

# 舰船遥感图像数据集 DSTD 的构建研究

何维娟<sup>1</sup>, 江涛<sup>1</sup>, 王林飞<sup>2</sup>, 徐权峰<sup>1</sup>, 王欣<sup>1</sup>

(1. 云南民族大学 数学与计算机科学学院, 云南 昆明 650500;

2. 云南大学 信息学院, 云南 昆明 650091)

**摘要:**目标检测算法在 PASCAL VOC、COCO 等一系列数据集中都取得了一定的效果,但是都是面向自然场景下的多目标检测任务,而这些数据集中的单类目标场景往往很单一,相应目标数量也很少,针对特定场景和特定目标的数据集并不多见。而对于此类任务的数据集往往也是很有价值的,例如遥感场景下的舰船检测或者飞机检测。针对此问题,构建一种面向遥感场景的大规模水面舰船目标检测数据集,其数据集的主要来源为 DIOR、DOTA、NWPU-VHRdataset、TGRS-HRRSD-Dataset-master 等几个开源数据集,将其命名为 DSTD(dataset for ship target detection),数据集中包含 4 854 张舰船图片,87 076 个舰船实例。DSTD 数据集具有数量多、多尺度和成像差异大以及较高的类内多样性等特点。在构建数据集的基础上,进一步分析了遥感舰船图像的细节特征,评估了一些经典目标检测方法的性能,并进行了实验对比,验证了该数据集的可行性,同时发现了当前较适合舰船的检测方法:即 YOLOV5。使用 YOLOV5 算法进行舰船图像目标检测,在保证高检测精度的同时,能保持极快的检测速度。

**关键词:**目标检测;遥感;数据集;舰船;YOLOV5

**中图分类号:**TP7

**文献标识码:**A

**文章编号:**1673-629X(2022)07-0082-05

**doi:**10.3969/j.issn.1673-629X.2022.07.014

## Research on DSTD Construction of Ship Remote Sensing Image Dataset

HE Wei-juan<sup>1</sup>, JIANG Tao<sup>1</sup>, WANG Lin-fei<sup>2</sup>, XU Quan-feng<sup>1</sup>, WANG Xin<sup>1</sup>

(1. School of Mathematics and Computer Science, Yunnan Minzu University, Kunming 650500, China;

2. School of Information, Yunnan University, Kunming 650091, China)

**Abstract:**Target detection algorithms have achieved certain effects in a series of datasets such as PASCAL VOC and COCO, but they are all for multi-target detection tasks in natural scenes. The single-class target scenes in these data sets are often single and the number of corresponding targets is small. Datasets for specific scenes and targets are rare. Datasets for such tasks are often valuable, such as ship detection or aircraft detection in remote sensing scenarios. To solve this problem, a large-scale surface ship target detection dataset for remote sensing scene is proposed. The main sources of the dataset are DIOR, DOTA, NWPU-VHRdataset, TGRS-HRRSD-Dataset-master and so on, called DSTD (dataset for ship target detection). The dataset contains 4 854 ship pictures and 87 076 ship instances. DSTD are characterized by large number, large multi-scale and imaging differences, and high intra-class diversity. On the basis of the proposed dataset, we further analyze the detailed features of remote sensing ship images, evaluate the performance of some classical target detection methods, and carry out experiments to verify the feasibility of the dataset. At the same time, YOLOV5, a more suitable ship detection method, is found. The YOLOV5 algorithm is used to detect ship image targets, which ensures high detection accuracy while maintaining extremely fast detection speed.

**Key words:**object detection; remote sensing; dataset; ship; YOLOV5

## 0 引言

近年来,研究人员在图像识别<sup>[1]</sup>开发中做了很多工作。随着遥感技术的快速发展,公开的目标检测数

据集也变得更加,大大增加了遥感图像的数量和质量。遥感图像可以描述空中观测的一切物体,如飞机、油罐、车辆等。但是目前提出的目标检测图像数据集存

收稿日期:2021-06-07

修回日期:2021-10-09

基金项目:国家自然科学基金(61363022);云南民族大学研究生创新基金项目(SJXY-2021-002)

作者简介:何维娟(1995-),女,硕士,研究方向为计算机视觉、语音识别;通讯作者:江涛(1973-),男,博士,教授,研究生导师,研究方向为特定领域的视觉建模、建模形式化、模型验证、语音识别、计算机视觉等。

