

# 基于 SCNN 的红外弱小目标检测算法研究

刘芬<sup>1</sup>,程勇策<sup>2</sup>,郑尧<sup>3</sup>,赵涛<sup>2</sup>

(1. 中煤科工集团重庆研究院有限公司,重庆 400039;2. 中国电子科技集团公司第三研究所,北京 100015;  
3. 陆航研究所,北京 101121)

**摘要:**提出了基于结构化特征卷积神经网络(Structural Feature convolution neural network, SCNN)的红外弱小目标检测算法。通过将红外弱小目标结构化特征引入 CNN 网络中,去除 CNN 网络的池化层、改变卷积扩展性、并加入分层融合机制,克服了 CNN 会损失小目标本身的信息和无法检测弱纹理小目标的问题。实验结果表明:本文提出的方法相比同类方法具有更高的检测率,并且对于不同场景具有较强的鲁棒性。

**关键词:**红外弱小目标检测;SCNN;结构化特征

**中图分类号:**TP391.41 **文献标识码:**A **DOI:**10.3969/j.issn.1001-5078.2021.06.021

## Research on the algorithm of infrared dim small target detection based on SCNN

LIU Fen<sup>1</sup>, CHENG Yong-ce<sup>2</sup>, ZHENG Yao<sup>3</sup>, ZHAO Tao<sup>2</sup>

(1. CCTEG Chongqing Research Institute, Chongqing 400039, China; 2. The Third Research Institute of China Electronics Technology Group Corporation, Beijing 100015, China; 3. Land and Aviation Research Institute, Beijing 101121, China)

**Abstract:** An infrared dim small target detection algorithm based on structural feature convolution neural network (SCNN) is proposed. By introducing the structural features of infrared dim small target into CNN, the pooling layer of CNN is removed, the convolution expansibility is changed, and the layered fusion mechanism is added to overcome the problems that CNN will lose the information of dim small target itself and cannot detect dim small target. The experimental results show that the proposed method has higher detection rate and stronger robustness for different scenes.

**Keywords:** infrared dim small target detection; SCNN; structural feature

### 1 引言

随着红外成像技术的逐渐成熟,基于红外目标的检测或跟踪方法已被广泛应用于诸多领域中。其中,红外小目标的检测尤其重要<sup>[1-6]</sup>。在过去几十年中,学者们提出了许多不同的检测算法,但这些方法仅仅依赖的是图像底层特征。当背景变得复杂时,使用传统方法得到的结果可能并不理想,如图 1

所示。从图中可得知,当红外图像模糊不清,存在大量噪声时,有些方法不能检测出小目标。传统方法存在以下不足:1. 依赖的结构化特征缺乏鲁棒性,导致在复杂场景下不能区分背景和小目标;2. 虽然已有一些方法采用了局部对比,但这些对比是在低层的视觉特征上进行的,缺乏高层特征的认知和理解。

**作者简介:**刘芬(1984-),女,硕士,高级工程师,主要研究方向为仪器仪表设计及监测监控。

**通讯作者:**程勇策(1982-),男,学士,高级工程师,主要研究方向为光电探测系统设计与应用。E-mail:chengyongce@163.com

**收稿日期:**2020-03-18;**修改日期:**2021-04-08

近年来,深度学习技术已被应用到许多计算机视觉中,包括图像分类,目标跟踪和目标检测<sup>[7-8]</sup>。其中,深度卷积神经网络能够提取高层特征,大大提高了目标检测的性能。通过一定样本的训练,CNN能够学习到丰富的语义特征,相比于传统底层特征,这些特征更能准确代表物体的本质。因此将 CNN 应用到红外小目标检测是可行的。但由于红外小目标不是十分明显,并且只有灰度信息,与自然图像的目标存在巨大差异,所以一些基于 CNN 的自然图像目标检测方法并不适用于红外小目标的检测,这要求我们寻求新的方法。

