

基于改进 YOLO v2 的船舶目标检测方法

于 洋 李世杰 陈 亮 刘韵婷

(沈阳理工大学自动化与电气工程学院 沈阳 110159)

摘 要 针对船舶图像目标检测中存在的小目标检测准确率低、系统鲁棒性差的问题,提出一种改进的 YOLO v2 算法对船舶图像目标进行检测。通过目标框维度聚类、网络结构改进、输入图像多尺度变换等方法对传统 YOLO v2 算法进行改进,使其能够更好地适应船舶目标检测任务。测试结果表明,在输入图像尺寸为 416×416 时,该算法的平均精确率(mean Average Precision, mAP)达到 79.1%,检测速度为 64 帧/s(Frames Per Second, FPS)。所提方法可满足实时检测的需要,且具有小目标检测精度高、鲁棒性强的特点。

关键词 船舶目标检测,目标检测,卷积神经网络,改进 YOLO v2

中图分类号 TP183 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2019.08.055

Ship Target Detection Based on Improved YOLO v2

YU Yang LI Shi-jie CHEN Liang LIU Yun-ting

(School of Automation and Electrical Engineering, Shenyang Ligong University, Shenyang 110159, China)

Abstract Aiming at the problem of low target detection accuracy and poor system robustness in ship image target detection, an improved YOLO v2 algorithm was proposed to detect ship image targets. The traditional YOLO v2 algorithm is improved by clustering the target frame dimension, optimizing the network structure, multi-scale transformation of input image, so as to better adapt to the ship target detection task. The test results show that the mean Average Precision (mAP) of the algorithm is 79.1% when the input image size is 416×416 , and the detection speed is 64 frames per second (FPS), which can satisfy the real-time detection and exhibit high precision and strong robustness for small target detection.

Keywords Ship target detection, Target detection, Convolutional neural network, Improved YOLO v2

1 引言

近年来,邻国侵犯我国海域^[1]、非法捕捞等一系列海上安全问题频发,海洋监管任务日趋严峻,因而船舶目标检测技术再次成为研究热点。目前,常见的船舶目标检测方法主要分为非成像方法和成像方法^[2]。非成像方法利用雷达^[3]、声呐等技术获取船舶目标的相关信号,通过信号处理技术和模式识别技术对船舶目标进行检测^[4]。这类方法易受水下环境干扰,且因通信传输信道复杂,对船舶小目标的检测效果较差。成像方法主要利用红外^[5]、SAR^[6]和卫星遥感^[7]等技术,通过图像处理方法对船舶目标进行检测^[8],但浓雾、波浪噪声等外界环境对该类方法影响较大,对船舶小目标的检测误识率高。

众所周知,深度学习是机器学习领域中的变革性技术,尤其是在计算机视觉领域。近年来,凭借卷积神经网络(Convolutional Neural Network)在提取图像高层特征上的优势,深度学习在目标检测领域取得了巨大成功^[9]。例如,2015年

Girshick^[10]和 Ren 等^[11]分别提出了 Fast R-CNN(Fast Region-based Convolutional Neural Network)和 Faster R-CNN(Faster Region-based Convolutional Neural Network)算法,它们在 PASCAL VOC2012 数据集上的 *mAP* 分别提升至 68% 和 73%。2016 年 Redmon 等^[12]提出的 YOLO(You Only Look Once)算法,虽检测效果不及 Faster R-CNN,但检测速度较 Faster R-CNN 提升了 10 倍,实现了实时检测(45 FPS),且其将分类和定位整合在一起的思想为后续的研究提供了新思路。在 YOLO 算法的基础上,2016 年 Liu^[13]和 Redmon^[14]等相继提出了 SSD(Single Shot MultiBox Detector)和 YOLO v2 算法,它们在提升检测速度的同时,取得了令人满意的检测效果,其中 YOLO v2 算法在目标检测领域取得了最佳检测效果。

本文采用改进的 YOLO v2 算法对船舶目标进行检测,通过深度神经网络对船舶图像进行特征提取。针对船舶目标检测过程中易受外界干扰、小目标检测效果差、小目标误识率高

