

引用格式: SHI Wen-xu, JIANG Jin-hong, BAO Sheng-li. Ship Detection Method in Remote Sensing Image Based on Feature Fusion[J]. *Acta Photonica Sinica*, 2020, 49(7):0710004

史文旭,江金洪,鲍胜利. 基于特征融合的遥感图像舰船目标检测方法[J]. 光子学报, 2020, 49(7):0710004

基于特征融合的遥感图像舰船目标检测方法

史文旭^{1,2}, 江金洪^{1,2}, 鲍胜利^{1,2}

(1 中国科学院成都计算机应用研究所, 成都 610081)

(2 中国科学院大学, 北京 100049)

摘 要: 针对常用的目标检测算法对遥感图像中的舰船目标进行检测时存在检测精度与实时性兼顾不佳的问题, 提出了基于特征融合的遥感图像舰船目标检测算法来检测复杂场景下的多尺度舰船目标. 该算法以多尺度单发射击检测框架为基础, 增加反卷积特征融合模块和池化特征融合模块, 增强网络特征提取的能力. 同时设计聚焦分类损失函数来解决训练过程中正负样本失衡的问题. 在高分遥感舰船目标数据集上的实验结果表明, 所提方法能够有效地增强复杂场景下舰船目标的检测精度. 此外, 该算法对遥感图像中的模糊舰船目标的检测效果也优于多尺度单发射击检测框架.

关键词: 遥感图像; 舰船目标检测; 神经网络; 复杂场景; 深度学习

中图分类号: TP753

文献标识码: A

doi: 10.3788/gzxb20204907.0710004

Ship Detection Method in Remote Sensing Image Based on Feature Fusion

SHI Wen-xu^{1,2}, JIANG Jin-hong^{1,2}, BAO Sheng-li^{1,2}

(1 Chengdu Institute of Computer Application, Chinese Academy of Sciences, Chengdu 610081, China)

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

Abstract: Aiming at the problem of low accuracy and poor real-time performance of commonly used target detection algorithms, a novel ship target detection algorithm based on convolutional neural network with feature fusions is proposed to detect multi-scale ship targets in complex scenes. The proposed method inherits the network structure of SSD and introduces the deconvolution feature fusion module and the pooling feature fusion module into it to generate the new feature maps with richer semantic information for both ship classification and boxes regression. In addition, we used a focal classification loss function in the training strategy to deal with the imbalanced difficult and easy samples in the training process. The experiments tested on the ship detection dataset demonstrate that the proposed method shows a better adaptability to ship detection of different sizes in complex scenes. On the extended experiment, the proposed method performance over SSD in blurry object detection.

Key words: Remote sensing imaging; Ship detection; Neural network; Complex background; Deep learning

OCIS Codes: 100.4996; 100.2000; 150.0155

基金项目: 国家自然科学基金(No.11871059), 四川省新一代人工智能重大专项(No.2018GZDZX0036), 四川省科技厅重点研发项目(No.2018SZ0040)

第一作者: 史文旭(1995—), 男, 硕士研究生, 主要研究方向为深度学习、目标检测. Email: shiwenxu1995@126.com

导师: 鲍胜利(1973—), 男, 正研级高级工程师, 博士, 主要研究方向为深度学习、智能信息处理. Email: baohigh@casit.com.cn

收稿日期: 2019-10-28; 录用日期: 2020-02-12

<http://www.photon.ac.cn>

0 引言

遥感图像是指利用高分辨率遥感卫星对地面目标进行探测而生成的图像,其数据采集具有其独特的优势.近年来,遥感图像广泛的应用于各个领域^[1],吉林1号和高分系列遥感卫星的发射更是为高分遥感图像的获取提供了广泛的数据来源.作为重要的军事目标和海洋运输载体,高分遥感图像中的舰船目标检测在军用和民用领域都发挥着重要的作用,因此遥感图像中的舰船目标检测具有重要的研究意义和应用价值^[2].

在深度学习技术迅猛发展之前,传统的遥感图像舰船目标检测方法有基于模板匹配^[3-4]、知识引导^[5-6]以及监督分类^[7-8]等方法.但由于遥感图像受光照、云层、大气中颗粒等因素影响较大,采用传统的算法对遥感图像舰船目标进行检测不仅需要预先设定条件而且需要掌握一定的遥感图像的专业知识,此外传统的目标检测算法鲁棒性较差,检测效率低,难以满足大规模自动化作业的需求.

近年来,随着KRIZHEVSKY A等^[9]提出的AlexNet网络在2012年的ImageNet图像分类比赛中获得成功,深度学习便得到了迅猛的发展,尤其是卷积神经网络(Convolutional Neural Networks, CNN)凭借其参数共享的优势,在目标检测领域得到了广泛的应用^[10].基于CNN的目标检测算法按照提取特征层数量可以分为以Fast R-CNN^[11], Faster R-CNN^[12], 统一实时目标检测算法(You Only Look Once, YOLO)^[13], YOLO v2^[14], 为代表的单层检测算法和以YOLO v3^[15], 多尺度单发射击检测算法(Single Shot MultiBox Detector, SSD)^[16], 反卷积多尺度单发射击检测算法(Deconvolutional Single Shot Detector, DSSD)^[17], 特征融合的多尺度单发射击检测算法(Feature Fused Single Shot Detector, FSSD)^[18]为代表的多层检测算法.相比于单层目标检测算法,多层目标检测算法合理地利用了卷积网络中不同尺度的特征,提高了目标检测的精度.鉴于深度学习对自然场景下的目标检测取得了巨大成功,不少学者开始利用深度学习对遥感图像进行目标识别检测.文献[19]提出一种基于Faster R-CNN^[12]网络的遥感图像快速目标检测方法,实现了对海上舰船和飞机目标的检测;文献[20]基于Faster R-CNN^[12]引入特征聚合操作,提高了对光学遥感图像中舰船目标的准确率;文献[21]利用回归型深度卷积神经网络以抑制扫描图像杂波背景,该网络不含池化层,输出的背景抑制图像尺寸与输入图像一致;其次,对抑制后的图像进行门限检测,提取候选目标小区域原始数据;最后,将候选目标区域数据依次输入分类型深度卷积神经网络以进一步判别目标、剔除虚警,能够识别遥感图像中的目标;文献[22]在SSD算法的基础之上,结合特征金字塔(Feature Pyramid Networks, FPN)结构^[23]的优势,提出了一种基于多尺度特征融合的遥感图像检测算法,提高了遥感图像目标的检测精度.虽然利用深度学习的方法一定程度上可以提高对遥感图像舰船目标检测的精度和效率,但仍未达到预期效果,主要由于两个因素:一是目前应用于遥感图像舰船目标检测的深度学习方法难以适应遥感图像中的舰船目标尺寸的多样性,尤其是对较小的舰船目标,主要原因是网络提取过程中浅层网络提取的特征分辨率较高,可以实现精确的定位,较高层网络提取的特征具有更多的语义信息,可以实现精确的分类,而现有的网络特征关联程度较低,忽略了上下文特征的关联性;二是遥感图像场景复杂多变且容易受到光照、云层、大气中颗粒等因素影响,斑点噪声严重,在深度学习训练的过程中会因正负样本特征的分布不平衡导致漏检错检等情况.

