

# 基于深度学习的红外遥感目标超分辨率检测算法

黄硕<sup>1,2</sup>, 胡勇<sup>1,2\*</sup>, 顾明剑<sup>1</sup>, 巩彩兰<sup>1,2</sup>, 郑付强<sup>1,2,3</sup>

<sup>1</sup>中国科学院上海技术物理研究所, 上海 200083;

<sup>2</sup>中国科学院红外探测与成像技术重点实验室, 上海 200083

<sup>3</sup>中国科学院大学, 北京 100049

**摘要** 红外遥感图像受限于红外衍射极限,其分辨率普遍较低,这为红外目标的精准检测和识别带来了困难。针对此问题,提出基于深度学习的红外目标超分辨率目标识别(SROD)算法,该算法主要包括两部分:第一部分是利用 WDSR(Wide Activation for Efficient and Accurate Image Super-Resolution)对红外遥感图像进行超分辨率重建,将模拟传感器下采样方式处理的红外图像作为训练集。第二部分是基于 Faster RCNN 的目标检测,提出多尺度特征传递网络结构,将低层特征输入区域候选网络(RPN)层,降低了弱小目标像素被简化的概率,并利用可调节非极大值抑制方法,减少了对密集目标检测框的抑制作用。将该算法应用于整幅红外遥感图像,与相同训练集的 Faster RCNN 相比,目标检测的准确率提升了 5.33%,召回率提升了 12.22%,特别是小目标的召回率提升了 13.25%。

**关键词** 图像处理;深度学习;超分辨率;目标检测;红外遥感

中图分类号 TP391

文献标志码 A

doi: 10.3788/LOP202158.1610015

## Super-Resolution Infrared Remote-Sensing Target-Detection Algorithm Based on Deep Learning

Huang Shuo<sup>1,2</sup>, Hu Yong<sup>1,2\*</sup>, Gu MingJian<sup>1</sup>, Gong Cailan<sup>1,2</sup>, Zheng Fuqiang<sup>1,2,3</sup>

<sup>1</sup> Shanghai Institute of Technical Physics, Chinese Academy of Sciences, Shanghai 200083, China;

<sup>2</sup> Key Laboratory of Infrared System Detection and Imaging Technology, Chinese Academy of Sciences, Shanghai 200083, China;

<sup>3</sup> University of Chinese Academy of Sciences, Beijing 100049, China

**Abstract** Owing to the infrared diffraction limit, the resolution of infrared remote sensing images is generally low, which makes precise detection and recognition of infrared targets difficult. To address this problem, an infrared target super-resolution detection algorithm based on deep learning is proposed. The algorithm comprises two main parts. The first part implements Wide Activation for Efficient and accurate image super-resolution (WDSR) to reconstruct infrared remote sensing images, and uses infrared images processed by the downsampling method of the sensor as the training set. The second part involves target detection based on Faster region-based convolutional neural network (Faster RCNN). A multiscale feature transfer network structure is proposed. The low-level features are input to region proposal network (RPN), which reduces the simplification rate of weak and small target pixels. In addition, a nonmaximum suppression method is used to reduce the suppression of dense target detection frames. Compared with Faster RCNN using the same training set, the proposed algorithm increased target detection accuracy, the overall recall rate, and the recall rate of small targets by 5.33%, 12.22%, and 13.25%, respectively.

**Key words** image processing; deep learning; super resolution; target detection; infrared remote sensing

**OCIS codes** 100.2000; 100.2980; 100.4996; 100.3010; 100.5010

收稿日期: 2020-11-05; 修回日期: 2020-11-15; 录用日期: 2020-12-27

基金项目: 国家重点研发计划(2017YFC0602103)、上海市科委项目(17411952800, 18441904500)

通信作者: \*huyong@mail.sitp.ac.cn

# 1 引言

红外图像具有昼夜可视的优势,利用红外图像进行目标的检测和识别在军事和民用方面有广泛的应用<sup>[1]</sup>。但是,红外衍射极限限制了红外遥感图像的分辨率,导致红外遥感图像中目标的像元个数少、边缘模糊,影响了小目标的检测和识别精度。超分辨率(SR)重建是一种利用低分辨率图像重建出高分辨率图像的技术,可以提升图像的清晰度和细节纹理<sup>[2]</sup>。20世纪60年代,超分辨率重建的概念被首次提出,Harris利用插值的方法提高了单帧图像的空间分辨率。随着深度学习的发展,Dong等<sup>[3]</sup>首次将卷积神经网络用于超分辨率重建,提出SRCNN,大幅改善了超分辨率重建的效果。Yu等<sup>[4]</sup>提出了WDSR(Wide Activation for Efficient and Accurate Image Super-Resolution),去掉了冗余的卷积层,是目前超分辨率重建效果最好的网络之一。实验证实超分辨率重建可以提高分类和目标的检测精度<sup>[5-6]</sup>。虽然超分辨率重建对检测精度的影响已被关注,但真正意义上对超分辨率图像重建和目标检测的关系研究较少。宋文韬等<sup>[7]</sup>研究了不同分辨率变换下的目标几何特性和辐射特性,得出舰船目标的目标类型和方向在空间分辨率低于12 m的图像中很难判断。由此说明在分辨率低于12 m的图像中舰船目标的被检测能力有所下降。Koester等<sup>[6]</sup>研究了两种超分辨率预处理方法(EDSR<sup>[8]</sup>和最邻近插值)对目标检测的影响,证实分辨率提高可以有效提高目标检测率。但是此研究只使用了两种超分辨率重建方法进行对比实验,说服力不强。Bai等<sup>[5]</sup>提出了一种端到端的多任务生成对抗网络SOD-MTGAN,并在COCO数据集中进行了大量的实验,提高了模糊的较小目标的检测能力,整体提升了COCO数据集的检测准确率。该网络分为生成器和判别器两部分。生成器是一个超分辨率网络,判别器用于检测并将分类和边界框回归损失反向传播到生成器中。以上算法都取得了一定的研究成果,但研究的数据类型都是可见光图像,并没有涉及红外图像。