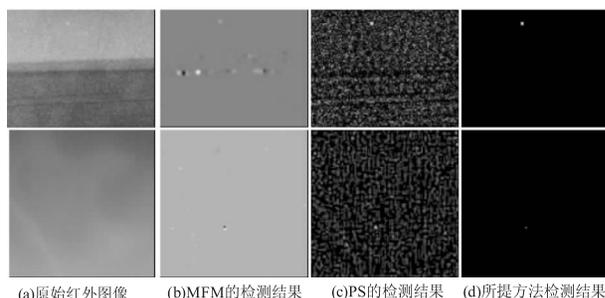


图1 不同方法的检测结果

Fig. 1 Detection results of different methods

## 2 算法概述

CNN 包括卷积层、池化层、全连接层、输出层四种网络结构,通过多层卷积和池化逐步扩大图像的局部特征,从而实现从底层到高层的认知过程。但是 CNN 由于池化层的存在,会损失目标本身的信息,所以 CNN 无法应用于小尺寸目标的识别问题。另外 CNN 网络体现的是机器认知过程,最终体现的特征也是目标高层纹理解析,并没有结合泛化良好的结构化特征,因此 CNN 网络不具备弱纹理目标的识别能力。

针对上述问题,需要研究基于结构化特征融合的 SCNN 算法,具体需要解决的关键问题包括:

(1) 针对目标尺度的变化,如何设计不同的卷积核,使得整个框架对弱小目标的尺度变化具有鲁棒性;

(2) 直接使用传统的池化操作会直接损失弱小目标的信息,如何在不使用池化操作的同时还可以扩大感受野而又不损失特征图分辨率;

(3) 泛化良好的结构化特征如何引入传统的 CNN 网络,如何与 CNN 自身获取的高层语义特征相融合而又不互相冲突。

将 SCNN 算法应用于红外弱小目标空间检测时,泛化良好的红外弱小目标结构化特征包括:

### (1) 局部对比度特征

复杂背景下的红外弱小目标周围  $9 \times 9$  区域的局部对比度如图 2 所示。

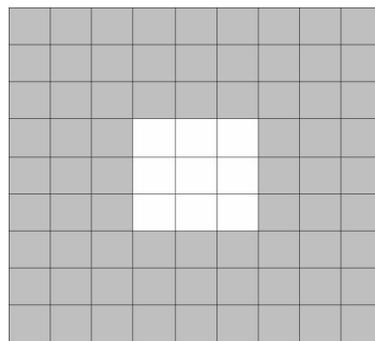


图2 红外弱小目标在局部对比度示意

Fig. 2 Local contrast of infrared small target

定义中心区域  $3 \times 3$  邻域的平均灰度为:

$$g_c = \frac{\sum_{i=1}^N p_i}{N} \quad (1)$$

则目标周围  $9 \times 9$  邻域内,目标局部对比度满足:  $g_c > g_i (i = 1, 2, \dots, 8)$ , 即在一定局部区域内目标平均灰度高于其邻域灰度。

### (2) 灰度空间分布的高斯特性

在实际工程应用中因目标或成像探测设备的姿态、位置的变化,红外成像图像中的点目标也很难有一个固定的形状。当成像距离足够远的时候,对于探测系统来说,它的灰度分布特性符合光学点扩散函数(PSF),即目标灰度分布符合高斯分布。

$$S(x, y) = \tau \cdot \exp\left\{-\frac{1}{2}\left(\frac{x^2}{\delta_x^2} + \frac{y^2}{\delta_y^2}\right)\right\} \quad (2)$$

在 SCNN 算法中引入上述结构化特征后,通过去除池化层、改变卷积扩展性、加入分层融合机制等算法,使其改进后的 SCNN 更加适合红外弱小目标的空间检测。

## 3 算法实现

本文提出一种双通道 CNN 的结构模型来进行弱小目标检测,通过将红外弱小目标的局部对比度等结构化空间特征在低层卷积阶段引入 CNN 网络,去除池化层,从而构建 SCNN 弱纹理小尺寸目标识别算法。