为解决高分遥感图像舰船目标检测所面临的问题,本文在SSD算法的基础上,提出了一种基于特征融合的多尺度舰船目标检测算法(Multi-Scale Ship Detection Networks, MSSDN).1)设计反卷积特征融合模块,通过反卷积操作使得浅层网络具备更多的深层特征的语义信息,增强网络的特征表达能力.2)设计池化特征融合模块,通过池化操作使得深层网络具备更多的浅层网络的细节信息,增强网络的细节表达能力.3)设计新的训练策略,在损失函数中增加聚焦分类损失,合理权衡正负样本.

1 SSD算法原理

SSD是由LIU W等^[16]提出的一种快速高效的检测算法,其以VGG16作为特征提取的骨干网络,在此基础上添加辅助卷积和池化层等结构对特征进行分类和定位,框架结构如图1.该算法得到由Conv4_3、FC7、Conv8_2、Conv9_2、Conv10_2、Conv11_2组成的多尺度特征层组,浅层网络的特征图尺寸较大含有较多细节信息,主要用于检测相对较小的目标,而尺寸较小的深层网络提取的特征图含有较多的抽象的语义信息,主要负责大目标检测.针对单张输入图像,SSD算法产生多个长宽比不同的候选框和框中目标的类别

得分,然后进行非极大值抑制(Non-Maximum Suppression, NMS)操作,得到最终的预测结果.

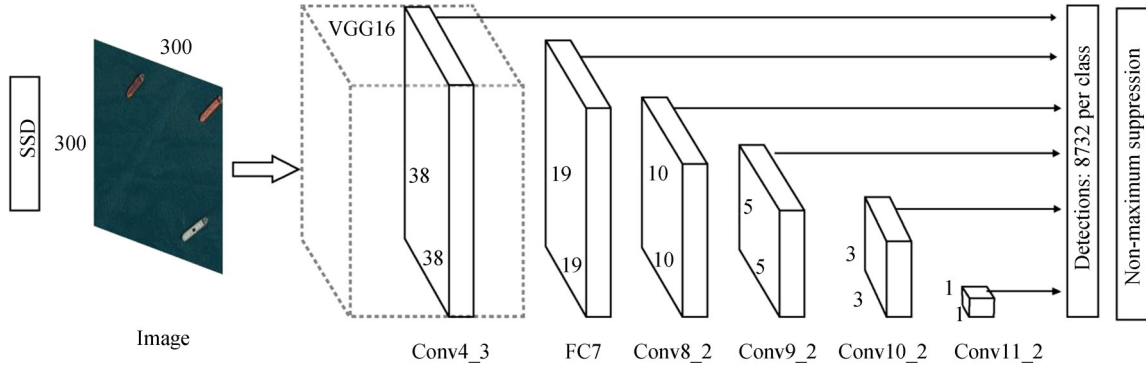


图1 SSD算法框架

Fig. 1 SSD overall framework

SSD算法的多尺度特征层组中的每一层特征都会与一组不同尺度的候选框相绑定,在每个特征点上,预测结果为相对于候选框的位置偏移和类别得分.图2为SSD算法的检测结构示例,其中图2(a)为含有真实坐标框的输入图像,图2(b)~(d)分别为在进行检测过程中 conv4_3层,FC7层和 Conv8_2层特征图匹配示例.在进行卷积操作时,特征图上的每个点都需要产生的 k 个不同长宽比的候选框(如图2(d)为 Conv8_2层提取的特征图上的每个点产生4个长宽比不同的候选框),并计算与真实框的坐标位置偏移量和 p 个类别的预测得分,分别用 $L_{loc}(\cdot)$ 和 $L_{conf}(\cdot)$ 表示.

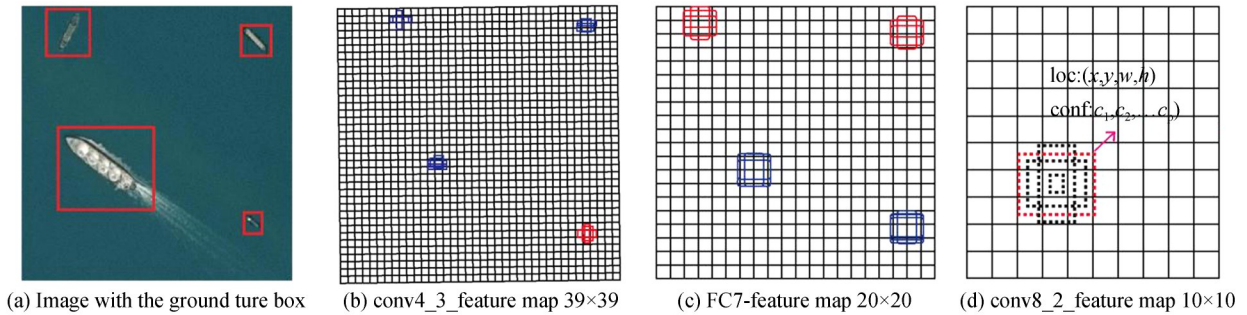


图2 SSD算法的检测结构

Fig. 2 SSD framework for detection

SSD算法将位置偏移损失和分类损失的加权和作为其损失函数,具体可表示为

$$L(x, c, l, g) = \frac{1}{N} [L_{conf}(x, c) + aL_{loc}(x, l, g)] \quad (1)$$

式中, N 代表与真实框相匹配的候选框的个数, $L_{conf}(x, c)$ 为分类损失, $L_{loc}(x, l, g)$ 为位置损失, a 为两者权重, x 输入图像, c 为目标类别, l 为预测边框, g 为真实边框.

2 MSSDN 舰船检测算法

2.1 MSSDN 网络结构

SSD算法虽然采用多尺度特征层组对目标进行分类和定位,但各层提取的不同尺度的特征之间缺少上下文信息交流,这使得SSD算法对遥感图像中的舰船目标进行检测时浅层网络因缺乏较为抽象的语义信息而导致出现小目标物体漏检问题和深层网络因含有较少的细节特征而导致出现重复检测的问题.

为解决这些问题,实现良好的检测性能,本文对SSD算法进行改进,提出了一种基于特征融合的多尺度舰船目标检测算法(Multi-Scale Ship Detection Networks, MSSDN),其网络结构如图3(a).主要对SSD算法做以下两个方面的改进:一是分别设计池化特征融合模块(Pool Feature Fusion Module, PFF)和反卷积特征融合模块(Deconvolution Feature Fusion Module, DFF)来提高各层特征之间的关联,二是在损失函数

中加入聚焦分类损失,改进训练策略,减少因正负样本不平衡对舰船检测的影响.