利用深度学习来解决遥感问题是目前的研究趋势,其中面临的主要问题有以下几点:1)遥感图像特征不同于自然图像,遥感图像多为空中拍摄,拍摄角度与常规自然图像不同,所以目标特征也不相同。2)遥感图像中的小目标检测难度大,小目标在图像中像素少、特征信息少,检测难度大,漏检率较高。且遥感目标经常会出现聚集和排列的特征,比如飞

机场、停车场、港口等场景,这对检测中的准确定位也是一个严峻的考验。3)卫星遥感图像尺寸较大。图像尺寸超过了一般检测模型的最大输入尺寸要求。常规检测手段可通过降采样来解决尺寸问题,但这种做法不适用于遥感图像小目标检测。

本文针对上述问题提出了红外目标超分辨率目标识别(SROD)系统:1)利用红外遥感图像训练网络参数,学习红外遥感目标的特征,并利用缩放和旋转处理提升训练集的丰富度。2)利用WDSR(Wide Activation for Efficient and Accurate Image Super-Resolution)算法对图像进行预处理,提升了红外遥感图像的分辨率,增强了目标边缘和纹理信息,并模拟传感器采样机制,对高分辨率图像进行下采样,代替双三次插值的方法,使SR图像更接近于真实的高分辨率图像。3)针对图像中小目标改进的Faster RCNN结构,加入多层特征参考,改变检测框筛选机制,提高了密集目标的检测精度。4)利用裁剪的方法减小输入的遥感图像尺寸,代替一般的重采样手段,可以防止小目标像素信息的减少。

## 2 SR网络学习传感器下采样过程

基于深度学习的超分辨率训练数据,普遍是利用高分辨率图像进行下采样来取得,常用的下采样方式包括最邻近、双线性和双三次插值。将对应的不同分辨率的图像同时输入到训练网络中,利用深度神经网络学习低分辨率和高分辨率的映射关系,此过程即图像退化的逆过程。但是一个真实的低分辨率图像,不是从高分辨率图像下采样获得的,而是从物体和环境的辐射通过光电转换得到的,故本研究利用传感器仿真模型来模拟图像从实际场景到成像的过程,利用神经网络来学习遥感成像的逆过程,由此可以重建出能够表达真实场景的高分辨率图像。

### 2.1 WDSR网络结构

WDSR是目前最先进的超分辨率重建算法之一。基于深度学习的超分辨率重建算法的基本思想是通过卷积神经网络提取图像的特征,并从特征图中学习图像的信息。卷积神经网络反映了高低分辨率的对应关系,能够尽可能地将有用信息传递到后端,有利于提升图像的重建效果。WDSR在EDSR<sup>[8]</sup>的基础上,改进网络结构,去除了冗余的卷积层,并将放大图像后的卷积层提前,这一操作既减少了计算量,也提高了计算精度。该算法用线性低秩卷积作为基本构件增加激活层前面的特征数,扩

大激活范围,进一步提升准确性;提出的权重归一化(WN)不会降低超分辨率精度,且加快了收敛速度;将所有卷积层都放在上采样之前,在上采样之前利

用 WDSR 提取了所有特征,超分辨率的准确性不受影响,重建速度得以提升。WDSR 的网络结构见图 1。

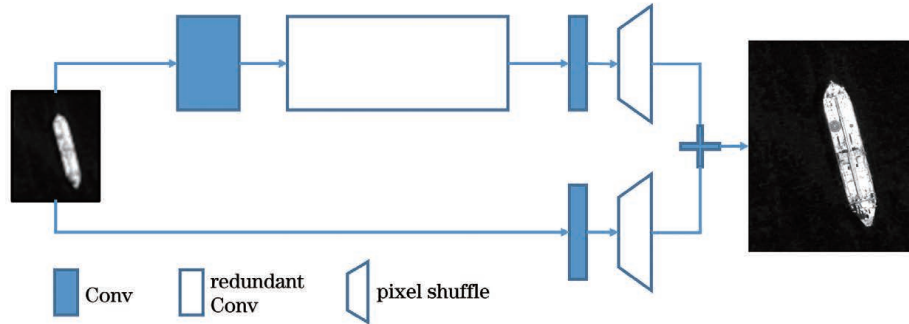


图 1 WDSR 结构

Fig. 1 Structure of WDSR

### 2.2 传感器采样模型仿真

传感器中探测元呈线性排列(见图 2),每个探测元对应一个图像像元,同时代表地面的瞬时视场。探测器通过扫描相机的摆动进行成像,在顺轨方向的垂直方向,每移动一个像元,瞬时视场的辐射就被在图像上记录一次。

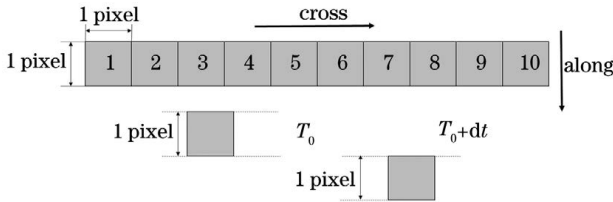


图 2 采样过程

Fig. 2 Process of sampling

点光源从拍摄场景到生成图像是一个退化过程,这个过程可以用点扩展函数(PSF,  $F_{PS}$ )来描述,PSF 代表了成像系统对点源或点对对象的响应,图像  $S_b(x, y)$  可以表示为

$$S_b(x, y) = \int_{\alpha_{\min}}^{\alpha_{\max}} \int_{\beta_{\min}}^{\beta_{\max}} \Phi_b(\alpha, \beta) F_{PS}(x - \alpha, y - \beta) d\alpha d\beta, \quad (1)$$

