



基于YOLO-v5的星载SAR图像海洋小目标检测

窦其龙, 颜明重, 朱大奇

引用本文:

窦其龙, 颜明重, 朱大奇. 基于YOLO-v5的星载SAR图像海洋小目标检测[J]. *应用科技*, 2021, 48(6): 1-7.

DOU Qilong, YAN Mingzhong, ZHU Daqi. Small marine target detection in space-borne SAR image based on YOLO-v5[J]. *Applied science and technology*, 2021, 48(6): 1-7.

在线阅读 View online: <https://dx.doi.org/10.11991/yykj.202105023>

您可能感兴趣的其他文章

Articles you may be interested in

改进U-Net网络的水下图像增强

Underwater image enhancement based on improved U-Net model

应用科技. 2021, 48(3): 34-40 <https://dx.doi.org/10.11991/yykj.202103025>

无人机视觉导航着陆标识检测与分割方法

Detection and segmentation method for the landmark based on UAV visual navigation

应用科技. 2020, 47(4): 1-7,13 <https://dx.doi.org/10.11991/yykj.202002013>

ISAR基带干扰系统设计与实现

Design and implementation of ISAR baseband jamming system

应用科技. 2020, 47(3): 24-29 <https://dx.doi.org/10.11991/yykj.201907012>

一种改进的卷积神经网络行人识别方法

An improved pedestrian detection method based on convolutional neural network

应用科技. 2019, 46(3): 51-57 <https://dx.doi.org/10.11991/yykj.201809014>

UUV海面红外侦察图像自适应归并直方图拉伸增强算法

Adaptive merging histogram stretching enhancement algorithm for UUV's observing infrared images at sea

应用科技. 2017, 44(6): 1-4,40 <https://dx.doi.org/10.11991/yykj.201610012>

基于混合高斯模型的运动目标检测算法

Moving object detection algorithm based on Gaussian mixture model

应用科技. 2015(1): 19-21,27 <https://dx.doi.org/10.3969/j.issn.1009-671X.201405003>



微信公众平台



期刊网址

DOI: 10.11991/ykj.202105023

网络出版地址: <https://kns.cnki.net/kcms/detail/23.1191.U.20210630.1641.002.html>

基于 YOLO-v5 的星载 SAR 图像海洋小目标检测

窦其龙, 颜明重, 朱大奇

上海海事大学 智能海事搜救与水下机器人上海工程技术研究中心, 上海 201306

摘 要: 针对星载合成孔径雷达(SAR)图像中的小目标检测, 提出了一种基于 YOLO-v5 的算法。首先根据目标在图像中占比很小的特点, 对深度学习网络进行优化; 其次通过自适应锚点框算法重新设置锚点框大小, 加快模型训练的收敛速度; 最后嵌入 GDAL 模块, 使得检测头能够直接读取被检测目标的地理位置信息。实验结果表明本文提出的方法具有较强的泛化能力, 能够有效检测出大范围海面上的船舶目标, 检测速度较快, 漏检率低。

关键词: 合成孔径雷达; 星载 SAR 图像; 深度学习; 数据增强; 自适应锚点框; YOLO-v5; 小目标检测; GDAL

中图分类号: TP389.1

文献标志码: A

文章编号: 1009-671X(2021)06-0001-07

Small marine target detection in space-borne SAR image based on YOLO-v5

DOU Qilong, YAN Mingzhong, ZHU Daqi

Shanghai Engineering Research Center of Intelligent Maritime Search & Rescue and Underwater Vehicles,
Shanghai Maritime University, Shanghai 201306, China

Abstract: An algorithm based on YOLO-v5 is proposed for small target detection in synthetic aperture radar(SAR) images. Firstly, the deep-learning network is optimized according to the small proportion of the target in the image. Secondly, an adaptive anchor box algorithm is used to reset the anchor frame size to accelerate the convergence speed of the model training. Finally, the GDAL module is embedded to enable the detector to directly read the geographic location information of the detected targets. The experimental results show that the proposed method has a relatively strong generalization ability, which can effectively detect ship targets on a wide range of sea surfaces with relatively fast detection speed and low misdetection rate.

