

基于 YOLOv5 的海洋目标检测

沈纪元^{1,2} 杨扬^{1,2} 祝嘉诚^{1,2}

摘要 多类别海洋船舶目标检测识别对于船舶自动驾驶技术、智能船舶等领域发展有重要意义。本次比赛通过建立 YOLOv5 模型,对海洋船舶等目标进行检测,达到了较高的精度和速度,具备良好的实用性。测试了多种优化方法,其中标签平滑和测试时增强可以在原模型基础上有效提高准确率。

关键词 YOLOv5, 目标检测, 优化

Ocean Object Detection Based on YOLOv5

JiYuan Shen^{1,2} Yang Yang^{1,2} JiaCheng Zhu^{1,2}

Abstract The detection and recognition of multi-class ocean ship targets plays an important role in the development of ship auto-navigation technology and other related fields. In this competition, YOLOv5 model is established to detect targets such as ocean ships. It achieves high accuracy and speed, which leads to good practicability. A variety of optimization methods are tested. Among them, label smoothing and test time augment can effectively improve the accuracy based on the original model.

Key words YOLOv5, object detection, optimization

1 引言

2020年首届海洋目标智能感知国际挑战赛立足于打造声光、电融合的360度全方位、全天候的新型海面环境感知系统,为后续船舶自动驾驶技术、智能船舶领域发展的重要技术和基础设备。

该比赛既有很强实际工程意义,又能够培养大学生的创新能力,从大学生的创新点出发,为我国船舶航行态势智能感知领域的技术发展助力。

本次竞赛的任务是多类别海洋船舶目标检测识别。通过定位在图片中出现目标的位置,识别每个目标示例的类别和属于该类的置信度。

本次竞赛的数据集包括6类检测目标,分别为: liner (游轮), container ship (集装箱船), bulk carrier (散货船), island reef (岛礁), sailboat (帆船), other ship (其他船), 各类检测目标的分布情况如图1所示,各类别分布存在一些不平衡的情况,集装箱船类最少,数量不到1000个,而其他类别最多,数量超过7000个。

1.上海大学计算机工程与科学学院 上海 200444 2. 这三位作者对于本文贡献相同,排名按照首字母顺序,不分先后
1. School of Computer Engineering and Science, Shanghai University, Shanghai 200444 2. The three authors contribute equally to this paper. The rankings are in alphabetical order, regardless of order.

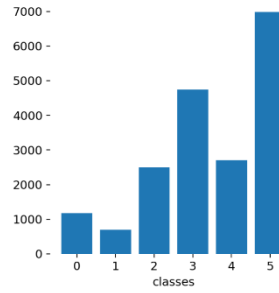


图1 训练集中6类检测目标的分布情况

Fig 1 Distribution of six categories of detection targets in training set

在比赛过程中,首先需要寻找合适的模型来作为baseline。常见的模型包括YOLO系列、RCNN系列等,其中YOLO系列在具有较高准确率的同时,保证了检测速度,有更强的实用性。然后可以尝试在数据增强做一些变化,包括使用标准化、几何变换等等。还可以尝试不同的损失函数、更换模型的backbone、调整训练策略和不同的NMS等方法来提升准确率。

2 模型和方法描述

2.1 模型和方法原理

本次比赛使用的模型为YOLOv5模型。YOLOv5模型虽是one-stage模型,但它比yoloV3更为强悍,能够在保证检测速度的同时也兼具精度,很适合落地应用以及实景检查,因此我们考虑该模型。

YOLOv5在输入端有三点改进,依次是Mosaic数据增强、自适应锚框计算、自适应图片缩放。传统的数据增强一般会对图像翻转一定角度,亮暗通道进行调整,遮盖,放缩, mixup等,而Mosaic在此基础上采用四张图片拼接,达到数据增强的作用。

