

# 基于全卷积网络的红外弱小目标检测算法

杨其利<sup>1,2</sup>, 周炳红<sup>1</sup>, 郑伟<sup>1</sup>, 李明涛<sup>1</sup>

(1. 中国科学院 国家空间科学中心, 北京 100190; 2. 中国科学院大学, 北京 100049)

**摘要:** 在小天体探测、导弹制导和战场侦察等航空航天领域, 由于目标信号较弱, 占有像素数少, 缺少目标形状和纹理信息, 使用手工特征提取的传统算法容易出现大量虚警, 而拥有强大特征提取能力的深度学习算法无法对微小且缺乏轮廓信息的目标训练。本文采用了滑动窗口取样训练, 它源自基于人类视觉特性的传统目标检测算法中嵌套结构的思想, 设计了一种使用递归卷积层的全卷积网络, 在不增加额外训练参数的情况下, 扩展了模型的网络深度, 该网络的并行卷积结构的多个分支网络模拟了传统算法的多尺度操作, 有利于在复杂环境中增强目标和背景之间的对比度, 并且设计使用了多种损失函数的组合, 以对抗正负样本严重不平衡的问题。实验结果表明: 该方法实现了比传统方法更好的检测效果, 为此领域的研究者们提供了一个新的思路和解决途径。

**关键词:** 红外图像; 弱小目标检测; 递归卷积; 全卷积网络; 背景抑制

中图分类号: TP391.4 文献标识码: A 文章编号: 1001-8891(2021)04-0349-08

## Small Infrared Target Detection Based on Fully Convolutional Network

YANG Qili<sup>1,2</sup>, ZHOU Binghong<sup>1</sup>, ZHENG Wei<sup>1</sup>, LI Mingtao<sup>1</sup>

(1. National Space Science Center, Chinese Academy of Sciences, Beijing 100190, China;

2. University of Chinese Academy of Sciences, Beijing 100049, China)

**Abstract:** In the field of aerospace research, such as in small celestial body detection, missile guidance, and battlefield reconnaissance, because the target signal is weak, the number of pixels occupied is small, and the target lacks shape structure and texture information, traditional algorithms with manual feature extraction are prone to false alarms, whereas deep learning methods with powerful feature extraction capabilities cannot train tiny targets that lack contour information. In this context, a sliding window sampling training method is adopted, which originates from the idea of nested structures in traditional algorithms based on human visual characteristics. A fully convolutional network using recursive convolutional layers is designed to extend the depth of the network without increasing the training parameters. The multi-branch structure of the network's parallel convolution structure simulates the multi-scale operation of the traditional algorithm, which can enhance the contrast between the target and the background. Additionally, various loss functions are designed to combat the serious imbalance between positive and negative samples. The results show that the algorithm achieves a better detection performance than the traditional algorithms.

**Keywords:** infrared image, dim and small target detection, recursive convolutional layers, fully convolutional network, background suppression

## 0 引言

红外弱小目标检测和跟踪是红外导引的一项关键技术, 在小天体探测、导弹制导和战场侦察等航空航天领域具有重要意义。

由于红外弱小目标探测距离远, 目标的成像尺寸非常小, 在成像平面上仅仅占几十甚至几个像素, 极大地增加了弱小目标检测的难度, 主要体现在: 目标信号弱, 无目标纹理、形状、尺寸等特征信息; 目标运动速度快, 机动性强, 难以获得速度、方向

收稿日期: 2020-02-13; 修订日期: 2020-03-26.

作者简介: 杨其利 (1992-), 男, 硕士研究生, 从事深度学习/弱小目标检测方面的研究。E-mail: yangqili17@mails.ucas.ac.cn.

通信作者: 周炳红 (1976-), 男, 研究员, 博士生导师, 从事飞行器设计/小天体探测与防御研究。E-mail: bhzhou@nssc.ac.cn.

基金项目: 北京市重大科技专项 (Grant Z181100002918004)。

等信息；背景灰度分布不均匀以及随机噪声和高亮度背景的干扰等都增加了弱小目标检测的难度。

中值滤波方法是一种经典的用于图像处理的非线性空间域滤波方法，Deshpand 等人<sup>[1]</sup>提出的最大中值滤波器，通过将红外图像与滤波图像差分运算，既能够有效抑制图像上起伏的背景信息，又能够抑制景象边缘纹理信息，但是这种方法只对高信噪比的小目标有较好的处理效果。Top-Hat<sup>[2]</sup>是一种实用的非线性背景估计方法，它对弱小目标的检测效果取决于结构元素的尺寸和形状，但远距离成像条件下无法获知小目标的先验信息，不能选取统一的结构元素。

基于频率域的弱小目标检测方法通过傅里叶变换将图像从空间域转换到频率域上，然后使用高通滤波器对其进行滤波操作，最后进行傅里叶逆变换得到预测图像。Yang 等人<sup>[3]</sup>将点目标作为图像的高频成分，提出了一种自适应巴特沃斯高通滤波器 (Butterworth high pass filter, BHPF)。Hilliard 等人<sup>[4]</sup>使用低通 IIR (infinite impulse response, IIR) 滤波器来预测杂波，该滤波器适用于多目标情况。Reed 等人<sup>[5]</sup>在已知速度、背景杂波和噪声的情况下，利用频域最优三维线性匹配滤波技术检测图像序列中的运动点目标。