Keywords: SAR; space-borne SAR image; deep learning; data enhancement; adaptive anchors; YOLO-v5; small target detection; GDAL

合成孔径雷达(SAR)是根据电磁散射回波的原理对区域进行成像, 其成像范围广、空间分辨率高, 因而在资源勘探、灾害监测、海洋管理及军事领域得到了广泛的应用。基于星载 SAR 图像, 对大范围的海上目标, 尤其是特定区域的船舶目标进行准确的检测、定位或识别, 具有重要的现实需求^[1]。

当今国内外的研究机构对于遥感卫星图像的目标检测做了大量的相关研究。传统的检测手段主要有区域选择、特征提取和分类器分类 3 个步

骤。区域选择即对可能存在目标的区域进行分割, 如在星载 SAR 图像里常用的 CA-CFAR^[2] 算法等。特征提取就是在待检索区域将鲁棒性强的特征进行保留, 在 SAR 图像中比较常见的有标准差特征提取法 (standard deviation, SD)^[3]、加权秩填充比特征提取法 (weighted-rank fill ratio, WRFR)^[4] 等。分类器分类就是将提取特征输入分类器, 和已知数据进行比对分类。传统的检测方法受图像噪声的影响较大, 存在选择区域冗余、特征鲁棒性较差、处理时间长等弊端。

自从 2014 年 Ross Girshick 提出卷积神经网络深度学习算法 (regions with convolutional neural network features, RCNN)^[5], 并在 PASCAL VOC^[6] 数据集取得突破性进展后, 深度卷积神经网络就进入了迅

收稿日期: 2021-05-18. 网络出版日期: 2021-07-01.

基金项目: 国家自然科学基金项目 (U1706224, 62033009).

作者简介: 窦其龙, 男, 硕士研究生.

颜明重, 男, 副教授, 博士.

通信作者: 颜明重, E-mail: mzyan@shmtu.edu.cn.

猛发展的时代,基于深度学习的神经网络模型也因其强大的特征提取能力而被广泛应用于各种实体目标的检测^[7-8]。在 RCNN 卷积神经网络提出之后,Shaoqing Ren^[9]和郭昕刚等^[10]又分别提出了 Fast RCNN、Faster RCNN,使得神经网络不断优化和拓宽,训练速度不断提高,误检率不断降低。华北电力大学的赵文清等^[11]就利用 Faster RCNN 算法对存在缺陷的绝缘子进行准确识别,从而对输电线路进行故障诊断和修复。上述算法一般都是通过神经网络提取候选框、分类处理、回归等操作进行目标特征的学习,从而进行目标检测与分类。近些年,基于端到端(end to end)学习的实例检测算法被提出,典型的代表为单步多框预测(single shot multibox detector, SSD)算法和 YOLO(you only look once)家族^[12]。文献[13]针对行人检测方法误检率高的问题提出了基于改进 SSD 网络的行人检测(pedestrian distinction, PDIS)模型。文献[14]提出了基于 YOLO-v2 和支持向量机(support vector machine, SVM)的船舶检测分类算法,在网络模型最后一步全局特征池化后,利用 SVM 实现船型的分类,有效地实现了不同种类船只的识别。中国科学院的陈科峻等^[15]提出了基于 YOLO-v3 模型压缩的卫星图像实时检测,采用 K-means 聚类算法选取初始锚点框(anchor),然后用多尺度金字塔图像进行模型训练,采用压缩后的模型大幅度减少了系统计算的时间,节约了计算机的计算空间。

SAR 图像目标检测的关键是加强对船舶等被检测目标的注意力,忽略无用信息的干扰。随着 SAR 图像分辨率的不断提高和不同工种模式下图像获取的极化方式、照射角度、干扰因子不同,因此对 SAR 图像目标的自适应检测并不理想^[16]。本文将 YOLO-v5 深度神经网络模型应用于星载 SAR 图像中的船舶目标检测。针对船舶目标在星载 SAR 图像中占比很小的特性,进行图像预处理和数据增强,采用 K-means 改进锚点框的尺寸大小,并优化神经网络模型,嵌入 GDAL 模块,对星载 SAR 图像目标的位置信息等进行读取。

1 YOLO-v5 的网络结构

1.1 输入端 Mosaic 数据增强

参考文献[17]的 CutMix 数据增强方式,将重新组合图像的数量由 2 张增加到 4 张。首先从数据集中取出不重复的 4 张图像;然后对图像依次进行随机的缩放、裁剪和拼接;最后需要将图片

进行灰色填充的操作,以获得符合网络特征训练的大小统一的检测图像。

图 1 为采取数据增强处理后的图像效果。图 1 中的灰色填充部分对神经网络的学习并没有帮助,因此灰色部分越少,训练时间越少,训练效果越好。我们采用式(1)~(3)的计算方法进行灰色填充。

