

# 基于深度学习的光学遥感影像舰船检测算法对比分析

夏文辉, 万剑华, 郑红霞, 许明明, 曲川萍

(中国石油大学(华东) 海洋与空间信息学院, 山东 青岛 266580)

**摘要:** 舰船目标检测是进行海洋环境监管, 保障海上权益的重要手段。基于深度学习的目标检测算法能在复杂环境下保持良好性能, 为测试不同深度学习目标检测算法在舰船检测中的效果, 构建了一个包含 3 893 张图像的数据集, 涵盖了复杂背景下不同类型的舰船, 基于此数据集分别采用 Faster RCNN、SSD、RetinaNet、YOLOv3、YOLOv4 算法进行实验, 结果表明, YOLOv4、YOLOv3、RetinaNet、Faster RCNN 平均精度均在 83% 以上, 其中 YOLOv4 最高达到 91.77%, Faster RCNN 误检较多, 而 SSD 平均精度最低, 只有 79.23%, 总的舰船检测数目偏少。将 5 种模型训练结果在高分二号影像上进行测试, 得到较好的检测效果, 对舰船检测未来理论研究的开展具有一定的指导意义。

**关键词:** 遥感; 目标检测; 舰船; 深度学习

中图分类号: P237

文献标识码: A

文章编号: 1000-3096(2021)05-0096-07

DOI: 10.11759/hyxx20201108009

舰船作为海上重要的运输载体, 其自动检测具有非常重要的现实意义<sup>[1]</sup>。现有的舰船研究工作大部分都是基于 SAR 影像<sup>[2-3]</sup>, 而光学遥感影像与 SAR 影像相比具有更高的分辨率和更直观的图像可读性, 能够提供更丰富的细节信息。因此, 基于光学遥感影像的舰船目标检测逐渐成为研究热点<sup>[4]</sup>。

传统舰船检测方法主要包括海陆分割、舰船预检与舰船识别 3 个流程, 其主要通过提取舰船目标的灰度特征、形状特征、纹理特征等进行多特征联合检测。Zhu 等基于形状和纹理特征并引入纹理算子局部多值模式(local multiple patterns, LMP)来增强特征集在特征提取中的表示能力<sup>[5]</sup>。彭敬涛等融合了舰船目标的形状、灰度、纹理和梯度等多特征信息<sup>[6]</sup>。王慧利等<sup>[7]</sup>提出一种新的表征舰船特性的边缘-方向梯度直方图特征对舰船目标进行描述。然而传统光学遥感手段无法对目标特性准确描述, 检测鲁棒性差, 虚警率高, 受云雾、波浪、杂波等因素影响严重。

目前目标检测领域的深度学习方法主要分为两类, 即双阶段目标检测算法和单阶段目标检测算法。作为双阶段检测算法代表, Faster RCNN<sup>[8]</sup>在 RCNN<sup>[9]</sup>、Fast RCNN<sup>[10]</sup>基础上创造性地使用 RPN 网络生成建议框, 速度跟精度都得到大幅改善, 并实现了端到端的检测, 取得较高精度。相较于带有候选区域提取的双阶段检测方法, 基于回归的单阶段方法如 YOLO<sup>[11]</sup>、SSD<sup>[12]</sup>等算法, 在检测速度上具有更

大的优势, 但准确性偏低。RetinaNet<sup>[13]</sup>提出 Focal loss 改进交叉熵损失来提升单阶段检测的精度。YOLOv3<sup>[11]</sup>保持检测速度的同时提高了该系列对小目标检测识别的准确率, 2020 年推出的 YOLOv4<sup>[14]</sup>引用各种优化策略, 形成一个高效而强大的模型。基于深度学习的目标检测算法, 能够充分提取检测目标的底层特征, 模型的泛化能力很强, 能在复杂环境下保持很好的鲁棒性。现有光学遥感影像舰船的公开数据集较少, Liu 等<sup>[15]</sup>制作的 HRSC2016 数据集在舰船识别领域应用较广, 该数据集主要包含军舰、商船等大型舰船, 可以较好地应用于大中型舰船的类型识别, 但该数据集缺乏小型舰船的截取标定, 而小目标的检测是衡量深度学习目标检测性能的一个重要指标。因此本文构建了复杂背景下不同类型舰船的数据集, 基于此数据集对 Faster RCNN、SSD、RetinaNet、YOLOv3、YOLOv4 五种算法进行实验对比分析, 并将模型训练结果在高分二号影像上进一步测试, 为舰船目标的实际检测应用提供参考。

