

一种改进的 YOLOv5 海洋船舶目标检测方法

侯旭佳¹ 张博强¹ 张飞虎¹

摘要 中文针对目前船舶样本数据集稀缺情况,设计了一种基于 YOLOv5 的海洋船舶目标检测方法。通过采用 Mosaic 数据增强方式,明显改善了小目标检测效果。采用自适应锚框计算及自适应图片缩放,提升了模型泛化性。2020 首届海洋目标智能感知国际挑战赛赛道公开测试集 A 榜测得平均检测精度为 62.19。为此后的海洋目标检测提供了一种新的检测思想,为实时检测船舶目标提供了有效方法。

关键词 深度学习, 船舶目标检测, 小样本训练

中图分类号: TP391.41;TP183

An improved YOLOv5 ocean ship target detection method

Hou Xujia^{1,2} Zhang Boqiang² Zhang Feihu¹

Abstract Aiming at the scarcity of ship sample data set, a method of ocean ship target detection based on YOLOv5 is designed. By using Mosaic data enhancement method, the detection effect of small targets was significantly improved. The model generalization is improved by using adaptive anchor frame calculation and adaptive image scaling. The average detection accuracy measured in the open test set A of the first Ocean Target Intelligent Perception International Challenge Race in 2020 was 62.19. It provides a new idea for ocean target detection and an effective method for real-time detection of ship targets.

Key words Deep learning, Ship target detection, Small sample training

1 引言

海洋领域智能技术是国家实施海洋强国战略的重要基础和支撑,也是国际海事界及船舶工业领域关注的热点。作为海上交通运输的主要交通工具--船舶,在智慧海洋背景下越来越受关注。为加速智能及信息新兴技术与海洋和航运领域的深度融合和技术赋能稿件,由中国造船工程学会、国际船舶与海洋工程创新与合作组织等组织举办了首届海洋目标智能感知国际挑战赛。通过完成多类船舶目标识别任务,将由更多机会深入了解国内外航海领域行业背景和广阔前景,对促进科技创新及实践能力高度融合具有重要意义。针对本次大赛,结合精度与速度综合考量,采用基于 YOLOv5 的目标检测算法进行参赛,在公开测试集 A 榜取得 62.19 分。

1. 西北工业大学航海学院

1. School of Marine Engineering, Northwestern Polytechnical University, Xi'an 710000

2 相关工作

2.1 YOLOv1-YOLOv4 算法介绍

传统目标检测方法:滑动窗口分类法,耗时,操作复杂,需要手动生成大量样本,检测性能依靠的是分类器的性能和遍历的性能。因此 yolo v1 实现了通过直接拟合坐标位置 x, y , 物体宽和高 w, h , 置信度的算法。存在很明显的缺陷坐标位置 x, y 可以任意值,也可以为没有意义的负值,降低了训练速度和预测的准确率。因此, yolo v2 改为偏移量预测:采用了 anchor 机制 (location prection) 局部预测,也就是坐标位置 x, y 是通过 gird 预测,然后宽和高 w, h 是依据 anchor 的宽和高预测。Yolo v2 的 backbone 改为了 darknet19, 训练不同 epoch 后的调整训练图,也就是 Multi-Scale Training, 连接检测头部分取消了 FC 的连接方式,改为卷积层。可是 yolo v2 还是存在小目标预测不准确的问题。随着残差网络 resnet 对 backbone 性能的提升, yolo v3 才有了新的性能提升。首先是因为 resnet 出现了,网络可以做的更深了, backbone 更改为 darknet53, 然后检测头也更改为多尺度,统计数据得来的 9 个尺度, 32 倍下采样, 16 倍下采样, 8 倍下采样, 分别用 3 个检测头来预测。到 yolo v4 在 backbone 更改为采用 CSPDarknet53, 同时采用了一些技巧进行提升性能, 首先数据增强采用了 cutmix and mosaic, class label smoothing, Mosaic 就是合并 4 张图,起到了跨越上下进行标检测的作用,同时相当于 minibatch 的数量扩大了 4 倍,同样的 GPU 内存训练提升了性能。损失函数改为 mish。网络结构采用了 SPP、PAN、SAM 的网络, DropBlock CmbN。YOLO v1-v3 (location prection) 因为激活用的 sigmoid 区间是 (0, 1) 开区间, yolo v4 对此进行了改进。在 sigmoid 乘上 1.01。