其中  $\Phi_b$  为入瞳处的辐射通量。场景是连续的,数字图像是非连续的,从光谱信号到数字图像经历了采样过程。系统的点扩展函数由光学扩散函数、探测器扩散函数、电子扩散函数组成,表达式为

$$F_{PS} = F_{PS_{opt}} * F_{PS_{det}} * F_{PS_{el}}, \quad (2)$$

式中  $F_{PS_{opt}}$  为光学扩散函数,  $F_{PS_{det}}$  为探测器扩散函数,  $F_{PS_{el}}$  为电子扩散函数。总的 PSF 函数为类高斯模型。光学 PSF 是光学系统对光源的响应,点光源的衍射效应通过光学系统在像面上产生衍射像斑。能量分布为类高斯函数,表达式为

$$F_{PS_{opt}}(x, y) = \frac{1}{2\pi ab} \exp(-x^2/a^2) \exp(-y^2/b), \quad (3)$$

式中,  $x$  和  $y$  分别为点光源在图像上的坐标位置,  $a, b$  为常数。

探测器 PSF 是指探测器获取辐射信息时,通过采样和滤波作用产生的空间模糊作用。单个探测元的点扩展函数可以表示为

$$F_{PS_{det}} = \text{rect}\left(\frac{x}{w}\right) \text{rect}\left(\frac{y}{w}\right), \quad (4)$$

其中,  $w$  代表像元宽度,  $(x, y)$  为像元目标,  $\text{rect}(\cdot)$  为矩形函数,表达式为

$$\text{rect}\left(\frac{x}{w}\right) = \begin{cases} 0, & |x/w| > 1/2 \\ 1/2, & |x/w| = 1/2 \\ 1, & |x/w| < 1/2 \end{cases}. \quad (5)$$

电子学 PSF 是在去除信号噪声时对像元产生的模糊作用,是一种低通滤波,表达式为

$$F_{PS_{el}} = \frac{2\nu_0}{1 + (2\pi\nu_0 x)^2} \frac{2\nu_0}{1 + (2\pi\nu_0 y)^2}, \quad (6)$$

其中  $\nu_0$  是空间频率,取值为 8 cycles/pixel。

## 3 红外遥感目标检测

本文目标检测部分是基于 Faster RCNN 网络结构,主要分为三部分:CNN 特征提取、区域候选网络(RPN)、分类和回归。针对于红外小目标特征,在 CNN 特征提取和 RPN 部分分别提出了多尺度特征传递和调节型非极大抑制方法。

### 3.1 多尺度特征传递网络结构

特征提取网络采用 ResNet101。残差神经网络(ResNet)是 2016 年提出来的网络结构,可以避免因网络结构加深引起的学习率下降的现象,提高了

网络的学习能力,是目标识别中最常用的网络结构之一。ResNet101 的具体结构见表 1,其中包括 5 个卷积模块和 1 个分类模块。conv1 是输入层为  $7 \times 7 \times 64$  的卷积,然后接着  $3 \times 3$  的池化层。conv2\_x、conv3\_x、conv4\_x、conv5\_x 分别包括 3、4、23、3 个残差块。在卷积层中间包含多个池化层,池化层的作用是对数据进行简化,提高算法效率,但是从另一方面减小了输出数据的尺寸,继而降低了图像的分辨率,对小目标的检测效果产生了负面影响。由表 1 可以看出,每个卷积层之间进行了最大池化(max-pooling),输出尺寸为原来的  $1/2$ ,在 conv5\_x 中,数据尺寸降为原图大小的  $1/32$ ,小目标特征将无法传递。卷积神经网络每一层提取的特征不同。Zeiler 等<sup>[9]</sup>对多层卷积神经网络的每层特征进行了可视化提取,得出由底层卷积层提取的特征较为简单,包括边缘、纹理等,信息量丰富,随着层级的增高,分辨率降低,感受野变大,特征变得抽象化和全局化。卷积层每层提取的目标代表了不同的感受野特征,层次越高,全局性越强,但局部细节就越弱,所以在对小目标进行检测和识别时不能只考虑高层特征。在 Faster R-CNN 中,将 conv4\_x 的输出特征作为 RPN 和 ROI pooling 的共享部分,然后将输出结果

输入到 conv5\_x 中,见图 3(a)。为了保证低层特征被充分利用,对 Faster RCNN 的残差网络部分进行了改进,将低层特征也作为 RPN 的共享特征,见图 3(b)。

表 1 ResNet101 结构

Table 1 Structure of ResNet101

Layer name	Output size	Config
Conv1	$1/2$	$7 \times 7, 64, \text{stride } 2$
$3 \times 3 \text{ max pooling, stride } 2$		
Conv2_x	$1/4$	$\begin{bmatrix} 1 \times 1 & 64 \\ 3 \times 3 & 64 \\ 1 \times 1 & 256 \end{bmatrix} \times 3$
Conv3_x	$1/8$	$\begin{bmatrix} 1 \times 1 & 128 \\ 3 \times 3 & 128 \\ 1 \times 1 & 512 \end{bmatrix} \times 4$
Conv4_x	$1/16$	$\begin{bmatrix} 1 \times 1 & 256 \\ 3 \times 3 & 256 \\ 1 \times 1 & 1024 \end{bmatrix} \times 23$
Conv5_x	$1/32$	$\begin{bmatrix} 1 \times 1 & 512 \\ 3 \times 3 & 512 \\ 1 \times 1 & 2048 \end{bmatrix} \times 3$
Classifier	$1 \times 1$	Average pooling, 1000-dFC, Softmax