收稿日期: 2020-11-08; 修回日期: 2020-12-02

基金项目: 国家重点研发计划项目(2017YFC1405600)

[Foundation: National Key Research and Development Program Plan, No. 2017YFC1405600]

作者简介: 夏文辉(1996—), 男, 山东青岛人, 硕士研究生, 主要从事舰船检测识别研究, E-mail: 1721760100@qq.com; 万剑华(1966—), 通信作者, 男, 山东单县人, 教授, 主要从事 3S 技术应用方面的研究, E-mail: wjh66310@163.com

# 1 方法

## 1.1 Faster RCNN

Girshick 在 2016 年提出 Faster RCNN, 创造性地使用 RPN 网络, RPN 网络会对预设锚点框进行初步筛选, 使得建议框数量大幅减少且质量大幅提高, 剔除过多的背景目标, 并和目标检测网络共享卷积网络, 采用非极大值抑制的方式, 抑制高重叠候选区域, 选取质量较高的建议框进行训练。RPN 损失函数  $L(\{p_i\}, \{t_i\})$  由二分类损失函数  $L(p_i)$  和边框回归损失函数  $L(t_i)$  相加组成, 定义如下所示:

$$L(\{p_i\}, \{t_i\}) = L(p_i) + L(t_i), \quad (1)$$

$$L(p_i) = \frac{1}{N_{cls}} \sum_i L_{cls}(p_i, p_i^*), \quad (2)$$

$$L(t_i) = \lambda \frac{1}{N_{reg}} \sum_i p_i^* L_{reg}(t_i, t_i^*), \quad (3)$$

式中,  $i$  表示锚点框的序号,  $p_i$  表示任意锚框  $i$  是目标的概率,  $p_i^*$  代表标注样本, 正样本时其为 1, 负样本其 0,  $t_i$  代表预测候选区域坐标,  $t_i^*$  代表对应为正样本对应的坐标。不同大小、尺寸的建议框映射回原图特征大小不一样, 通过池化层将其生成固定大小特征输入后续网络进行分类和回归。

## 1.2 SSD

作为单阶段检测器的代表, 有明显的速度优势。图像输入到 VGG16 骨干网络获取不同大小的特征图, 从中抽取 6 层特征图用于检测不同尺度的目标, 特征图经过卷积感受野不断增大, 对不同尺度特征图预测分类从而达到多尺度检测的目的。

## 1.3 RetinaNet

RetinaNet 由特征提取网络、特征金字塔网络 (FPN)、分类和边框回归网络组成, 在特征提取残差网络 ResNet 基础上利用 FPN 来构成多尺度卷积特征金字塔, 同时提出了 Focal loss 来解决单阶段检测器存在的类别失衡问题, Focal loss 计算方式如下:

$$FL(p_i) = -\alpha_i (1 - p_i)^\gamma \log(p_i), \quad (4)$$

式中  $p_i$  表示不同类别的分类概率,  $\gamma$  和  $\alpha_i$  是固定值, 不参与训练。通过调整  $p_i$  参数对难分样本赋予更大的损失,  $p_i$  越大,  $(1 - p_i)^\gamma$  越小, 有效提升了整体的目标检测精度, 以此来消除样本间的不平衡问题, 能保持速度的同时达到较高的精度。

## 1.4 YOLOv3

YOLOv3 基础网络采用 Darknet-53, 网络结构层非常深入, 没有使用池化层、全连接层, 在一些层之间设置了快捷链路, 可有效抑制算法的过拟合现象, 减少了计算量, 提高了速度, 利用多尺度特征结合不同大小的感受野对各种大小对象进行检测。