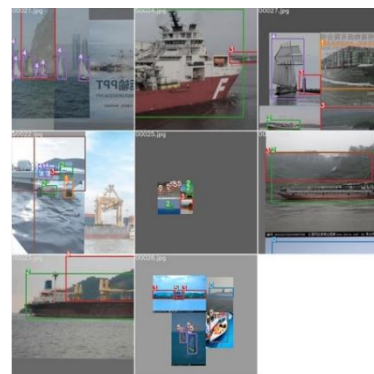


图2 数据增强样例

Fig 2 Example of Data Augmentation

YOLOv5 改变了以往要额外计算锚框的步骤, 采用遗传算法, 自适应锚框计算。在图片缩放上, 也抛弃了以往会产生黑边过多的现象, 采用对原始图像自适应, 添加最少的黑边。

在 backbone 上沿用 v4 的 CSPDarknet53, 但在此基础上加入了 Focus 结构, 同时在 neck 部分也用到了 CSP 结构。

在 neck 部分, 沿用了 FPN+PAN 的结构。在 FPN 层自顶向下的, 将高层的特征信息通过上采样的方式进行传递融合, 得到进行预测的特征图。在 FPN 层的后面还添加了一个自底向上的特征金字塔。其中包含两个 PAN 结构。这样结合操作, FPN 层自顶向下传达强语义特征, 而特征金字塔则自底向上传达强定位特征, 两两联手, 从不同的主干层对不同的检测层进行参数聚合, 能够更好地进行目标检测。

在最后的输出端, YOLOv5 中采用其中的 GIOU_Loss 做 Bounding box 的损失函数。nms 上, yolov4 在 DIOU_Loss 的基础上采用 DIOU_nms 的方式, 而 YOLOv5 中采用加权 nms 的方式, 使一些被遮盖的物体也能检测出来, 而且还不产生额外的计算成本。

2.2 模型训练细节

在数据集划分上, 总共有 9800 张图片, 按照 8:2 的比例划分训练集和验证集, 训练集有 7840 张图片, 验证集有 1960 张图片。

在训练模型的时候, 我们使用了在 coco 数据集上预训练的权重做迁移学习, 采用了 warm up 训练策略, 优化器选择的是 SGD, 初始学习率设置为 0.0032, batch size 设置为 8。

训练环境为 linux 系统, 显卡为 RTX 2080Ti, 显存为 11G。

2.3 在现有技术之上的创新之处

2.3.1 标签平滑

我们在训练模型时使用了标签平滑[1]的方法。标签平滑改变了真实标签的编码, 将将各类别真实标签的概率从 0,1 变为了如(1)所示的概率分布

$$P_i = \begin{cases} 1 - \varepsilon, & i = y \\ \frac{\varepsilon}{K-1}, & i \neq y \end{cases} \quad (1)$$

其中, i 代表某个类别, P_i 为真实类别的概率, y 为样本真实类别, K 为类别数量, ε 是一个较小的常数, 我们在训练时取 0.1。

使用标签平滑能够减少错误标注的干扰, 起到了一定的正则化作用, 增强了模型的泛化能力。

2.3.2 测试时增强

除此之外, 我们采用了测试时增强。测试时增强就是在测试的时候, 对图像进行多种数据增强, 让模型对于原图和各个增强后的图片进行预测, 对

结果取平均。这种方法可以增加预测结果的稳定性。

我们在测试的时候, 对图片进行了尺寸缩放、水平翻转等增强, 多次预测, 对结果进行集成。

2.3.3 调整损失函数

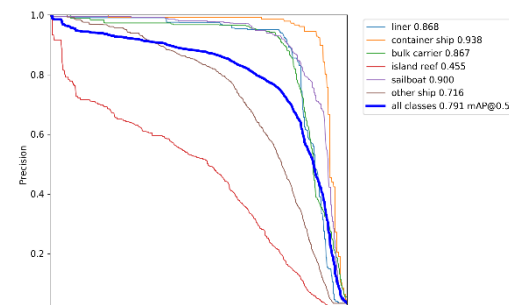


图 3 各类别验证准确率

Fig 3 Validation accuracy of each category

在验证集上, baseline 模型各个类别的验证准确率如图 3。观察发现, 各个类别准确率存在较大差异, 最低的类别为岛屿, 得分仅为 0.455, 最高的为集装箱船, 准确率为 0.938。因此我们认为准确率的差异, 可能与类别间本身的难度差异、类别数量不平衡、部分类别标注存在标准不统一等问题有关, 考虑在损失函数上做出改进。