在一些普遍的问题,大多都是自然场景下的数据集。并且这些数据集中单类目标场景通常很单一,目标数量也较少,对于特定场景和目标的数据集并不多见,但是这类数据集却是有价值的。研究人员广泛地研究致力于航空图像中的目标检测<sup>[2-8]</sup>,利用计算机视觉的最新进展,并考虑到地球视觉应用的高要求,这些方法中的大多数<sup>[4,8]</sup>试图将针对自然场景开发的目标检测算法转移到航空图像领域。此外,目前的目标检测虽然取得了很大的成功,但是在舰船遥感图像<sup>[9]</sup>中却存在很多难题,从当前舰船遥感图像目标检测的研究来看,存在以下几个普遍的问题:一是舰船目标较小,导致舰船占像素点较少;二是舰船的种类多、大小不一,检测难度相对较高;三是舰船图像的成像受天气、岸边建筑和光照等因素的影响引起舰船物体视觉的改变。

对于舰船目标检测和舰船数据集的研究已有显著成果。例如,徐芳<sup>[10]</sup>提出了一种梯度方向特征方法;余东行等人<sup>[11]</sup>采用基于频率域相位谱的多尺度显著性检测方法抑制干扰因素;宋明珠等人<sup>[12]</sup>针对难以辨认的弱小舰船目标,设计了不同图像层的权重,借此合成了基于多类特征的显著性图;李健伟等基于深度学习的 SAR 图像舰船检测数据集及性能分析提出了 SSDD 数据集;Li 等<sup>[13]</sup>提出了 DIOR 数据集等等。舰船检测的传统方法是基于自动识别系统和舰船特征<sup>[14]</sup>,Li 等提出了一种改进的空间聚类算法来识别舰船的异常行为<sup>[15]</sup>,Zhang 等使用 AIS 数据来识别碰撞舰船,虽然这些研究都获得了不错的效果,但是却普遍存在识别精度低和检测速度低等一系列问题,因此,传统的舰船检测难以达到理想的检测效果。近些年,基于深度学习的方法在各种图像识别中占据主导地位,基于深度学习的方法主要包括两种:一种是一阶段基于回归的方法,包括 SSD<sup>[16]</sup>、YOLOV1<sup>[17]</sup>、YOLOV2<sup>[18]</sup>、YOLOV3<sup>[19]</sup>等;另一种是两阶段基于区域建议的方法,包括 R-CNN<sup>[20]</sup>、Fast R-CNN<sup>[21]</sup>、Faster R-CNN<sup>[22]</sup>(Faster Regions with CNN features)等。

为了更好地解决以上问题,该文构建一种舰船数据集,将其命名为 DSTD。从开源数据集中采集了 4 845 张图像,其中包含不同方向、形状和比例的对象,共有 87 076 个舰船实例,每个实例都由一个 bounding box 标记。主要贡献(Contribution)如下:(1)构建舰船数据集,构建的 DSTD 数据集是一组规模较大的舰船遥感图像目标检测数据集;(2)以 DSTD 舰船数据集为基准,选择了单次目标检测的 YOLOV3 算法和两次目标检测 R-FCN 算法与 YOLOV5 算法对该数据集进行实验对比,最终在当前几种主流的目标检测方法中找到了最适合舰船检测的方法:即 YOLOV5。

## 1 相关工作

遥感图像是以航空拍摄为基础,开始为航空遥感,随着遥感数据集的发展,数据类型越来越丰富,数据量也越来越大,目前而言,遥感图像数据集已经有了很好的发展。

### 1.1 遥感图像数据集的发展

近些年,有许多人热衷于遥感图像数据集的研究,并公开发表了他们的数据集(部分数据集见表 1),这些数据集简要描述如下:

表 1 遥感图像数据集

Dataset	Categories	Images	Instances	Year
TAS	1	30	1 319	2008
SZTAKI-INRIA	1	9	665	2012
NWPU VHR-10	10	800	3 775	2014
VEDAI	9	1 210	3 640	2015
UCAS-AOD	2	910	6 029	2015
DLR 3K Vehicle	2	20	14 235	2015
HRSC2016	1	1 070	2 976	2016
RSOD	4	976	6 950	2017
DOTA	15	2 806	188 282	2017
SSDD	1	1 160	2 456	2018
DIOR	20	23 463	190 288	2018
ITCVD	1	173	29 088	2018