## 1.5 YOLOv4

YOLOv4 在 YOLOv3 基础上做出较大改进, 为了提升准确度, 在主干网络、激活函数、损失函数等多个方面进行优化。核心主干网络采用 CSPDarknet53, 借鉴 CSPNet 结构简洁准确度高的经验, 在 Darknet53 残差块上加入 CSP, 将基础层的特征映射划分为两部分并通过跨阶段结构合并, 在减少计算量的同时保证准确率。在骨干网络上添加了 SPP 块扩大感受野分离出最显著的上下文特征, 在 3 个有效特征层上使用 PANet 作为参数聚合方法, 充分利用特征提取信息, 具体结构如图 1 所示。

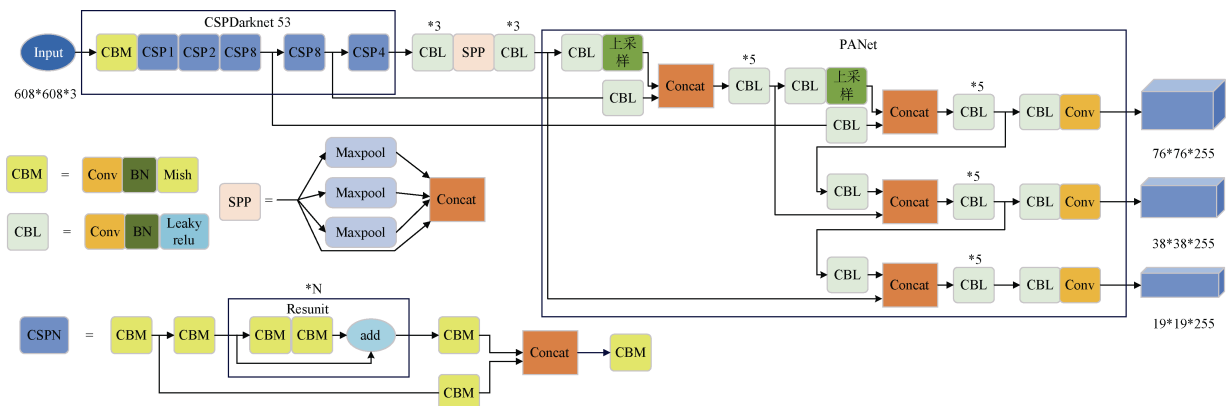


图 1 YOLOv4 结构示意图

Fig. 1 Schematic of the YOLOv4 structure

## 2 实验

### 2.1 实验环境与数据

本实验采用的是 64 位 Win10 系统, 计算机配置 GPU 为 NVIDIA GeForce GTX 1080TI, 显存为 11 GB, 深度学习框架使用 PyTorch 和 Keras。丰富样本集的学习是支撑深度学习检测算法发挥高性能的基础,

公开的 HRSC2016 样本集舰船大部分处于海域港口, 船舶类型多为军舰、大中型船只, 本文构建的舰船数据集包含全球范围内各海域, 舰船目标的场景、种类更加丰富。样本集的 3 893 张影像来源于 Arcgis online 的 world imagery, 分辨率达到 0.3 m 到 1 m, 影像尺寸均为 1 000×600, 涵盖了各种复杂背景下不同类型的舰船, 样本数据集整体情况如图 2 所示。