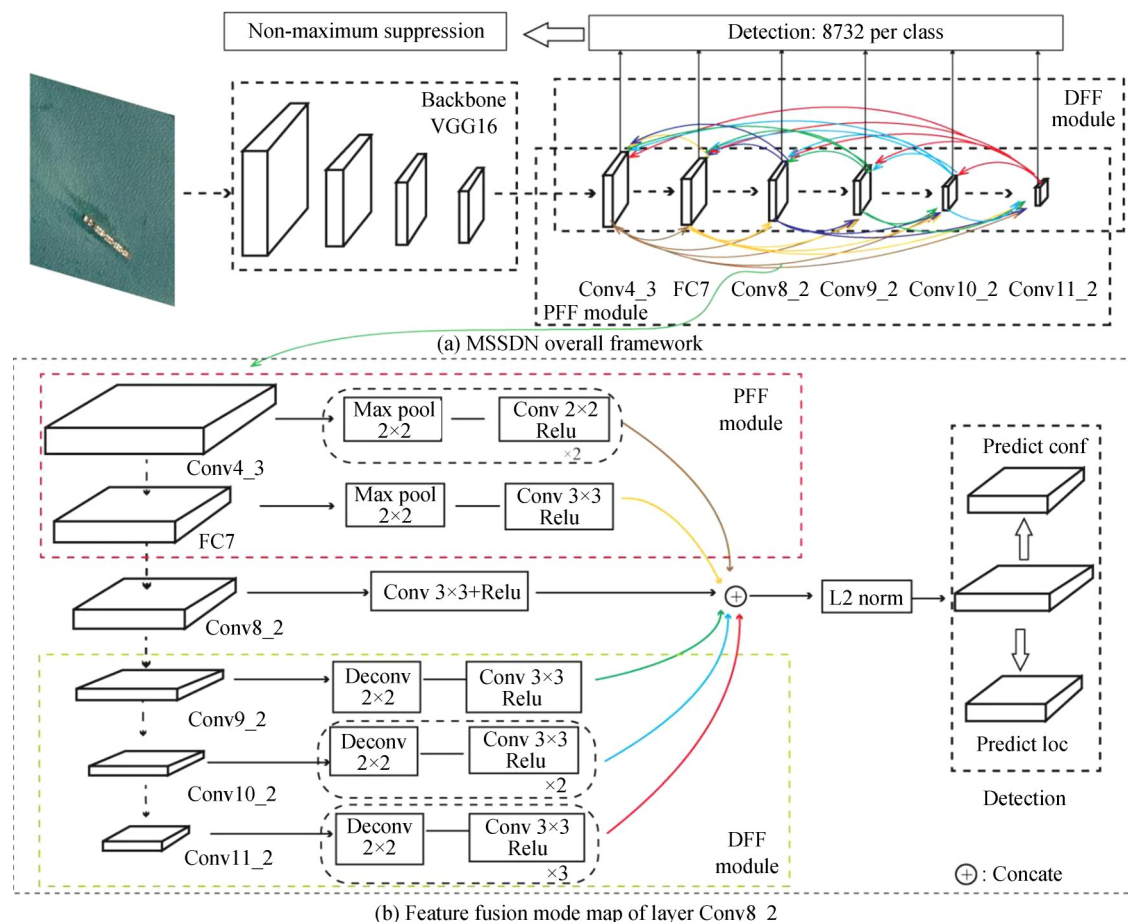


图3 MSSDN算法框架和对Conv8_2层特征融合流程图
Fig.3 MSSDN overall framework and the feature fusion mode map of layer Conv8_2

MSSDN网络结构如图3(a),主要由SSD算法主干网络、DFF module、PFF module三个部分组成,共提取出6组不同尺度的特征图.图3(b)为MSSDN算法中的DFF module和PFF module对SSD主干网络中Conv8_2层进行特征融合操作的流程示意图,首先使用DFF module模块中的Max Pool 2×2、3×3卷积和Relu激活函数操作对相对较浅的Conv4_3、FC7层特征进行处理,在保留浅层网络显著的细节特征、降低特征维度的同时使浅层网络特征学习到更多非线性关系,与Conv8_2层融合后使得Conv8_2层获取到更多的Conv4_3、FC7层网络的边缘细节信息;然后使用PFF module模块中的Deconv 2×2、3×3卷积和Relu激活函数操作对相对较深的Conv9_2、Conv10_2、Conv11_2层进行处理,在填充特征内容、提取敏感特征信息的同时使深层网络特征学习到更多非线性关系,与Conv8_2层融合后,使得Conv8_2层具备更多Conv9_2、Conv10_2、Conv11_2层网络的特征信息.图3(b)中的×2,×3分别代表将相应的操作进行2次或3次,最后利用Concat操作将上述提取特征进行融合并做L2 norm归一化操作,组成新的Conv8_2层特征.同理,利用DFF module和PFF module对SSD算法主干网络中的其它层做相同的操作,具体可表示为

$$\begin{cases} X_i = w(x_i) \\ [X_p] = \Pi_{i-p}\{w[\partial(x_p)]\} \\ [X_d] = \Pi_{d-i}\{w[\ell(x_d)]\} \\ Y_i = L_2[C_{con}(X_p, X_i, X_d)] \end{cases} \quad 1 \leq i \leq 6, 1 \leq p < i, i < d \leq 6 \quad (2)$$

式中, x_i 表示SSD算法主干网络中的第*i*层特征, $w(\cdot)$ 表示进行3×3卷积和Relu激活操作, $\partial(\cdot)$ 表示进行下采样操作,采用的是最大池化操作, $\Pi_{i-p}(\cdot)$ 表示进行*(i-p)*次下采样操作, $[X_p]$ 表示下采样特征集合, $\ell(\cdot)$ 表示上采样操作,采用的是最近邻插值法, $\Pi_{d-i}(\cdot)$ 表示进行*(d-i)*次上采样操作, $[X_d]$ 表示上采样特征集合,

$C_{\text{con}}(\cdot)$ 表示进行Concat特征融合操作, $L_2(\cdot)$ 表示进行L2 norm 归一化操作, Y_i 表示对原始网络中 x_i 层特征进行上下文特征融合后生成的特征图.

相比于SSD算法,MSSDN算法增强了各层之间的关联,其利用DFE module使深层网络特征具备更多浅层特征的细节信息,可弥补SSD算法因深层网络细节信息丢失导致目标定位不准确的问题.另一方面,MSSDN算法通过PFE module使浅层网络特征具备更多的深层特征的语义信息,弥补SSD算法因浅层网络语义信息较差导致小目标物体分类失败的问题.

图4为本文所提MSSDN算法与SSD算法在Conv4_3、FC7、Conv8_2、Conv9_2层输出的特征对比图,图4(b)~(e)为SSD算法输出的特征图,图4(f)~(i)为MSSDN算法输出的特征图.对于较浅层网络,显然图4(b),图4(c)的特征图相对于图4(f),图4(g)中的细节信息更为丰富,较小目标的舰船特征仍高亮显示.对于较深层网络,由于特征分辨率降低,舰船特征信息降低,但图4(h),图4(i)的特征图相对于图4(d),图4(e)中较大舰船目标的特征信息更为丰富.理论上,在图4(f)~(i)上的预测效果会优于图4(b)~(e),尤其是大多数的小目标物体都由浅层网络识别,因此可以提升对遥感图像中小目标舰船的检测精度.