(1)TAS:TAS 数据集是用于航空图像里的汽车检测。包含 30 张图像共 1 319 个汽车目标,图像的空间分辨率相对较低。

(2)SZTAKI-INRIA:SZTAKI-INRIA 数据集包含 9 张图像,共 665 个建筑目标,该数据集用于对各种建筑检测方法进行基准测试。

(3)NWPU VHR-10:NWPU VHR-10 数据集由 800 张图片构成,有 10 个对象类,其中包括舰船、油罐、飞机等等。

(4)VEDAI:VEDAI 数据集由 1 210 张图像构成,其中含有的目标为车辆,分辨率 0.12 m,RGB+N(近红外),尺寸大小:1 024 \* 1 024。

(5)UCAS-AOD:中国科学院大学的实验成果,UCAS-AOD 数据集由 910 张图像构成,其中含有的目标为汽车和飞机,还有部分背景样本。

(6)DLR 3K Vehicle:DLR 3K Vehicle 数据集由 20 张航拍图像,共 14 235 个实例构成,这些图片是用 DLR 3K 摄像系统拍摄的。

(7)HRSC2016:HRSC2016 数据集由 1 070 张图像构成,用于舰船检测,该数据在位置、外观和形状等上都有很大的变化。(8)RSOD:武汉大学团队标注,RSOD 数据集由 976 张图像构成,包含 6 950 个实例,

有四个类别,分别为油箱、飞机、桥和操场。

(9)DOTA:DOTA 数据集一共有 15 种土地利用类型,总样本数高达 18 万,包含 2 806 张图片,其中含有 15 个对象类别,共计 188 282 个实例目标。

(10)SSDD:SSDD 是国内外公开的第一个专门用于 SAR 图像舰船目标检测的数据集,SSDD 中长宽比的分布范围比较广,从 0.4 到 3,该数据集用于舰船检测。

(11)DIOR:DIOR 数据集由 20 个类别,23 463 张图片包含 190 288 个实例构成,该数据集在类别数量、对象数量和图像数量都是最大的。

(12)ITCVD:ITCVD 数据集是由飞机在高空约 330 米高度拍摄的,该数据集用于车辆检测。

## 1.2 参与评估的目标检测方法

(1)R-FCN:R-FCN 与 Faster R-CNN 是一个体系的,但是与 Faster R-CNN 比较,R-FCN 的速度快得多,并且将部分卷积特征图的计算实现了共享。R-FCN 的关键思想:分类需要特征具有平移不变性,检测则要求对目标的平移做出准确响应;消耗的卷积都尽量移到前面共享的 subnetwork 上,因此,与 Faster RCNN 中用的 ResNet 策略有所不同。

(2)YOLOV3:YOLOV3 引入了 FPN 结构。从 YOLOV1 和 YOLOV2 中的 Darknet19 网络升级到 Darknet-53<sup>[23]</sup>实现特征提取,借鉴残差网络<sup>[24]</sup>的思想,交替使用  $3 \times 3$  与  $1 \times 1$  的卷积神经网络,但该方法借鉴了 SSD 方法中的多尺度思想,在保留了 YOLO 系列方法采用单阶段网络进行高帧率检测的同时,进一步具备了检测微小目标和多尺度目标的能力。它与 SSD 一样准确,但是比 Faster R-CNN 和 SSD 检测速度更快。

(3)YOLOV5:YOLOV5 推理速度是目前最强的,它在 YOLOV1-YOLOV4 的基础上进行优化。为了达到更快的目标检测效果,YOLOV5 在输入端采用了 Mosaic 数据增强、自适应锚框计算和自适应图片缩放,运行速度大大提高,最快的速度达到每秒 140 帧,在 Backbone 中使用了 Focus 结构和 CSP 结构,针对 Focus 结构,这种结构的主要操作是切片,目的是最大程度地减少信息损失而进行下采样操作,保证了特征图的语义多样性。与 YOLOV1-YOLOV4 相比,YOLOV5 具有更高的准确率和更好的识别小物体的能力。

## 2 DSTD 数据集

实验数据是由 TGRS-HRRSD-Dataset-master、NWPU-VHRdataset、DIOR、DOTA 等几个开源遥感图像数据集构成,对图片进行剪切、修改以及筛选,并通