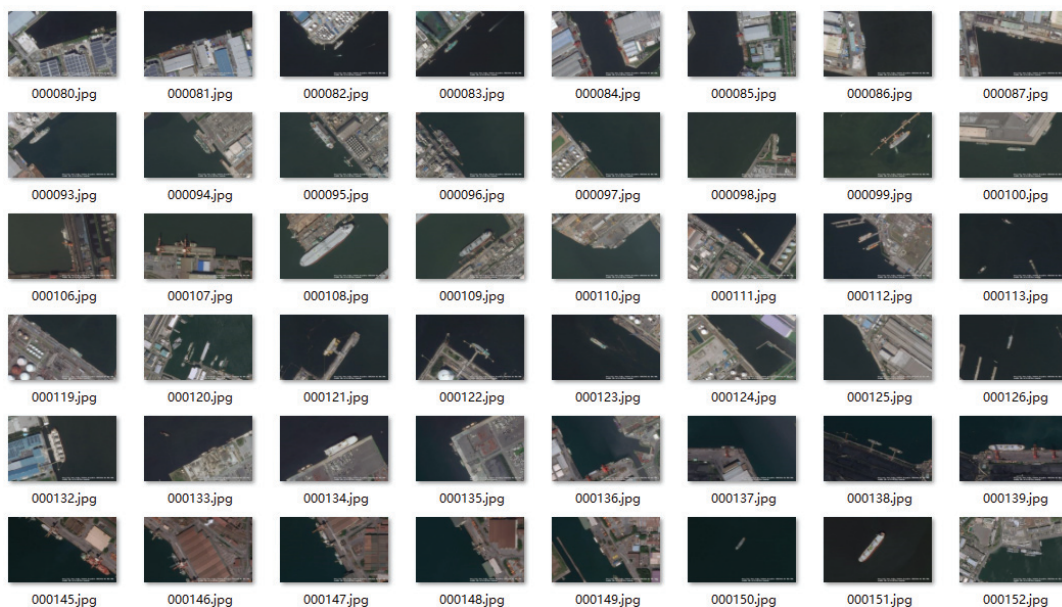


图 2 样本数据集影像示意图  
Fig. 2 Schematic of a sample dataset

使用 Labelimg 工具进行标注, 影像中包含舰船 8 098 只, 舰船大小对应个数统计结果如图 3 所示, 可以发现数据集包含舰船大小、种类丰富, 因此可以较好地深度学习算法性能进行测试。

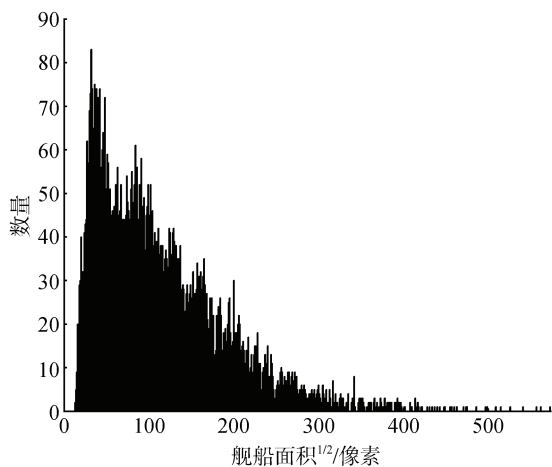


图 3 舰船大小统计示意图  
Fig. 3 Statistical diagram of the ship size

### 2.2 实验结果与分析

将数据集 3 116 张用于训练, 777 张用于测试, 采用该数据集对 Faster RCNN、SSD、RetinaNet、YOLOv3、YOLOv4 进行测试, 为了准确预估算法检测效果, 取置信度为 0.3 对舰船检测数目进行统计。不同置信度下获得多组精确率(precision)和召回率(recall)绘制 PR 曲线, 计算平均精度 mAP 值进行精度评价<sup>[16-17]</sup>。

检测结果统计如表 1 所示, YOLOv4 平均精度达到 91.77%, 远高于其他算法, 能准确检测出大部分船只, 并且误检数较少。YOLOv3、RetinaNet、Faster RCNN 平均精度接近, 均在 83% 以上, 其中 Faster RCNN 能检测出更多的舰船, 但误检偏多。而 SSD 平均精度只有 79.23%, 检测性能较差, 总的检测数目偏少。

选取了几张复杂背景下包含多种尺寸舰船的影像进行测试, 从上到下依次为 a、b、c、d 4 个场景, 影像测试结果如图 4 所示:



表 1 各算法检测结果统计表

Tab. 1 Statistical table of the detection results of each algorithm

	骨干网络	总目标数/个	正确检测数/个	错误检测数/个	mAP/%
Faster RCNN	VGG16	1 647	1 430	616	83.71
SSD	VGG16	1 647	1 073	131	79.23
RetinaNet	ResNet50	1 647	1 212	232	83.98
YOLOv3	Darknet53	1 647	1 333	185	85.75
YOLOv4	CSPDarknet53	1 647	1 465	172	91.77

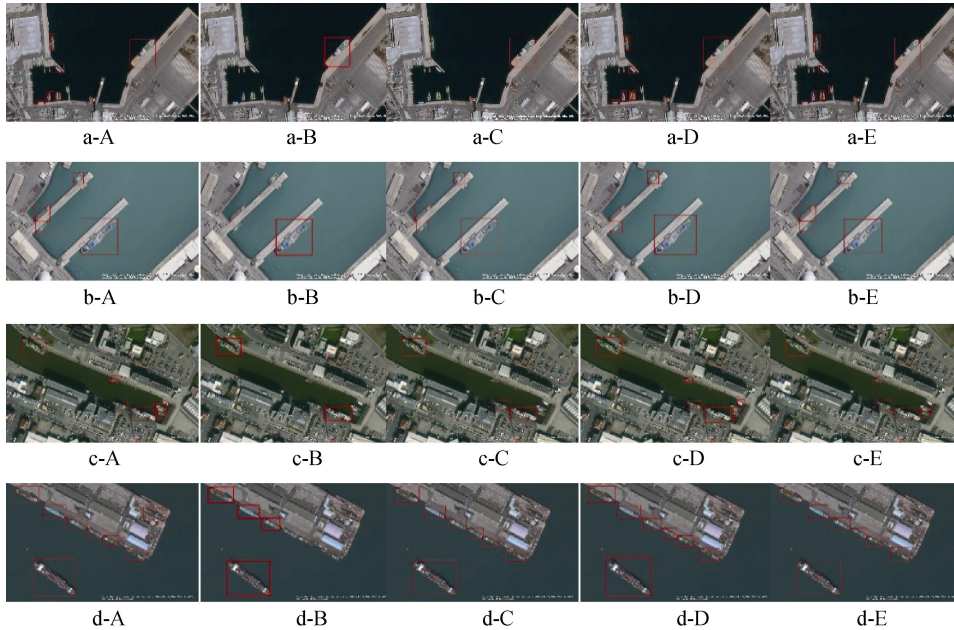


图 4 复杂背景下不同算法检测效果示意图

Fig. 4 Schematic of the detection effect of different algorithms for the case of a complex background

注: a—d: 4 个复杂背景下的影像; A: Faster RCNN; B: SSD; C: RetinaNet; D: YOLOv3; E: YOLOv4

SSD、YOLO 作为单阶段检测器，具有明显的速度优势，都采用多尺度结构，但测试中 YOLOv3、YOLOv4 检测性能明显优于 SSD。这是由于 SSD 网络层数较少，不能对舰船特征信息进行有效提取，同时训练时会受到负样本影响，导致训练时学习特征效果不好，输入尺寸较小也会影响到舰船特征的表达。相比之下，YOLO 系列深层卷积网络结合快捷链路能对特征信息有效提取，背景信息及干扰因素较好区分出来，多层次的结构对小物体也能进行有效检测。YOLOv4 作为一个高效而强大的目标检测算法，其在原有 YOLOv3 目标检测算法的基础上采用多种优化策略，进一步提升了该系列的效果和精度，因而取得不错的效果，图 4 四个场景中的舰船均能准确检测。