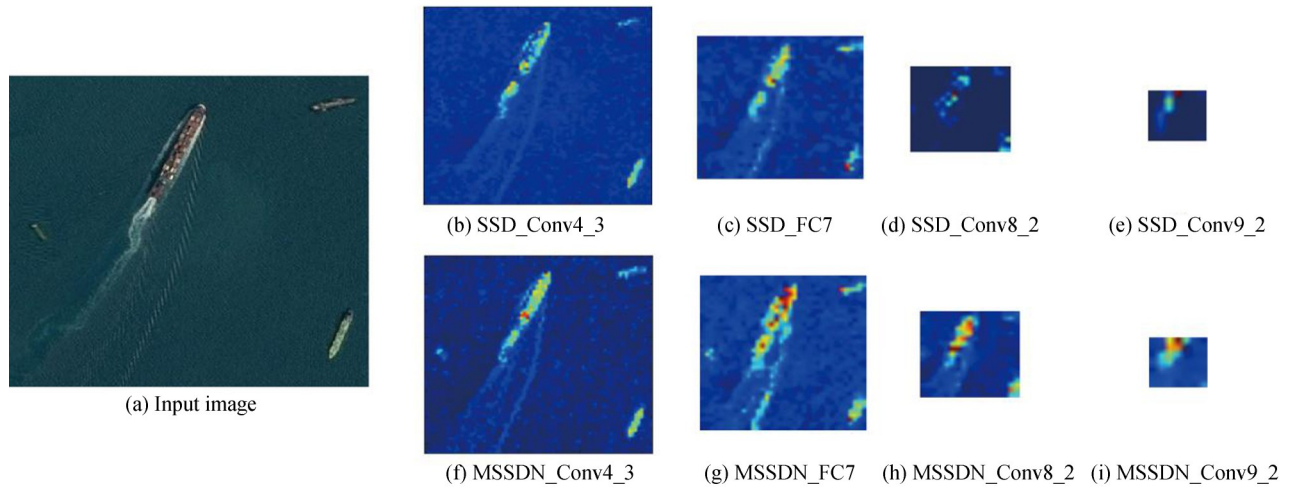


图4 特征融合前后特征图对比

Fig. 4 Comparison of feature maps before and after feature fusion

2.2 候选框匹配和损失函数设计

为实现对遥感图像中不同尺度的舰船目标进行检测,设计不同纵横比的候选框进行匹配以适应不同尺度大小的遥感图像目标,根据算法1得到 p 个卷积层所得到的候选框 $D=\{d_1, d_2, \dots, d_n\}$ 其中 d_i 由 (c_x, c_y, w, h) 4个坐标值组成, (c_x, c_y) 为中心点坐标, w, h 分别为候选框的宽和高.将上述得到的候选框与真实的标签框按照算法2进行匹配,得到候选框的坐标及其对应的目标类别,具体可表示为 $\hat{Y} \in \mathbb{R}^{n \times (l+4)}$, l 表示类别种类,为简单标记可以令 $\hat{Y}=[\hat{Z}, \hat{B}]$,其中 $\hat{Z}=\mathbb{R}^{n \times l}$ 表示预测类别向量集合,其中 $\hat{B}=\mathbb{R}^{n \times 4}$ 表示预测坐标向量集合,当候选框与真实的标签框匹配大于阈值时将其记为正例样本用 P_{os} 表示,当候选框与真实的标签框匹配小于阈值时将其记为负例样本用 N_{eg} 表示.

此外,为避免因正负样本分布不均导致模型训练退化的问题,MSSDN算法引入聚焦分类损失函数(focal classification loss)^[24],对模型训练进行监督,可表示为

$$L(\hat{Y}, Y) = \frac{1}{N} [L_{\text{class_fl}}(\hat{Z}, Z) + \alpha L_{\text{loc}}(\hat{B}, B)] \quad (3)$$

式中, N 为在算法2中与真实框正匹配的候选框个数, $L_{\text{loc}}(\hat{B}, B)$ 为MSSDN算法的定位损失函数,其借鉴Faster R-CNN的位置回归函数即Smooth L1函数,可表示为

$$L_{\text{loc}}(\hat{B}, B) = \sum_{i \in P_{\text{os}}} \sum_{m \in \{c_x, c_y, w, h\}} s_{\text{smooth_L1}}(\hat{B}_i - B_i) \quad (4)$$

$$s_{\text{smooth_L1}}(x) = \begin{cases} 0.5x^2 & \text{if } |x| < 1 \\ |x| - 0.5 & \text{otherwise} \end{cases} \quad (5)$$

式中, $L_{\text{class_fl}}(\hat{Z}, Z)$ 为 MSSDN 算法的分类损失函数, 采用交叉熵计算损失, 可表示为

$$L_{\text{class_fl}}(\hat{Z}, Z) = - \sum_{c=1}^l \sum_{i \in P_{\text{os}}}^N a_i (1 - Z_{i,c})^r \log(\hat{Z}_{i,c}) - \sum_{c=0}^l \sum_{j \in N_{\text{eg}}} (1 - a_i) \log(\hat{Z}_{j,c}) \quad (6)$$

式中, $\hat{Z}_{j,c}$ 表示正确且类别为背景预测框的概率, a_i 和 r 为超参数, 且 $a_i \in [0, 1]$, $r \in [0, 5]$, 当 $r > 0$ 时意味着正样本的损失相对减少, 模型将更聚焦于负样本的训练. 因此加入聚焦分类损失函数有效地解决正负样本分布不均衡的问题, 提高了模型的优化效率.

Algorithm 1: Initialize set of default box D

Inputs:
 p : Number of feature maps
 $|f_k|$: Dimension of k^{th} square feature map
 S_{min} : Minimum scale value-default 0.2
 S_{max} : Maximum scale value-default 0.95
Initialize:
 $D = \{\}$: Set of default boxes
Process:
For $k \in [1, p]$ **do:**
 Set of default boxes in k^{th} feature map
 $\forall i, j \in [0, |f_k|]: C_k = \left(\frac{i+0.5}{|f_k|}, \frac{j+0.5}{|f_k|} \right)$
 Scale values of default boxes in k^{th} feature map
 $s_k = s_{\text{min}} + \frac{s_{\text{max}} - s_{\text{min}}}{p-1} (k-1)$, $s_k' = \sqrt{s_k \cdot s_{k+1}}$
 For $\forall (c_x, c_y) \in C_k, \forall a \in \{1, 2, 3, 1/2, 1/3\}$
 $D = D \cup \{[c_x, c_y, s_k \sqrt{a}, s_k' / \sqrt{a}]\}$
 End for
 $D = D \cup \{[c_x, c_y, s_k, s_k'], [c_x, c_y, s_k', s_k]\}$
End for
Output
 D : Set of default boxes

