

基于深度学习的海上遥感图像目标检测算法综述

杨宽¹ 吴利强¹ 王玉志¹ 王万历²

1 武警承德支队 2 航天工程大学

DOI:10.12238/gmsm.v7i7.1890

[摘要] 目标检测是海上遥感图像目标信息提取的关键环节,广泛应用于海上舰船检测、环境监测、岛礁变化等领域。近年来,随着深度学习的快速发展,基于深度学习的海上遥感图像目标检测准确性和实时性得到显著提升。本文首先介绍了基于深度学习的目标检测相对于传统的目标检测算法的优势,然后梳理了当前主流的目标检测算法发展演变和应用特点,接着介绍了当前目标检测的最新成果以及未来的发展趋势,结合海上遥感图像特点,归纳了一些对于海上遥感图像的目标检测研究成果。最后总结未来海上遥感图像面临的问题并提出相对应的解决思路。

[关键词] 深度学习; 目标检测; 海上; 遥感图像; 算法

中图分类号: TJ765.4 **文献标识码:** A

A review of target detection algorithms for maritime remote sensing images based on deep learning

Kuan Yang¹ Liqiang Wu¹ Yuzhi Wang¹ Wanli Wang²

1 Chengde Detachment of the Armed Police 2 Aerospace Engineering University

[Abstract] Target detection is a key link in the extraction of target information from marine remote sensing images, which is widely used in the fields of marine ship detection, environmental monitoring, and island and reef changes. In recent years, with the rapid development of deep learning, the accuracy and real-time performance of target detection in marine remote sensing images based on deep learning have been significantly improved. This paper first introduces the advantages of deep learning-based object detection algorithms compared with traditional object detection algorithms, then sorts out the development and application characteristics of the current mainstream object detection algorithms, and then introduces the latest achievements and future development trends of target detection. Finally, the problems faced by marine remote sensing images in the future are summarized and corresponding solutions are proposed.

[Key words] deep learning; object detection; Maritime; remotely sensed imagery; algorithm

引言

近年来,随着航天遥感技术的快速发展,遥感图像在海洋资源开发、海上交通管理、海洋环境保护和海上灾害监测等方面的作用日益突出。目标检测算法对于航天遥感图像在海上应用具有重要意义。

航天遥感图像目标检测从发展脉络来说分为传统目标检测算法与基于深度学习的目标检测算法。传统目标检测只能处理一些简单场景的目标,对于海上遥感图像中港口背景复杂、舰船方向不定、目标多样、数据不均衡等特点,传统目标检测算法检测识别能力较为薄弱。

基于深度学习的目标检测算法是一种通过模拟人脑神经元的原理,通过对图像、视频、文字等数据的自动学习,从而实现对图像与视频目标的识别、标记、提取以及对文字

的理解、分析、应答。近十年来,基于深度学习的目标检测技术在无人驾驶、生物医疗、安防、军事、航天遥感等方面都有广泛应用,相比于传统的目标检测方法,基于深度学习的目标检测方法,有更高的准确性、更强的泛化能力、更好的自动迭代能力。

1 基于深度学习的目标检测算法

目前,基于深度学习的目标检测算法主要分为三个发展方向:一是以R-CNN^[1-3]为代表的两阶段目标检测算法模型,先通过第一阶段检测对于可能出现目标的区域选取候选区,然后再通过第二阶段的检测对候选区域中的目标进行精确判别,两阶段目标检测耗时长精确度高。二是以YOLO^[4-6]、SSD^[7]为代表的单阶段目标检测模型,直接通过一次检测筛选出目标,检测速度快,精度低。三是以Transformer^[8]为代表自注意力机制目标