Faster RCNN 和 RetinaNet 平均精度较高，检测性能较好，对于图 4b、4d 场景的舰船检测效果较好，

但在图 4a、4c 场景测试中小型舰船的检测效果一般。Faster RCNN 由于其骨干提取网络提取能力不足，RPN 网络采用的是多次卷积后形成的特征图，像素感受野较大，对于面积较小的物体检测效果不好。同时不同大小候选框池化到统一尺寸，对目标特征描述也会存在偏差。RetinaNet 作为单阶段检测方法针对类别不均衡问题做出改进，由于通过层层卷积与融合形成多层金字塔，不同金字塔层锚框大小从  $32 \times 32$  到  $512 \times 512$ ，最小锚框对应原影像  $32 \times 32$  大小的窗口，对于过小目标不能进行有效检测。同时 RetinaNet 测试中误检较少，说明骨干网络 ResNet 结合 FPN 网络形成的多尺度卷积特征金字塔对影像特征能够有效提取，将复杂背景干扰因素很好的区分出来。

作为效果最优的 YOLOv4 在各背景环境下均能准确检测出舰船，然而对于并排舰船，也未能取得理想的效果，如图 5 所示。

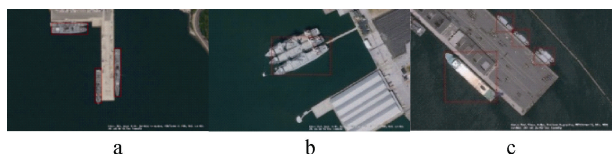


图 5 YOLOv4 并排舰船检测效果示意图

Fig. 5 Schematic of the detection effect of YOLOv4 for side-by-side ships

注: a—c: 3 个包含并排舰船的复杂背景影像

### 2.3 高分二号影像测试

将 5 种模型训练结果对高分二号影像数据进行测试, 本文采用的是 2020 年 7 月 25 日高分二号卫星拍摄青岛沿海某区域的影像数据, 空间分辨率为 1 m, 影像情况如图 6 所示。

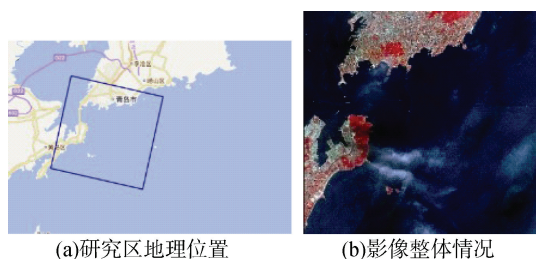


图 6 高分二号影像示意图

Fig. 6 Schematic of a GF-2 image

影像中包含舰船 40 只, 取置信度为 0.3 对舰船数

目进行统计, 并使用品质因数  $Q$  进行定量评价分析:

$$Q = \frac{N_{tt}}{N_{fa} + N_{gt}}$$

式中:  $N_{tt}$  为正确检测的舰船数目;  $N_{fa}$  为虚警数目;  $N_{gt}$  为测试集中真实舰船数目。  $Q$  能同时考虑到检测率和虚警率, 对应值越高表示检测性能越好。测试结果如表 2 所示, 可以看出, YOLOv4 品质因数远高于其他几种算法, 达到 90.48%, YOLOv3、RetinaNet、Faster RCNN 品质因数接近保持在 80% 以上, SSD 品质因数较低, 仅有 70.45%。

表 2 高分 2 号影像测试结果

Tab. 2 Test results of a GF-2 image

	正确检测数/个	错误检测数/个	品质因数/%
Faster RCNN	37	6	80.43
SSD	31	4	70.45
RetinaNet	36	3	83.72
YOLOv3	33	0	82.50
YOLOv4	38	2	90.48

图 7 展示了该高分二号影像中各模型实际测试效果, 可以看出, YOLOv4 几种场景中均能准确检测到舰船, YOLOv3、RetinaNet、Faster RCNN 能检测到较多舰船, 但 Faster RCNN 误检偏多, SSD 准确检测到舰船数较少。