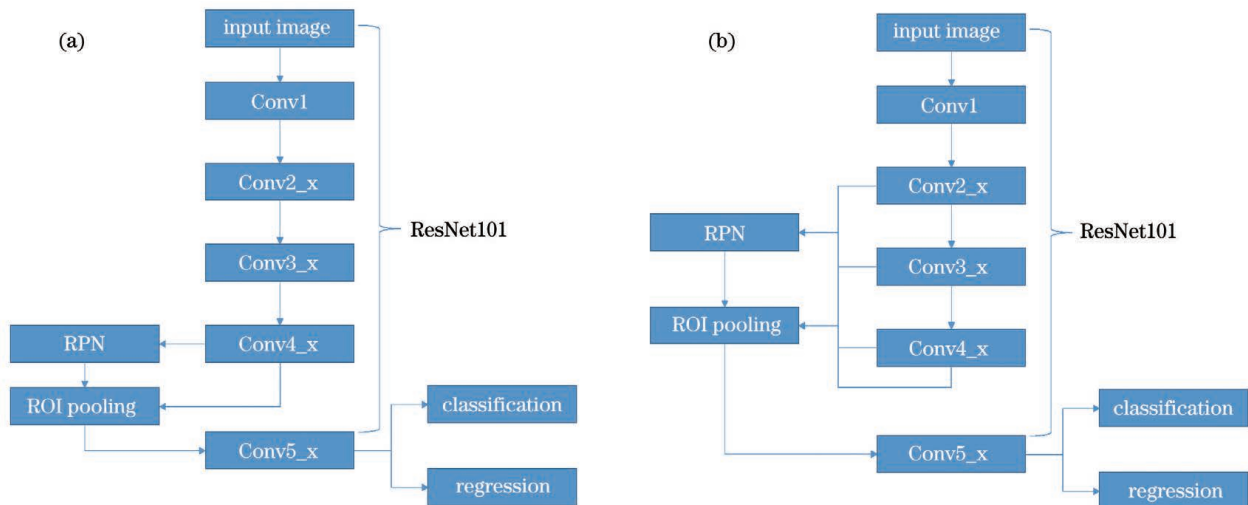


图 3 多尺度特征提取网络。(a) Faster RCNN 的特征提取网络;(b)多尺度特征提取网络

Fig. 3 Diagrams of multiscale feature extraction networks. (a) Feature extraction network in Faster RCNN; (b) multiscale feature extraction network

为了避免影响运算效率,舍弃 conv1 的初级特征,将 conv2\_x、conv3\_x、conv4\_x 与 RPN 层和 ROI Pooling 的特征共享。输出的特征层经过 conv\_5 输入到全连接层中进行分类和回归,由于在输入之前经过均值池化,因此不需要考虑每层输出的特征图的尺寸问题。在每一层的 RPN 中,锚框

尺寸会根据层级发生调整,在低级层锚框尺寸减小,利于小目标检测。将输入层后和 ROI pooling 中 max-pooling 的采样间隔由  $3 \times 3$  转换成  $5 \times 5$ ,降低了小目标被简化的概率。

### 3.2 调节型非极大抑制

非极大值抑制(NMS)方法是用来筛选超过阈

值的检测框,最终选择最合适的检测框作为输出结果。此方法会将交并比(IoU,  $\eta_{IoU}$ )较大的框进行剔除。但是在遥感图像中的目标会出现聚集排列的现象,如飞机场、码头和停车场。目标排列紧密不易区分,相邻的不同目标的检测框容易重叠,超过交并比的阈值利用 NMS 方法进行抑制,易造成目标漏检的现象。本文提出调节型非极大抑制,对交并比较大的检测框没有采取直接删除的方法,而是利用衰减函数对阈值进行调整。衰减函数的表达式为

$$s'_i = \begin{cases} s_i, \eta_{IoU}(M, b_i) < N_t \\ s_i \exp\left[-\frac{\eta_{IoU}^2(M, b_i)}{\sigma}\right], \eta_{IoU}(M, b_i) > N_t \end{cases}, \quad (7)$$

其中  $s_i$  为置信阈值,  $N_t$  为交并比参数阈值。改进后的方法有利于聚集目标的检测,降低了漏检率。从图 4 中可以看出,排列紧密的目标会出现检测框筛选错误,使得检测结果精度较低。改进调节型非极大值抑制后,检测精度明显上升。