2.2 YOLOv5 网络分析

YOLOv5 按照网络深度大小和特征图宽度大小分为 Yolo v5s、Yolo v5m、Yolo v5l、Yolo v5x。这四个模型精度逐次变高,相应的速度也逐渐变慢。本文采用了 yolo v5x 作为使用模型。Yolo v5 的结构分为 input, backbone, Neck, 预测层。

(1) 在输入端使用了 Mosaic 的数据增强方式,随机调用 4 张图片,随机大小和分布,进行堆叠,

丰富了数据，增加了很多小目标，提升小物体的识别能力。可以同时计算 4 张图片，相当于增加了 Mini-batch 大小，减少了 GPU 内存的消耗。Yolov5 首先也可以通过聚类设定 anchor 大小，然后还可以在训练过程中，在每次训练时，计算不同训练集中的 anchor 值。然后在预测时使用了自适应图片大小的缩放模式，通过减少黑边，提高了预测速度。

(2) 在 Backbone 上的主要是采用了 Focus 结构，CSPnet 结构。

(3) 在 Neck 上采用了 FPN 结构和 PAN 结构。

(4) 在损失函数则使用了 GIoU_Loss。

$$GIoU = IoU - \frac{|C \setminus (A \cup B)|}{|C|}$$

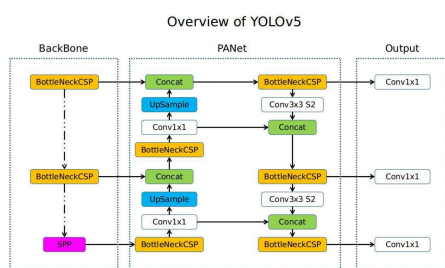


图 1 YOLOv5 结构框图

Fig.1 Structure block diagram of YOLOv5

3 模型和方法描述

3.1 模型和方法原理

为完成海洋船舶目标检测的任务，我们使用 YOLO-V5 目标检测算法，使用给出数据集，进行训练。我们在输入端采用 Mosaic 数据增强方式，同时对数据集中的样本使用随机旋转、随机水平翻转、尺度变换的图像增强方式，提高模型的泛华性能。

在训练时，我们合理设置学习率，采用学习率指数衰减的衰减方式，保证模型性能达到较高水平。

在数据测试时，我们采用测试时增强(Test Time Augmentation)。这里会为原始图像造出多个不同版本，包括不同区域裁剪和更改缩放程度等，并将它们输入到模型中；然后对多个版本进行计算得到平均输出，作为图像的最终输出分数。这种技术很有效，因为原始图像显示的区域可能会缺少一些重

要特征，在模型中输入图像的多个版本并取平均值，能解决上述问题。

3.2 在现有技术之上的创新之处

(1) 使用最新目标识别算法框架 YOLOV5。

(2) 对数据集实施多种类型的数据增强方式，提高算法的泛化能力。

(3) 使用测试时增强，提高使用时目标检测准确率。

(4) 使用预训练模型参数，加快训练速度，提高目标识别精度。

4 实验结果

实验中，首先根据数据集进行训练集与测试集的划分，我们将 9800 张数据集照片以 9000:800 的比例划分训练集与数据集。此后将组委会提供的标签从 xml 格式转为训练需要的 txt 文件格式。对训练集进行统计分析可得整个数据集小样本最多，container ship 数量最少，other ship 数量最多。且目标大多分布在图像中间。

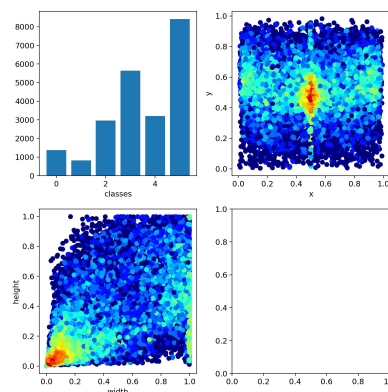


图 2 数据集分布

Fig.2 Data set distribution

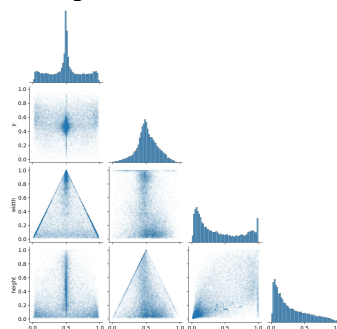


图 3 标签相关性

Fig.3 Label correlation

划分好数据集后，开始准备训练，图形显卡采用 NVIDIA GeForce RTX 2080 Ti。训练参数 epochs 为 70，batch_size 为 4，将图像尺寸进行 $\pm 50\%$ 的缩