到稿日期:2018-06-05 返修日期:2018-10-22 本文受国家重点研发计划(2017YFC0821001),国家自然科学基金(61373089),辽宁省自然科学基金(201602652),辽宁省教育厅基本科研項目(LG201707)资助。

于 洋(1963—),男,硕士,教授,主要研究方向为智能检测与控制、故障诊断与监控技术等,E-mail:781784300@qq.com(通信作者);李世杰(1994—),女,硕士,主要研究方向为智能检测与信息处理、人工智能;陈 亮(1979—),男,博士,主要研究方向为机器学习、嵌入式仪表与信息处理技术等;刘韵婷(1983—),女,博士,主要研究方向为人工神经网络、无线传感器网络与数据分析等。

的问题,通过目标框维度聚类、网络结构改进、输入图像多尺度变换等方法对传统 YOLO v2 算法进行改进,使改进算法更好地适应船舶目标检测任务,协助船检人员及早发现潜在的危险并做出正确判断。

2 YOLO v2 原理

2.1 特征提取网络 Darknet-19

在 YOLO v2 算法问世前,大多数检测框架将 VGG-16 网络^[15]作为基本特征提取器。VGG-16 网络分类准确,但模型复杂、计算量大。YOLO 的检测框架结构类似于 Google net^[16],检测速度比 VGG-16 快,但准确性比 VGG-16 略差。因此,Redmon 提出了一种新的分类网络 Darknet-19 作为 YOLO v2 的基础网络。

Darknet-19 由 19 个卷积层和 5 个最大池化层组成。与 VGG-16 网络类似,Darknet-19 网络中大多使用 3×3 的卷积核,并在每次池化之后使通道数量加倍。Darknet-19 也借鉴了 Network in Network^[17]的思想,使用了全局平均池化,将 1×1 的卷积核置于 3×3 的卷积核之间,用来压缩特征;在每一个卷积层后使用批量归一化方法稳定网络模型训练,加速收敛并泛化网络模型。在 Imagenet 数据集上,Darknet-19 网络的分类达到了 72.9% 的 Top-1 准确率和 91.2% 的 Top-5 准确率,优于 Alexnet^[18] 和 VGG-16 网络在 Imagenet 数据集上的分类表现。3 种网络的性能如表 1 所列,综合来看,Darknet-19 网络的性能更优异。

表 1 Darknet-19, VGG-16 和 Alexnet 网络的性能

Table 1 Performance of Darknet-19, VGG-16 and Alexnet network (单位: %)

Model	Top-1	Top-5
Alexnet	53.8	84.7
VGG-16	70.5	90
Darknet-19	72.9	91.2

2.2 YOLO v2 算法结构及其改进

YOLO v2 算法对 Darknet-19 网络进行调整,删除了 Darknet-19 网络的最后一个卷积层,在后面增加了 3 个卷积核尺寸为 3×3 、卷积核数量为 1024 的卷积层,在这 3 个卷积层的最后一层的后面增加了一个卷积核尺寸为 1×1 、输出维度为检测所需维度的卷积层,并在检测网络中加入了直通层 (passthrough layer),得到待检测目标的细粒度特征 (fine-grained features)。与 YOLO 使用卷积层之后的全连接层来直接预测 bounding boxes 的坐标不同,YOLO v2 去掉了全连接层,借鉴了 Faster R-CNN 的做法,使用 anchor box 预测 bounding boxes。不同于 Faster R-CNN 手动精心挑选 anchor box,YOLO v2 使用 K-means^[19] 方法对数据集中的目标框进行维度聚类,得到 anchor box 的尺寸和数量。YOLO v2 沿用了 YOLO 的损失函数,将定位误差 (localization error) 和分类误差 (classification error) 整合在一起,采用不同的权值来平衡两种误差对模型稳定性的影响。损失函数的公式为:

$$loss = \lambda_{\text{coord}} \sum_{i=0}^2 \sum_{j=0}^B \prod_{ij} [(x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2] + \lambda_{\text{coord}} \sum_{i=0}^2 \sum_{j=0}^B \prod_{ij} [(\sqrt{w_i} - \sqrt{\hat{w}_i})^2 + (\sqrt{h_i} - \sqrt{\hat{h}_i})^2] +$$

$$\sum_{i=0}^2 \sum_{j=0}^B \prod_{ij} (C_i - \hat{C}_i)^2 + \lambda_{\text{noobj}} \sum_{i=0}^2 \sum_{j=0}^B \prod_{ij} (C_i - \hat{C}_i)^2 + \sum_{i=0}^2 \prod_{i \in \text{classes}} \sum_{ij} (p_i(c) - \hat{p}_i(c))^2 \quad (1)$$

