



电光与控制

Electronics Optics & Control

ISSN 1671-637X, CN 41-1227/TN

《电光与控制》网络首发论文

题目：基于改进 YOLO V3 的舰船目标检测算法
作者：姜文志，李炳臻，顾佼佼，刘克
网络首发日期：2021-03-23
引用格式：姜文志，李炳臻，顾佼佼，刘克. 基于改进 YOLO V3 的舰船目标检测算法. 电光与控制. <https://kns.cnki.net/kcms/detail/41.1227.TN.20210322.1258.003.html>



网络首发：在编辑部工作流程中，稿件从录用到出版要经历录用定稿、排版定稿、整期汇编定稿等阶段。录用定稿指内容已经确定，且通过同行评议、主编终审同意刊用的稿件。排版定稿指录用定稿按照期刊特定版式（包括网络呈现版式）排版后的稿件，可暂不确定出版年、卷、期和页码。整期汇编定稿指出版年、卷、期、页码均已确定的印刷或数字出版的整期汇编稿件。录用定稿网络首发稿件内容必须符合《出版管理条例》和《期刊出版管理规定》的有关规定；学术研究成果具有创新性、科学性和先进性，符合编辑部对刊文的录用要求，不存在学术不端行为及其他侵权行为；稿件内容应基本符合国家有关书刊编辑、出版的技术标准，正确使用和统一规范语言文字、符号、数字、外文字母、法定计量单位及地图标注等。为确保录用定稿网络首发的严肃性，录用定稿一经发布，不得修改论文题目、作者、机构名称和学术内容，只可基于编辑规范进行少量文字的修改。

出版确认：纸质期刊编辑部通过与《中国学术期刊（光盘版）》电子杂志社有限公司签约，在《中国学术期刊（网络版）》出版传播平台上创办与纸质期刊内容一致的网络版，以单篇或整期出版形式，在印刷出版之前刊发论文的录用定稿、排版定稿、整期汇编定稿。因为《中国学术期刊（网络版）》是国家新闻出版广电总局批准的网络连续型出版物（ISSN 2096-4188，CN 11-6037/Z），所以签约期刊的网络版上网络首发论文视为正式出版。

基于改进YOLO V3的舰船目标检测算法

姜文志¹, 李炳臻², 顾佼佼¹, 刘克¹

(1. 海军航空大学岸防兵学院, 山东 烟台 264001 2.中国人民解放军95668部队, 昆明 650000)

摘要: 针对YOLO V3算法中对于小目标检测精度不高、容易出现漏检误检的问题进行改进, 提出了一种基于改进YOLOV3的舰船目标检测算法, 首先通过在YOLO V3原网络结构基础上额外从主干网络引出一个输出尺度, 将其与上一个输出尺度中的特征信息进行特征拼接, 得到具有更丰富语义信息的特征向量; 其次基于数据集进行聚类改进, 改进度量距离公式、重新设置anchor boxes的个数与相应参数; 最后优化改进YOLO V3的损失函数, 提高模型的整体性能。通过对测试数据集进行分析实验, 结果表明改进后的检测算法平均精确度达到83.98%, 较之于原YOLO V3, 精确度提升了6.72%。

关键词: 深度学习; 目标检测; 舰船目标; 小目标; YOLO V3

中图分类号: TP399 **文献标识码:** A

Ship Target Detection Algorithm based on Improved YOLO V3

JIANG Wenzhi¹, LI Bingzhen², GU Jiaojiao¹, LIU Ke¹

(1. Naval Aviation University, Coast Guard Academy, Yantai 264001, China;

2. No. 95668 Unit of PLA, Kunming 650000, China)

Absrtact: Aiming at the problem of low accuracy of small target detection and easy to miss detection in YOLO V3 algorithm, a ship target detection algorithm based on improved YOLO V3 is proposed. Firstly, an additional output scale is derived from the backbone network on the basis of original YOLO V3 network structure, and the feature vector with richer semantic information is obtained by splicing it with the feature information in the previous output scale. Secondly, it improves the clustering based on the data set, improves the measurement distance formula, and resets the number of anchor boxes and the corresponding parameters. Finally, the loss function of the improved YOLO V3 is optimized to improve the overall performance of the model. Analysis and experimental results on the test data set show that average accuracy of the improved detection algorithm is 83.98%, which is 6.72% higher than that of original YOLO V3.