### 3.1 算法框架流程

模型的输入是一幅红外图像,输出是其对应的

检测结果图像,其中每个像素的大小表示该像素属于小目标的概率,值越大,越可能属于小目标。框架流程图如图 3 所示。每个通道的前部分所使用的卷积核大小是不同的,这样每个通道能够适应不同大小的小目标,从而使得整个框架能够对小目标的尺度变化具有鲁棒性。由于弱小目标的尺度很小,若使用池化操作,会损失关于小目标的信息,这样会导致后续的检测不准确。因此在中间部分,每个通道使用了不同的膨胀率进行扩张卷积,扩张卷积的作用是扩大感受野而又不会损失特征图的分辨率,从而能够有效地保留小目标的信息。通过进行不同膨胀率的扩张卷积操作,学习到了不同的语义特征,为了在不同层次上突出小目标,提出一种多层次局部对比(multi-level local contrast, MLLC)的模型,如图

4 中 MLLC 部分所示。之所以计算局部对比,是因为小目标与周边背景差异大,若所得局部对比越大,越可能是小目标。在得到不同层次的局部对比之后,需要把他们进行融合起来,本文提出一种分层次融合(hierarchical merge, HM)的方法,如图 4 中 HM 部分所示。分层次融合即是把两个通道中相同层次的局部对比进行两两融合。然后把两两融合后的多层次结果再次融合,从而得到最终的融合特征。框架的后部分首先整合了分层次融合的结果以及两个通道的浅层特征,这样做是为了保留更多的语义信息,然后经过多个卷积操作生成最后的检测图像。