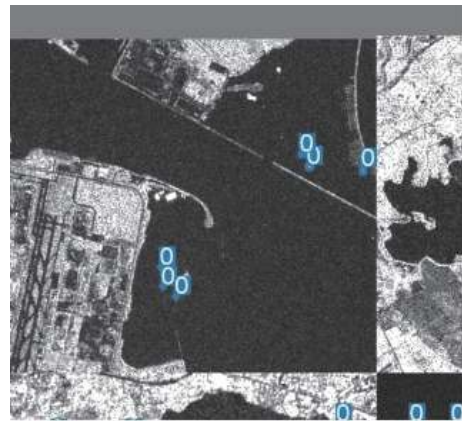


图1 Mosaic 数据增强(包含灰色填充)

$$z = \min\left(\frac{x_1}{x_0}, \frac{y_1}{y_0}\right) \quad (1)$$

$$R = \text{np.mod}\left(\frac{x_1}{z}, \frac{y_1}{z}\right) \quad (2)$$

$$Y = y_1 + R \quad (3)$$

式中: x_0 、 y_0 ($x_0 \geq y_0$) 为原始图像尺寸的长度与宽度; x_1 、 y_1 为图像缩放后的长度与宽度; z 为缩放系数; R 是 x_1/z 整除 y_1/z 的余数, R 表示灰色填充的数值。最终得到宽度为 Y 的图像。

这种方法根据图像原尺寸和缩放尺寸中较小的缩放系数来使得填充的灰色尽可能的小,从而减少计算时的内存占用,达到加快训练速度的目的。

在 YOLO 算法中,初始锚框是针对 VOC^[12] 等数据集计算得到的,本文中 YOLO-v5 的锚点框大小的选择与星载 SAR 图像中的目标尺寸有关,为此,通过数据集和实验来进行设定。网络在初始锚框(anchor)的基础上输出预测框,进而和真实框(ground truth)进行对比,计算损失函数,再反向更新,迭代网络参数。由于 YOLO-v5 采用的 CNN 卷积网络对特征图像分别进行了 32 倍、16 倍和 8 倍的下采样,每次下采样对应 3 个锚点框。32 倍的下采样用较大的锚点框去检测感受视野较大的特征图像,16 倍和 8 倍的下采样分别采用中等的和较小的锚点框去检测感受视野中等和较小的特征图像,从而降低模型在训练的时候寻找被检测目标的盲目性,有助加快模型寻找被检测目标的速度。

在 YOLO-v5 中加入了自适应锚框的计算, 在进行每次训练之前, 通过 K-means 算法在训练集中对所有样本的真实框进行聚类, 从而找出高复杂度和高召回率中最优的那组锚点框。K-means 算法步骤如下。

1) K-means 聚类法即先输入 k 的值, 即我们所希望得到的 k 个类别。

2) 从数据集中随机选取 k 个二维数组作为质心 (centroid)。

3) 对集合中的每一个二维数组进行计算, 与哪一个质心接近则与其分为一组。

4) 在每一组中选出一个新的质心使其与每个点的距离更接近。

5) 当新的质心与旧的质心的直线距离小于设定的阈值时, 则算法收敛, 聚类区域稳定。

6) 当新的质心与旧的质心差距较大时, 重复迭代步骤 3) ~ 5)。

针对本文的数据集, 经过 YOLO-v5 自适应锚点框计算后所得锚点框设置如表 1 所示, 四舍五入取整后得到表 2 的锚点框数据, 更新模型中的原始锚点框进行神经网络训练。从表 2 中可见, 海面船舶小目标框也有明显的大小分别, 主要分布在 $5 \times 5 \sim 25 \times 21$ pixel。

表 1 K-means 聚类结果

序号	宽/pixel	高/pixel
1	4.6334	5.3697
2	7.1158	4.7732
3	5.0434	8.7878
4	7.0174	7.0368
5	6.3257	12.5903
6	11.8290	7.9728
7	8.7401	23.3960
8	17.7000	12.9680
9	25.3290	20.9221

表 2 星载 SAR 图像锚点框聚类

下采样	Anchor box 1/pixel	Anchor box 2/pixel	Anchor box 3/pixel
32	9×23	18×13	25×21
16	7×7	6×13	12×8
8	5×5	7×5	5×9

1.2 YOLO-v5 主干网络 (Backbone and Neck)

图 2 为 YOLO-v5 的主干网络结构, 该模型主要由注意力 (FOCOUS) 模块、卷积归一化 (convolution and batch normalization, CBL) 模块、跨阶段局部网络 (cross stage partial, CSP) 模块、跨阶段缩放局部网络 (scaling cross stage partial, SPP) 模块和张量拼接 (Concat) 模块构成。