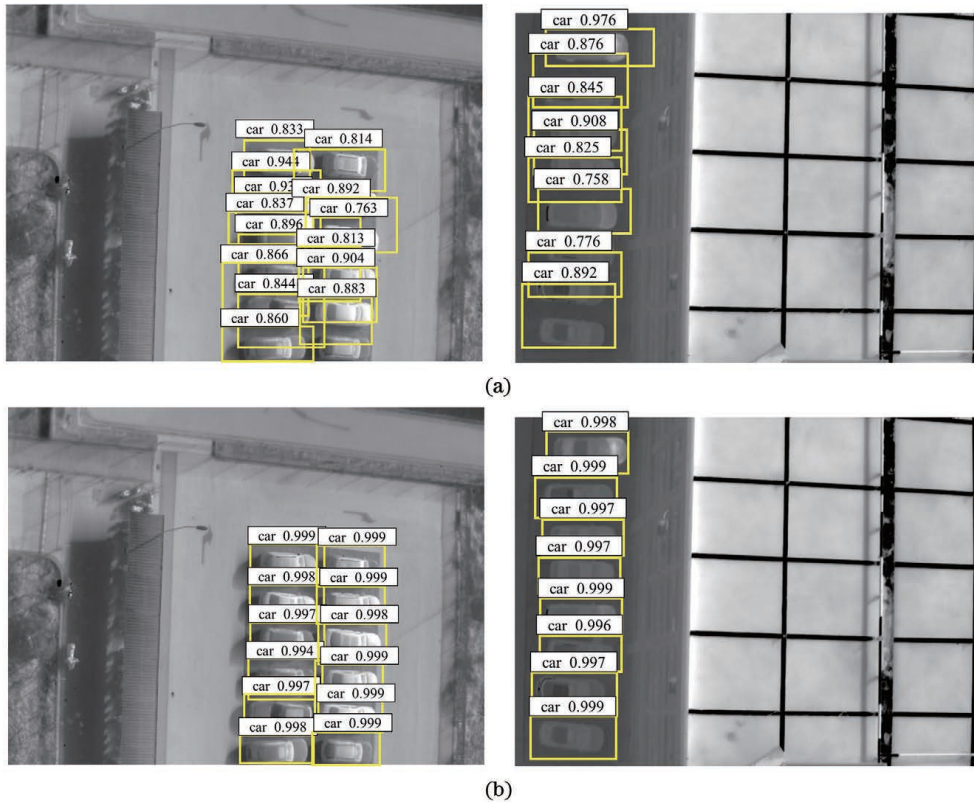


图 4 目标检测结果对比。(a)改进前;(b)改进后

Fig. 4 Comparison of target detection results. (a) Before improvement; (b) after improvement

#### 4 超分辨率目标检测系统

硬件设备限制了红外遥感图像的分辨率,目标纹理特征不明显,可利用的特征信息较少,直接利用深度卷积网络检测的效果不佳,而且目前比较常用的几种网络结构用于小目标的检测研究不多。超分辨率算法可以提高图像的分辨率,利于后续的目标检测和识别。基于此,本文提出了红外遥感图像 SROD 算法,将超分辨率网络和目标识别网络相结合,利用前者的输出作为后者输入,提高了小目标的识别率和识别精度。

SROD 算法系统主要分为两部分:预处理部分和目标检测识别部分,算法流程见图 5。预处理部

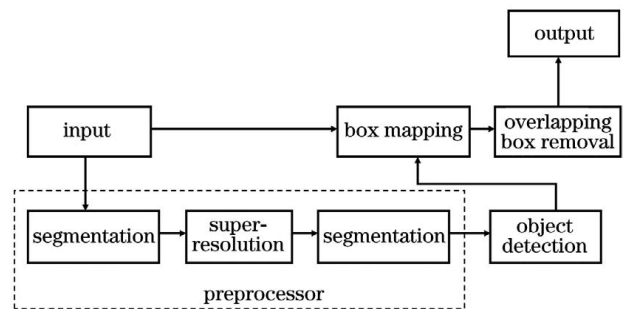


图 5 SROD 算法流程图

Fig. 5 Digram of SROD

分的目的是解决遥感图像尺寸大、目标小的“缺点”。首先将输入图像按照 15% 的重叠率等距分割并编号,然后利用改进的 WDSR 网络对图像进行重建,

然后按照超分辨率倍数对图像进行分割,即扩大两倍,图像就按照长宽的 1/2 进行分割。预处理后的图像尺寸变小,分辨率提高。检测网络利用 Fast RCNN 结构对预处理后的目标进行检测和识别,输出目标位置和类别,然后将检测框映射到原始图像的位置,避免了分割图像融合。

## 5 实验分析

### 5.1 超分辨率重建对目标检测的影响

依据宋文韬等<sup>[7]</sup>得出的结论,红外船舶目标分辨率低于 12 m 时,红外船舶的几何和辐射特性发生明显变化。在选取实验数据时,将哨兵 2 号卫星搭载的多光谱成像仪(MSI)的图像数据作为红外目标检测的数据源。MSI 的波段涉及可见光、近红外、短波红外波段,地面分辨率分别包括 10 m、20 m 和 60 m,多光谱图像幅宽为 290 km。将实验 Band 12-SWIR(中心波长为 2.190 μm,分辨率为 20 m)的图像作为船舶目标的研究图像。选用中国科学院上海技术物理研究所研制的面阵摆扫型无人机载红外光谱成像仪进行目标检测,波长范围为 7.5~14.0 μm,分辨率大小为 0.1 m。无人机数据参数见表 2。

表 2 无人机数据参数

Table 2 UAV data parameters

Parameter	Value
Array size	640×480
Wavelength range /μm	7.5~14.0
Pixel size /μm	17
Focal length /mm	35
Angular resolution /mrad	0.68
Data format	U16(Unsigned 16 bits)
Number of pictures taken when the imager sweeps a line	5

目标检测算法的评价指标采用精确率和召回率。假设图像中有真实目标数  $T$ ,检测出来的目标个数为  $N$ ,此处目标是没有核实的疑似目标。经核验,检测出的目标为真实目标的个数为  $T'$ ,漏检的目标个数  $\Delta T$  为

$$\Delta T = T - T', \quad (8)$$

则虚假目标个数为

$$\Delta N = N - T'. \quad (9)$$

通过以上数据就可以得到目标检测算法的精确率为

$$P = \frac{T'}{N}, \quad (10)$$

召回率(Recall)计算公式为

$$R = \frac{T'}{T}. \quad (11)$$

为了验证分辨率对目标检测精度的影响,本研究首先利用自然图像数据集来测试分辨率和检测准确度的关系。测试数据为 Pascal VOC2007,共有 4952 张图像,其中包含 20 种目标 12032 个。将测试数据下采样分辨率降为原来的 1/2 作为降采样测试集,利用 Faster RCNN<sup>[10-12]</sup>进行目标检测,得到 20 种目标的检测结果如图 6 所示,从图中可以看出平均精度由 75.8% 降为 65.5%,这说明分辨率降低使得检测精度明显降低。

利用 VOC2007 直接训练的 Faster RCNN 进行检测,并不能检测出小的船舶目标,见图 7(a)。利用红外遥感图像数据集训练的 Faster RCNN 网络进行检测,只能够检测出较大的船舶,见图 7(b)。利用 WDSR 进行超分辨率重建之后再行目标检测,检测效果得到明显提升,见图 7(c)。