Algorithm 2: Create ground truth matrix Y

Inputs:
 τ : Overlap threshold-default 0.5
 $D = \{d_1, d_2, \dots, d_n\}$ Set of default boxes
 $G = \{g_1, g_2, \dots, g_h\}$ Set of ground truth boxes
Initialize:
 $Y \in \mathbb{R}^{n \times (l+4)}$: such that $Y \in [Z, B]$
 $p_{\text{os}}, n_{\text{eg}} = \{\}$: Sets to store indexes of positively and negatively matched default boxes
Process:
For $i \in [1, n]$ **do:**
 For $j \in [1, h]$ **do:**
 $g = g_j, d = d_i$
 $\text{iou} = 1 - \frac{|g \cap d|}{|g \cup d|}$
 If $\text{iou} \geq \tau$ **then**
 $\text{class} = \arg \max(g_{1:l})$
 $p_{\text{os}} = p_{\text{os}} \cup \{i\}$
 Else
 $n_{\text{eg}} = n_{\text{eg}} \cup \{i\}$
 End if
 End for
End for
Output
 $Y \in [Z, B]$

3 实验结果与分析

3.1 实验数据

为验证本文所提算法对遥感图像舰船目标的检测性能, 使用含有渔船、集装箱货轮、驱逐舰等类别舰船的公开的 kaggle 舰船识别比赛提供的图像作为模型的训练集, 选取其中 2 000 幅含有舰船目标的图片作为实验数据, 将其大小统一为 500 pixel \times 375 pixel 并按照 7:3 的比例分为训练集和测试集. 相比于有 9 000 多幅样本 20 类目标的 PASCAL VOC 数据集^[25], 本文数据集足够达到可以对相应遥感图像进行检测的需求.

图 5 为本文数据集的样本示例, 并采用开源软件 LabelImg (版本: windows_v1.8.0) 对数据进行标注, 该软件很容易就能对遥感图像中的舰船目标进行标记 (x, y, w, h) , 其中 (x, y) 为目标所在矩形框的左上角坐标, w 是矩形框的宽度, h 为矩形框的高度. 所标注样本真实框长宽比及面积统计如图 6 所示.



图5 数据集部分样本示例
Fig. 5 Samples of the dataset

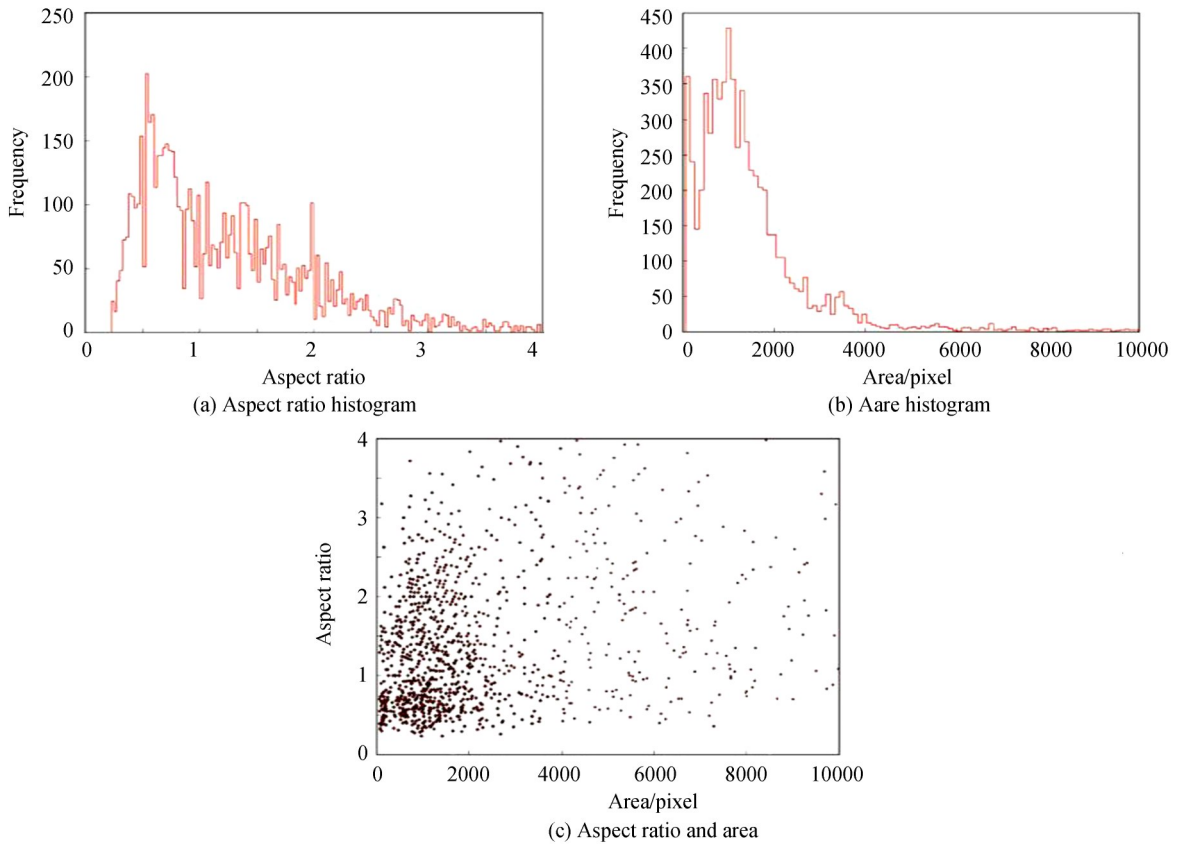


图6 标注真实框面积大小和纵横比分布直方图
Fig. 6 Histograms of the distribution of the area size and aspect ratio of the bounding boxes

3.2 评价指标

采用已被广泛应用于各类目标检测的评估中^[11-18]的精度均值(Average Precision, AP)^[26]和每秒帧数(Frame Per Second, FPS)^[26]作为目标检测结果的评价指标.其中AP的计算表达式为

$$AP = \int_0^1 p(r) dr \quad (7)$$

其计算了准确率(precision)和召回率(recall)在 $[0, 1]$ 范围内绘制的曲线与坐标轴所围成的面积,准确率和召回率的分别定义为

$$\begin{cases} P_{\text{precision}} = \frac{TP}{TP + FP} \\ R_{\text{recall}} = \frac{TP}{TP + FN} \end{cases} \quad (8)$$

式中,TP代表真正例,FN代表假反例,FP代表假正例.

此外, FPS 是目标检测的另一个重要的参考指标, 其数值的大小直接代表着检测速度的快慢, 具体可以表示为

$$F_{\text{fps}} = \frac{N_{\text{test}}}{T_{\text{time}}} \quad (9)$$

式中, N_{test} 为对测试集合的样本数量, T_{time} 为对测试集合进行检测时消耗的时间.