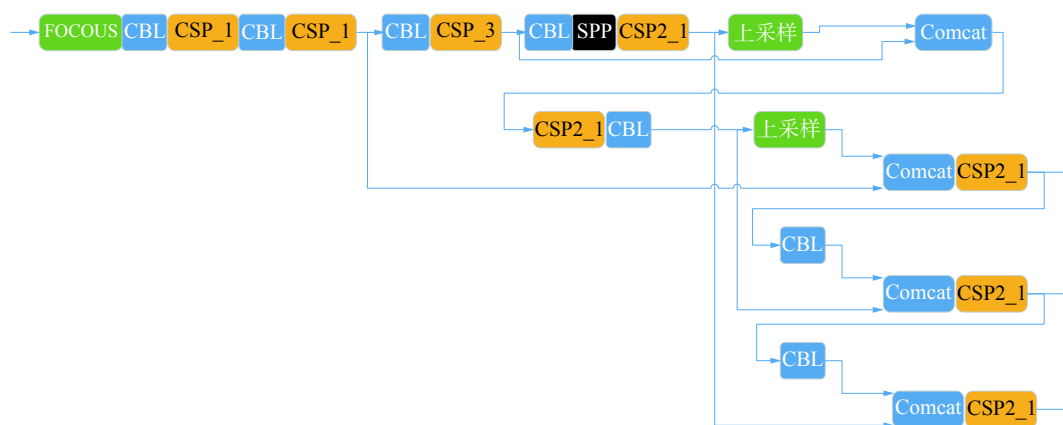


图 2 YOLO-v5 思维 Backbone 和 Neck 结构

其中 FOCOUS 是图像切片操作。如图 3 所示, 将图像按照像素格进行分割再融合。当原始图像输入为 $608 \times 608 \times 3$ pixel 时, 通过切片操作进一步提取特征变成 $304 \times 304 \times 12$ pixel 的图像, 再经过 32 个卷积核的操作最终变成 $304 \times 304 \times 32$ pixel 的特征图像。

CBL 是由卷积 (convolution)、批量归一化 (batch normalization, BN) 和激活函数 (Leaky_relu) 等 3 部分构成。由于输入的分布逐渐向非线性函数的两端靠拢, 神经网络收敛速度较慢, BN 层将输入的分布通过式 (4) 拉回到均值为 0、方差为 1 的正态

布, 从而使输入激活函数的值在反向传播中产生更明显的梯度, 避免了梯度消失的问题。

$$\hat{x}^{(k)} = \frac{x^{(k)} - E[x^{(k)}]}{\sqrt{\text{Var}[x^{(k)}]}} \quad (4)$$

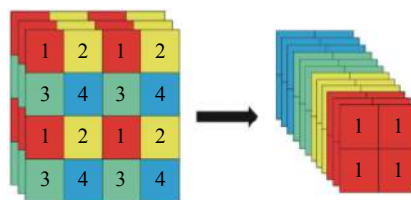


图 3 切片操作

式中: $x^{(k)}$ 为经过该层线性变换后的损失函数值, Var 为均方差操作符。

将输入分布变为标准状态分布后, 输入的值靠近中心的概率变大, 此时采用 sigmoid^[18] 函数, 即使输入存在微小的变化, 也能够反向传播时产生较大的变化。Leakyrelu 激活函数如式 (5) 所示。

$$f(x) = \begin{cases} x, & x \geq 0 \\ ax, & x < 0 \end{cases} \quad (5)$$

式中 a 采取一个很小的数值, 本文设置为 0.01。传统激活函数直接取值为 0, 这样可以保证输出小于 0 的神经元也进行小幅度的更新。

SPP 采用 1×1 、 5×5 、 9×9 和 13×13 的最大池化方式, 进行多尺度融合操作。Concat 为扩充维度的张量拼接。YOLO-v5 与 YOLO-v4 一样采用了 CSP Darknet53 的网络结构, 与 YOLO-v5 不同的是, YOLO-v4 中只有主干网络中设计了 CSP (由卷积层和 Res unit 模块张量拼接而成) 结构, 而 YOLO-v5 在主干网络 Backbone 和 Neck 中设计了 2 种不同的 CSP 结构。CSP 结构主要优点是在网络模型轻量化的同时保证准确性, 同时降低了对计算机设备的要求。新增的 CSP2 进一步加强了网络特征融合的能力。