过标注工具 labelImg 进行手动标注,一共得到了 3 900 张训练集和 954 张测试集,制作了该舰船目标遥感数据集 DSTD(见图 1)。

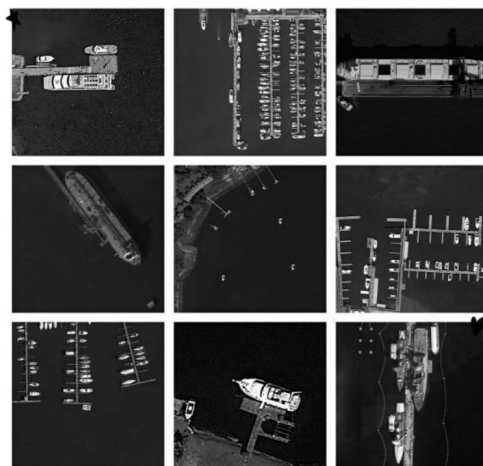


图 1 DSTD 数据集部分舰船图像数据示例

数据集文件中包括 annotation、image 两个文件,其中 annotation 包含 4 854 个 txt 文件,是使用开源工具 LabelImg 对 DSTD 进行人工标注,生成后缀为 .txt 的文件,每个 txt 文件主要为目标舰船的 bounding box,格式为  $x_1, y_1, x_2, y_2$ ,其中  $(x_1, y_1)$  与  $(x_2, y_2)$  分别是舰船 bounding box 的左上角坐标和右下角坐标。在训练数据集之前,需要对数据通过人工标记标签,这是非常费时费力的步骤,因此要获得训练数据集变得非常困难。

DSTD 遥感图像目标检测数据集有以下显著特征:

(1)针对性强、规模较大。DSTD 数据集由 4 854 张最优舰船遥感图像和 87 076 个舰船目标实例组成,该数据是从开源数据集中采集的。针对目前的遥感图像数据集,DSTD 针对性很强,只有舰船遥感图像并且在图像数量和目标数量上规模都较大,进一步提高了水上目标检测技术水平。

(2)目标尺寸变化大。目标尺寸变化是目标检测的一个重要特征,因为它对图像的分辨率有很大影响,在 DSTD 数据集,舰船的大小变化范围较大,为了增加目标的尺寸变化,该数据集收集了分辨率不同的图像,两张舰船图片的目标大小与数量均有较大差异。

(3)图片变化大。对于目标检测而言,需要图像具有较好的鲁棒性,DSTD 数据集就是在不同天气、不同成像条件、不同水上环境下拥有丰富的图像变化,因此 DSTD 数据集在光照、物体姿态和背景等方面都有丰富的变化,从而具有一定的鲁棒性。

(4)类内多样性高。为了增加类内多样性,在收集舰船数据时,考虑了舰船比例、颜色、大小等因素,因此 DSTD 数据集具有较高的类内多样性。



### 3 实验及结果

在这一节中,用 DSTD 数据集训练了几种经典目标检测方法的性能,验证了该数据集的可行性,同时找到了最适合舰船的检测方法:即 YOLOV5。进一步进行实验分析并讨论了验证集上的可视化结果。

#### 3.1 实验数据集

为保证进行实验时训练和测试数据分布的相似性,在 4 854 张 DSTD 舰船遥感图像中,随机获得 3 900 幅舰船图像当作训练集,剩下的 954 幅图片当作测试集(见表 2)。由于每幅图片至少包含一艘舰船,有的一幅图片包含多艘舰船,因此,舰船目标数量不等于总的图片数。

表 2 舰船目标数据集

参数	数值
总数据集	4 854
训练集	3 900
测试集	954

#### 3.2 实验过程及平台

该文使用 python 语言进行实验,实验使用的操作系统为 Red Hat 4. 8. 5-39,其中,处理器型号为 Intel (R) Xeon(R) CPU E5-2620 v4 @ 2.10 GHz,显卡型号为 GeForce GTX 1080,采用 NVIDIA CUAD11.0 加速工具箱。具体配置如表 3 所示。

表 3 实验环境

参数	配置
CPU	Intel(R) Xeon(R) CPU E5-2620 v4 @ 2.10 GHz
GPU	GeForce GTX 1080
系统	Red Hat4. 8. 5-39
语言	Python
加速环境	CUAD11.0 python3. 6

除了特别说明以外,对于 YOLOV5 算法在 DSTD 数据集的训练阶段使用的其他部分实验参数如表 4 所示。

表 4 实验参数设置

参数	值
输入图片的批大小	64
调整输入图片尺寸	640×640
采样步长	64
迭代次数	300
初始学习率	0.004 7
结束学习率	0.000 044 7
动量参数	0.873

