和 (m_1, n_1) 处于连通区域,且连通区域中的像素属于相同的目标。一旦确定了图像中的所有目标,就可以根据其坐标计算质心。

2 实验与分析

2.1 数据集

数据的质量、数量和场景多样性显著影响检测方法的性能。考虑到真实场景下红外小目标采集成本较高,因此本研究使用真实红外背景合成红外小目标。红外背景图像主要采集自无人机航拍图像,包括城市、田野等。数据集中每张图像都是在各种 SCR 和丰富姿态下,从真实背景合成的各种目标。

实验时,按 8:1:1 将数据集分为训练集、验证集和测试集。训练集和验证集用于实验检测网络调参,测试集用来验证不同检测方法的性能。

2.2 实验设置

实验时环境设置如下:硬件为 Ubuntu 18.04 操作系统,1 块 NVIDIA GeForce RTX 2080Ti GPU,AMD Ryzen 7 1700, 3.0 GHz, 32 GB

RAM CPU;软件环境为 pycharm2019 社区版编译,红外小目标检测网络由 Python 3.7、CUDA 11.3、Pytorch 1.12.0 搭建并执行。Python 中主要使用的模块如下: pandas 0.25.1, numpy 1.17.2, matplotlib 3.1.1。预测网络训练时超参数设置如下:初始学习率为 0.001,学习率衰减率为 0.001,衰减周期为 500,批量归一化数量为 128, dropout 概率设置为 0.25,迭代总次数为 20000,控制函数的 dropout 超参数设置为 0.3。

2.3 结果与分析

2.3.1 消融实验

本节对所提红外小目标检测方法的 HGCK、DNIM、CSAM、FPFM、NCM 等进行消融实验。消融实验结果如表 1 所示。

表 1 消融实验结果/%			
名称	IOU	POD	FAR
基础 U-Net	75.13	96.74	0.3484
HGCK	77.48	97.47	0.2353
DNIM	76.94	97.71	0.1203
CSAM	81.03	97.03	0.1269
FPFM	83.75	96.96	0.0656
NCM	85.89	97.29	0.0565
所提检测方法	88.10	98.75	0.0422

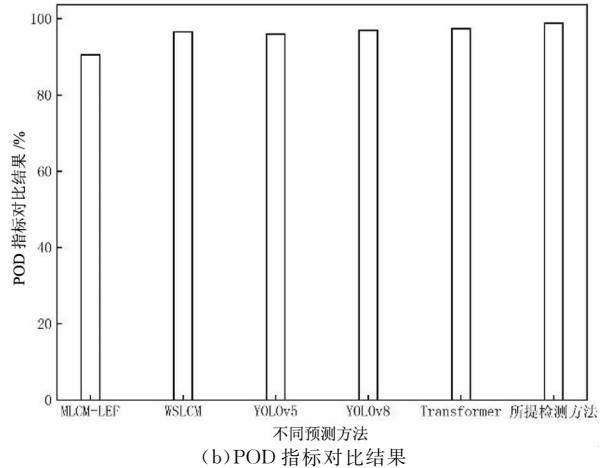
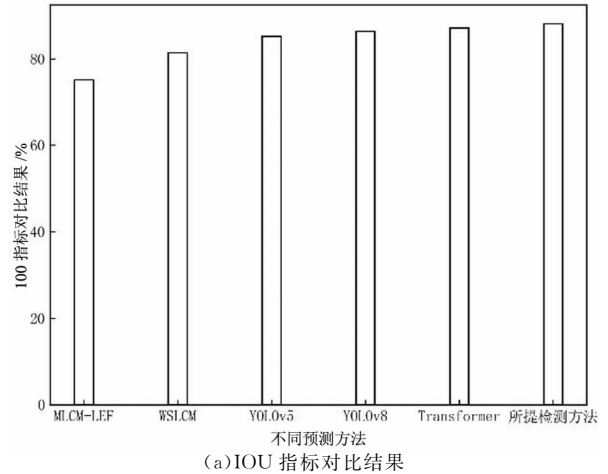
消融实验可以看出,与基础 U-Net 相比,采用 HGCK 后 IOU、POD 指标值分别提升了 3.13%、0.75%,FAR 指标降低 32.46%;采用 DNIM 后 IOU、POD 指标值分别提升了 2.41%、1%,FAR 指标降低 65.47%;采用 CSAM 后,IOU、POD 指标值分别提升了 7.85%、0.301%,FAR 指标降低 63.57%;采用 FPFM 后,IOU、POD 指标值分别提升了 11.47%、0.228%,FAR 指标降低 81.17%;采用 NCM 后,IOU、POD 指标值分别提升了 14.32%、0.57%,FAR 指标降低 83.78%;融合 HGCK、DNIM、CSAM、FPFM、NCM 等策略后,所提检测方法 IOU、POD 指标值分别提升了 17.26%、2.08%,FAR 指标降低 87.89%。