1.3 损失函数计算

损失函数计算公式如式 (6) 所示, 表示 2 个方框所在区域的交并比 (intersection over union, IoU)。

$$X_{IoU} = \frac{|A \cap B|}{|A \cup B|} \quad (6)$$

如图 4 所示, 2 个方框完全重合时, $X_{IoU}=1$; 2 个方框交集为空时, $X_{IoU}=0$; 2 个方框重叠一部分时, X_{IoU} 的值在 0 和 1 之间。但是当 2 个框的交集为 0 时, 不管 2 个框相距多远, IoU 损失函数值恒等于 0, 无法表示该情况下的损失大小。在 YOLO-v5 中采用了 IoU 的损失函数公式为

$$X_{GloU} = X_{IoU} - \frac{|C/(A \cup B)|}{|C|}$$

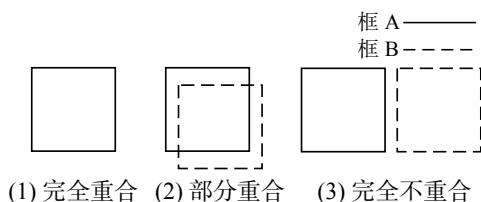


图4 IoU 损失函数算法

在 IoU 的基础上衍生出的 GloU^[19] 表示先计算 2 个方框的最小闭包区域面积 (也就是 2 个框重合的交集)。用 C 和 $A \cup B$ 比值的绝对值除以 C 的绝对值得到闭包区域中不属于 2 个框的区域的比重, 最后计算 IoU 与比重的差值, 最终得到 GloU

的值。在 2 个框无限趋近重合的情况下 $X_{GloU}=X_{IoU}=1$ 。

如图 5 所示, 采用 IoU 损失函数时, 当 2 个框不重合时, 无论差距多大, 损失函数都为 0。与 IoU 不同的是 GloU 算法不仅关注 2 个框重叠区域的大小, 也加入了非重合区域, 因此 YOLO-v5 避免了上述问题。

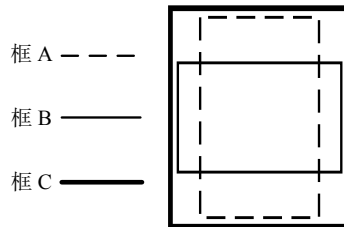


图5 GloU 损失函数算法

2 GDAL 模块

在星载 SAR 图像海洋目标检测时, 本文在检测头部分 (detect) 中嵌入了 GDAL 模块, 直接读取 TIF 图像中每个目标的经纬度坐标, 使得目标数据更为直观、清晰。

GDAL 是一个对遥感卫星等地理图像进行读取、写入和转换的库。遥感图像是一种带大地坐标的栅格数据, 每个栅格点所对应的数值为该点的像元值, 像元值包含了该点的大地坐标等空间投影信息, GDAL 通过仿射矩阵的坐标变换将栅格数据转换为经纬度等信息。

首先使用 GDAL 模块对遥感卫星图像进行图像校正 (需要指定 3 个已知正确的空间坐标点), 然后再进行目标经纬度的读取与显示。

在截取星载 (SAR) 图像的子图像时保留 TIF 格式, 从而保留遥感卫星图像的像素格属性。利用式 (7) 和式 (8) 进行栅格数据转换, 从而读出图像经纬度信息。

$$x_p = \text{padTransform}[0] + P \times \text{padTransform}[1] \quad (7)$$

$$y_p = \text{padTransform}[3] + L \times \text{padTransform}[5] \quad (8)$$

在经过校正的北向的星载 SAR 图像中, $\text{padTransform}[1]$ 是像素的宽度, $\text{padTransform}[5]$ 是像素高度, $\text{padTransform}[0]$ 和 $\text{padTransform}[3]$ 分别是星载 SAR 图像左上角的经纬度坐标。

3 实验结果与分析

3.1 实验平台

本文中的实验模型在 Pytorch0.8 框架上, 采用 Darknet53 学习网络, 在 PyCharm Community Edition 2020.2.1 x64 平台上实现。模型训练在 Titan2080Ti (显存 12GB) GPU, CUDA10.0 实验环境下完成, 操作系统为 Windows.x64。YOLO-v5 随

着神经网络宽度和深度的增加分为 YOLO-v5s、YOLO-v5m、YOLO-v5l 和 YOLO-v5x 这 4 个模型,本次实验采用 YOLO-v5l,并加入与 YOLO-v3 神经网络训练模型的对比。

3.2 数据集

数据集采用了 2020 年 1 月—2020 年 11 月期间由欧洲航天航空局 (European Space Agency, ESA) 拍摄的 Sentinel-1 星载 SAR 图像,选取长江三角洲、新加坡樟宜港口等船舶较多的码头港口

地区。数据集包含 41 张原始比例尺为 1:3000 的 SAR 图像,每张图像包含 1~40 个不同的船舶目标,对其中 100 张图像用 Labelimg 工具软件进行坐标标记。该数据集包含了不同分辨率以及不同背景下(港口、近海、远海)不同尺寸的舰船目标,引入不同的岛屿港口背景是为了增加要素,增强训练效果。