近年来，基于人类视觉系统 (human visual system, HVS) 的弱小目标检测方法主要是通过局部区域中构造一个包含内部窗口和外部窗口的嵌套结构，用于计算内部窗口与外部窗口的对比度，增强局部的目标特征。弱小目标的检测是通过在整副图像上逐像素滑动嵌套结构得到预测图像，最后使用自适应阈值分割方法来实现的。Wei 等人<sup>[6]</sup>提出了一种基于多尺度局部对比度测量 (multiscale patch-based contrast measure, MPCM) 方法，MPCM 方法考虑目标边缘和云边缘的分布差异，利用对角线方向上的最小乘积作为最终增强结果，该方法能够同时增强亮目标和暗目标，在云边缘去除方面取得了很好的效果，但在复杂背景和强杂波干扰下其检测性能会下降。Deng 等人<sup>[7]</sup>提出多尺度绝对灰度均差 (multiscale average absolute gray difference, MS-AAGD) 法，它使用局部平均法抑制背景噪声，通过局部对比度增强目标区域。该算法计算复杂度较低，适合实时处理图像，但是当场景中包含高强度边缘和结构背景比较混乱时，非目标区域也得到增强。Kim 等人<sup>[8]</sup>利用高斯拉普拉斯算子 (Laplacian of Gaussion, LoG) 提出了一种同时增强目标和抑制背景的弱小目标检测算法，该方法关注于抑制低频

杂波，虽然能够检测到低对比度的小目标，但滤波器的二阶导数部分会增强背景噪声，不能滤除高频成分的强杂波。沈等人<sup>[9]</sup>在视觉注意机制的基础上，提出了一种结合尺度自适应的局部对比度测量的红外弱小目标检测方法，该方法能够在增强目标对比度的同时抑制背景杂波，但对于复杂云层下的目标无法有效检测。

对于复杂环境下的弱小目标检测结果容易出现大量虚警的问题，表明这些传统算法手工提取的特征不够充分，而深度学习算法拥有着强大的特征提取和信息抽象能力。深度学习应用于红外图像弱小目标检测的难度在于目标微小，缺乏轮廓特征信息，这对目标检测网络的设计带来了非常大的困难，且目标所处的背景多样性和复杂性，以及目标自身灰度、大小的起伏变化也加大了检测难度。

受传统的基于 HVS 弱小目标检测算法中使用滑动嵌套窗口的启发，本章提出了采用滑动窗口取样训练的全卷积神经网络 (fully convolutoional neural network, F-CNN)，此外，它使用了恒等映射以及递归卷积层，在不增加额外参数的情况下，扩展了模型的网络深度，其并行卷积结构的多分支网络还能模拟传统算法的多尺度表达，本文还设计了多种损失函数的组合用于解决正负样本严重不平衡的问题。

## 1 基于全卷积网络的弱小目标检测方法

### 1.1 滑动窗口取样方案

由于红外弱小目标形状微小，所占像素数不足 80 个像元，对于 256×256 的成像平面，它仅占图像的 0.122%，使得一张图像中正样本和负样本像素严重失衡，极大增加了网络训练的难度。为了改善这一情况，本文采用滑动窗口取样训练的方式，这种做法源自基于人类视觉特性传统弱小目标检测算法的滑动嵌套结构思想，滑动步长设置为略小于取样窗口大小，避免被窗口边界切割造成目标边缘信息丢失，图 1 展示了本文滑动取样窗口的示意图。

通过对取样窗口大小的选择，可以控制网络对整副图像的搜索次数，其计算公式为：

$$N = \frac{W - S + 1}{\text{stride}} \times \frac{H - S + 1}{\text{stride}} \quad (1)$$

式中： $W$ 、 $H$  分别表示为红外图像的宽和高； $S$  表示取样窗口大小； $\text{stride}$  为搜索步长。本文选择了尺寸为 48×48 的取样窗口，步长设为 47，因此对于尺寸为 256×256 的红外图像，需要的搜索次数为 25 次，远小于逐像素遍历需要的 65536 次，在检测

效率上提高了约 2621 倍。

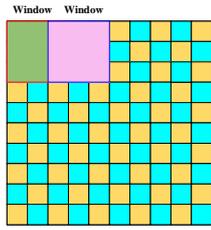


图 1 滑动窗口为 3, 滑动步长为 2 的取样窗口示意图  
Fig.1 The illustration of sampling window with size 3 and sliding step with 2

### 1.2 残差单元

从理论上讲, 非常深的网络可以学习许多不同抽象层次的特征。在深度残差神经网络模型 (deep residual network, ResNet) [10] 中, 通过跳跃连接, 模型可以从某一层获取激活输出, 然后迅速反馈给另外一层。ResNet 的主要思想是使用残差学习来简化深层网络的训练, 作者认为这些网络层通过残差映射可以更容易优化模型。