其中, λ_{coord} 代表定位预测误差的权值系数, λ_{noobj} 代表分类预测误差的权值系数, s 代表输入图像区域被划分的行数或列数, B 代表划分得到的每个格子 (grid) 中的边界框 (bounding box) 数目, \prod_{ij} 用于判断第 i 个网格中的第 j 个 bounding box 是否负责某个目标 (object), x_i 和 y_i 是指当前格子预测得到的物体的 bounding box 的中心位置的坐标, w 和 h 是 bounding box 的宽度和高度, \hat{x}_i 和 \hat{y}_i 为每个目标的真实标注框的中心位置坐标, \hat{w}_i 和 \hat{h}_i 为标注框的宽度和高度, C_i 和 \hat{C}_i 分别代表真实标注框和边界框间的 IOU 值、bounding box 含有物体的置信度, $p_i(c)$ 代表预测该格子中包含某一类目标的概率, $\hat{p}_i(c)$ 代表该格子属于某一类目标的真实概率。

YOLO v2 的网络只用到卷积层和池化层,故可以动态调整输入图像的尺寸,每隔几次迭代就微调网络的输入尺寸,这种策略使得 YOLO v2 对于不同尺寸的图像有较好的检测鲁棒性。在检测低分辨率的图像时,其精度略低但速度非常快,在输入图像尺寸为 288×288 时,检测速度达到了 90FPS,而其 mAP 几乎与 Faster R-CNN 算法的 mAP 相同;在检测大尺寸图像时,在 PASCAL VOC2007 数据集上的 mAP 达到了 78.6%。

利用评价参数 mAP 和 FPS 进行主流检测框架的对比。 mAP 的计算公式为:

$$mAP = \frac{\sum \text{Average Precision}}{N(\text{classes})} \quad (2)$$

其中, Average Precision 代表平均准确率, $N(\text{classes})$ 代表物体类别的个数。

主流检测框架的对比结果如表 2 所列。

表 2 主流检测框架的对比

Table 2 Comparison of mainstream detection frameworks

Detection Frameworks	$mAP/\%$	检测速度
Fast R-CNN	70.0	0.5
Faster R-CNN VGG-16	73.2	7
Faster R-CNN Resnet	76.4	5
YOLO	63.4	45
SSD500	76.8	19
YOLO v2 288×288	69.0	91
YOLO v2 416×416	76.8	67
YOLO v2 544×544	78.6	40

3 改进 YOLO v2 算法

YOLO v2 算法虽然在 PASCAL VOC2007 数据集和 COCO 等标准检测任务上取得了目前最佳的检测效果,但对于小目标物体的检测效果却不尽人意。为了弥补 YOLO v2 算法的不足,针对船舶目标检测问题,对 YOLO v2 算法主要做如下改进。

1) 目标框维度聚类:通过 K-means 算法对自制船舶数据集的目标框进行维度聚类,重新定义 anchor box 的数目和大小。

2)网络结构的改进:针对降低船舶小目标检测误识率的问题,对YOLO v2算法的基础网络Darknet-19进行改进,减少Darknet-19的一些冗余网络层,在网络中增加了直通层,增强对船舶小目标的检测。

3)输入图像多尺度变换:微调网络的输入图像尺寸,训练时每经过10 epoch就改变模型的输入图像尺寸,使网络可以胜任不同分辨率的检测任务。

3.1 目标框维度聚类

与Faster R-CNN算法相比,YOLO算法会产生较多的bounding boxes的定位错误,因此YOLO v2算法借鉴了Faster R-CNN算法的思想,引入anchor box机制。Faster R-CNN的anchor box存在不确定的人为误差,故Redmon等提出使用K-means方法在训练集中聚类得到anchor box的参数:以聚类个数K作为anchor box的个数,以K个聚类中心的宽高维度作为anchor box的宽高。YOLO v2在COCO和PASCAL VOC 2007数据集的聚类结果都为5,故YOLO v2网络中的anchor个数是5。