Key words: deep learning; target detection; ship target; small target; YOLO V3

0 引言

舰船目标检测在船舶船只救援、舰船类军事目标侦察等领域方面有着非常重要的研究意义与应用价值。近年来, 基于深度学习的目标检测算法在检测效果方面取得了很大的突破, 同传统方法相比检测性能有着质的飞跃。深度学习目标检测算法根据设计思想的不同可以分成两大类, 分别是One-stage与Two-stage算法, Two-stage算法将检测过程分成两大部, 首先生成待检测区域, 然后基于此进行目标检测, R-CNN^[1]是最早被提出的

Two-stage 检测算法, 随后出现的 Fast-RCNN^[2]、Mask-RCNN^[3]与 Faster-RCNN^[4], 使检测精度不断提高, 但速度相对较慢。One-stage 算法直接产生物体的类别概率和位置坐标, 单次检测就得到最终的检测结果, 典型的算法有 2015 年提出的 YOLO V1^[5]、2016 年提出的 YOLOV2^[6]、SSD^[7], 2017 年提出的 DSSD^[8]以及 2018 年提出的 YOLOV3^[9], 性能最为突出的就是 YOLO V3, 其核心思想是将目标检测看做回归问题求解, 通过图片的全部像素点直接得到 Boundingbox, 检测速度快是其最大特点, 缺点在于检测小目标物体时会出现漏检误检的问题。

作者简介: 姜文志 (1964-), 男, 山东莱州人, 博士, 教授, 博导, 研究方向为武器系统与应用研究、人工智能深度学习技术。

目标检测算法现在已经被应用于生产生活的各个方面,包括舰船目标检测:于洋在文献[10]中通过改进网络结构、变换输入尺度的方法对YOLOV2进行改进,应用在船舶目标检测;苏娟在文献[11]中通过特征增强对SSD进行改进并应用在舰船目标检测上;李靛在文献[12]中,结合FPN思想对Faster RCNN进行改进并应用在遥感图像检测中。但上述文献所用算法均存在着小目标检测效果差,容易出现漏检误检的问题,并不完全适用于舰船目标检测,为解决上述问题,本文提出了一种改进的YOLO V3目标检测算法,通过改进算法的网络结构、进行聚类改进以及优化模型损失函数的方法,使改进后的算法更加适用于舰船目标检测。

1 YOLO V3 算法原理

在YOLO V3算法中,将一张图片划分成 $S \times S$ 个网格,如果待检测目标物体存在于某网格中,那么此网格就负责检测该目标。每个网格负责检测 B 个Boundingboxes信息,包括五个元素: $x, y, w, h, confidence$,其中 (x, y, w, h) 分别代表Boundingbox的中心坐标与宽高值相对于所在网格的相对值,confidence为置信度,

$$confidence = Pr(Object) * IOU_{pred}^{truth} \quad (1)$$

式(1)中, $Pr(Object)$ 表示该网格中是否存在待检测物体,值为0或1;IOU为Boundingbox与Groundtruthbox的交并比。

YOLO V3的主干网络是Darknet-53^[9],因为包含53个卷积层,所以称为Darknet-53,相比于YOLO V2^[6]的Darknet-19,网络增加了残差单元,并且卷积层数从19层增加到了53层,对图片的特征提取能力更强。但网络结构的加深,带来的问题就是梯度爆炸与梯度消失^[13],即反向传播时梯度值会随着连乘而变得不稳定。因此,YOLO V3采用了残差模块Residual来解决梯度爆炸与消失的问题,此时学习目标不再是一个完整的输出,而是目标值 $H(x)$ 和输入 x 的差值,即所谓的残差 $F(x)=H(x)-x$,因此,训练目标改变为将残差结果

逼近于0,使得网络不断加深的情况下,也能达到很高的准确率。

YOLO V3采用三个尺度的特征图目标检测,每个特征图采用anchorboxes作为先验框,借鉴了Faster RCNN的anchor^[4]思想,其实就是一组长宽比固定的候选框,相当于一组模板,用这组模板进行后续的检测。anchor的参数设定采用k-means聚类算法得到,共9个聚类中心,按照大小分给三个不同的尺度,输出三个不同尺度的feature map用于进行后续的目标检测。YOLO V3中最小特征图尺寸为 13×13 ,对于小目标而言仍然偏大,容易出现漏检、误检等问题,因此针对这一不足之处进行改进,得到改进后的YOLO V3算法。