放，采用 YOLOV5x 作为预训练模型。
训练过程中的训练过程如图

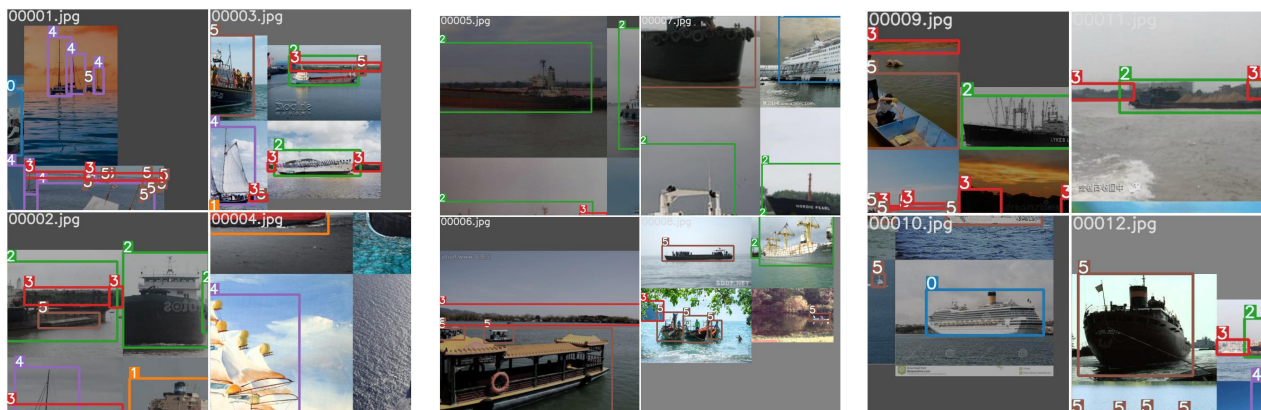


图 4 多尺度混合训练
Fig.4 Multi-scale mixed training

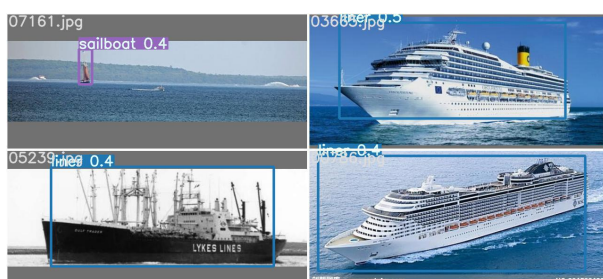


图 5(a) 首轮检测结果
Fig.5(a)First epoch detection result



图 5(b) 首轮标签
Fig.5(b)First epoch label



图 6 模型检测结果
Fig.6 ours results

5 结论

本文针对海洋船舶数据集缺少的情况，改进了基于 YOLOv5 的目标检测模型。为了充分利用数据集，在输入端使用了 Mosaic 的数据增强方式，随机调用 4 张图片，随机大

小和分布，进行堆叠，丰富了数据，增加了很多小目标，提升小物体的识别能力。在最终的测试集得到了较为良好的检测结果，在未来的工作中，将尝试通过注意力机制来改进，争取得到泛化性能更好的默认，更好的完成复杂的视觉任务。

References

1. 马小陆,方洋,王兵,吴紫恒.一种改进的 YOLO v3 红外图像行人检测方法 [J]. 湖北理工学院学报,2020,36(06):19-24+38.
2. 唐熔钊,伍锡如.基于改进 YOLO-V3 网络的百香果实时检测 [J]. 广西师范大学学报 (自然科学版),2020,38(06):32-39.
3. S.-H. Bae, "Object detection based on region decomposition and assembly," arXiv preprint arXiv:1901.08225, 2019.
4. Ba, V. Mnih, and K. Kavukcuoglu, "Multiple object recognition with visual attention," arXiv preprint arXiv:1412.7755, 2014.

作者一 侯旭佳, 西北工业大学航海学院 2018 级本科生, 专业为信息工程。获得第九届全国海洋航行器设计与制作大赛全国一等奖、二等奖各一项, 全国大学生计算机设计大赛西北赛区二等奖。

FIRST Author Hou Xujia is an undergraduate of the Navigation College of Northwest University of Technology, majoring in information engineering. He has won one first and one second prize respectively in the ninth National Ocean Vehicle Design and Production Competition, and the second prize in the Northwest Division of The National College Students Computer Design Competition.