YOLO v2的anchor box参数具有普适性,针对特定问题,需要在自制船舶数据集上重新进行维度聚类操作。采用K-means方法对自制船舶数据集中的目标框进行维度聚类,得到最适合本数据集的anchor box的宽高维度和数目。

使用Average IOU(Intersection-Over-Union)作为评价参数,且为了避免IOU scores的大小受到anchor box尺寸的影响,使用YOLO v2中定义的新的距离公式:

$$d(box, centroid) = 1 - IOU(box, centroid) \quad (3)$$

使用一系列K值对自制数据集中的目标框进行维度聚类,得到相应的不同Average IOU;随着K值的增大,Average IOU的值也增大,如图1所示。当 $K < 4$ 时,Average IOU变化得较快;当 $K > 4$ 时,Average IOU变化得慢。综合考虑网络模型的复杂度和检测速度,选择最佳anchor box的个数为4,并在配置文件中更改anchor box参数为(1.8896, 1.6538), (5.5329, 4.8666), (8.82124, 10.3572)和(10.6766, 5.9791)。

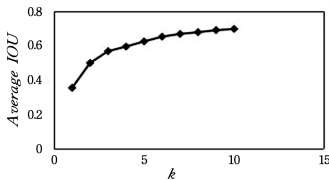


图1 K-means 维度聚类结果

Fig.1 K-means dimension clustering results

3.2 网络结构的改进

YOLO v2中的直通层,类似于ResNet中的恒等映射^[20],通过将相邻特征堆叠到不同通道,将高分辨率特征映射变成低分辨率特征映射,细粒度化特征来增强对小目标的定位。因此,本文在YOLO v2的基础上又增加了一个直通层,强化了直通层对小目标的定位作用。

船舶目标检测问题,本质上是船与非船的二分类问题和对图像上所有船舶的定位问题,因此本文删减了YOLO v2网络中冗余的卷积层和池化层。

表3详细对比了YOLO, YOLO v2和改进网络New-Model的网络框架。

表3 YOLO, YOLO v2和改进网络New-Model的网络框架对比
Table 3 Frame comparison of YOLO, YOLO v2 and improved

		New-Model	
	YOLO v1	YOLO v2	New-Model
	Conv7/2-64	Conv3-32	Conv3-32
	Maxpool/2	Maxpool/2	Maxpool/2
	Conv3-192	Conv3-64	Conv3-64
	Maxpool/2	Maxpool/2	Maxpool/2
	Conv1-128	Conv3-128	Conv3-128
	Conv3-256	Conv1-64	Maxpool/2
	Conv1-256	Conv3-128	Conv3-256
	Conv3-512	Maxpool/2	Conv1-128
	Maxpool/2	Conv3-256	Conv3-256
	Conv1-256	Maxpool/2	Maxpool/2
	Conv3-512	Conv3-512	Conv3-512
	Conv1-256	Conv1-256	Maxpool/2
	Conv3-512	Conv3-512	Conv3-1024
	Conv1-256	Maxpool/2	Route
Bottom	Conv3-512	Conv3-1024	Reorg/Route
	Conv1-512	Conv1-512	Conv3-1024
	Conv3-1024	Conv3-1024	Maxpool/2
	Maxpool/2	Conv3-1024	Route
	Conv1-512	Conv3-1024	Reorg/Route
	Conv3-1024	Route	Conv3-1024
	Conv1-512	Conv1(64)	Conv1-24
	Conv3-1024	Reorg/Route	Detection
	Conv3-1024	Conv3-1024	—
	Conv3/2-1024	Conv1	—
	Conv3-1024	Detection	—
	Conv3-1024	—	—
	Local	—	—
	Dropout	—	—
	Conn	—	—
Top	Detection	—	—