3.3 实验步骤

MSSDN 算法在预训练阶段采用 VGG16 网络对 ImageNet 数据集进行训练, 微调阶段采用 2.1 节数据集的训练集进行训练, 最后利用测试集对 MSSDN 算法进行测试. 参照文献[16]的实验设置以及鉴于本文实验设备的显存限制, MSSDN 算法训练过程中 batchsize 设置为 8, 动量因子设置为 0.9, 权重衰减系数为 0.000 5, 采用小批量 SGD 优化方式^[27]对损失函数式(4)进行优化, 初始学习率设置为 10^{-3} 并梯度衰减至 10^{-5} , 整个算法训练迭代 1.2×10^5 次.

此外, 为了验证本文所提 MSSDN 算法有效性和迁移性, 在第 3.5 节的拓展实验中, 将直接利用训练好的 MSSDN 算法和 SSD 算法检测模型对国产高分辨率光学遥感卫星传回的舰船影像数据进行目标检测, 直观地表明了 MSSDN 算法优于 SSD 算法.

3.4 算法性能分析

为了更好地评估所提 MSSDN 算法对光学遥感图像舰船目标的检测性能, 同时利用训练集训练了多种算法并与之对比, 对比结果如表 1. 从检测精度上看, 当匹配阈值设置为 0.5 时, MSSDN 的检测精度达到 81.7%, 与基于 VGG16 的 FSSD 网络算法相比提高 0.9%, 与基于 ResNet101 的 DSSD 网络相比提高了 2.6%, 与 SSD 算法相比提高 5.3%, 相较于其他算法, 本文所提的 MSSDN 算法检测精度具有更大的优势; 当匹配阈值设置为 0.7 时, MSSDN 算法的检测精度到达 72.2%, 相较于 FSSD、DSSD、SSD 算法分别提高 0.6%、1.8%、3.5%. 从检测速度上看, 因为本文 MSSDN 算法在 SSD 算法的基础之上添加 DFF module 和 PFF module, 一定程度上增加了网络的计算量, 使得算法的检测效率有所下降, 但相对于以 ResNet101 作为骨干网络的 DSSD 算法, MSSDN 算法在检测效率上仍具有较大的优势, 且能够满足实时检测的需求.

表 1 实验结果对比
Table 1 Comparison of experimental results

Method	Backbone	Inputsize	AP(%)@0.5	AP(%)@0.7	Speed /FPS
FastR-CNN[11]	VGG16	1 000×600	65.6	48.2	0.07
FasterR-CNN[12]	VGG16	1 000×600	73.1	53.2	4.4
YOLO[13]	GoogleNet	448×448	68.1	50.1	70.2
YOLOV2[14]	DarkNet19	544×544	74.2	64.6	59.2
SSD[16]	VGG16	300×300	76.4	68.7	63.6
DSSD[17]	ResNet101	321×321	78.1	70.4	9.5
FSSD[18]	VGG16	300×300	80.8	71.6	47.1
MSSDN(ours)	VGG16	300×300	81.7	72.2	43.2

3.5 检测结果对比

由 3.4 节可知, 本文所提的 MSSDN 算法对光学遥感图像舰船目标的检测精度优于其它算法, 本节将分别展示针对不同尺度的目标和复杂场景下算法的检测效果, 检测对比结果如图 7 所示.

图 7 为分析 MSSDN 算法对不同尺度的目标与检测精度相对较高的 SSD、DSSD、FSSD 算法下可视化的检测效果示意图, 图 7(a) 为典型的港口靠岸中型尺度舰船目标, 从图中可以看出原始 SSD 算法对港口靠岸舰船目标检测存在漏检情况, DSSD、FSSD 算法对靠岸舰船目标的检测效果有一定程度的提升, 但仍出现漏检情况, 本文 MSSDN 算法对港口靠岸舰船目标的检测效果取得了较大的提升; 图 7(b) 为典型的港口小尺度舰船目标, 通过直观的对比可以发现 SSD、DSSD、FSSD 算法均存在严重的漏检现象, 而本文 MSSDN 对港口小尺度舰船目标的检测效果有所提升; 图 7(c) 为在开阔海域大尺度与小尺度舰船紧密相邻