图 3 中,卷积层参数表示为:conv - (卷积核大小) - (特征图数量) - (膨胀率);“C”表示特征图合并。

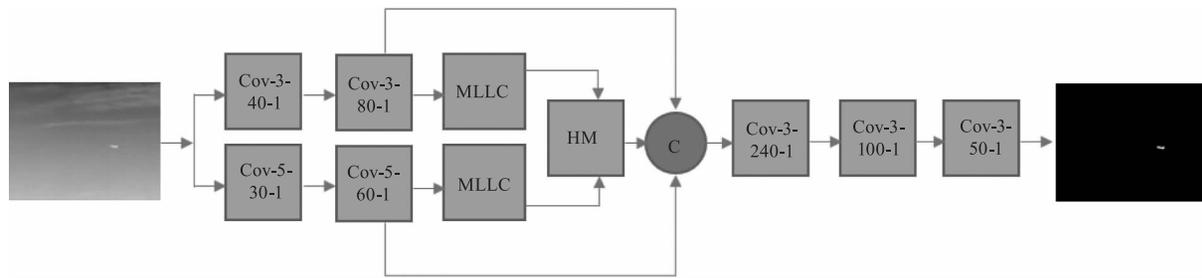


图 3 方法流程图

Fig. 3 Flow chart of the algorithm

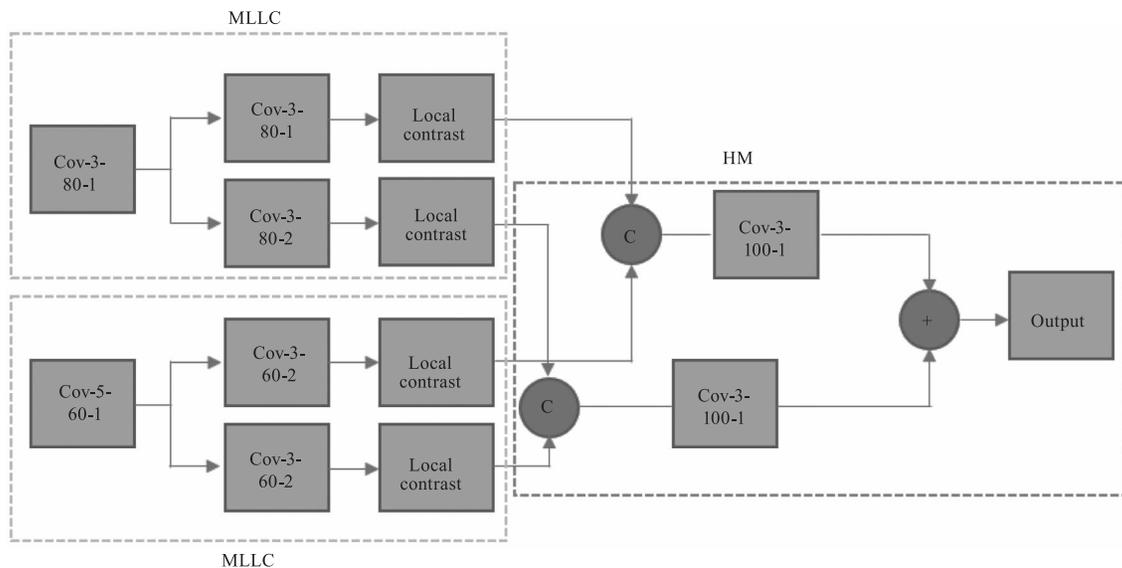


图 4 MLLC 和 HM 结构

Fig. 4 The structure of MLLC and HM

图 3 中,实际上每个通道计算 4 个层次的局部对比,即膨胀率取 1,2,3,4,为了演示方便,这里只取膨胀率为 1,2 作为示例,“+”表示特征图

相加。

### 3.2 计算操作

其中本算法框架所涉及的计算操作如下:

(1) 卷积操作可表示为:

$$F^l = W * F^{l-1} + b \quad (3)$$

其中,  $F^l$  表示第  $l$  层的特征图;  $F^{l-1}$  表示第  $(l-1)$  的特征图;  $W$  表示卷积核;  $b$  表示偏置项;  $*$  表示卷积操作。在卷积操作之后, 一般采用非线性激活函数, 而使用的比较多的是修正线性单元 (Rectified Linear Unit, ReLU)。

$$\text{ReLU}(x) = \max(0, x) \quad (4)$$

(2) 扩张卷积可表示为:

$$y(u, v) = b + \sum_{i=1}^M \sum_{j=1}^N x(u+r \times i, v+r \times j) w(i, j) \quad (5)$$

其中,  $y(u, v)$  表示通过扩张卷积后  $x(u, v)$  位置的对应输出;  $M, N$  分别表示卷积核的高和宽;  $r$  表示膨胀率。当  $r = 1$  时, 扩张卷积变回传统的卷积操作。图 5 展示了使用不同膨胀率时对应的卷积感受野。

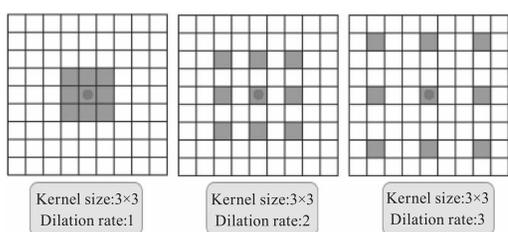


图 5 膨胀率为 1, 2, 3 的 3 × 3 扩张卷积  
(灰色块为有效的卷积位置)

Fig. 5 Expansion convolution with expansion rate of 1, 2, 3  
(the blue dot is the effective convolution position)

(3) 局部对比<sup>[9]</sup> (local contrast) 表示为:

$$F_c = F - \text{AvgPool}(F) \quad (6)$$

其中,  $\text{AvgPool}(\cdot)$  表示计算特征图  $3 \times 3$  范围的均值。

(4) 分层次融合表示为:

$$H^i = \sigma(W * \text{Concat}(F_c^{1,i}, F_c^{2,i}) + b) \quad (7)$$

其中,  $H^i$  表示第  $i$  层次的融合结果;  $\sigma(\cdot)$  表示激活函数;  $\text{Concat}(\cdot)$  表示特征图的合并;  $F_c^{1,i}, F_c^{2,i}$  分别表示第一和第二通道中第  $i$  层次的局部对比,  $i = 1, 2, 3, 4$ 。