残差神经网络是由残差块构建的, 在图 2 中, 左侧的图像显示了通过网络的“主路径” [10]。右侧的图像为主路径添加了跳跃连接快捷方式。通过将些残差块堆叠在一起, 可以形成一个非常深的网络。将输入表示为  $x$ , 基本残差单元表示为:

$$\hat{x} = \sigma(F(x, W) + h(x)) \quad (2)$$

式中:  $\hat{x}$  是残差单元的输出;  $h(x)$  是一个恒等映射:  $h(x) = x$ ;  $W$  是一组权重参数;  $\sigma$  表示 ReLU 激活函数;  $F(x, W)$  是要学习的残差映射。对于堆加两个卷积层的基本残差单元, 则有:

$$F(x, W) = W_2 \sigma(W_1 x)$$

式中:  $W_1$  和  $W_2$  分别为两个卷积层的权重参数。通过堆叠这些结构, 构建一个 152 层的网络, ResNet 在 ILSVRC 2015 分类竞赛中获得第一名。

### 1.3 递归模块

深度递归卷积网络 (deeply recursive convolutional network, DRCN) [11] 的作者认为, 增加更多的权重层会引入更多的参数, 这些参数会导致模型

过拟合。为了解决这个问题, 他首次提出了递归模块, 并在网络中添加了 16 个递归层, 每个递归层使用相同的卷积参数, 这样模型在递归层中执行卷积操作时就不会额外增加参数。

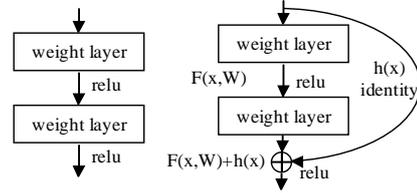


图 2 普通网络 (左) 和残差块网络 [10] (右)

Fig.2 Left: general network, Right: residual network [10]

对于具有  $D$  个递归层的递归模块, 我们对所有卷积操作使用相同的权重  $W$  和偏值参数  $b$ 。递归模块的展开结构如图 3 所示, 定义  $g$  为递归模块中单个递归层的卷积函数,  $H$  为递归层的输入。

$$g(H) = \max(0, W * H + b) \quad (3)$$

第  $d$  个递归层的递归模块输出为:

$$H_d = g(H_{d-1}) = \max(0, W * H_{d-1} + b) = g^{d-1}(H_1) = g^d(H) \quad (4)$$

式中:  $g^d$  表示函数的  $d$  次迭代。

### 1.4 F-CNN 网络

F-CNN 是一种全卷积网络, 用于检测红外图像中的暗弱小目标, 它可以将网络自主提取的特征转换为绝大部分激活值为 0 的稀疏特征映射。网络的输入图像尺寸为  $48 \times 48$ , F-CNN 的结构如图 4 所示, 它由特征提取模块、递归模块和重建模块组成。适当缩小卷积核尺寸同时增加卷积核数量, 有助于网络稳定, 减少参数数量, 也能够提取足够多的特征。因此本文所有卷积层参数设置为: 卷积核大小  $f$ 。卷积操作的步长  $s=1$  和零填充  $p=1$ , 网络各层的参数设置如表 1 所示, 定义  $f_1, f_2, f_3$  分别为特征提取模块、递归模块、重建模块的函数。

特征提取模块: 它由 5 个卷积步长为 1, 卷积核大小为 3 的卷积层组成。网络结构采用多个级联的小滤波器, 用于提取足够多的图像特征。式中,  $x$  为尺寸为  $48 \times 48$  的输入图像,  $W_i, b_i, (i=1, 2, \dots, 5)$  分别为各层卷积核参数。

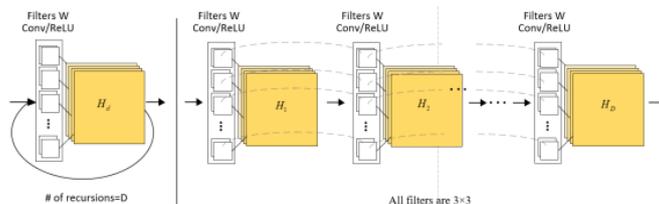


图 3 递归模块的展开结构, 相同卷积参数应用不同递归层

Fig.3 Unfolding recursive module, the same filter  $W$  is applied to feature maps recursively