2 改进的YOLO V3目标检测算法

2.1 改进的网络结构

一般来说,在卷积神经网络中生成的低层网络特征图主要用来反映待检测目标物体的细节特征,低层特征图主要用来检测小目标物体;而高层特征图主要反映的是抽象的语义信息^[14],如轮廓、大体特征等,因此主要用来检测大目标物体。不同层级的可视化特征对比如图1所示,



图1 不同层级卷积后特征对比图

Fig. 1 Features after convolution at different levels

其中,图1(a)为原图,图1(b)为Darknet-53中经过第二块卷积模块Res2后的特征图,图1(c)为第三块卷积模块Res8后的特征图,从图1可明显看出,Res2的特征图较Res8相比细节信息更丰富,特征也更加明显。

因此在原网络结构的基础上额外引入这一层的输出特征图,并与上一个尺度中经过上采样的特征图进行拼接,增加一个输出尺度特征图用于小目标检测,提高对小目标的检测能力。改进后网络结构如图2所示,

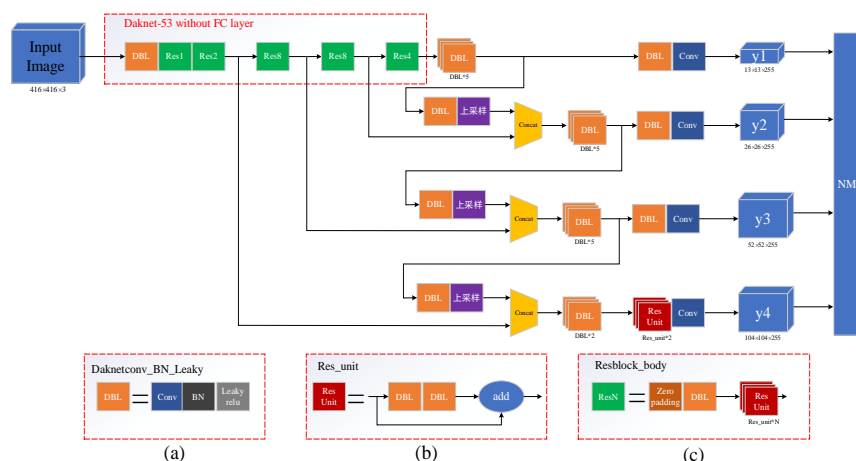


图 2 改进后网络结构图

Fig. 2 Improved network structure diagram

图 2 下方的(a)、(b)、(c)是解释图；图 2 (a)中的 DBL 是网络结构中的基本组件，DBL 就是卷积层+BN(归一化指数层)+Leaky relu(激活函数层)；ResN 是 YOLO V3 中的大组件，N 代表此组件中的 Res_unit 个数；Concat 代表张量拼接，此操作可以扩充张量的维度；

改进后的网络结构中增加了一个尺度的输出，将尺度三中经过 5 层 DBL 的特征图同主干网络中第二个残差块卷积后的特征图进行拼接，如图 3 所示，

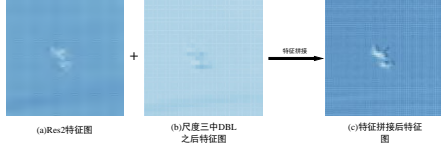


图 3 特征拼接后特征图

Fig. 3 Feature map after feature stitching

图 3 表明，特征拼接后相比于拼接前，待检测目标的细节信息要更加丰富。为了避免卷积增加带来参数过多的问题，在特征拼接之后，将 DBL 由 5 层减为 2 层，受 DSSD^[8]的启发，DBL 后加入两个 ResNet 避免梯度消失或爆炸问题的发生。

2.2 基于数据集的聚类改进

采用 K-means 聚类算法对数据集的目标框进行聚类，目的是使 anchor boxes 和临近的 Ground truth 有更大的 IOU，更小的距离。而常用的距离公式如曼哈顿距离、欧式距离等在这里往往并不能取得理想的效果，因此自定义距离度量公式

$$d(box, centroid) = 1 - IOU(box, centroid) \quad (2)$$

在公式(2)中， $d(box, centroid)$ 代表 anchorbox 到聚类中心的距离， $IOU(box, centroid)$ 代表 anchorbox 与 ground truth box 的交并比。

anchorboxes 个数是通过数据集的候选框做聚类得到的，一般来说 anchor 的类别越多，预测框越接近真实框，但同时会带来更庞大的计算参数，因此选值对于平衡精度与速度来说就很关键，YOLO V3 算法采用 AvgIOU (平均重叠度) 作为聚类分析的度量，可表示为

$$f = \arg \max \frac{\sum_{i=1}^k \sum_{j=1}^{n_i} I_{iou}(B, C)}{n} \quad (3)$$