3.3 输入图像多尺度变换

本文借鉴YOLO v2动态调整输入图像尺寸的多尺度训练思想,在训练过程中,每经过10 epoch,就随机选择新的输入尺寸。因为模型的下采样因子为32,所以输入图像的尺寸均为32的倍数。输入尺寸的计算公式为:

$$S = 32 \times (7 + N) \quad (4)$$

其中,S为输入图像的尺寸,N为0~12中随机产生的自然数。

这种改进让网络模型可以胜任不同分辨率的图像检测任务,对低分辨率的图像检测速度更快,对高分辨率的图像能够做到实时检测且检测准确率更高。

4 对比实验及结果分析

4.1 实验数据

在网络上下载船舶图片来自制船舶数据集,数据集样本总数为700。为解决样本数量少的问题,通过旋转图片角度、调整色调、改变饱和度和曝光度等方法来扩充样本数量。

4.2 实验配置与训练结果

实验配置:显卡为2× GeForce GTX 1080, CPU为Inter Xeon E5-1603 v4,主频为2.8GHz,内存为32GB,操作系统为Windows 10,框架为Tensorflow。

网络参数如表4所列。

表 4 网络参数

Table 4 Network parameters

Parameter	Value
Momentum	0.9
Learning_rate	0.0001
Steps	200,1000,10000,15000
Max_batches	25000
Decay	0.0005
police	steps
scales	10,0.1,0.1,0.1
batch	64

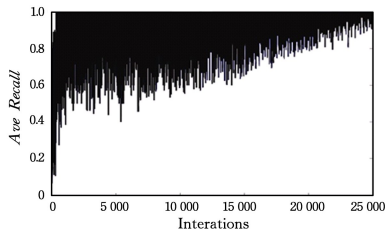
网络训练的评价参数的可视化图如图 2 所示。Avg Recall 和 Avg IOU 的计算公式为:

$$Avg\ Recall = \frac{true\ positive}{true\ positive + false\ negative} \quad (5)$$

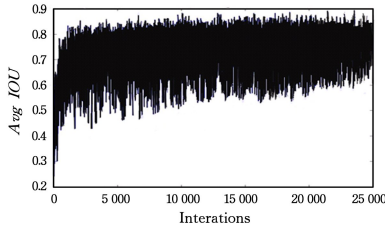
其中, *true positive* 代表正类被判定为正类的数量, *false negative* 代表正类被判定为负类的数量。

$$Avg\ IOU = \frac{Area\ of\ Overlap}{Area\ of\ Union} \quad (6)$$

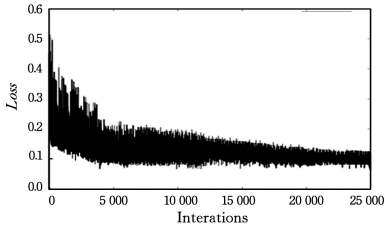
其中, *Area of Overlap* 代表真实框 (truth ground box) 与 bounding box 的区域交集, *Area of Union* 代表 truth ground box 与 bounding box 的区域并集。



(a) 平均召回率可视化图



(b) 平均 IOU 可视化图



(c) 损失值可视化图

图 2 网络训练的评价参数可视化图

Fig. 2 Visualization of evaluation parameters for network training

由图 2 可知,当迭代步数达到 2.5 万次时,平均召回率逐渐趋近于 1,平均 IOU 稳定在 0.6~0.9 之间,损失值下降至 0.1 左右。从各评价参数的收敛情况可以得出:网络的训练结果是比较理想的。

4.3 性能对比

4.3.1 目标框维度聚类前后的性能对比

在 3.1 节中,采用 K-means 方法对自制船舶数据集中目

标区域对应的目标框进行维度聚类,得到本数据集的最佳 anchor box 的尺寸和数目。与 Faster R-CNN, YOLO v2 的 anchor box 生成方案进行对比,结果如表 5 所列。通过重新对自制数据集进行目标框维度聚类,减少了 anchor box 的数量,降低了计算资源的耗费,在提高检测速度的同时保证了较高的平均 IOU。

表 5 Anchor box 重新生成方案的结果

Table 5 Results of the anchor box regeneration scheme

Anchor box 生成方案	Anchor box 数量	Avg IOU/%
Faster rcnn	9	60.9
YOLO v2	5	60.1
New-Model	4	72.2