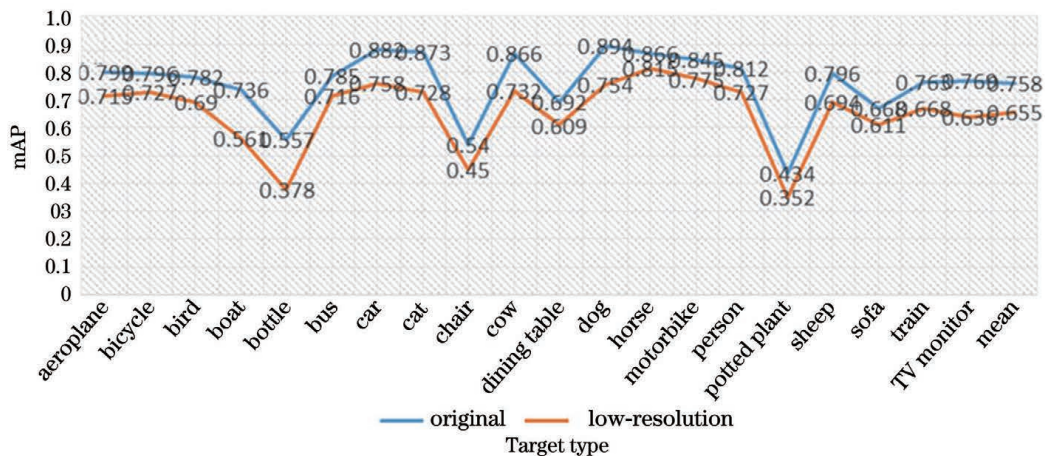


图 6 分辨率对数据集识别精度的影响

Fig. 6 Impact of resolution on the accuracy of dataset recognition

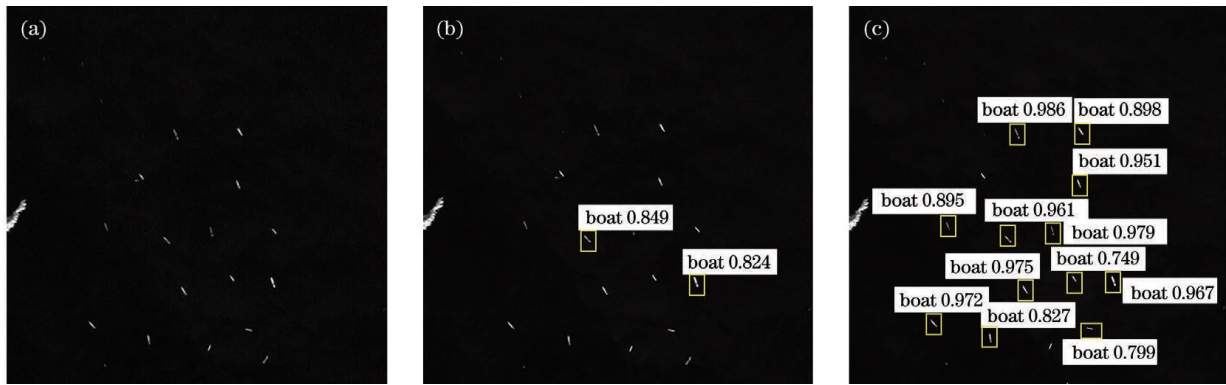


图 7 检测结果对比。(a)直接检测结果;(b)利用红外数据训练;(c)超分辨率重建后

Fig. 7 Comparison of results. (a) Direct detection results; (b) training with infrared data; (c) after super-resolution reconstruction

为了验证超分辨率重建算法对目标检测结果的影响,利用不同算法处理的图像进行检测试验,见图 8。算法包括:双三次插值、稀疏编码(ScSR)、SRCNN、WDSR。对比结果可以看出由 WDSR 算