(5) 两通道合并表示为:

$$M = \sigma(W * \text{Concat}(H, C^{1,2}, C^{2,2}) + b) \quad (8)$$

其中,  $H = \sum_{i=1}^4 H^i$  表示多层次局部对比的整合结果,  $C^{1,2}$  表示第一通道的第二个卷积操作,  $C^{2,2}$  同

理, 这是为了结合浅层的空间特征细节。

(6) 损失函数使用交叉熵损失, 表示为:

$$L = - \sum_{z_i \in Z} z_i \log \Pr(z_i = 1 | X; \theta) + (1 - z_i) \log \Pr(z_i = 0 | X; \theta) \quad (9)$$

其中,  $X = \{x_i, i = 1, \dots, |X|\}$  为图像训练集;  $Z = \{z_i, i = 1, \dots, |Z|\}$  为对应的真值图像;  $\Pr(z_i = 1 | X; \theta)$  由计算框架的输出预测结果的 sigmoid 激活函数而得;  $\theta$  表示整个框架所有权重参数的集合。

### 3.3 网络训练

对 300 幅图像 (真实图像以及合成图像) 进行数据增广, 共 2400 张图像进行训练, 训练直到损失函数收敛, 共迭代了 150 次。图 6 展示了部分原始图像及对应的人工标注图像示例。

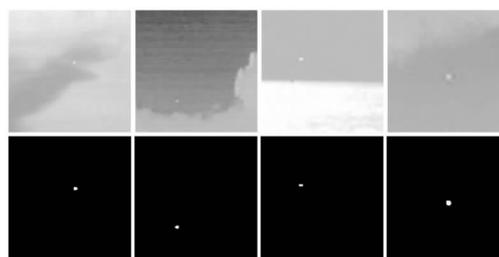


图 6 红外图像及其人工标注

Fig. 6 Infrared image and its manual annotation

## 4 实验结果及分析

图 7 展示了不同方法在测试集上的部分代表场景下的检测结果。第 1 列是原图像, 最后一列是本文方法的检测结果, 第 2 ~ 5 列是对比方法的检测结果。

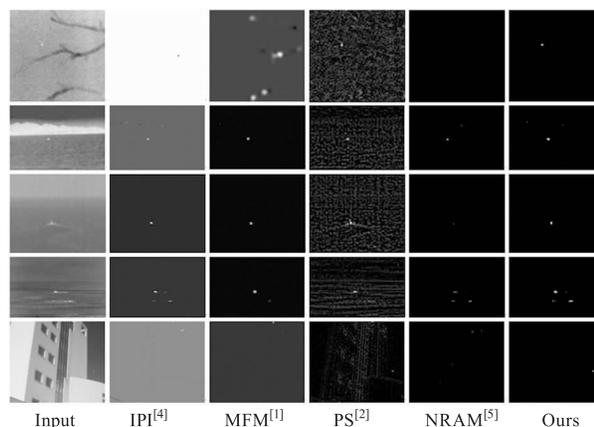


图 7 不同方法的检测结果示例

Fig. 7 Results of different methods

表 1 给出了不同方法在部分图像上的检测率, 可以看到, 本文给出的算法性能最优, 且具有良好的场景适应能力。

表 1 不同方法的检测率

Tab. 1 Detection rate of different methods

图像	IPI	MFM	PS	NRAM	ours
1	0.92	0.73	0.85	0.76	0.94
2	1.00	0.91	0.88	0.83	1.00
3	0.95	0.84	0.94	0.76	0.96
4	0.87	0.89	0.87	0.91	0.93
5	0.96	0.94	0.85	0.93	0.99