在 YOLOV5 算法下,训练阶段设置 batch 为 64,输入图像为 640×640,初始阶段学习率为 0.004 7,结束学习率为 0.000 044 7,训练迭代次数为 300,loss 值

稳定在 0.105 左右,训练过程得到正常收敛效果,Loss 曲线如图 2 所示。

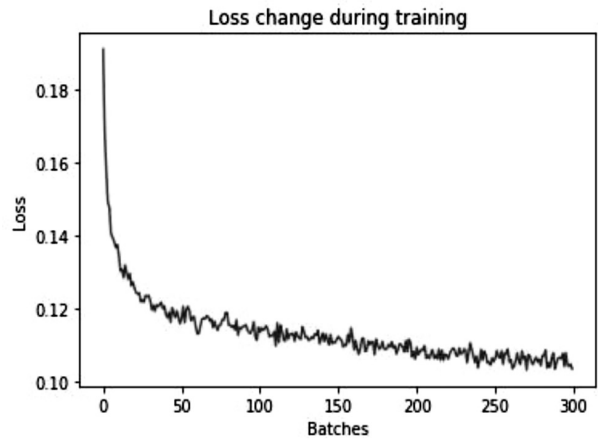


图 2 损失值函数

#### 3.3 实验评价指标

实验评价指标主要是验证实验结果好坏的一种评比方法,该文选择的评价指标为:精准率、召回率、平均精度和帧率。

(1) 精准率(Precision,  $P$ )和召回率(Recall,  $R$ )。

$$P = \frac{N_{tp}}{N_{tp} + N_{fp}}$$

$$R = \frac{N_{tp}}{N_{tp} + N_{fn}}$$

其中, $P$ 是舰船的精准率, $N_{tp}$ 是准确检测舰船的数量, $N_{fp}$ 是错误检测舰船的数量, $R$ 是舰船的召回率, $N_{fn}$ 是将舰船误检为背景的数量。

(2) 平均精度(mean Average Precision, mAP)。

$$mAP = \frac{\sum_{i=1}^n AP_i}{n}$$

其中,mAP 表示多类别精度。

(3) 帧率(FPS):目标检测中每秒钟能够检测的图片数量,用该指标评价目标检测的速度。

#### 3.4 实验结果及对比

为了验证 YOLOV5 算法在 DSTD 舰船遥感图像的检测性能,另外选了 2 种目标检测方法作为对比,这两种方法都是目前在深度学习中很有效的方法,一种为基于区域建议的方法:R-FCN,另一种为基于回归的方法:YOLOV3。在进行 R-FCN、YOLOV3 模型的训练时,使用官方代码和默认参数。使用 YOLOV3、R-FCN 和 YOLOV5 对数据集训练 300 次 Batches。

三种算法在 DSTD 数据集上的精度、速度和召回率的对比如表 5 和图 3 所示。图 3 为三种算法的 mAP 图像对比,其中 YOLOV5、R-FCN、YOLOV3 的图像依次向下。YOLOV5 目标检测算法在检测时间上每张图片大约花费 0.005 4 s, YOLOV3 算法检测一张图片

需要 0.014 2 ms, R-FCN 算法检测一张图片需要 0.056 5 ms, YOLOV5 的检测速度是最快的。在检测精度方面, YOLOV5 的检测精度为 99.4%, YOLOV3 的检测精度为 81.2%, R-FCN 的检测精度为 86.2%, YOLOV5 比 YOLOV3 和 R-FCN 的检测精度提高了 22.4% 和 18.8%。R-FCN 的召回率为 92.3%, YOLOV3 的召回率为 87.2%, YOLOV5 的召回率为 99.2%, YOLOV5 较 R-FCN 及 YOLOV3 提高了 7.5% 和 13.8%。

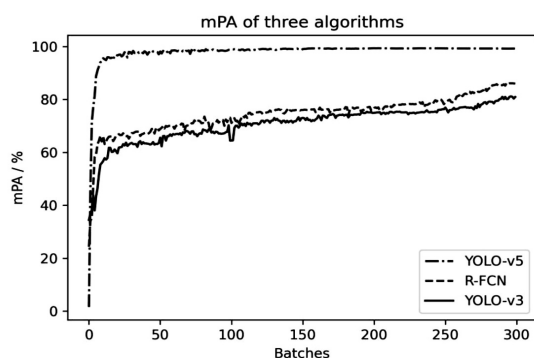


图 3 三种算法 mAP 对比结果

表 5 实验对比结果

算法	mAP/%	FPS	Recall/%
R-FCN	86.2	0.056 5	92.3
YOLOV3	81.2	0.014 2	87.2
YOLOV5	99.4	0.005 4	99.2