在公式(3)中， n 表示样本总数， k 为类别数， n_i 代表第 k 个聚类中心中样本的个数， $I_{iou}(B, C)$ 代表 anchor 与真实框的交并比。选取 k 为 1~9，分别对数据集中样本进行聚类分析，得到 k 值与 AvgIOU 的关系如图 4 所示，

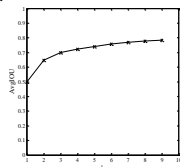


图4k- AvgIOU 值坐标关系图

Fig.4 k- AvgIOU coordinate relation diagram

图 4 中可以看出，在 $k=4$ 之后，函数曲线趋于平稳。在原 YOLO V3 算法中，作者认为 k 取 5 可以达到最佳的效果，但是由于前文在网络结构中增加了一个输出尺度，共有 4 个尺度，因此 k 值选择最好为 4 的倍数，设置 $k=8$ ，即 anchor boxes 个数为 8，根据 python 脚本读取训练集中保存的.txt 文件中对应目标宽高值信息，进行 K-means 聚类，聚类结果如图 5 所示，设置 anchor boxes 的宽高分别为(12,16)，(16,30)，(33,23)，(30,61)，(82,30)，(116,90)，(156,197)，(224,203)。

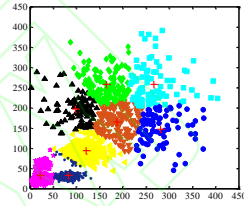


图 5 K-means 聚类分析结果图

Fig.5 K-means cluster analysis result chart

2.3 损失函数

在 YOLO V3 目标检测算法中，损失函数共由三部分构成，分别是回归损失、分类损失和置信度损失，根据前文对网络结构的改进，对 YOLO V3 中的回归损失与分类损失进行优化：采用宽高方差误差表达回归框误差；分类损失则采用二分类交叉熵函数表达。具体如公式(4)所示，

$$\begin{aligned} Loss = & \lambda_{coord} \sum_{i=0}^{S^2} \sum_{j=0}^B I_{ij}^{obj} (2 - w_i \times h_i) \left[(x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2 \right] + \\ & \lambda_{coord} \sum_{i=0}^{S^2} \sum_{j=0}^B I_{ij}^{obj} (2 - w_i \times h_i) \left[(w_i - \hat{w}_i)^2 + (h_i - \hat{h}_i)^2 \right] - \\ & \sum_{i=0}^{S^2} \sum_{j=0}^B I_{ij}^{obj} \left[\hat{C}_i^j \log(C_i^j) + (1 - \hat{C}_i^j) \log(1 - C_i^j) \right] - \\ & \lambda_{noobj} \sum_{i=0}^{S^2} \sum_{j=0}^B I_{ij}^{noobj} \left[\hat{C}_i^j \log(C_i^j) + (1 - \hat{C}_i^j) \log(1 - C_i^j) \right] - \\ & \sum_{i=0}^{S^2} I_{ij}^{obj} \sum_{c \in classes} \left[\hat{P}_i^j \log(P_i^j) + (1 - \hat{P}_i^j) \log(1 - P_i^j) \right] \end{aligned} \quad (4)$$

在公式(4)中， (x, y, w, h) 分别代表 Boundingboxes 的参数值， S, B 代表图片的网格数及 Boundingboxes 数， I_{ij}^{obj} 用于判断第 i 个网格中第 j 个 Boundingboxes 是否包含检测对象；公式前两行代表回归损失，为了加强对小目标的误差表达，在该项前面加了系数 $(2 - w_i \times h_i)$ ；第三、四行是置信度损失，分别代表存在与不存在目标的 Boundingboxes 置信度误差 \hat{C}_i^j 代表参数置信度；最后一行为分类损失，选择交叉熵函数作为损失函数，能把所有目标物体的分类归结于二分类问题，避免因多个目标重叠造成的漏检问题。