3.3 实验参数

本次实验部分实验参数如表 3 所示。

表3 训练模型参数

训练模型	训练批次(epochs)	单次训练图片数(batch-size)	图片尺寸	权重文件大小/MB
YOLO-v5l	3 000	8	640×640	90.8
YOLO-v3	3 000	8	640×640	117

3.4 实验结果与分析

训练过程参数如图 6 所示, YOLO-v3-spp 训练如图 7 所示。图 6 和图 7 中 (a) 为各类别 AP 均值 (mean average precision, mAP), 作为衡量网络模型训练的一个重要参数, 其中 R_{AP} 是以准确率

$R_{precision}$ 和召回率 R_{recall} 为两轴作图后围成的面积, R_{mAP} 表示平均, @后面的数表示判定 IoU 正负样本的阈值。图 6 和图 7 中 (b) 为找对的目标数量与找到的目标数量比值, 图 6 和图 7 中 (c) 为找对的目标数量与实际待检测目标数量的比值。

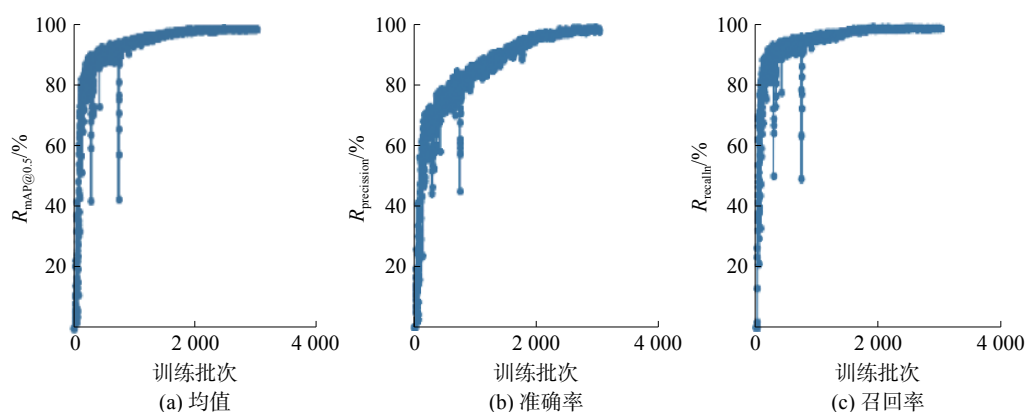


图6 YOLO-v5l 训练过程参数

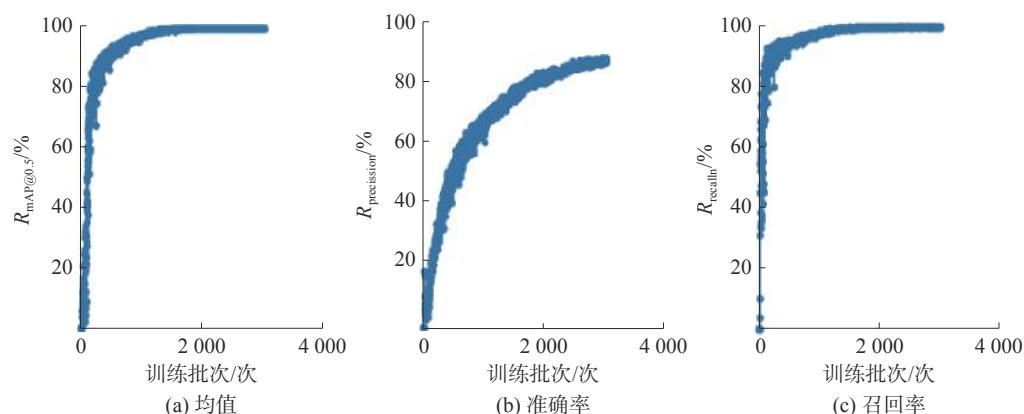


图7 YOLO-v3 训练过程参数

从图 6 和图 7 对比可以看出, YOLO-v3 在训练 2000 epochs 时, 准确率逐渐趋于平缓, 而 YOLO-v5l 的准确率有着进一步的训练上升空间。YOLO-

v3 经常出现错误的函数迭代方向, 如图 7 中的竖型分支, 主要原因为:

1) 进行了 Mosaic 数据增强, 使得增强后的数

据集目标分布更加均衡,并且这种重新组合图像的方式增强了数据集的丰富性,使得神经网络训练的鲁棒性更好。

2) 加入了 CBL 等模块的 YOLO-v5 在反向传播更新模型参数时失误较少,有着更为明显和正确的梯度,训练参数曲线较为平滑。

最终训练结果的具体参数如表 4 所示。YOLO-v5l 模型的召回率 R_{recall} 为 0.9945,相比于 YOLO-v3 提高了 1.78%; YOLO-v5l 的召回率 $R_{\text{mAP}@0.5}$ 为 0.9945,相比 YOLO-v3 提高了 1.18%。