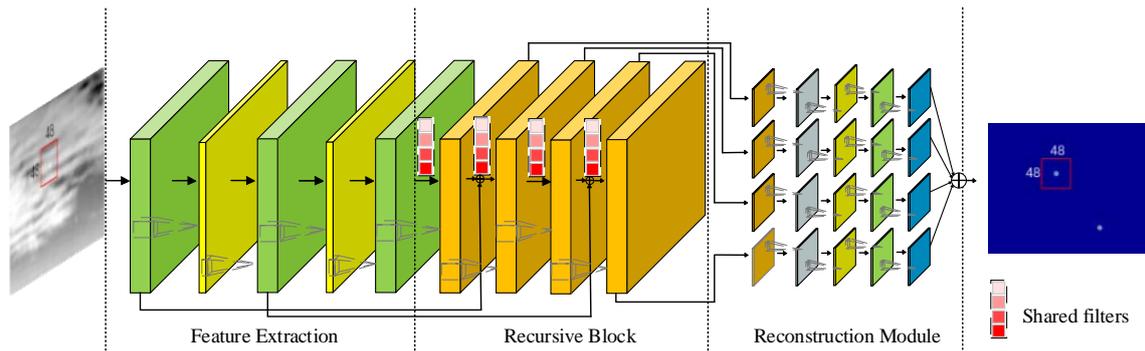


图4 F-CNN网络结构图

Fig.4 An illustration of the F-CNN architecture

$$\begin{aligned} X_1 &= \max(0, W_1x + b_1) & (5) \\ X_2 &= \max(0, W_2X_1 + b_2) & (6) \\ X_3 &= \max(0, W_3X_2 + b_3) & (7) \\ X_4 &= \max(0, W_4X_3 + b_4) & (8) \\ X_5 &= \max(0, W_5X_4 + b_5) & (9) \\ f_1(x) &= X_5 & (10) \end{aligned}$$

表1 本文使用的F-CNN网络结构

Table 1 F-CNN architecture for semantic segmentation

Layers	Output size	Layer configurations
Feature extraction	48×48	$\begin{bmatrix} 3 \times 3, 32 \\ 3 \times 3, 16 \end{bmatrix} \times 2$ $[3 \times 3, 32] \times 1$
Recursive block	48×48	$[3 \times 3, 32] \times 4$
Reconstruction module	48×48	$\begin{bmatrix} 3 \times 3, 32 \\ 3 \times 3, 32 \\ 3 \times 3, 16 \\ 3 \times 3, 2 \end{bmatrix} \times 1$

递归模块：如图4所示，它由4层递归卷积层组成，每个卷积层使用同一组网络层的卷积参数，为了解决梯度和最优递归问题，本文利用残差网络的跳跃连接，将特征提取模块中的第1、3层卷积层分别连接到递归模块中的第1、3层递归层，对这两层递归层进行信息监督，以减轻梯度消失或爆炸的影响。在递归模块的所有递归卷积层中，由于共享一个网络层的卷积参数，只需要训练递归模块的一组卷积核参数  $W, b$ ，就可以加速模型的收敛。

$$H_1 = \max(0, WX_5 + b) + X_1 = g^1(X_5) \quad (11)$$

$$H_2 = \max(0, WH_1 + b) = g^2(X_5) \quad (12)$$

$$H_3 = \max(0, WH_2 + b) + X_3 = g^3(X_5) \quad (13)$$

$$H_4 = \max(0, WH_3 + b) = g^4(X_5) \quad (14)$$

$$f_2(X_5) = H_4 \quad (15)$$

重建模块：它是一种并行卷积结构，用于加权不同深度网络层的抽象特征，并输出预测结果。本

文将递归模块中所有递归层的层间输出作为重建模块并行卷积结构不同分支的输入。在重建模块中，并行卷积结构有4个分支网络，所有网络分支共同使用同一组网络分支参数，其参数设置如表1所示，每个分支经过4层卷积操作，由于不同分支的信息输入来自不同网络深度递归层的输出，使得图像的感知野不同，这意味着并行卷积结构的不同分支相当于多尺度表达。

$$\hat{y}_d = f_3(g^d(X_5)) = f_3(g^d(f_1(x))) \quad (16)$$

式中： $d=1,2,3,4$ 。之后将卷积操作得到的特征图在深度方向进行权重  $w_d$  加权平均，这个参数是在网络训练中得到的。网络的预测结果为：

$$\hat{y} = \sum_{d=1}^4 w_d \cdot \hat{y}_d \quad (17)$$

### 1.5 损失函数

基于像素点的交叉熵损失函数是一种语义分割问题中最常用的损失函数。这种损失函数单独检查每个像素，其数学表达式为：

$$\begin{aligned} CE(y, \hat{y}) &= -(y \ln(\hat{y}) + (1-y) \ln(1-\hat{y})) = \\ &\begin{cases} -\ln(\hat{y}), & y = 1 \\ -\ln(1-\hat{y}), & y = 0 \end{cases} \end{aligned} \quad (18)$$

式中： $y$  是真实样本的标签， $\hat{y}$  是重建模块的预测结果（数值在0~1之间）。为了减少易分类样本的贡献，使网络更关注难分类样本，本文引入了Focal函数（FL），它是在交叉熵损失函数中加入调制系数  $(1-\hat{y})^\gamma$ ，能够控制易分类样本和难分类样本的权重，让模型更关注于困难的、错分的样本，并引入一个平衡因子  $\alpha$ ，用来平衡正负样本的数量比例不均匀。

$$\begin{aligned} \text{FL}(y, \hat{y}) &= -(\alpha(1-\hat{y})^\gamma y \ln(\hat{y}) + (1-\alpha)\hat{y}^\gamma (1-y) \ln(1-\hat{y})) \\ &= \begin{cases} -\alpha(1-\hat{y})^\gamma \ln(\hat{y}), & y=1 \\ -(1-\alpha)\hat{y}^\gamma \ln(1-\hat{y}), & y=0 \end{cases} \end{aligned} \quad (19)$$