检测模型,通过多头注意力机制,可以让模型去关注不同维度的信息,提高目标检测的精度。

1.1 R-CNN

Ross Girshick^[1]等人于2014年提出了R-CNN算法,首先通过候选区选择性搜索(Selective Search)方法对可能包含物体的区域选取多个候选框,然后通过卷积(Convolutional Neural Networks, CNN)对每一个候选框生成特征图,最后再通过支持向量机(Support Vector Machines, SVM)分类器对特征图进行分类处理,通过交并比(Intersection over Union, IoU)对比以及非极大值抑制(Non-Maximum Suppression, NMS)的方法将多余的边框去掉,实现对目标物体的检测。由于训练过程中步骤繁琐,存在训练速度较慢的问题。在2015年,Ross Girshick^[2]等人又提出了Fast RCNN算法,通过将卷积对每个候选区计算改变成为对整张图计算,采用感兴趣区域池化(Region of Interest Pooling, ROI pooling)进行特征的尺寸变换,利用softmax代替SVM分类器等手段,使Fast RCNN训练速度提升了9倍,测试速度提升了213倍,在VOC 2012数据集上mAP提升到68.4%。但还存在候选框选取过程中耗费过多时间的问题。2016年Ross Girshick等人又提出了更快的Faster RCNN^[3]网络模型,他提出了区域生成网络(Regional Proposal Network, RPN)方法,将传统的候选框提取方法替换成网络训练。通过RPN网络生成候选框,同时把图像经过CNN得到一个特征图(feature map),把候选框在feature map的区域经过RoI Pooling生成特征向量,通过对特征向量进行数据处理来实现预测。Faster RCNN解决了多尺度、小目标问题,在Fast RCNN基础上实现了预测精度的提高。二阶段目标检测算法模型适用于精度和鲁棒性要求高的复杂场景目标检测,由于计算量较大,不适用于时效性强的场景。

1.2 YOLO

为了解决二阶段目标检测算法实时性检测效果差等问题,Redmon^[4]等人从2016年开始先后提出了YOLOv1、YOLOv2、YOLOv3算法,将图片划分为 $N \times N$ 个网格,每一个网格去检测那些中心落在该网格上的目标,同时预测多个边界框及边界框的置信度,通过NMS只留取置信度最大的边界框,再通过预测的边界框与标注的边界框对比IoU计算Loss(损失函数)值,通过后向反馈进行参数学习,成功解决了二阶段目标检测速度慢的问题。随后BochKovskiy等人于2020年提出了YOLOv4模型,YOLOv4以CSPDarknet 53^[5]作为主干网络,采用Mish激活函数,添加了Dropblock模块、FPN^[6]模块,极大的融合了当时的各种优化技巧,大大提高了模型的预测精度,降低了训练成本,具有里程碑意义。在YOLOv4推出2个月之后,Glen Jocher便推出了YOLOv5模型,在其中添加了自适应锚框,在每次训练数据集之前,都会通过K-means算法自动计算最适合此数据集的锚框。YOLOv5采用了Focus结构,对原始图像进行切片操作,此操作可以有效的提高运算速度,减少浮点运算数目。美团于2022年6月份推出YOLOv6,采用EfficientRep替换CSPDarknet,Neck通过结合Rep和PAN构建了Rep-PAN,借鉴了RepVGG思想,构造了RepBlock模块,并且将

其加入Neck中,进一步对YOLOv5进行了优化。BochKovskiy等人于2022年7月份又推出了YOLOv7版本,主要借鉴了YOLOv5的组织结构,和YOLOX中的动态标签分配策略,提出ELAN和E-ELAN模型结构,在MS-coco数据集中训练精度可达到56.8%,检测速度最高可达160FPS,相对于前几代版本的YOLO系列有了大幅度提升。2023年1月,Edison G推出了YOLOv8,将耦合头变成解耦头,将Anchor-Based变成Anchor-Free,YOLOv8使用了TOOD的Task Aligned Assigner分配策略,解决了样本分配不均衡的问题,减少了计算量,提高了检测速度,增强了YOLO实时检测性能。