表4 YOLO-v5l 与 YOLO-v3 模型训练结果对比

模型	R_{recall}	$R_{\text{mAP}@0.5}$
YOLO-v5l	0.9945	0.9945
YOLO-v3	0.9767	0.9827

截取长江口和新加坡樟宜港区域的星载 SAR 图像,采用改进后的 YOLO-v5 检测模型并嵌入 GDAL 模块后,水面船舶目标的检测结果分别如图 8 和图 9 所示。

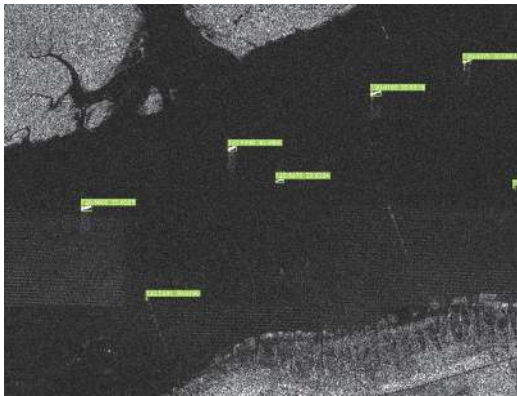


图8 长江口区域的目标检测结果



图9 樟宜港区域的目标检测结果

在表 5 和表 6 中,包括了目标在图像中的坐标以及在地理位置上的经纬度信息。表 5 和表 6 中的经纬度信息截取小数点后 4 位,用 GDAL 模

块读取的星载 SAR 图像经纬度信息与目标实际地理位置信息一致。目标的经纬度采用目标在图像中的检测框的中心位置。

表5 长江口区域目标位置信息

目标编号	图像坐标x/pixel	图像坐标y/pixel	目标框长度/pixel	目标框宽度/pixel	经度/(°)	纬度/(°)
1	936	331	4	5	120.6417	30.6114
2	260	533	5	7	120.5681	30.6190
3	499	318	5	7	120.5976	30.6034
4	411	258	6	10	120.5898	30.5968
5	844	100	15	7	120.6375	30.5893
6	141	367	10	10	120.5600	30.6023
7	674	157	19	10	120.6190	30.5919

表6 樟宜港区域目标位置信息

目标编号	图像坐标x/pixel	图像坐标y/pixel	目标框长度/pixel	目标框宽度/pixel	经度/(°)	纬度/(°)
1	579	549	10	6	102.2831	0.1695
2	28	214	4	9	102.2408	0.1288
3	770	28	8	33	102.3094	0.1272
4	767	281	6	10	102.3045	0.1489
5	636	541	8	6	102.2882	0.1697
6	55	83	5	8	102.2456	0.1175
7	0	151	6	16	102.2395	0.1230
8	40	191	10	5	102.2426	0.1269
9	392	274	10	6	102.2719	0.1411
10	807	446	10	7	102.3051	0.1644
11	640	606	9	8	102.2873	0.1758
12	55	155	9	8	102.2444	0.1241
13	75	152	7	9	102.2462	0.1241
14	327	113	8	8	102.2691	0.1254
15	2	182	10	9	102.2394	0.1255
16	685	579	9	9	102.2918	0.1742
17	84	184	11	10	102.2465	0.1273
18	61	228	13	10	102.2438	0.1309
19	206	113	4	10	102.2587	0.1233
20	291	38	18	15	102.2678	0.1183
21	81	99	18	13	102.2482	0.1198
22	8	226	12	12	102.2391	0.1298
23	119	134	17	16	102.2507	0.1238
24	207	25	21	13	102.2607	0.1155
25	256	69	18	19	102.2641	0.1206
26	919	72	21	16	102.3224	0.1333
27	784	87	9	31	102.3096	0.1327
28	17	95	15	20	102.2424	0.1185
29	666	87	19	19	102.2996	0.1304
30	850	77	17	32	102.3159	0.1332
31	598	19	26	29	102.2953	0.1231

可以看出,在图8中,对于这种没有港口、海面开阔和船舶目标数量较少的星载SAR图像,改进后的YOLO-v5检测模型漏检率很低,即使目标在图像中占比较小也能准确地检测出来。在图9中,港口沿岸情况复杂、干扰因素较多、目标检测效果仍然较好,并且该模型能够有效检测到动态船舶目标,能够满足对于进出港口船舶的动态检测、定位和跟踪等应用需求。