式中:  $\alpha=0.25$ ,  $\gamma=2$ 。为了解决弱小目标所在图像中正负样本比例严重失衡的问题, 使目标轮廓被精确定位, 本文引入 Dice 系数差异函数用于衡量实际小目标和预测小目标轮廓区域的相似程度, 这种损失函数可以使弱小目标轮廓部分的正样本正确分类。Dice 系数 (DC) 是一种基于两个样本之间相似度的度量函数, 其数学表达式为:

$$\text{DC} = \frac{2|X \cap Y|}{|X| + |Y|} \quad (20)$$

式中:  $|X \cap Y|$  表示集合  $X$ 、 $Y$  之间的共同元素,  $|X|$  表示集合  $X$  中的元素,  $|Y|$  表示集合  $Y$  中的元素。Dice 系数差异损失函数为:

$$\text{DL}(y, \hat{y}) = 1 - \frac{2 \sum_{i=1}^N y_i \cdot \hat{y}_i}{\sum_{i=1}^N y_i^2 + \sum_{i=1}^N \hat{y}_i^2} \quad (21)$$

式中:  $\hat{y}_i = \sum_{j=1}^4 w_{j,i} \cdot \hat{y}_{j,i}$ ,  $y_i, \hat{y}_i$  分别为真值和预测在不同通道的特征图。由于是二分类问题, 网络的预测输出通道数为 2, 因此  $N=2$ ,  $\cdot$  表示点积。因此, 模型的优化目标是 minimized 损失函数:

$$L(\theta) = \alpha_1 \text{FL}(y, \hat{y}) + \alpha_2 \text{DL}(y, \hat{y}) + \beta \|\theta\|_2 \quad (22)$$

式中:  $\alpha_1, \alpha_2$  表示不同损失函数的权重;  $\beta$  表示网络参数  $\theta$  的惩罚因子, “ $\|\theta\|_2$ ” 表示  $\theta$  的 L2 范数。

## 2 实验结果和讨论

本实验使用 Google 开源框架 Tensorflow 构建 F-CNN 模型进行训练。使用的数据集为仿真的图像样本, 包括 8000 张大小为  $48 \times 48$  的训练图像, 测

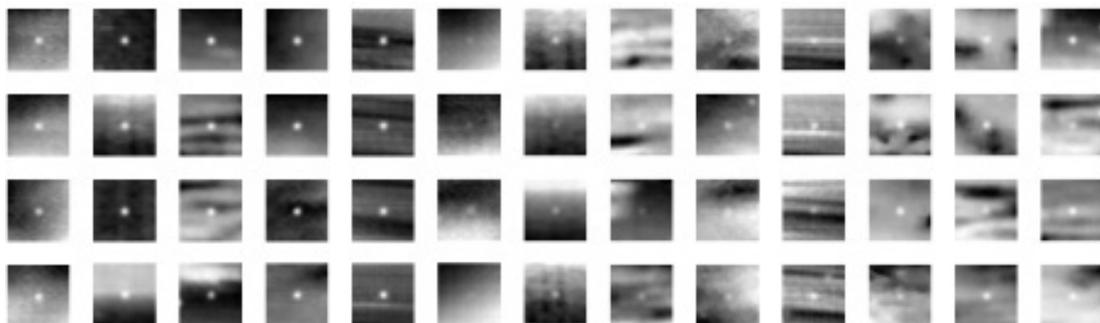


图 5 弱小目标仿真数据集

Fig. 5 Simulated dataset of infrared small targets

试样本来自无人机拍摄的可见光图像<sup>[12]</sup>和其他红外图像。实验设置 batch 为 32, 初始学习率为 0.0001, 权重衰减设为 0.0001, 采用 Adam 模型优化, 共训练 100 个 epoch。

### 2.1 数据集建立

由于在复杂云层背景或海平面中掺杂的暗弱目标近似于点, 本文将二维高斯强度模型作为目标点添加到背景图像中。二维高斯模型如下:

$$s(i, j) = s_E \exp\left(-\frac{1}{2} \left[ \frac{(i-x_0)^2}{\sigma_x^2} + \frac{(j-y_0)^2}{\sigma_y^2} \right]\right) \quad (23)$$

式中:  $(x_0, y_0)$  是图像目标的中心;  $s(i, j)$  是目标在图像位置  $(i, j)$  的灰度值。生成目标的强度  $s_E$  是介于 0~1 之间的随机数。  $\sigma_x$  和  $\sigma_y$  是高斯分布标准差。利用不同的参数可以产生不同信杂比的弱小目标, 本文生成的目标信杂比在 2~12 之间, 训练样本如图 5 所示。

### 2.2 度量指标

信杂比增益 (the signal-to-clutter ratio gain, SCRG)<sup>[13]</sup> 和背景抑制因子 (the background suppression factor, BSF)<sup>[14]</sup> 能够描述目标增强以及背景抑制的性能, 它们是评价最终检测效果的重要指标, SCRG 定义为:

$$\text{SCRG} = \frac{\text{SCR}_{\text{out}}}{\text{SCR}_{\text{in}}} \quad (24)$$