1.3 SSD

Wei Liu^[7]等人于2016年提出了SSD算法,SSD结合了YOLO中的回归思想和Faster-RCNN中的Anchor机制,SSD首先对图片进行随机裁剪、随机翻转、光度扭曲等数据增强变换,再经过VGG16主干网络进行特征提取,得到6个特征层,分别将特征图通过Detector&Classifier的作用对目标进行分类判断,识别出图像中的目标。SSD算法缺陷对于小物体的检测效果不好,难以实现分类预测的效果。同时需要人工设置每一层的预测框大小,针对SSD存在的问题,有学者相继提出DSSD、RSSD、RefineDet等算法。DSSD为了解决浅层网络的特征信息不足的问题,将浅层与深层的特征按通道维度进行拼接,最终实现对小目标的检测性能的提升。RSSD算法针对原始SSD算法中没有考虑特征图之间的关联性问题,出现多个特征图的预测框与一个真实框相匹配的情况以及小物体检测效果差等问题,通过特征融合,将浅层的特征图与下层特征图进行通道拼接,在提高模型的目标检测精度的同时减少了参数量。RefineDet为了解决SSD漏检的问题,在SSD基础上结合Faster RCNN网络中边框由粗到细的思想,由锚框再精细化模块(Anchor Refinement Module, ARM)、传输连接模块(Transfer Connection Block, TCB)与目标检测模块(Object Detection Module, ODM)这3个模块组成,先通过网络粗略的定位到固定框,然后再对网络进行细化微调。同时采用FPN提高对小物体的检测效果,进而达到改善模型,提升检测能力的效果。

1.4 Transformer

2017年6月谷歌团队Ashish Vaswani^[8]等人提出了Transformer模型,用于自然语言处理(Natural Language Processing, NLP)领域,大大刺激了NLP领域的发展,Transformer整个网络采用注意力机制代替了CNN和循环神经网络(Recurrent Neural Network, RNN),Transformer的核心模块是编码-解码(Encoder-Decoder)模块,在Encoder中会通过Attention(Q, K, V)处理产生一个特征向量Z,特征向。

量会被送入到下一个模块,即前馈神经网络(Feed Forward Neural Network, FFN),经过FFN处理后的特征向量进入Decoder(Decoder比Encoder多一个Encoder-Decoder Attention层),经过多轮的Encoder和Decoder迭代最后产生特征向量进行目标检测。Transformer提供了加入自注意力机制的思路,成功解决了像素点之间相互关联的问题,提高了目标检测算法对图像的全局感知能力。