法重建的目标清晰度最高,得到的检测结果也相对较好。实验结果证明,超分辨率重建的精度影响目标检测的效果,重建精度越高,检测结果越好。

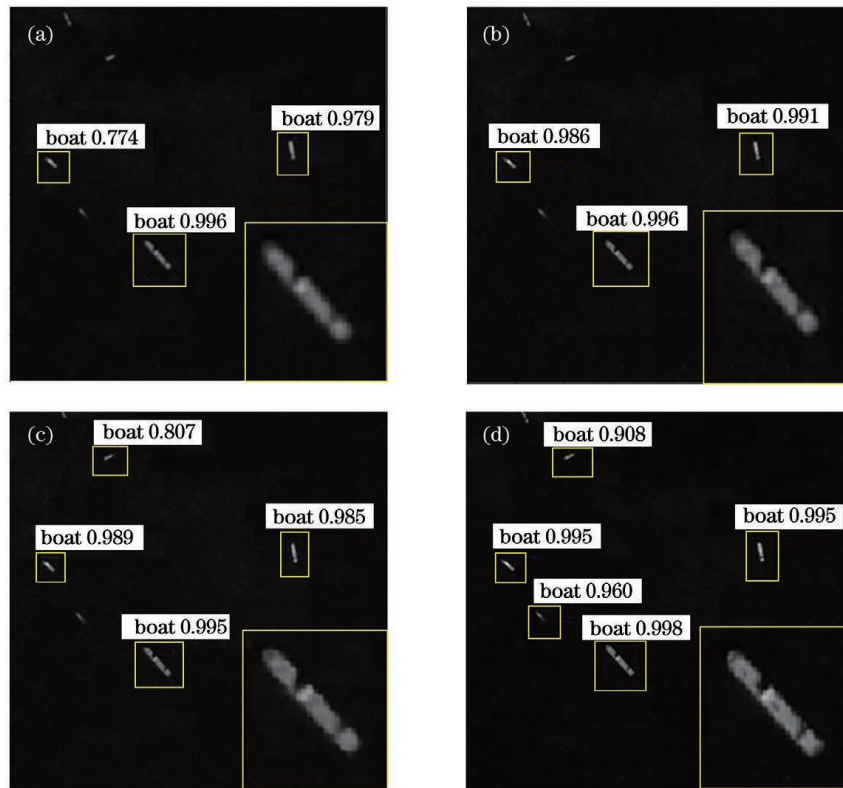


图 8 重建后的图像检测结果。(a)双三次插值;(b)稀疏字典;(c)SRCNN;(d)WDSR

Fig. 8 Detection results of super-resolution reconstructed image. (a) Bicubic; (b) ScSR; (c) SRCNN; (d) WDSR

### 5.2 SROD 算法检测红外遥感目标实验

利用整幅的哨兵 2 号卫星搭载的 MSI 短波红外(第 11 波段,中心波长为  $1.610 \mu\text{m}$ ,分辨率 20 m)遥感图像作为实验数据,拍摄时间为 2015-12-29T02:58,图像大小为  $5490 \text{ pixel} \times 5490 \text{ pixel}$ ,图 9 为整幅图像的 SROD 算法检测结果,被框出的

区域为对比研究区域。利用 Itti 显著度算法、红外目标训练的 Faster RCNN 算法和 SROD 算法分别对图像中的目标进行检测。利用人工标注的方式得到图像中共有真实目标数  $T=248$ ,检测结果见表 3。

选取整幅图中的局部进行对比。从图 10 对比可以看出,图 10(b)为显著度分割的检测结果,显著

表 3 检测结果

Table 3 Detection results

Algorithm	$T'$	$N$	$\Delta T$	$\Delta N$	$P / \%$	$R / \%$
Saliency algorithm	196	289	52	93	67.80	79.03
Faster RCNN	180	214	68	34	84.11	72.58
SROD	202	228	46	26	88.59	81.45

度算法基于异常检测的原理,可以将区域内的异常目标检测出来,但是无法区分形状和大小与船舶类似的异常目标,会将一部分小岛屿和云当作目标被检测出来,虚警率较高。图 10(d)中标注出了 Faster RCNN 的两个错检(将云和岛屿当作船只目标)和两个漏检。从图 10(c)可以看出利用 SROD 的检测准确度和检出度都最高,检测效果最好。

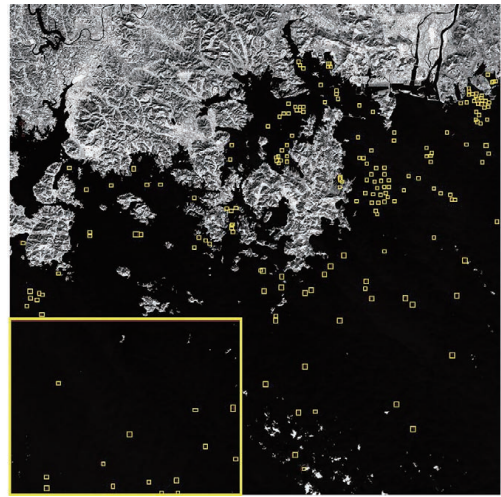


图 9 SROD 对整幅图像船舶检测结果

Fig. 9 SROD ship detection results of the entire image

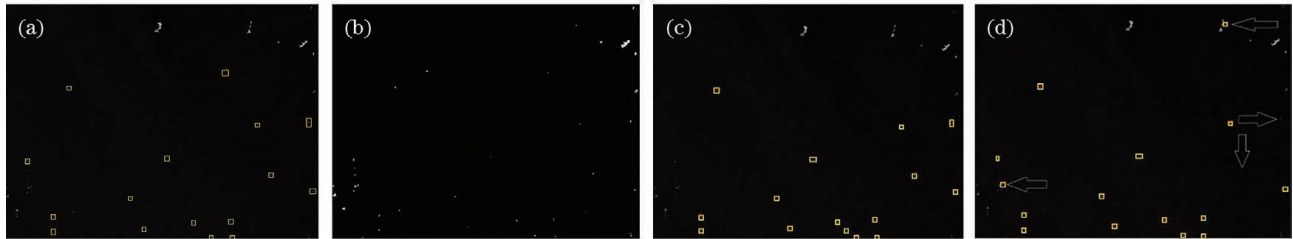


图 10 不同方法的检测结果对比。(a)真目标;(b)显著度分割的检测结果;(c)SROD 检测结果;(d)Faster RCNN 的检测结果

Fig. 10 Comparison of detection results of different methods. (a) Real target; (b) detection result of saliency segmentation; (c) detection result of SROD; (d) detection result of Faster RCNN

将检测出来的目标进行统计,利用目标的像元个数作为分类标准,将船舶分为大(大于 50 pixel),中(20~50 pixel),小(小于 20 pixel)三种(见图 11),真实的目标个数分别是 49,61,138。对 Faster RCNN 和 SROD 检测出的目标进行统计,结果见表 4。

表 4 目标类型统计

Table 4 Target type statistics

Algorithm	Parameter	Number of detected targets			Total
		Large target	Medium target	Small target	
SROD	$T'$	48	58	96	202
	$\Delta T$	1	3	42	46
	$\Delta N$	10	8	8	26
	$N$	58	66	104	228
	$P / \%$	82.75	87.87	92.30	88.59
	$R / \%$	97.91	95.08	69.56	81.45
Faster RCNN	$T'$	43	51	86	180
	$\Delta T$	6	8	54	68
	$\Delta N$	19	9	6	34
	$N$	62	60	92	214
	$P / \%$	69.35	85	93.47	84.11
	$R / \%$	87.75	86.44	61.42	72.58