4.3.2 网络结构改进前后的性能对比

船舶目标检测问题的难点在于小目标检测效果差、小目标误识率高。针对这一问题,本文对 YOLO v2 的原网络做出改进,删减一些冗余的卷积层和池化层,增加了直通层来细粒度化特征,增强对小目标的定位。由表 6 可知,通过改进网络结构, *mAP* 有了提高,但检测速度略有降低。

表 6 YOLO v2 网络改进前后的性能对比

Table 6 Comparison of performance of original and improved

YOLO v2 network		
Model	<i>mAP</i> /%	检测速度
YOLO v2 416×416	76.8	67
New-Model 416×416	79.1	64

4.3.3 输入图像尺度改变前后的性能对比

采用多尺度输入尺寸进行网络训练,训练每迭代 10 次,就会随机选择新的输入图像尺寸,这种策略使得改进网络不得不学着去适应不同尺寸的图像输入,这意味着同一个网络可以胜任不同分辨率的检测任务。如图 3 所示,通过对不同尺寸的验证数据集进行检测,得到多尺度训练的网络和单一尺度训练的网络的性能对比,其中单一尺度训练网络时输入图像的尺寸为 416×416。由图 3 得知,随着输入图像尺寸的增大,多尺度训练的网络和单一尺度训练的网络的 *mAP* 均在增长。与单一尺度训练的网络相比,多尺度训练的网络在检测低分辨率图像时,检测速度更快、准确率更高;在检测高分辨率图像时,多尺度训练的网络的 *mAP* 值已经达到了 83.3%,取得了显著的改进效果。

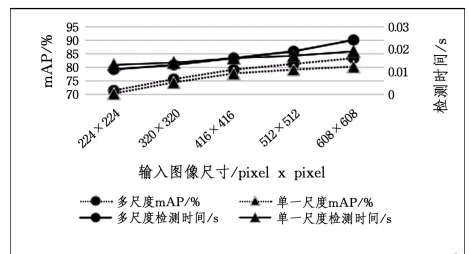


图 3 多尺度网络与单一尺度网络的性能对比结果

Fig. 3 Performance comparison of multi-scale networks and single-scale network

4.4 检测结果的对比

本文在船舶目标检测验证集上对 YOLO v2 和 New-Model 进行对比实验,检测验证集包括 150 张图片且其尺寸不一。在网络输入图像尺寸为 416×416 时,检测结果对比如

表7所列。由表7可知,New-Model的检测速率比YOLO v2略低,但其mAP较YOLO v2提高了1.2%。

表7 检测结果的对比

Table 7 Comparison of test results

Model	Detection time/s	mAP/%
YOLO v2 416×416	0.015	77.9
New-Model 416×416	0.016	79.1

图4展示了YOLO v2网络和改进网络New-Model在船舶目标检测数据集上的部分检测结果。可以看出,New-Model与YOLO v2在船舶大目标检测方面的效果不相上下;而对于船舶小目标,与YOLO v2相比,New-Model检测更准确,定位更精准;对于船舶大小目标同时出现的情况,YOLO v2的小目标检测误识率高,New-Model检测效果更好。

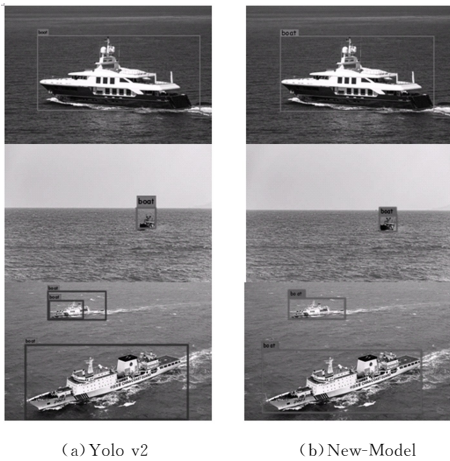


图4 YOLO v2和New-Model在船舶检测数据集上的检测结果
Fig. 4 Results of YOLO v2 and New-Model on the ship detection data set

结束语 本文以YOLO v2算法为基础,通过目标框维度聚类、网络结构改进、输入图像多尺度变换等方法来改进YOLO v2算法的船舶目标检测性能。YOLO v2算法改进前后的对比实验结果表明:与YOLO v2原网络相比,改进后的YOLO v2算法在船舶目标检测任务中对小目标的检测准确率高、误识率低、鲁棒性好;检测速度满足船舶目标检测系统的实时性要求。但改进后的YOLO v2算法还存在多个小目标检测效果不好的问题,下一步将拓展研究范围,继续研究以目标检测网络为基础的船舶目标检测算法。