实验证明了所提 HGCK、DNIM、CSAM、FPFM、NCM 策略的有效性及合理性,通过改进策略有助于降低网络对各种干扰的敏感性,提升网络性能。

2.3.2 综合对比

本节将所提检测方法与几种最先进的基准方

法进行了比较,包括多尺度局部能量因子检测方法(MLCM-LEF)、改进加权增强局部对比算法(Weighted Strengthened Local Contrast Measure, WSLCM)的传统检测方法以及基于 YOLOv5、YOLOv8、多级回归 Transformer 等深度学习的红外小目标检测方法。此外,为保证实验公平性,所有基准方法均运行于同一硬件环境。综合指标对比图如图 6 所示。其中图 6(a)为 IOU 指标对比结果;图 6(b)为 POD 指标对比结果;图 6(c)为 FAR 指标对比结果。可以看出,所提出的检测方法 IOU、POD、FAR 指标值较优。此外,传统检测方法各指标略低于深度学习模型,表明传统方法在特征提取方面弱于深度学习方法,造成部分细节特征识别失败。与 Transformer 相比,所提方法 IOU、POD 分别提升了 1.16% 和 1.40%,FAR 降低 19.16%。综合以上分析,所提红外小目标检测方法可有效学习具备特征,具备较好的识别性能。所提方法可为红外小目标检测技术的发展提供一定借鉴作用。



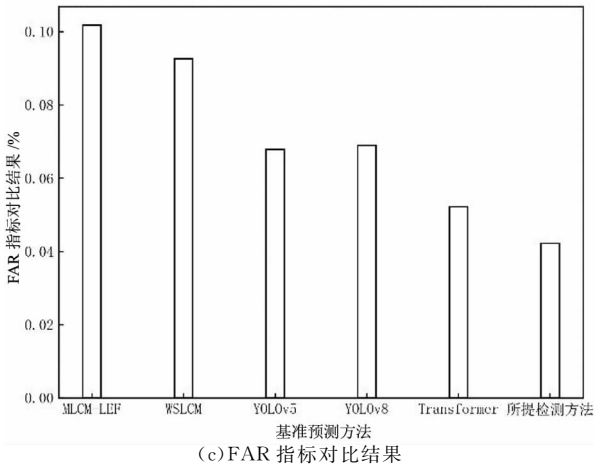


图 6 综合指标对比图

3 结 论

红外小目标检测因其对光照条件不敏感、能够穿透烟雾等特点,在复杂环境中具有独特优势。然而,由于红外图像中目标与背景对比度低、噪声干扰大,小目标检测面临巨大挑战。针对这些问题,本文提出了一种基于改进 U-Net 的红外小目标检测方法,以期提高检测性能。同时,设计了 HGCK、DNIM、CSAM、FPFM、NCM 等改进模块,可有效解决小目标在深层丢失的问题,降低小目标检测虚警率。

未来研究可以进一步探索更强大的特征提取和融合机制,以及引入更多的上下文信息来提高检测精度。此外,可针对模型压缩、轻量化等技术进行研究,将检测算法移植至嵌入式智能加速芯片,并搭载至无人机,从而推广所提红外小目标检测方法应用范围,提升检测方法工程化实用价值。

参考文献

[1] 王家豪,付雨昕,李怡,等. 复杂背景下空中小目标动态红外成像仿真技术研究[J]. 航空科学技术, 2024, 35(4): 112—118.

[2] 余祉祺. 红外图像的目标检测研究[J]. 数字通信世界, 2024(1): 34—37.

[3] 刘洁,云利军,龙晖,等. 一种改进的基于频域的红外图像细节增强方法[J]. 云南师范大学学报(自然科学版), 2020, 40(1): 52—56.

[4] ZHAO M, LI W, LI L, et al. Single-frame infrared small-target detection: a survey[J]. IEEE Geoscience and Remote Sensing Magazine, 2022, 10(2): 87—119.

[5] JAKOBSEN P, FERRUIT P, DEOLIVEIRA C A, et al. The near-infrared spectrograph (nirspec) on the James Webb Space Telescope- I . overview of the instrument and its capabilities[J]. Astronomy & Astrophysics, 2022, 661: A80.

[6] 饶俊民,穆靖,刘士建,等. 基于聚类思想的红外弱小目标检测[J]. 红外与毫米波学报, 2023, 42(4): 527—537.

[7] 唐凌霄,黄昶. 基于双层局部能量因子的红外小目标检测方法[J]. 华东师范大学学报(自然科学版), 2024(2): 97—107.

[8] 王磊,郭宏林,潘明然,等. 一种改进 WSLCM 的红外小目标检测方法研究[J]. 弹箭与制导学报, 2023, 43(6): 1—7.

[9] KRICHEN M. Convolutional neural networks: a survey[J]. Computers, 2023, 12(8): 151.

[10] 赵松璞,杨利萍,赵昕,等. 基于自适应注意力机制的红外与可见光图像目标检测[J]. 红外技术, 2024, 46(4): 443—451.

[11] 刘芷汐,周春桂,崔俊杰,等. 基于改进 YOLOv5s 轻量化模型的红外场景目标检测方法研究[J]. 兵器装备工程学报, 2024, 45(3): 323—330.

[12] 吴合凤,王国伟,万造君,等. 基于改进 YOLOv8s 的配电设备红外目标检测模型[J]. 电气技术, 2024, 25(3): 18—23.

[13] 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.