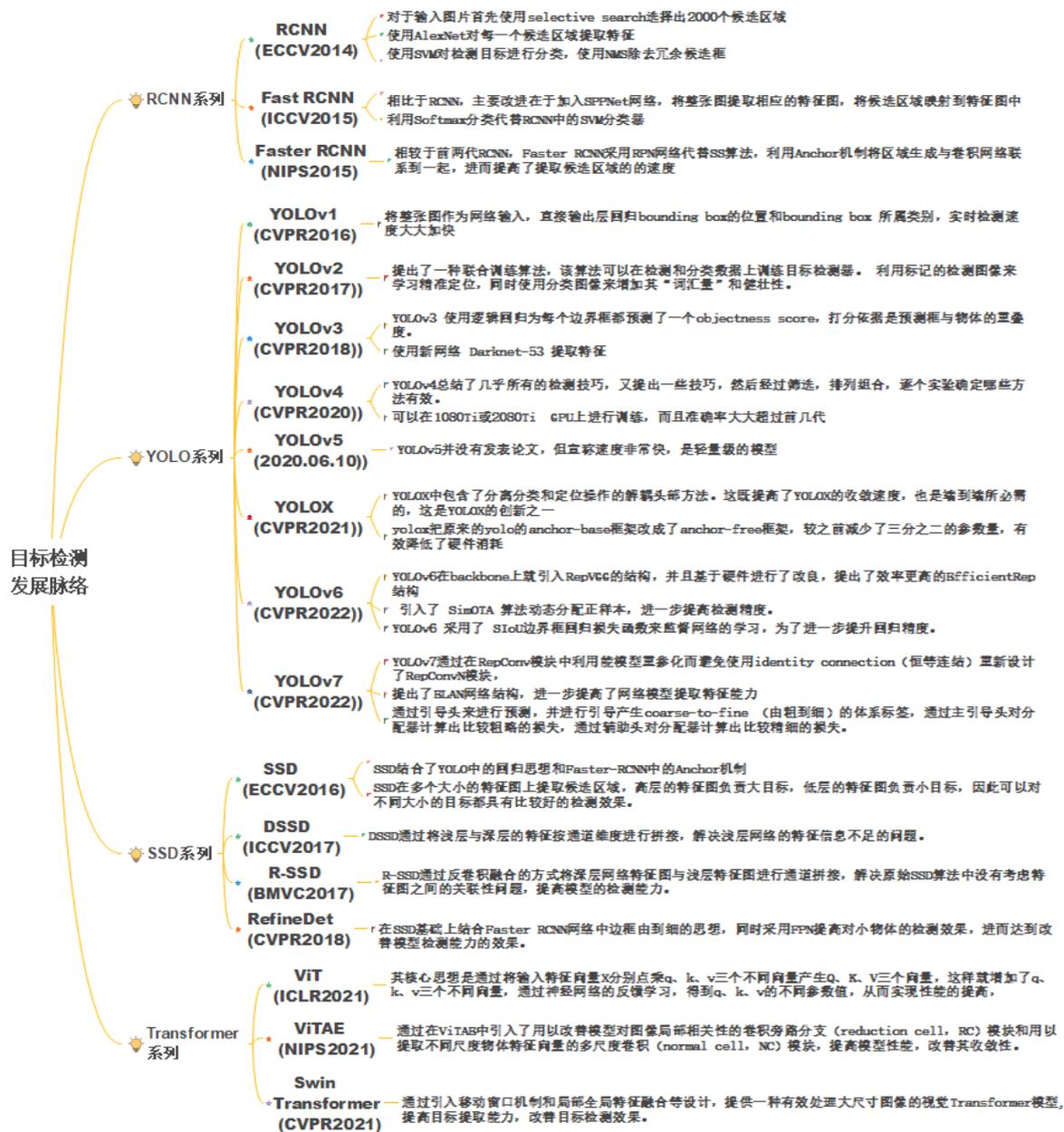


图1 主流目标检测算法模型

2 基于深度学习的海上目标检测

由于遥感影像数据量大且易受海上天气及光照条件的影响, 对于海面海上雾气、云层等自然环境的干扰问题也会使得遥感影像中的目标检测带来困难, 遥感影像在海上舰船检测过程中存在检测区域冗余、方向不确定、尺度变化大、尺度小、特征量少、场景复杂等问题。

针对海上遥感影像检测区域冗余、检测目标方向不确定、尺度变化大等问题, 2018年6月中国科学院的杨雪^[9]等人提出一种旋转密集特征金字塔网络 (Rotation Dense Feature Pyramid Networks, R-DFPN) 框架, 用于解决船舶尺寸狭长的问题, 设计了

旋转框来限定最小目标区域, 减少冗余区域, 提高了召回率。北京理工大学的 Qi Ming^[10]等人提出了对于任意方向的舰船目标, 动态锚框学习 (Dynamic Anchor Learning, DAL) 策略。该方法利用新定义的匹配度来综合评估锚点的定位潜力, 以此来合理地进行标签分配, 提高标签分配效率。

针对光学遥感影像目标尺寸小、特征量少、海上雾气、云层干扰等问题, 文献^[11]提出通过对海上遥感影像校正、改进 IoU 阈值的设计、多尺度特征融合等手段, 实现对海上舰船目标精度的提高。杭州电子科技大学的孙乔^[12]等人提出一种基于全卷积网络 (Region based Fully Convolutional Networks, RFCN) 网

络模型,通过联合变换、柔性非极大值抑制、多尺度特征融合等方式提高对小尺寸与重叠舰船目标的检测能力。空军工程大学成倩^[13]等人通过对聚合网络(Path Aggregation Network, PANet)结构进行优化,加入语义信息增强模块(Semantic Information Enhancement module, SIEM)层使得浅层网络变成深层网络,从而扩大了特征图的感受野,提高特征表达能力。

针对高分辨率光学遥感影像背景复杂、舰船尺度多样等问题,军事科学院的周旗开^[14]等人提出基于高效通道注意力(efficient channel attention, ECA)模和双向特征金字塔网络(bidirectional feature pyramid network, BIFPN)模块的目标检测网络模型,通过将ECA模块融入YOLOv5s网络结构中提升网络提取特征能力,利用加权BIFPN代替路径增强网络(Path Aggregation Network, PANet)特征金字塔,加强网络特征融合的能力,在模型轻量化的同时实现对舰船目标的高速与高精度检测。