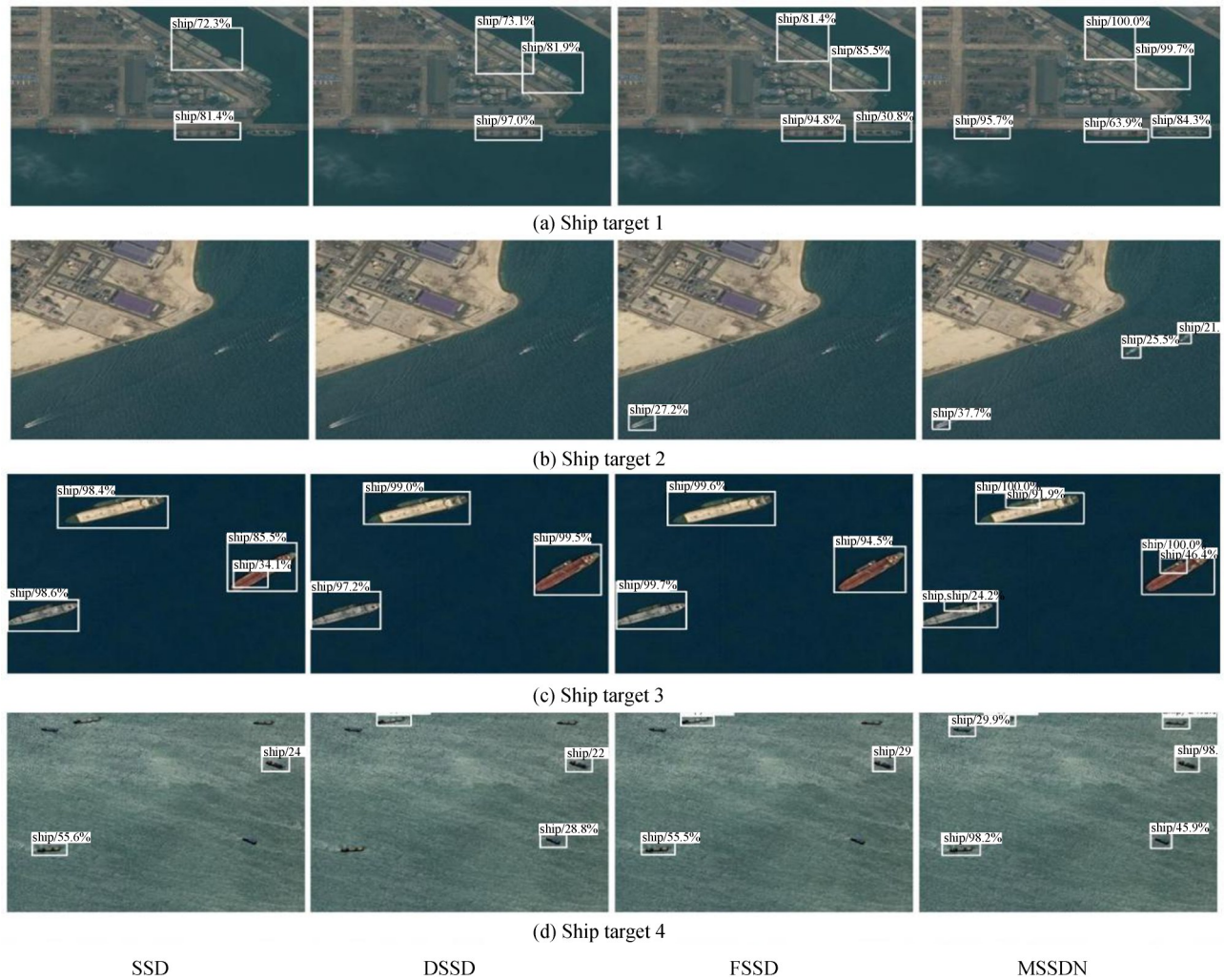


图7 四种算法检测结果对比
Fig. 7 The detection results of four methods

的舰船图像,从图中可以看出,SSD、DSSD、FSSD均将紧邻的两艘舰船视为一艘舰船,出现小目标舰船漏检,且SSD算法的检测结果出现重复检测的情况即存在一个较小的检测框只检测出舰船的某一部分;图7(d)是复杂海况下小尺度舰船目标,受环境干扰严重,从图中可以看出SSD、DSSD、FSSD算法对图中的舰船目标均存在漏检,本文算法能够精确的检测出复杂海况下小尺寸舰船目标,说明本文所提MSSDN算法对复杂场景下遥感图像舰船目标的适应性强。

为了验证本文所提MSSDN算法对复杂场景下遥感图像目标的检测效果,除了在上述采集的训练集和测试集进行算法性能的对比之外,增加拓展实验.对因云层、大气颗粒等问题导致的成像模糊的图像,利用上述训练模型,直接对国产高分辨率光学遥感卫星传回的影像数据进行检测,并与SSD算法进行对比,相应的检测结果如图8.通过直观地对比可以发现,MSSDN算法对受干扰目标的检测效果明显优于SSD算法,这是因为在进行特征融合的过程中,深层特征与浅层特征融合除了引入丰富的语义信息,在一定程度上也增强了网络对受干扰目标的特征相应,增强特征的泛化能力,进而提升检测的性能.上述实验证明了本文所提MSSDN算法对复杂场景下遥感图像检测的有效性,且MSSDN算法的迁移性优于SSD算法,训练的模型具有一定的通用性。

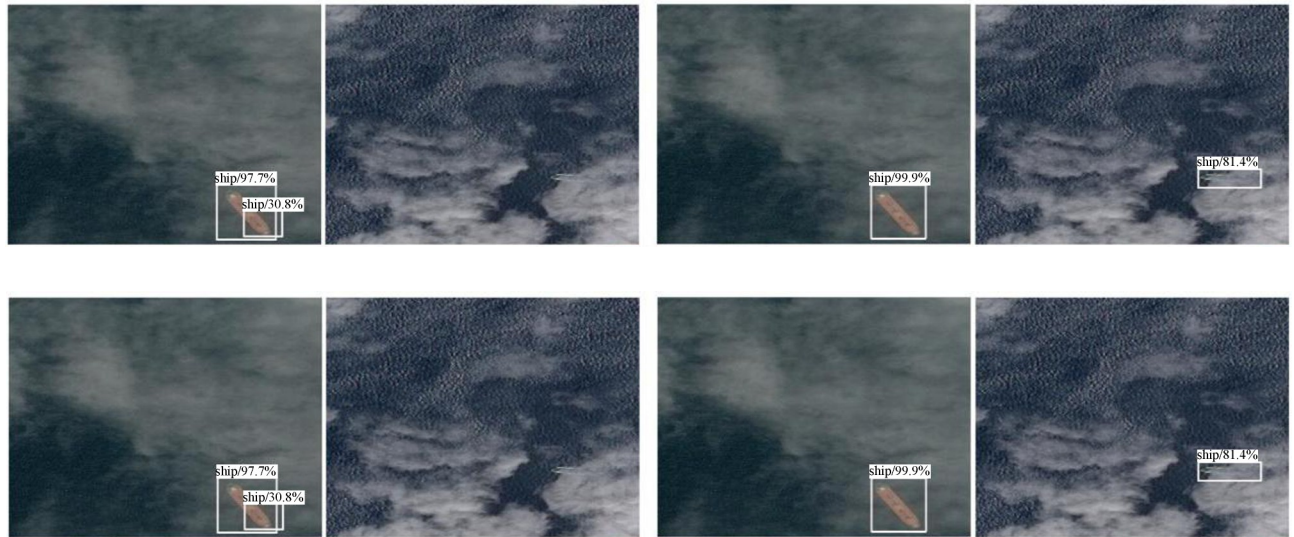


图8 MSSDN算法与SSD算法对模糊目标物体的检测结果示例
Fig.8 Example of detection results of fuzzy ship targets by MSSDN and SSD

4 结论

本文提出了一种基于特征融合的多尺度舰船目标检测算法.设计反卷积特征融合模块(DFE module)和池化特征融合模块(PFE module),增强各层特征之间的关联,同时在损失函数中加入聚焦分类损失,平衡训练过程中的正负样本分布,使得本文算法对复杂场景下的多尺度舰船目标检测具有更好的鲁棒性.对以中小目标为主的舰船目标进行检测实验分析,结果表明:1)与当前多种深度学习算法相比,本文所提的MSSDN算法的结构设计合理,检测精度更高且能够满足实时检测需求.2)从检测精度上看,当IoU阈值设置为0.5时,MSSDN算法较SSD,DSSD,FSSD算法分别提高5.3%,2.6%,0.9%;当IoU阈值设置为0.7时,MSSDN算法较SSD,DSSD,FSSD算法分别提高3.5%,1.8%,0.6%.3)从检测速度上看,MSSDN算法相较于SSD算法降低20.2 FPS,相比于FSSD算法降低3.9 FPS,但相比于Fast R-CNN,Faster R-CNN和DSSD算法,MSSDN算法检测速度具有显著优势.4)从检测效果上看,MSSDN算法的检测效果要优于其他算法,且对模糊目标的检测更准确.

本文所提MSSDN算法利用深度学习技术对遥感图像舰船目标进行检测与识别,对遥感图像目标检测具有一定的参考价值,尤其是对复杂场景下小目标物体检测具有较大优势.尽管MSSDN算法能够达到实时检测的应用需求,但相对于原始SSD算法耗时依然过多.未来将进一步优化算法,扩充遥感图像舰船目标数据集,以提升对复杂场景下的舰船目标的检测效果.