式中:  $\text{SCR}_{\text{in}}$  和  $\text{SCR}_{\text{out}}$  分别表示原始图像和预测图像的信杂比。信杂比 (the signal-to-clutter ratio, SCR) 定义为:

$$\text{SCR} = \frac{|\mu_t - \mu_b|}{\sigma_b} \quad (25)$$

式中:  $\mu_t$  表示目标的平均灰度值;  $\mu_b$  表示目标邻域像素的平均灰度值;  $\sigma_b$  表示目标周围像素的标准差。邻域的大小为  $(a+2d) \times (b+2d)$ , 目标的大小为  $a \times b$ , 实验设置  $d=15$ 。BSF 定义为:

$$BSF = \frac{\sigma_1}{\sigma_0} \quad (26)$$

式中： $\sigma_1$ 和 $\sigma_0$ 分别表示原始图像和处理后图像像素的标准差。通常来讲，弱小目标图像的 SCR<sub>G</sub> 和 BSF 值越高，越容易检测到目标。

### 2.3 实验结果与对比方法

为了测试本文方法检测弱小目标的性能，在 200 张测试图像中对比了 F-CNN、MS-AAGD、LoG 和 MPCM 方法。该测试集来自文献[12]中 3 个分辨率为 1920×1080 的视频序列，帧速率为 30 fps，它们是由安装在三角翼无人机的 GoPro3 摄像机拍摄的。为了获取有效的测试图像，我们截取了视频的部分区域，制作了 200 张尺寸为 720×540 的复杂云背景图像作为测试集。

图 6 展示了不同方法对其中 5 张测试图像的背景抑制结果。第一列图像是 5 张具有代表性的原始图像，背景复杂多变，每张图像各有一个小目标，它们都淹没在厚重的云层中，且目标占有的像素数

少，缺少目标形状结构和纹理信息，弱小目标相应的信杂比和尺寸信息如表 2 所示。图 6 第二列展示了各测试图像对应的三维灰度分布图。图 6 第 3 列至第 6 列分别表示 MS-AAGD、LoG、MPCM 和本文方法对应的背景抑制图像。基于 MS-AAGD 方法的滤波图像中，虽然消除了大量的杂波和噪声残留，但也检测出许多虚警的目标，且虚警目标的信号强度接近真实目标强度。基于 LoG 方法的滤波图像展示了明显的真实目标信号，能够降低虚警率，但是背景中依然存在大量的噪声和杂波。基于 MPCM 的方法要优于 MS-AAGD 方法，滤波结果没有过多的背景边缘纹理，且接近真实目标强度的虚警目标数量要少于 MS-AAGD 的方法。本文的方法 F-CNN 得到的滤波结果显示，背景噪声接近为 0，无论是虚警目标的数量或者是信号强度，都远远低于对比方法。这种方法能够使背景噪声得到极大抑制，使目标信号极大增强，实现了目标和背景杂波区分开的理想效果，性能远远高于对比方法。

表 2 无人机测试图像的 SCR 值和目标像素数

Table 2 The SCR and target size of UAV test images.

Test images	Image 1	Image 2	Image 3	Image 4	Image 5
SCR	4.009	2.337	8.378	4.411	2.976
Target size/Pixel	15	25	12	13	28

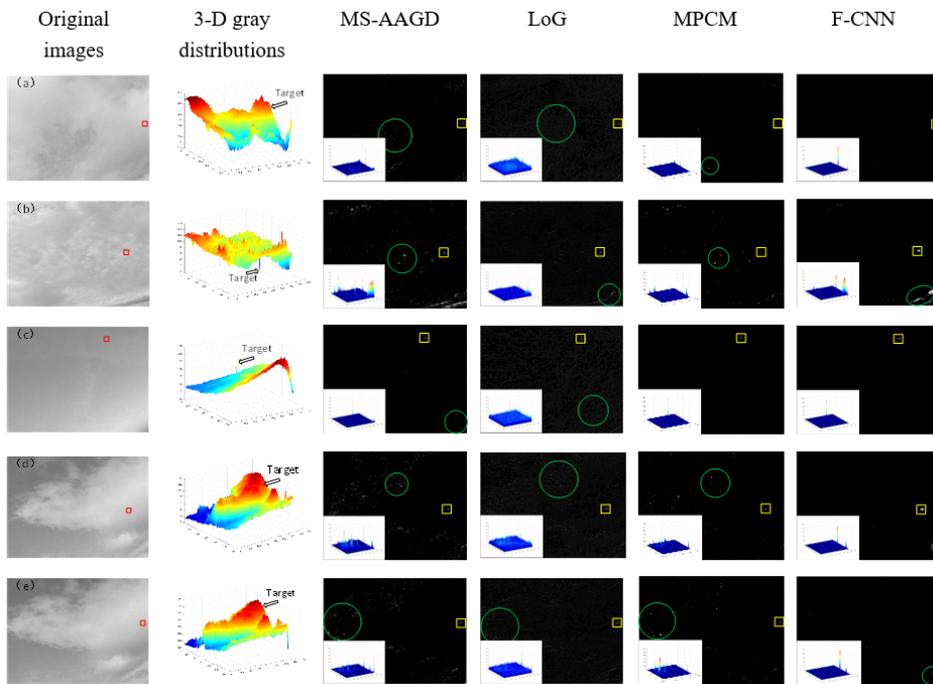


图 6 不同检测方法在 5 张测试图像上的滤波结果，矩形框表示目标，圆圈表示滤波后的噪声

Fig.6 The representative results of different methods on five test images, the rectangles denote the targets and the circles are representative examples of noise