参考文献

- [1] ZHANG Y. Image Recognition of Ship Target Based on Cloud Computing [D]. Guangzhou: South China University of Technology, 2016. (in Chinese)
张羽. 基于云计算的舰船目标图像识别[D]. 广州: 华南理工大学, 2016.
- [2] ZHAO Y. A brief talk on the methods and techniques of ship target recognition [J]. Ship Science and Technology, 2016, 38(1A): 163-165. (in Chinese)
赵友. 浅谈舰船目标识别的方法和技术[J]. 舰船科学技术, 2016, 38(1A): 163-165.
- [3] PIDGEON V W. Frequency Dependence of Radar Ducting[J]. Radio Science, 2016, 5(3): 541-549.
- [4] SHEN G N. Research on ship target recognition technology [D]. Harbin: Harbin Engineering University, 2012. (in Chinese)
沈广楠. 舰船目标识别技术研究[D]. 哈尔滨: 哈尔滨工程大学, 2012.
- [5] ZHANG S, YAN Y Y, LI Y F. Moving target detection based on Video Fusion of thermal infrared and visible light[J]. Computer Science, 2015, 42(8): 86-89. (in Chinese)
张笙, 严云洋, 李郁峰. 热红外与可见光视频融合的运动目标检测[J]. 计算机科学, 2015, 42(8): 86-89.
- [6] ZHANG X, DONG G, XIONG B, et al. Refined segmentation of ship target in SAR images based on GVF snake with elliptical constraint[J]. Remote Sensing Letters, 2017, 8(8): 791-800.
- [7] YU J Y, HUANG D, WANG L Y, et al. A real-time on-board ship targets detection method for optical remote sensing satellite [C] // International Conference on Signal Processing. New York: IEEE Press, 2017: 204-208.
- [8] ZHU R. Application of Computer Identification and Location Algorithm in Small far Infrared Target Recognition of Ship Under Surge Interference [J]. Polish Maritime Research, 2017, 24(S1): 171-181.
- [9] LECUN Y, BENGIO Y, HINTON G. Deep learning[J]. Nature, 2015, 521(7553): 436.
- [10] GIRSHICK R. Fast R-CNN [C] // IEEE International Conference on Computer Vision. New York: IEEE Computer Society, 2015: 1440-1448.
- [11] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: towards real-time object detection with region proposal networks [C] // International Conference on Neural Information Processing Systems. Massachusetts: MIT Press, 2015: 91-99.
- [12] REDMON J, DIVVALA S, GIRSHICK R, et al. You Only Look Once: Unified, Real-Time Object Detection [C] // Computer Vision and Pattern Recognition. New York: IEEE Press, 2016: 779-788.
- [13] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single Shot MultiBox Detector [C] // European Conference on Computer Vision. Springer: Cham, 2016: 21-37.
- [14] REDMON J, FARHADI A. YOLO9000: Better, Faster, Stronger [C] // IEEE Conference on Computer Vision and Pattern Recognition. New York: IEEE Computer Society, 2017: 6517-6525.
- [15] SIMONYAN K, ZISSERMAN A. Very Deep Convolutional Networks for Large-Scale Image Recognition [J]. Computer Science, 2014. <https://arxiv.org/abs/1409.1556>.
- [16] SZEGEDY C, LIU W, JIA Y, et al. Going deeper with convolutions [J/OL]. <https://arxiv.org/abs/1409.4842>.
- [17] LIN M, CHEN Q, YAN S. Network In Network [J/OL]. Computer Science. <https://arxiv.org/abs/1312.4400>.
- [18] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks [C] // International Conference on Neural Information Processing Systems. New York: ACM, 2012: 1097-1105.
- [19] ZHOU W B, SHI Y X. Optimization algorithm of K-means clustering center of selection based on density [J]. Application Research of Computers, 2012, 29(5): 1726-1728.
- [20] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition [M] // Computer Vision and Pattern Recognition. New York: IEEE, 2016: 770-778.