

融合改进密集连接和分布排序损失的遥感图像检测

袁磊¹ 刘紫燕¹ 朱明成¹ 马珊珊¹ 陈霖周廷²¹ 贵州大学大数据与信息工程学院 贵阳 550025² 贵州理工学院航空航天工程学院 贵阳 550003

(17508561920@163.com)

摘要 针对遥感图像中小目标尺寸较小、样本分布不均匀、特征不明显等问题,提出一种改进的 YOLOv3 目标检测算法。在使用 Stitcher 数据增强解决小目标样本分布不均匀的问题后,提出 VOVDarkNet-53 基础网络,将 DarkNet-53 基础网络中第 4 次下采样后的 8 个残差模块减少为 4 个残差模块。然后采用 VOVNet 的密集连接方式,使网络利用更多的浅层小目标特征信息,增加网络感受野。最后,采用分布排序损失改进 YOLOv3 中的分类损失,解决单阶段目标检测器正负样本不平衡的问题。实验使用 YOLOv3 目标检测算法和改进后的 YOLOv3 算法在 HRRSD 遥感数据集上进行对比。结果表明,改进后的 YOLOv3 算法对小目标和中目标的检测精确度分别提升了 7.2% 和 2.1%,尽管对大目标的检测精度下降了 1%,但在平均单张图片处理时间几乎不变的情况下,平均检测精度均值(mAP)提升了 4.1%,召回率和准确率也有所提升。

关键词: 遥感图像;目标检测;YOLOv3;基础网络;样本不平衡

中图法分类号 TP391.4

Improved YOLOv3 Remote Sensing Target Detection Based on Improved Dense Connection and Distributional Ranking LossYUAN Lei¹, LIU Zi-yan¹, ZHU Ming-cheng¹, MA Shan-shan¹ and CHEN Lin-zhou-ting²¹ College of Big Data and Information Engineering, Guizhou University, Guiyang 550025, China² School of Aerospace Engineering, Guizhou Institute of Technology, Guiyang 550003, China

Abstract Aiming at solving the problems of small object size, uneven sample distribution, and unclear features in remote sensing images, an improved YOLOv3 object detection algorithm is proposed. The Stitcher data enhancement method is used to solve the problem of uneven distribution of small object samples. The VOVDarkNet-53 is proposed. The residual modules of the fourth downsampling in DarkNet-53 are reduced from eight to four. And then the dense connection mode of VOVNet is adopted to extract lower features of small objects to increase the network receptive field. The distributional ranking loss is used to improve the classification loss in YOLOv3 to solve the problem of imbalance between positive and negative samples in single-stage object detector. Comparative experiments are carried out on HRRSD remote sensing datasets by using YOLOv3 object detection algorithm and improved YOLOv3 algorithm. The results demonstrate that the proposed algorithm can achieve better performance of higher detection accuracy of the improved YOLOv3 algorithm for small objects and medium objects are improved by 7.2% and 2.1%, respectively. Although the detection accuracy for large objects is reduced by 1%, the average detection accuracy (mAP) is improved by 4.1%, and the recall and accuracy are also improved.

Keywords Remote sensing image, Object detection, YOLOv3, Baseline, Sample imbalance

到稿日期:2020-08-01 返修日期:2020-09-10 本文已加入开放科学计划(OSID),请扫描上方二维码获取补充信息。

基金项目:贵州省科学技术基金资助项目(黔科合基础[2016]1054);贵州省联合资金资助项目(黔科合 LH 字[2017]7226);贵州大学 2017 年度学术新苗培养及创新探索专项(黔科合平台人才[2017]5788);贵州省科技计划项目(黔科合基础[2017]1069);贵州省教育厅创新群体重大项目(黔教合 KY 字[2018]026);贵州省普通高等学校工程研究中心(黔教合 KY 字[2018]007);贵州省科技计划重点项目([2019]1416)

This work was supported by the Science and Technology Foundation of Guizhou Province, China([2016] 1054), Joint Funding Project of Guizhou Province, China([2017]7226), Special Project on the Training and Innovation of New Academic Seedlings of Guizhou University in 2017, China([2017] 5788), Science and Technology Program of Guizhou Province, China([2017] 1069), Major Research Projects on Innovative Groups of Guizhou Provincial Department of Education([2018]026), Engineering Research Center of Ordinary Colleges and Universities of Guizhou Province([2018]007) and Key Projects of Science and Technology Program of Guizhou Province([2019]1416).