我们测试了多种损失函数, 包括 BCELoss、Focal Loss[2]、QFocal Loss, BCEBlurLoss, 主要想增强模型的健壮性, 加强它对于部分类别的检测能力。

2.3.4 其他改进

我们发现分类结果中包含不少预测框位置正确, 但是结果分类错误的情况。因此我们尝试通过外接分类器, 来减少分类错误。

除此之外, 我们发现数据集中包含一些标注, 我们尝试手动剔除有错误的标记类。

3 实验结果

标签平滑起到了不错的效果, 让我们的结果提升了 1 分左右。

测试时增强十分有效, 让结果提升了 3 分。

外接分类器的效果不是很好, 在测试集上分数反而下降了一些。我们查看了一些图片, 发现问题主要在检测到的框内有时候包含两个物体, 比如包含船的方框内, 背景有很多是岛屿, 由于岛屿面积比较大, 分类器往往把这张情况分为岛屿, 而实际标签是船。

剔除部分数据后, 训练出的模型表现也略有下降, 可能是由于手动剔除错误标记的数据存在一定

偏差,反而剔除了正确标记的数据,从而导致数据量的减少。

损失函数的测试结果显示,经典的 BCELoss 效果相对更好,而 Focal Loss 和 QFocal Loss 等损失函数效果比较一般,这和 yolov3 论文中提到的结果类似。而 YOLOv5 中新提出的 BCEBlurLoss 在我们做的测试里没有明显的提升。

4 结论

本次比赛中,我们基于 YOLOv5 建立了目标检测模型,实现了对于多类别海洋目标的检测,在精度和速度上都有不错的表现。除此之外,我们采用了标签平滑和测试时增强有效提升了检测准确率。

在我们测试的优化方案中,有一些方法没有提高精度,但是我们认为再进行一定的改进是可以起作用的。首先是数据集中存在少量标注错误的图片,手动去除的效果不是很好,可以尝试采用一些自动的方法,比如使用神经网络对于图片进行特征提取,然后对特征向量进行聚类,从而发现异常数据。此外,外接分类器效果不好,是由于分类器对于图片中的重叠类别区分不好,而目前有一种 3D 的神经网络,能区分图片中的前景和背景,这种方法或许能够提升分类的准确率。这几点是我们之后可以尝试继续提高检测精度的潜在方法。

References

1. Muller R, Kornblith S, Hinton G. When does label smoothing help?. In: *Advances in Neural Information Processing Systems*, 2019. 4694--4703
 2. Lin T, Goyal P, Girshick R, He K, Dollár P. Focal loss for dense object detection. In: *Proceedings of the IEEE international conference on computer vision*, 2017. 2980-2988
- 沈纪元** 上海大学计算机工程与科学学院本科大三学生,修读专业为智能科学与技术,主要研究方向为机器学习、计算机视觉。
Email: 815462629@qq.com
- JiYuan Shen** Junior students at School of Computer Engineering and Science, Shanghai University. He majors in intelligent science and technology. His research interest covers machine learning and computer vision.
- 杨扬** 上海大学计算机工程与科学学院本科大三学生,修读专业为智能科学与技术,主要研究方向为机器学习、计算机视觉。
Email: tony0821@shu.edu.cn
- Yang Yang** Junior students at School of Computer Engineering and Science, Shanghai University. He majors in intelligent science and technology. His research interest covers machine learning and computer vision.
- 祝嘉诚** 上海大学计算机工程与科学学院本科大三学生,修读专业为智能科学与技术,主要研究方向为机器学习、计算机视觉。
Email: 1398359669@qq.com
- Jiacheng Zhu** Junior students at School of Computer Engineering and Science, Shanghai University. He majors in intelligent science and technology. His research interest covers machine learning and computer vision.

1. Muller R, Kornblith S, Hinton G. When does label