本文采用局部 SNR 增益 (SNR gain, SNRG) 来评估不同方法的检测性能。局部 SNR 增益定义为:

$$\text{SNRG} = \frac{\text{SNRC}_{\text{out}}}{\text{SNRC}_{\text{in}}} \quad (10)$$

$$\text{SNR} = \frac{|\mu_t - \mu_b|}{\sigma} \quad (11)$$

其中,  $\mu_t$  和  $\mu_b$  分别是在检测结果图像中目标区域和其周围局部邻域的平均像素值。

表 2 展示了不同方法在测试图像上的 SNRG 结果,从表中可以看出本项工作取得的效果基本上优于其他所对比的方法,说明本项工作的检测性能得到较大的提升。

表 2 不同方法的 LCG 对比

Tab. 2 LCG of different methods

图像	1	2	3	4	5
IPI <sup>[4]</sup>	0	1.8258	4.1393	1.7612	0.4243
MFM <sup>[1]</sup>	0.0257	3.8560	4.2727	1.8511	0
PS <sup>[2]</sup>	8.9343	4.1041	3.6256	1.9360	4.7404
NRAM <sup>[5]</sup>	-	4.5186	2.9113	2.1590	-
Ours	12.6532	4.5680	5.0870	2.1513	4.9489

表 2 中“-”表示不存在,即小目标以及其局部邻域像素值均为 0,导致计算检测图像的 LC 时出现分子和分母为 0 的情况。

## 5 结论

提出了一种基于结构化特征卷积神经网络 (SCNN) 的红外弱小目标检测算法。通过将红外弱小目标灰度高斯分布,局部对比度强烈这两个结构化特征作为先验知识引入 CNN 网络中,并对网络进行训练和测试,实现了对红外弱小目标的高可靠检测。实验结果表明:与 4 种代表性红外小目标检测方法相比,本文提出的方法能达到了更优的效果,能够满足稳定、准确的要求。

复杂背景下的红外弱小目标所面临的应用场景非常复杂多变,面临的干扰类型包括:多样性地物、云层边缘、云层的亮斑、传感器坏元及闪元等。所以依靠起始训练的数据集很难实现所有场景的覆盖,因此如何进一步研究分布式迁移学习机理和较好地应用深度学习方法进行红外弱小目标检测是未来研究工作的重点。

## 参考文献:

- [1] Bai X, Bi Y. Derivative entropy-based contrast measure for infrared small-target detection [J]. IEEE Transactions on Geoscience and Remote Sensing, 2018, 56(4): 2452 - 2466.
- [2] Cao W, Huang S. A two-dimensional adaptive target detection algorithm in the compressive domain [J]. Sensors, 2019, 19(3): 567.
- [3] Yang C, Ma J, Zhang M, et al. Multiscale facet model for infrared small target detection [J]. Infrared Physics and Technology, 2014, 67: 202 - 209.
- [4] Bai K, Wang Y, Song Q. Patch similarity based edge-preserving background estimation for single frame infrared small target detection [C]//IEEE International Conference on Image Processing. IEEE, 2016.
- [5] Gao C, Meng D, Yang Y, 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.
- [6] Zhang L, Peng L, Zhang T, et al. Infrared small target detection via non-convex rank approximation minimization joint l2,1 Norm [J]. Remote Sensing, 2018, 10(11): 1821.
- [7] Hou Q, Cheng M M, Hu X, et al. Deeply supervised salient object detection with short connections [C]//IEEE Conference on Computer Vision and Pattern Recognition, 2017: 5300 - 5309.
- [8] Zhang L, Dai J, Lu H, et al. A bi-directional message passing model for salient object detection [C]//IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2018: 1741 - 1750.
- [9] Luo Z, Mishra A, Achkar A, et al. Non-local deep features for salient object detection [C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE Computer Society, 2017.