F-CNN 方法之所以能够实现这种效果, 在于递归模块中使用了递归卷积结构, 这是它将原始图像中的目标和背景完全区分开的重要一步。同时, 重建模块中继承了递归模块多个递归层的输出预测, 构造成多分支网络结构, 这相当于 F-CNN 学习到多个弱小目标背景抑制方法, 通过对多个分支的预测结果赋予不同权重, 能够增强真实目标信号的强度, 同时对有噪声杂波的预测赋予较低的权重, 能够极大抑制杂波信号和背景结构纹理。

为了客观地评价不同方法的背景抑制效果, 本文还计算了信杂比增益 SCR<sub>G</sub> 和背景抑制因子 BSF, 如表 3 所示。结果表明, F-CNN 方法能够显著提高图像的 SCR 值和 BSF 值, 且此方法在 SCR 值和 BSF 值指标均优于或接近对比方法。本文方法能够有效地抑制各种复杂噪声背景, 并且在滤波结果图中具有非常少的杂波和噪声残留。这些实验结果证明了我们的方法在增强目标、抑制背景杂波和噪声方面优异的性能。因此, 我们的方法有助于探测嵌入在不同程度云层背景杂波中的弱小目标。

为了进一步验证本文方法的有效性, 对 4 张红

外图像进行对比实验, 除第三张测试图像中含有两个红外目标外, 其余图像均只含有一个目标。如图 7 所示, 在第一张图像的滤波结果对比图中, 本文方法提取的目标信号足够强烈, 在背景杂波抑制方面要弱于 MPCM 方法, 但优于其他方法。第二张图像有一条明显的海天分割线, 对比方法均出现了海天分割线的杂波, 且滤波结果存在过多的杂波噪声, 杂波信号强度近似目标信号, 而本文方法背景抑制效果要优于对比方法。第三张图像包含两个红外目标和一个灯塔建筑, 除 MPCM 方法外, 其他方法均出现了建筑边缘的滤波噪声, 而 MPCM 出现了信号明显的虚警目标。第四张红外图像的滤波结果表明, 本文提取的弱小目标信号弱于对比方法, 但保持了原始图像中目标的尺寸和纹理结构信息, 而不是对比方法中点目标的滤波结果。本文还计算了 4 张红外图像的信杂比增益和背景抑制因子, 表 4 显示本文方法在客观指标上要优于或接近对比方法, 这为今后的红外图像弱小目标检测任务提供了解决问题的新途径。

表 3 不同方法对图 6 第一列测试图像滤波结果的 SCR<sub>G</sub> 和 BSF 值

Table 3 The evaluation results of SCR<sub>G</sub> and BSF of different methods for images in the first column in Fig.6

Methods	Image 1		Image 2		Image 3		Image 4		Image 5	
	BSF	SCR <sub>G</sub>								
MS-AAGD	0.596	6.908	0.639	5.272	0.298	3.440	0.690	6.764	0.952	6.178
LoG	0.193	1.526	0.773	4.517	0.099	1.030	0.243	1.791	0.306	2.379
MPCM	0.793	8.577	1.395	8.202	0.358	3.547	0.732	7.201	0.985	6.401
F-CNN	0.875	8.655	1.419	13.684	0.305	3.694	0.764	8.821	1.195	11.137

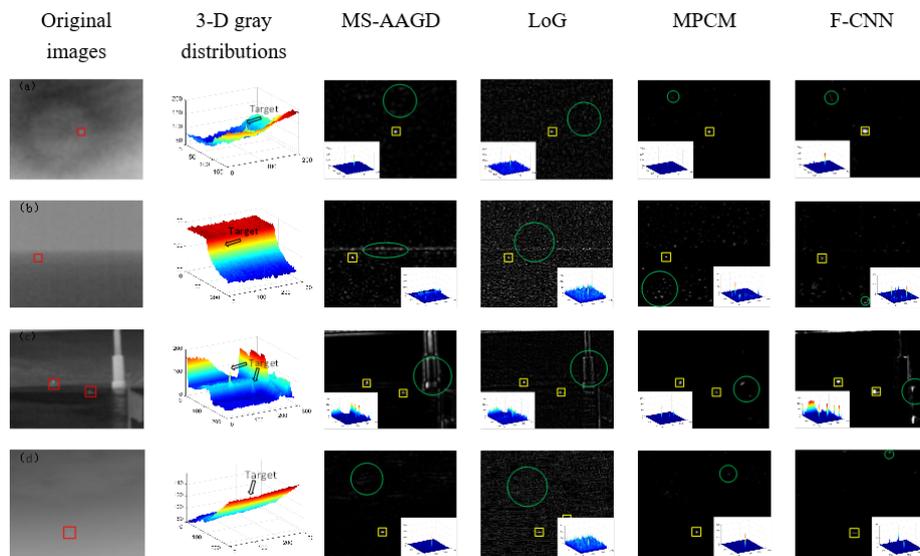


图 7 不同检测方法在红外图像上的滤波结果, 矩形框表示目标, 圆圈表示滤波后的噪声  
Fig.7 Different methods on infrared images, the rectangles denote the targets and the circles are representative examples of noise