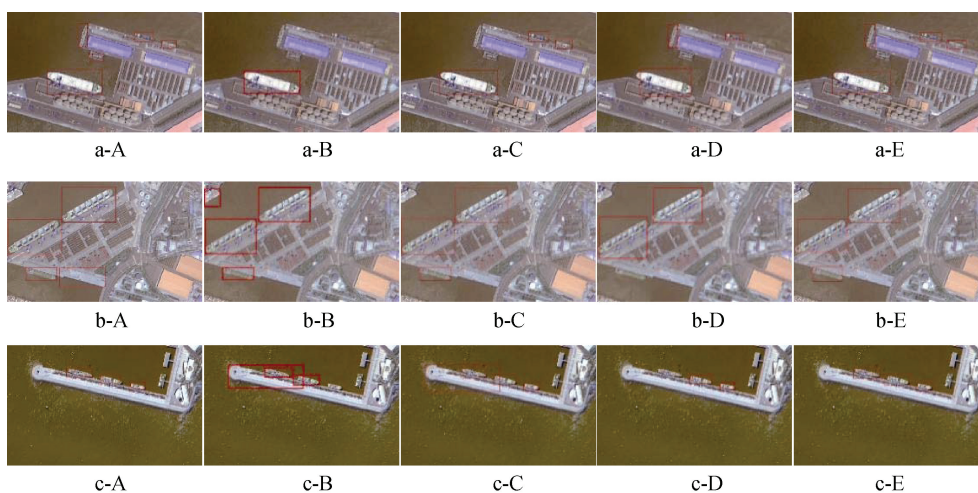


图 7 高分二号影像不同算法检测效果示意图

Fig. 7 Schematic of different algorithm detection results of a GF-2 image

注: a—c: 3 个复杂背景下的影像; A: Faster RCNN; B: SSD; C: RetinaNet; D: YOLOv3; E: YOLOv4

## 3 结论

本文构建了复杂背景下不同类型舰船的数据集,

在对深度学习中的目标检测算法理论总结与分析基础上, 分别采用 Faster RCNN、SSD、YOLOv3、RetinaNet、YOLOv4 五种算法在构建数据集上进行

实验对比, 得出以下结论:

1) YOLOv4 相比其他算法能达到较高的精度, 在 YOLOv3 基础上采用多种优化策略进一步提高了该系列精度。同时作为单阶段检测器具有速度优势, 鲁棒性较好, 该系列深层卷积网络结合快捷链路能对特征信息有效提取, 将背景信息及干扰因素较好区分出来, 多层次结构能有效检测小目标。SSD 效果较差, 浅层卷积网络不能有效地提取目标特征信息, 在复杂背景下对舰船不能有效地检测。

2) 几种算法都存在不同程度的误检, 复杂背景会对目标检测器效果产生影响, 通过优化检测器对特征信息有效提取可以减少误检。同时对于密集排列的舰船, 几种目标检测算法都不能很好地解决, 效果不佳, 如何对并排舰船进行有效提取是一个重要研究方向。