3 实验与分析

在实验与分析部分,训练数据集共计 1000 张,均为互联网上搜集和自己拍照搜集的各类舰船目标图像;测试集为分布在各类码头、港口、海洋等背景下的多种舰船目标图像,共 200 张。选择精度评价指标 mAP(mean average precision)与速度评价指标 FPS(frame per second)对训练好的模型进行性能评估。

3.1 模型训练

由于训练数据集仅有 1000 张,担心数量较少不足以支撑模型的训练,因此通过对数据集进行左右与上下镜像处理来增强训练数据集,运行数据处理文件 left_right.py 与 up-down.py 对训练数据集进行镜像翻转处理操作,得到数据增强后的图片共 4000 张。

本次实验中,操作系统为 Windows10,采用的深度学习框架为 tensorflow1.9.0。训练过程中 batch_size 设置为 16,图 6 为通过 tensorflow 可视化工具 tensorboard 查看的训练过程中的损失函数图像,



图 6 改进的 YOLO V3 损失函数曲线

Fig. 6 Improved YOLOV3 loss function curve

从图 6 中可以看出,当训练次数迭代至 7000 次左右时,损失函数已经收敛至 6 附近,继续进行模型的训练至 7500 次,loss 值上升至 10 左右,由于前文在 YOLO V3 的网络结构中增加了一个输出尺度,考虑到会引入额外的计算参数,继续训练可能会产生过拟合的现象,于是在迭代次数到达 7500 次时停止训练,训练结束得到最终的检测模型。

3.2 模型测试

通过在测试集上测试模型的检测效果,如图 7 所示,

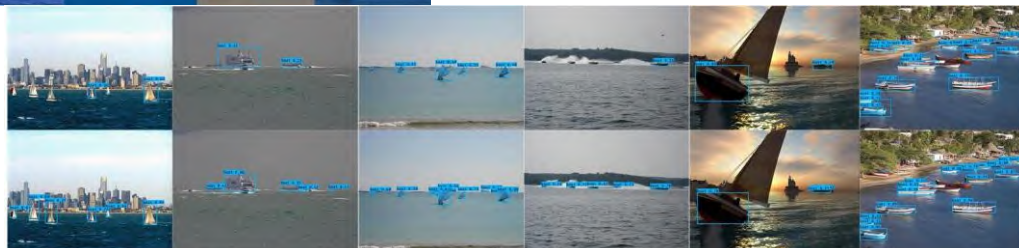


图 7 舰船图像检测结果

Fig. 7 Results of ship image detection

从图 7 中可以看出检测效果非常好,训练好的模型可以准确识别出不同图片的舰船目标。

选取一些光照不足、目标被遮挡以及目标等复杂环境下的图像进行测试,分析改进算法的环境适用性,如图 8 所示,



图 8 复杂环境下部分检测结果

Fig. 8 Partial test results in complex environment

在图 8 中,从左至右的背景环境依次是雨雾背景、光照不足、船体被遮挡以及海浪背景下,从图中可以看出,改进算法在不良条件环境下依然可以做出正确的判断。

3.3 结果分析

通常使用 mAP 来评估检测准确度,mAP 值越大证明检测效果越好,在本次实验中,因为检测目标只有舰船这一类物体,因此在数值上,mAP 值等于 AP 值,AP 值为召回率(Recall)与准确率(Precision)所围成的 R-P 曲线面积;使用原 YOLO V3 算法对同一数据集进行 7500 次训练,分别运行 python 工程文件 mAP-master,对比检测效果,得到改进前后检测结果的平均准确率 AP 值,画出 R-P 曲线,如图 9 所示,

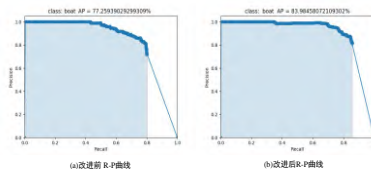


图 9 改进前后在舰船测试集上的 R-P 曲线

Fig. 9 Comparison of R-Pcurve before and after

图 9 中横坐标为召回率,纵坐标为相应的准确率,原 YOLO V3 的平均准确率 AP=77.25%,经过改进之后在同一测试集上的 AP=83.98%为了验证改进方法针对小目标检测的有效性,选取小目标物体图像数据进行对比实验,如图 10 所示,图 10 中原图片的分辨率为 500×375 ,均包含多个 32×32 以内的小目标。