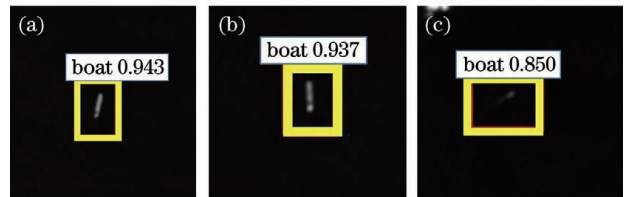


图 11 不同大小的船舶目标。(a)大型;(b)中型;(c)小型

Fig. 11 Ship targets of different sizes. (a) Large; (b) medium; (c) small

分析表 4 中数据可知:从召回率来看,大型船舶的值最高,而小型船舶的值最低,说明大型目标容易被检测出来,小型的船只漏检率较高。目标的大小与分辨率相关,分辨率越高,目标识别的召回率就越高。从准确率来看,大型船舶的准确率较低,说明在检测识别中出现了较多与大型船舶相似的虚假目标。这是由海洋的特殊环境造成的,岛屿与大型船舶的形态特征相类似,容易被当作目标。比较两种检测方法,可以看出,SROD 对小目标识别的召回率比 Faster RCNN 高出 13.25%,这说明 SROD 算法对小目标(小于 20 pixel)的检测效果有明显提升。



## 6 结 论

本文围绕红外遥感目标检测与识别算法展开研究,针对红外小目标的检测难点提出了解决方案,并基于解决方案和深度学习算法提出了 SROD 算法。算法主要分为两部分,第一部分采用“分割-超分辨率-分割”的结构,将分割后的图像进行超分辨率重建,然后再进行分割后输出。第二部分是目标检测模块,输入分割后的超分辨率图像,这一部分改进了 Faster RCNN,利用多尺度特征提取和可调型非极大值抑制方法构建了利于检测小目标的深度学习网络。

对比双三次插值、ScSR、SRCNN、WDSR 算法,其中 WDSR 算法的超分辨率重建效果最好,将重建后的图像用于目标检测,结果证实 WDSR 重建后目标检测的准确率也最高。结果证实 SROD 算法优于 Faster RCNN 算法和 Itti 显著度分割算法,准确率相比于 Faster RCNN 算法提升了 5.33%,召回率提升了 12.22%。SROD 对小目标(小于 20 pixel)检测的召回率相比于 Faster R-CNN 高出 13.25%,说明 SROD 算法对小目标的检测效果有明显提升。

### 参 考 文 献

- [1] Song W T, Hu Y, Kuang D B, et al. Detection of ship targets based on CFAR-DCRF in single infrared remote sensing images [J]. *Journal of Infrared and Millimeter Waves*, 2019, 38(4): 520-527.  
宋文韬, 胡勇, 匡定波, 等. 基于 CFAR-DCRF 红外遥感舰船单帧目标检测方法 [J]. *红外与毫米波学报*, 2019, 38(4): 520-527.
- [2] Huang S, Hu Y, Gong C L, et al. Saliency region super-resolution reconstruction algorithm for infrared images based on sparse coding [J]. *Journal of Infrared and Millimeter Waves*, 2020, 39(3): 388-395.  
黄硕, 胡勇, 巩彩兰, 等. 基于稀疏编码的红外显著区域超分重建算法 [J]. *红外与毫米波学报*, 2020, 39(3): 388-395.
- [3] Dong C, Loy C C, He K M, et al. Image super-resolution using deep convolutional networks [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2016, 38(2): 295-307.
- [4] Yu J H, Fan Y C, Yang J C, et al. Wide activation for efficient and accurate image super-resolution [EB/OL]. (2018-08-27)[2020-11-01]. <https://arxiv.org/abs/1808.08718v1>.
- [5] Bai Y C, Zhang Y Q, Ding M L, et al. SOD-MTGAN: small object detection via multi-task generative adversarial network [M] // Ferrari V, Hebert M, Sminchisescu C, et al. *Computer vision-ECCV 2018. Lecture notes in computer science*. Cham: Springer, 2018, 11217: 210-226.
- [6] Koester E, Sahin C S. A comparison of super-resolution and nearest neighbors interpolation applied to object detection on satellite data [EB/OL]. (2019-07-08)[2020-11-01]. <https://arxiv.org/abs/1907.05283>.
- [7] Song W T, Hu Y, Gong C L, et al. Spatial characteristics analysis of multi-scale ship target in scanning detection system [J]. *Acta Optica Sinica*, 2019, 39(7): 0728010.  
宋文韬, 胡勇, 巩彩兰, 等. 扫描探测系统多尺度舰船目标的空间特性分析 [J]. *光学学报*, 2019, 39(7): 0728010.
- [8] Lim B, Son S, Kim H, et al. Enhanced deep residual networks for single image super-resolution [C] // 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 1132-1140.
- [9] Zeiler M D, Fergus R. Visualizing and understanding convolutional networks [M] // Fleet D, Pajdla T, Schiele B, et al. *Computer vision-ECCV 2014. Lecture notes in computer science*. Cham: Springer, 2014, 8689: 818-833.
- [10] Girshick R, Donahue J, Darrell T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation [C] // 2014 IEEE Conference on Computer Vision and Pattern Recognition, June 23-28, 2014, Columbus, OH, USA. New York: IEEE Press, 2014: 580-587.
- [11] Girshick R. Fast R-CNN [C] // 2015 IEEE International Conference on Computer Vision (ICCV), December 7-13, 2015, Santiago, Chile. New York: IEEE Press, 2015: 1440-1448.
- [12] Ren S Q, He K M, Girshick R, et al. Faster R-CNN: towards real-time object detection with region proposal networks [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 39(6): 1137-1149.