表4 不同方法对图7中红外图像滤波结果的 SCRG 和 BSF 值

Table 4 The evaluation results of SCRG and BSF of different methods for infrared images in Fig.7

Methods	Image 1		Image 2		Image 3				Image 4	
	BSF	SCRG	BSF	SCRG	BSF	SCRG	BSF	SCRG	BSF	SCRG
MS-AAGD	2.110	11.483	0.712	18.757	1.872	3.998	2.56	6.298	1.091	16.575
LoG	0.853	4.967	0.473	15.559	1.279	3.042	1.358	3.408	0.241	4.298
MPCM	1.758	12.399	2.011	96.279	1.984	4.088	4.142	8.333	0.966	18.781
F-CNN	2.485	16.337	3.314	63.100	1.710	5.407	2.465	10.528	2.404	15.440

### 3 结论

传统的小目标检测基本处理过程为先进进行图像预处理, 然后进行目标增强, 最后依靠一定的判决条件检测目标。本文首次将计算机视觉中语义分割的问题与红外图像弱小目标检测的任务联系起来, 提出了一种基于卷积神经网络的弱小目标检测方法 F-CNN, 这是一种结合残差结构和递归结构使用的全卷积网络, 它使用了滑动窗口的方式取样训练, 由特征提取模块、递归模块和重建模块构成, 可以将图像中的目标从背景杂波中完全分割出来, 将真实目标和干扰目标(如云的边缘)分离。方法还设计使用了多种损失函数的组合, 以对抗正负样本严重不平衡的问题。本文还将 F-CNN 方法与 MS-AAGD、LoG、MPCM 方法在实际场景中进行检测效果的对比。实验表明, 本文方法在目标增强、背景抑制方面均优于对比方法。

#### 参考文献:

- [1] DESHPANDE S D, ER M H, VENKATESWARLU R, et al. Max-mean and max-median filters for detection of small targets[C]//*Proceedings of SPIE*, 1999, **3809**: 74-83.
- [2] ZENG M, LI J X, PENG Z, et al. The design of top-hat morphological filter and application to infrared target detection[J]. *Infrared Physics and Technology*, 2006, **48**(1): 67-76.
- [3] YANG L, YANG J, YANG K, et al. Adaptive detection for infrared small target under sea-sky complex background[J]. *Electronics Letters*, 2004, **40**(17): 1083-1085.
- [4] HILLIARD C I. Selection of a clutter rejection algorithm for real-time target detection from an airborne platform[C]//*Proceedings of SPIE*, 2000, **4048**: 74-84.
- [5] REED I, GAGLIARDI R, STOTTS L, et al. Optical Moving Target Detection with 3-D Matched Filtering[J]. *IEEE Transactions on Aerospace and Electronic Systems*, 1988, **24**(4): 327-336.
- [6] WEI Y T, YOU X G, LI H, et al. Multiscale patch-based contrast measure for small infrared target detection[J]. *Pattern Recognition*, 2016, **58**: 216-226.
- [7] DENG H, SUN X P, LIU M L, et al. Infrared small-target detection using multiscale graydifference weighted image entropy[J]. *IEEE Transactions on Aerospace and Electronic Systems*, 2016, **52**(1): 60-72.
- [8] KIM S H, LEE J Y. Scale invariant small target detection by optimizing signal-to-clutter ratio in heterogeneous background for infrared search and track[J]. *Pattern Recognition*, 2012, **45**: 393-406.
- [9] 沈旭, 程小辉, 王新政. 结合视觉注意力机制基于尺度自适应局部对比度增强的红外弱小目标检测算法[J]. *红外技术*, 2019, **41**(8): 764-771.
- [10] SHEN X, CHENG X H, WANG X ZH. Infrared dim-small object detection algorithm based on adaptive scale local contrast enhancement combined with visual attention mechanism[J]. *Infrared Technology*, 2019, **41**(8): 764-771.
- [10] HE K M, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition[C]//*Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition*, New York: 2016: 770-778.
- [11] KIM J, LEE J K, Lee K M, et al. Deeply-recursive convolutional network for image super-resolution[C]//*Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition*, New York: 2016: 1637-1645.
- [12] LI J, YE D H, CHUNG T, et al. Multi-target detection and tracking from a single camera in unmanned aerial vehicles (UAVs)[C]//*Proceedings of 2016 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2016: 4992-4997.
- [13] GAO C, ZHANG T, LI Q, et al. Small infrared target detection using sparsening representation[J]. *IEEE Aerospace and Electronic Systems Magazine*, 2012, **27**(3): 21-30.
- [14] QU X J, CHEN H, PENG G H, et al. Novel detection method for infrared small targets using weighted information entropy[J]. *Journal of Systems Engineering and Electronics*, 2012, **23**(6): 838-842.