图 10 改进前后小目标检测效果对比

Fig. 10 Comparison of small target detection before and after improvement

改进后的 YOLO V3 算法对于小目标物体检测的准确度明显高于原算法。这说明算法改进后提取到了更丰富的细节特征,改善了原有算法中在检测小目标情况下会出现的漏检误检、检测率低等现象。

对改进前后的两种检测算法,分别运行 python 工程文件 mAP-master,得到检测 200 张图片所花费的时间,原算法花费时间为 10.18s,改进后时间为 12.21s,其具体的性能对比如表 1 所示,

表 1 改进前后算法性能对比

模型名称	AP/%	mAP/%	Time/s
YOLO V3	77.26	77.26	0.0509
改进后 YOLO V3	83.98	83.98	0.0532

从表中可以看出,原 YOLO V3 算法的 mAP 值为 77.26%,而改进后算法的 mAP 值为 83.98%,精度提升了 6.72%,检测速度方面 FPS 可达到约 19 帧/秒,基本满足实时性的要求。

4 结论

提出了一种改进 YOLO V3 的舰船目标检测算法,对原算法中的网络结构和 K-means 聚类进行改进,增加了一个输出尺度用于检测小目标物体,同时对损失函数进行了优化。应用于舰船目标检测时,检测精度可以达到 83.98%,较原算法提高了 6.72%,速度达到 19 帧/秒;在雨雾背景、光照不足等复杂背景下也能取得准确的检测结果,能够满足在干扰条件下对舰船检测的要求,并且通过对比实验,验证了改进方法的可行性。本文提出的改进后 YOLO V3 目标检测算法可以大大提升舰船目标的检测准确率,因此可以将此算法应用于工业或者军事领域,其中工业领域可应用于码头或港口,进行船舶检测与监控,对海上救援、海上交通有十分重要的应用价值;而军事领域不仅可应用于海战场的敌我舰船检测,还可应用于反舰武器系统,提升武器装备对于海上舰船目标的识别与打击精度,从而提高我军部队的战斗力

参考文献:

[1] Uijlings J R R, Sande K E A V D, Gevers T, et al. Selective search for object recognition[J]. International Journal of Computer Vision, 2013, 104 (2) :154-171.
[2] Girshick R. Fast R-CNN[C]//Proceedings of

ICCV 2015, 2015: 1440-1448.
[3] Kaiming H, Georgia G, Piotr D, et al. Mask R-CNN[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2018, 18(11):1-1.
[4] Ren S, He K, Girshick R, Sun, J. Faster r-cnn: towards real-time object detection with region proposal networks. [C]//IEEE Transactions on Pattern Analysis and Machine Intelligence, 39(6), 1137-1149.
[5] Redmon J, Divvala S, Girshick R, et al. You only look once: unified, real-time object detection[C]//Proceedings of CVPR 2015, 2015: 779-788.
[6] Redmon J, Farhadi A. YOLO9000: Better, Faster, Stronger [C]//IEEE Conference on Computer Vision and Pattern Recognition. 2017: 6517— 6525.
[7] Liu W, Anguelov D, Erhan D, et al. SSD: single shot multibox detector[C]//European Conference on Computer Vision. 2016: 21-37.
[8] Fu C Y, Liu W, Ranga A, et al. DSSD: deconvolutional single shot detector[J]. arXiv preprint arXiv:1701.06659, 2017.
[9] Redmon J, Farhadi A. YOLOv3: An Incremental Improvement [C]//IEEE Conference on Computer Vision and Pattern Recognition. 2018.
[10] 于洋,李世杰,陈亮,刘韵婷.基于改进 YOLO v2 的船舶目标检测方法[J].计算机科学,2019,46(8):332-336.
[11] 苏娟,杨龙,黄华,金国栋.用于 SAR 图像小目标舰船检测的改进 SSD 算法[J].系统工程与电子技术,2020,42(5):1026-1034.
[12] 李靓,贺柏根,霍家道. FPN 在遥感图像检测中的应用[J]. 指挥控制与仿真, 2019, 41(2):130-136.
[13] 公明,刘妍妍,李国宁.改进 YOLO-v3 的遥感图像舰船检测方法[J].电光与控制,2020,27(5):102-107.
[14] 郑志强,刘妍妍,潘长城,李国宁.改进 YOLO V3 遥感图像飞机识别应用[J].电光与控制,2019,26(4):28-32.