4 结论

本文基于Darknet神经网络,提出了利用优化的YOLO-v5网络模型的目标检测的算法。

1)本算法泛化性较强,对于大范围的水面船舶小目标检测效果较好,并且检测时间较短,分辨率为 720×720 pixel的图像平均检测时间小于1 s。

2)加入CBL模块的YOLO-v5比YOLO-v3在训练中的反向传播更为稳定,更新梯度更为平滑。

3)优化后YOLO-v5的模型降低了漏检率,在提高检测效果的同时并没有增加检测模型文件的内存。

在未来的研究中,需要进一步优化网络模型结构,加快检测速度,并且用于与其他检测目标的手段进行实时的数据融合。

参考文献:

- [1] 马啸,邵利民,金鑫,等.舰船目标识别技术研究进展[J].科技导报,2019,37(24): 65–78.
- [2] KUANG Chunming, WANG Caijun, WEN Biyang, et al. An improved CA-CFAR method for ship target detection in strong clutter using UHF radar[J]. IEEE signal processing letters, 2020, 27: 1445–1449.
- [3] FERRETTI A, PRATI C, ROCCA F. Analysis of permanent scatterers in SAR interferometry[C]// Proceedings of IEEE 2000 International Geoscience and Remote Sensing Symposium. Taking the Pulse of the Planet: The Role of Remote Sensing in Managing the Environment. Honolulu, USA, 2000.
- [4] 张世辉,王红蕾,陈宇翔,等.基于深度学习利用特征图加权融合的目标检测方法[J].计量学报,2020,41(11): 1344–1351.
- [5] 王莉莉,张晓.基于Faster R-CNN的卫星图像污水处理厂识别[J].计算机应用,2019,39(S2): 50–54.
- [6] LIU Gang, WANG Chuyi. A novel multi-scale feature fusion method for region proposal network in fast object detection[J]. International journal of data warehousing and mining, 2020, 16(3): 132–145.
- [7] DRUZHKOVA P N, KUSTIKOVA V D. A survey of deep learning methods and software tools for image classification and object detection[J]. Pattern recognition and image analysis, 2016, 26(1): 9–15.
- [8] HE Kaiming, ZHANG Xiangyu, REN Shaoqing, et al. Deep residual learning for image recognition[C]// Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, USA, 2016.
- [9] REN Shaoqing, HE Kaiming, GIRSHICK R, et al. Faster R-CNN: towards real-time object detection with region proposal networks[J]. IEEE transactions on pattern analysis and machine intelligence, 2017, 39(6): 1137–1149.
- [10] 郭昕刚,张培栋,梁锦明,等.改进Faster R-CNN目标检测方法[J].长春工业大学学报,2020,41(5): 474–480.
- [11] 赵文清,程幸福,赵振兵,等.注意力机制和Faster RCNN相结合的绝缘子识别[J].智能系统学报,2020,15(1): 92–98.
- [12] 张慧,王坤峰,王飞跃.深度学习在目标视觉检测中的应用进展与展望[J].自动化学报,2017,43(8): 1289–1305.
- [13] 伍鹏瑛,张建明,彭建,等.多层卷积特征的真实场景下行人检测研究[J].智能系统学报,2019,14(2): 306–315.
- [14] 段敬雅,李彬,董超,等.基于YOLOv2的船舶目标检测分类算法[J].计算机工程与设计,2020,41(6): 1701–1707.
- [15] 陈科峻,张叶.基于YOLO-v3模型压缩的卫星图像船只实时检测[J].液晶与显示,2020,35(11): 1168–1176.
- [16] 刘畅,朱卫纲.基于卷积神经网络的SAR图像目标检测综述[J].兵器装备工程学报,2021,42(3): 15–21.
- [17] 涂豫.基于贝叶斯卷积神经网络与数据增强的SAR图像目标分类方法[J].探测与控制学报,2020,42(6): 43–48.
- [18] 张舞杰,李迪,叶峰.基于Sigmoid函数拟合的亚像素边缘检测方法[J].华南理工大学学报(自然科学版),2009,37(10): 39–43.
- [19] REZATOFIGHI H, TSOI N, GWAK J Y, et al. Generalized intersection over union: a metric and a loss for bounding box regression[C]//Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach, USA, 2019.

本文引用格式:

窦其龙,颜明重,朱大奇.基于YOLO-v5的星载SAR图像海洋小目标检测[J].应用科技,2021,48(6): 1–7.

DOU Qilong, YAN Mingzhong, ZHU Daqi. Small marine target detection in space-borne SAR image based on YOLO-v5[J]. Applied science and technology, 2021, 48(6): 1–7.