#### 参考文献:

- [1] 王彦情, 马雷, 田原. 光学遥感图像舰船目标检测与识别综述[J]. 自动化学报, 2011, 37(9): 1029-1039.  
WANG Yanqing, MA Lei, TIAN Yuan. Overview of ship target detection and recognition in optical remote sensing images[J]. Acta Automatica Sinica, 2011, 37(9): 1029-1039.
- [2] ELDHUSET K. An automatic ship and ship wake detection system for spaceborne SAR images in coastal regions[J]. IEEE Transactions on Geoscience and Remote Sensing, 1996, 34(4): 1010-1019.
- [3] WANG Y, LIU H. A hierarchical ship detection scheme for high-resolution SAR images[J]. IEEE Transactions on Geoscience and Remote Sensing, 2012, 50(10): 4173-4184.
- [4] 尹雅, 黄海, 张志祥, 等. 基于光学遥感图像的舰船目标检测技术研究[J]. 计算机科学, 2019, 46(3): 82-87.  
YIN Ya, HUANG Hai, ZHANG Zhixiang, et al. Research on ship target detection technology based on optical remote sensing image[J]. Computer Science, 2019, 46(3): 82-87.
- [5] ZHU Changren, ZHOU Hui, WANG Runsheng, 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.
- [6] 彭敬涛. 光学遥感图像舰船目标检测技术的研究[D]. 上海: 东华大学, 2017.  
PENG Jingtao. Research on ship target detection technology in optical remote sensing image[D]. Shanghai: Donghua University, 2017.
- [7] 王慧利, 朱明, 蔺春波, 等. 光学遥感图像中复杂海背景下的舰船检测[J]. 光学精密工程, 2018, 26(3): 723-732.  
WANG Huili, ZHU Ming, LIN Chunbo, et al. Ship detection in complex sea background in optical remote sensing image[J]. Optics and Precision Engineering, 2018, 26(3): 723-732.
- [8] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2017, 39(6): 1137-1149.
- [9] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]//computer vision and pattern recognition, Columbus, 2014: 580-587.
- [10] GIRSHICK R. Fast R-CNN[C]//2015 IEEE International Conference on Computer Vision (ICCV), Santiago Chile, December 7-13, 2015: 1440-1448.
- [11] REDMON J, FARHADI A. YOLOv3: an incremental improvement[C]//IEEE Conference on Computer Vision and Pattern Recognition, 2018. <https://arxiv.org/pdf/1804.02767.pdf>.
- [12] LIU W, ANGELOV D, ERHAN D, et al. SSD: single shot multiBox detector[C]//European Conference on Computer Vision, Amsterdam, Netherlands, 2016: 21-37.
- [13] LIN T, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection[C]//Proceedings of the IEEE International Conference on Computer Vision, Venice, 2017: 2999-3007.
- [14] BOCHKOVSKIY A, WANG C, LIAO H. YOLOv4: Optimal Speed and Accuracy of Object Detection[C]//International Conference on Computer Vision and Pattern Recognition, 2020. <https://arxiv.org/abs/2004.10934>.
- [15] LIU Z, YUAN L, WENG L, et al. A high resolution Optical satellite image dataset for Ship recognition and some new baselines[C]//Proceedings of the 6<sup>th</sup> International Conference on Pattern Recognition Applications & Methods-ICPRAM, 324-331, Porto, Portugal, 2017.
- [16] 余东行, 郭海涛, 张保明, 等. 级联卷积神经网络的遥感影像飞机目标检测[J]. 测绘学报, 2019, 48(8): 1046-1058.  
YU Dongxing, GUO Haitao, ZHANG Baoming, et al. Aircraft target detection in remote sensing image based on cascade convolution neural network[J]. Acta Geodae-tica et Cartographica Sinica, 2019, 48(8): 1046-1058.
- [17] 李宸尧, 郭海涛, 马东洋, 等. 深度学习遥感影像油罐检测算法精度对比分析[J]. 海洋测绘, 2020, 40(2): 53-56.  
LI Chenyao, GUO Haitao, MA Dongyang, et al. Accuracy comparison and analysis of oil tank detection algorithm based on deep learning remote sensing image[J]. Hydrographic Surveying and Charting, 2020, 40(2): 53-56.



# Comparative analysis of ship detection algorithms based on deep learning in optical remote sensing images

XIA Wen-hui, WAN Jian-hua, ZHENG Hong-xia, XU Ming-ming, QU Chuan-ping  
(China University of Petroleum, Qingdao 266580, China)

**Received:** Nov. 8, 2020

**Key words:** remote sensing; target detection; ship; deep learning

**Abstract:** Ship target detection is an important means to supervise a marine environment and protect maritime rights and interests. A target detection algorithm based on deep learning can maintain good performance in a complex environment. This study describes the construction of a dataset containing 3, 893 images, covering different types of ships in a complex background. Based on this dataset, the performances of Faster RCNN, SSD, RetinaNet, YOLOv3, and YOLOv4 algorithms are examined. The results show that the mean average precision of YOLOv4, YOLOv3, RetinaNet, and Faster RCNN are above 83%. Furthermore, YOLOv4 reached 91.77%, Faster RCNN produced more false detections, and SSD achieved the lowest mean average precision, only 79.23%, with a low total number of ship detections. The training results of the five models are tested on a GF-2 image, with good detection results obtained; these results have certain guiding significance for the future theoretical research of ship detection.

(本文编辑: 赵卫红)