参考文献

- [1] YI Wei, ZENG Yong, YUAN Zheng. Fusion of SAR and optical images based on the nonsubsampled contourlet transform [J]. *Acta Optica Sinica*, 2018, **38**(11): 1110002.
易维,曾湧,原征.基于NSCT变换的高分三号SAR与光学图像融合[J].*光学学报*,2018,**38**(11):1110002.
- [2] SUN Yue-jiao, LEI Wu-hu, HU Yi-hua, *et al.* Ship target detection method based on multi-source remote sensing satellite [J]. *Laser & Infrared*, 2018, **48**(2): 267-272.
孙越娇,雷武虎,胡以华,等.基于多源遥感卫星的海面舰船目标检测方法[J].*激光与红外*,2018,**48**(2):267-272.
- [3] ZHU C, ZHOU H, WANG R, *et al.* A novel hierarchical method of ship detection from spaceborne optical image based on shape and texture features [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2010, **48**(9): 3446-3456.
- [4] WANG Wei-wei, XI Deng-yan, YANG Gong-peng. Marineship target detection using structural texture decomposition [J]. *Journal of Xidian University*, 2012, **39**(4): 131-137.
王卫卫,席灯炎,杨堃鹏.利用结构纹理分解的海洋舰船目标检测[J].*西安电子科技大学学报*,2012,**39**(4):131-137.
- [5] QI S, JIE M, JIN L, *et al.* Unsupervised ship detection based on saliency and S-HOG descriptor from optical satellite images [J]. *IEEE Geoscience & Remote Sensing Letters*, 2015, **12**(7): 1451-1455.
- [6] TANG J, DENG C, HUANG G B, *et al.* Compressed-domain ship detection on spaceborne optical image using deep neural network and extreme learning machine [J]. *IEEE Transactions on Geoscience & Remote Sensing*, 2014, **53**(3): 1174-1185.

- [7] DONG Y, DU B, ZHANG L. Target detection based on random forest metric learning[J]. *IEEE Journal of Selected Topics in Applied Earth Observations & Remote Sensing*, 2017, **8**(4): 1830-1838.
- [8] ZHANG Di-fei, ZHANG Jin-suo, YAO Ke-ming. Infrared ship target recognition based on SVM classification[J]. *Infrared and Laser Engineering*, 2016, **45**(1): 167-172.
张迪飞,张金锁,姚克明.基于SVM分类的红外舰船目标识别[J].红外与激光工程,2016,**45**(1):167-172.
- [9] KRIZHEVSKY A, SUTSKEVER I, HINTON G. ImageNet classification with deep convolutional neural networks[C]. NIPS, Curran Associates Inc, 2012.
- [10] WANG Wen-xiu, FU Yu-tian, DONG Feng, *et al.* Infrared ship target detection method based on deep convolutional neural network[J]. *Acta Optica Sinica*, 2018, **38**(7):0712006.
王文秀,傅雨田,董峰,等.基于深度卷积神经网络的红外船只目标检测方法[J].光学学报,2018,**38**(7):0712006.
- [11] GIRSHICK R. Fast R-CNN[C]. Proceedings of the IEEE International Conference on Computer Vision, Los Alamitos: IEEE Computer Society 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[C]. Proceedings of the 28th International Conference on Neural Information Processing Systems, Cambridge: MIT Press, 2015, 1: 91-99.
- [13] REDMON J, DIVVALA S, GIRSHICK R, *et al.* You only look once: unified, real-time object detection[C]. Proceedings of the IEEE International Conference on Computer Vision, Los Alamitos: IEEE Computer Society Press, 2016: 779-788.
- [14] REDMON J, FARHADI A. YOLO9000: better, faster, stronger[C]. Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition(CVPR), Hawaii: IEEE, 2017: 6517-6525.
- [15] REDMON J, FARHADI A. YOLOv3: An incremental improvement[J]. *arXiv*:1804.02767, 2018.
- [16] LIU W, ANGELOV D, ERHAN D, *et al.* SSD: Single shot multibox detector[C]. Proceedings of the European Conference on Computer Vision. Amsterdam: Springer International Publishing, 2016: 21-27.
- [17] FU C, LIU W, RANGAA, *et al.* DSSD: Deconvolutional single shot detector[J]. *arXiv*:1701.06659, 2017.
- [18] CAO G M, XIE X M, YANG W Z, *et al.* Feature-fused SSD: fast detection for small objects[J]. *arXiv*:1709.05054, 2017.
- [19] FENG Ke-yao, RAO Peng, LU Fu-xing, *et al.* High resolution and fast target detection based on neural network[J]. *Electronic Design Engineering*, 2018, **26**(22): 175-179.
冯珂垚,饶鹏,陆福星,等.基于神经网络的高分辨率快速目标检测方法[J].电子设计工程,2018,**26**(22):175-179.
- [20] LI Jian-wei, QU Chang-wen, PENG Shu-Juan, *et al.* Ship detection in SAR images based on convolutional neural network[J]. *Systems Engineering and Electronics*, 2018, **40**(9): 1953-1959.
李健伟,曲长文,彭书娟,等.基于卷积神经网络的SAR图像舰船目标检测[J].系统工程与电子技术,2018,**40**(9): 1953-1959.
- [21] LIN Liang-kui, WANG Shao-you, TANG Zhong-xing. Point target detection in infrared over-sampling scanning images using deep convolutional neural networks[J]. *Journal of Infrared and Millimeter Waves*, 2018, **37**(2): 219-226.
林两魁,王少游,唐忠兴.基于深度卷积神经网络的红外过采样扫描图像点目标检测方法[J].红外与毫米波学报,2018, **37**(2):219-226.
- [22] ZHU Min-chao, FENG Tao, ZHANG Yu. Remote sensing image multi-target detection method based on FD-SSD[J]. *Computer Applications and Software*, 2019, **36**(1): 238-244.
朱敏超,冯涛,张钰.基于FD-SSD的遥感图像多目标检测方法[J].计算机应用与软件,2019,**36**(1):238-244.
- [23] LIN T Y, DOLLAR P, GIRSHICK R, *et al.* Feature pyramid networks for object detection[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Hawaii: IEEE, 2017: 936-944.
- [24] LIN T Y, GOYAL P, GIRSHICK R, *et al.* Focal loss for dense object detection[C]. 2017 IEEE international conference on computer vision(ICCV), October 22-29, Venice, Italy, New York: IEEE, 2017: 2980-2988.
- [25] EVERINGHAM M, ESLAMI S M A, GOOL L V, *et al.* The pascal visual object classes challenge: a retrospective[J]. *International Journal of Computer Vision*, 2015, **111**(1): 98-136.
- [26] LIN T Y, MAIRE M, BELONGIE S, *et al.* Microsoft coco: Common objects in context[C]. Proceedings of the European Conference on Computer Vision, Zurich: Springer International Publishing, 2014: 740-755.
- [27] WANG Z, CRAMMER K, VUCETIC S. Breaking the curse of kernelization: budgeted stochastic gradient descent for large-scale SVM training[J]. *Journal of Machine Learning Research*, 2012, **13**(1): 3103-3131.

Foundation item: The National Natural Science Foundation of China(No.11871059), the New Generation of Artificial Intelligence Major Program in Sichuan Province(No.2018GZDZX0036), the Key Program of Sichuan Science and Technology Department (No.2018SZ0040)