对比 YOLOV3 和 R-FCN, YOLOV5 能够保持较高分辨率的主要原因是具有较强的语义信息, 具有开阔的感受野, YOLOV5 中的 Mosaic 数据增强, 随机缩放、随机裁剪、随机排布的方式进行拼接等, 提高了对舰船小目标的检测能力。YOLOV5 在保持一定精度的同时检测速度最快, 所以 YOLOV5 对舰船遥感图像检测效果最好。

#### 4 结束语

该文构建了一种规模较大、舰船目标较多的 DSTD 舰船遥感图像数据集。该数据集可以帮助水上舰船目标检测有进一步的发展, 并且利用构建的数据集对基于深度学习的方法进一步探索和验证。对三种具有代表性的目标检测方法的性能进行了评价, 最终在当前几种主流的检测方法中发现了最适合舰船遥感图像检测的方法, 即 YOLOV5。将 YOLOV5 应用到遥感舰船检测领域中, 更好地实现了 DSTD 目标的检测。YOLOV5 虽然效果很好, 但是网络在目标框构建的时候存在明显的不足, 使得模型完美地屏蔽了一些小目标的训练, 原因是通过比较目标框和锚框的比例, 过滤掉了一部分锚框比例差距较大的目标框, 训练不好, 并

且进行过滤。该实验结果可作为今后研究舰船目标检测的一个有用性能。

#### 参考文献:

- [1] 王 慧, 季 雪. 图像识别处理技术在农业工程中的应用[J]. 传感器与微系统, 2018, 37(6): 158-160.
- [2] LIN Y, HE H, YIN Z, et al. Rotation-invariant object detection in remote sensing images based on radial-gradient angle[J]. IEEE Geoscience and Remote Sensing Letters, 2015, 12(4): 746-750.
- [3] LIU Z, WANG H, WENG L, et al. Ship rotated bounding box space for ship extraction from high-resolution optical satellite images with complex backgrounds[J]. IEEE Geoscience and Remote Sensing Letters, 2016, 13(8): 1074-1078.
- [4] LONG Y, GONG Y, XIAO Z, et al. Accurate object localization in remote sensing images based on convolutional neural networks[J]. IEEE Transactions on Geoscience and Remote Sensing, 2017, 55(5): 2486-2498.
- [5] MORANDUZZOT, MELGANI F. Detecting cars in UAV images with a catalog-based approach[J]. IEEE Transactions on Geoscience and Remote Sensing, 2014, 52(10): 6356-6367.
- [6] OK A O, SENARAS C, YUKSEL B. Automated detection of arbitrarily shaped buildings in complex environments from monocular VHR optical satellite imagery[J]. IEEE Transactions on Geoscience and Remote Sensing, 2013, 51(3-2): 1701-1717.
- [7] PORWAY J, WANG Q, ZHU S C. A hierarchical and contextual model for aerial image parsing[J]. International Journal of Computer Vision, 2010, 88(2): 254-283.
- [8] ZHANG F, DU B, ZHANG L, et al. Weakly supervised learning based on coupled convolutional neural networks for aircraft detection[J]. IEEE Transactions on Geoscience and Remote Sensing, 2016, 54(9): 5553-5563.
- [9] 魏 玮, 蒲 玮, 刘 依. 改进 YOLOv3 在航拍目标检测中的应用[J]. 计算机工程与应用, 2020, 56(7): 17-23.
- [10] 徐 芳. 可见光遥感图像海面目标自动检测关键技术研究[D]. 北京: 中国科学院大学(中国科学院长春光学精密机械与物理研究所), 2018.
- [11] 余东行, 张保明, 郭海涛, 等. 联合显著性特征与卷积神经网络的遥感影像舰船检测[J]. 中国图象图形学报, 2018, 23(12): 1947-1958.
- [12] 宋明珠, 曲宏松, 金 光. 含噪光学遥感图像海面弱小舰船目标检测[J]. 光学学报, 2017, 37(10): 1011004.
- [13] LI Ke, WAN Gang, CHENG Gong, et al. Object detection in optical remote sensing images: a survey and a new benchmark[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2020, 159: 296-307.
- [14] BYE R J, AALBERG A L. Maritime navigation accidents and risk indicators: an exploratory statistical analysis using

(下转第 92 页)