3 总结与展望

基于深度学习的目标检测对于解决海上遥感影像中目标识别问题是一个可行的解决方案,近年来,基于深度学习的目标检测虽然在海上遥感影像中取得了一定的应用效果,但是针对海上遥感图像特点还存在着一些问题尚未解决。

(1)目标检测轻量化问题。为了应对突发事件或应急响应的实战需求,需要在嵌入式设备中运行轻量化目标检测算法以保证实时性。

(2)多源数据目标检测问题。单一数据源的目标识别全天候、全天候应对能力不足。光学遥感影像数据量大且易受海上天气及光照条件的影响,SAR影像会易受到斑点环境噪声干扰。多源数据目标检测可以提高对不同环境的适应能力,增强目标检测的鲁棒性。

(3)海上目标尺度、形状不定问题。遥感影像的目标尺度大小不一、形状比例变化大、方向不确定等问题增加检测难度,方向自适应算法、形状自适应锚框等成为有效解决海上目标方向、尺度不定等问题的重要突破口。

(4)小目标难以检测问题。遥感影像由于成像设备距离物体较远,导致分辨率较低,出现大规模的小目标时很容易出现漏检的问题。数据增强、时空语义关联、多尺度特征提取等算法可以有助于小目标的检测。

本文通过对当前几种主流的目标检测算法进行梳理,分析当前主流目标检测算法的优缺点,然后结合深度学习在目标检测中的最新研究成果和一些专家学者针对海上遥感影像目标检测研究,最后总结海上目标检测所面临的问题,并针对不同问题提出了相对应的研究建议。

[参考文献]

[1] Ross G, Jeff D, Trevor D, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]. Proceedings of the IEEE conference on computer vision and pattern recognition.2014:580-587.

[2] Girshick R. Fast r-cnn[C]. Proceedings of the IEEE international conference on computer vision.2015:1440-1448.

[3] Ren S, He K, Girshick R, et al. Faster r-cnn: Towards real-time object detection with region proposal networks[J]. Advances in neural information processing systems,2015,28.

[4] Redmon J, Divvala S, Girshick R, et al. You only look once: Unified, real-time object detection[C]. Proceedings of the IEEE conference on computer vision and pattern recognition. 2016:779-788.

[5] Wang C Y, Liao H Y M, Wu Y H, et al. CSPNet: A new backbone that can enhance learning capability of CNN[C]. Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops.2020:390-391.

[6] Lin T Y, Dollár P, Girshick R, et al. Feature pyramid networks for object detection[C]. Proceedings of the IEEE conference on computer vision and pattern recognition.2017: 2117-2125.

[7] Liu W, Anguelov D, Erhan D, et al. Ssd: Single shot multibox detector[C]. European conference on computer vision. Springer, Cham,2016:21-37.

[8] Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need[J]. Advances in neural information processing systems, 2017,30.

[9] Yang X, Sun H, Fu K, et al. Automatic ship detection in remote sensing images from google earth of complex scenes based on multiscale rotation dense feature pyramid networks [J]. Remote Sensing,2018,10(1):132.

[10] Ming Q, Zhou Z, Miao L, et al. Dynamic anchor learning for arbitrary-oriented object detection[C]. Proceedings of the AAAI Conference on Artificial Intelligence.2021,35(3): 2355-2363.

[11] 张玉莲.光学图像海面舰船目标智能检测与识别方法研究[D].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),2021.

[12] 孙乔.基于深度卷积神经网络的舰船目标识别技术研究[D].杭州电子科技大学,2019.

[13] 成倩,李佳,杜娟.基于YOLOv5的光学遥感图像舰船目标检测算法[J/OL].系统工程与电子技术:1-9[2022-11-28].

[14] 周旗开,张伟,李东锦,牛福.基于改进YOLOv5s的光学遥感图像舰船分类检测方法[J].激光与光电子学进展,2022,59(16): 476-483.

作者简介:

杨宽(1991—),男,汉族,河北沧州人,硕士研究生,研究方向:目标识别技术,计算机视觉技术。