通信作者:刘紫燕(Leizy@sina.com)

1 引言

遥感图像目标检测是目标检测领域的研究热点之一,被广泛应用于舰船监测、渔业管理、环境监控、敌方军事动态监控等民用和军用领域^[1]。从遥感图像中检测和定位物体的方法很多,但是,对于有噪声和低分辨率的遥感图像,尤其是当对象是小目标^[2]时,这些方法的检测效果较差,即使在高分辨率图像中,对小物体的检测性能远低于对大物体的检测性能^[3]。

当前的目标检测算法在基准数据集 ImageNet^[4],PASCAL VOC^[5]和 COCO^[6]上有着出色的表现。在仅关注精度的算法中,如 CBNet^[7],通过合并功能强大的基础网络(backbone),提出复合骨干网络(composite backbone network),以此提高目标检测器的性能,在 COCO 数据集上的 *mAP* 达到了 53.3%,因此该方法为当前目标检测精度最高的方法,但其需要极为庞大的计算量;EfficientDet^[8]算法以 52M 的参数量和 326B FLOPS,在 COCO 数据集上获得了 51.0% 的平均检测精度均值;Song 提出的 TSD^[9]算法,从空间维度和特征提取两方面同时对分类任务和定位任务进行拆解,在 COCO 数据集上获得了 51.02% 的平均检测精度均值;谷歌提出的 SpineNet^[10],通过 NAS 搜索获取效果更好的网络,具有多尺度排列特征和交叉尺度连接,在 COCO 数据集上达到了 49.2% 的平均检测精度均值。而在关注速度与平衡的算法中,RetinaNet^[11]采用新的分类损失 Focal Loss,解决正负样本不平衡的问题,以单张图像 90ms 的处理速度在 COCO 数据集上达到了 53.1% 的平均检测精度均值的效果;YOLOv3^[12]算法则采用 Darknet53 作为 backbone,采用多尺度预测,在 COCO 数据集以单张图像 51ms 的处理速度在 IOU 阈值 0.5 情况下达到了 57.9% 平均检测精度均值的效果;FCOS^[13]采用全像素卷积的方法,实现了无锚点、无提议的解决方案;CornerNet^[14]采用预测角点的方式代替预测框,获得了更快的处理速度。

由于遥感图像成像距离较远,导致图像中包含的目标,例如飞机、车辆、小房子、小型石油和天然气储罐等占据图像中较少的像素^[15],属于典型的小目标^[16]。因此,基于基准数据集的目标检测算法在遥感图像数据集上的表现不够理想,往往需要对这些基于基准数据集的算法进行改进以使其适应遥感数据。Zhang 等^[17]在 YOLOv2 目标检测算法的基础上进行改进,减少了卷积层数与维度,并结合特征金字塔的思想增加了检测尺度,达到了提高检测精度的目的。为了更好地检测遥感图像中的目标,本文通过改进 YOLOv3 目标检测算法来提高其遥感图像目标检测性能。本文的主要贡献如下。

(1)采用 Sticher 数据增强,将数据集中随机 4 张图片融合为一张进行训练,保证小目标样本分布均匀,提升小目标的检测精度。

(2)提出一种新的基础网络 VOVDarkNet-53,在 DarkNet-53 网络的基础上,通过引入改进的密集连接来提升网络感受野,获取更多小目标特征,提升 YOLOv3 对小目标的检测精度。

(3)改进 YOLOv3 中的分类损失,以 DR Loss 代替网络中的分类损失,解决正负样本不平衡的问题,提升检测精度。

2 YOLOv3 网络

YOLOv3 是 Joseph 等提出的一种目标检测器,其网络结构如图 1 所示。其基础网络采用 DarkNet-53,借鉴了 ResNet 的残差思想,在网络中形成多个残差单元(ResNet unit),可以很好地控制梯度的传播,避免出现梯度消失或者爆炸等不利于训练的情形,这使得训练深层网络的难度大大减小。如图 2 所示,每个残差单元由一个 1×1 的 DBL 模块、一个 3×3 的 DBL 模块以及一个跳过连接层组成,而 DBL 模块由一个卷积层、一个 BN 正则化模块和激活函数组成,其各个残差单元之间依序向后连接。

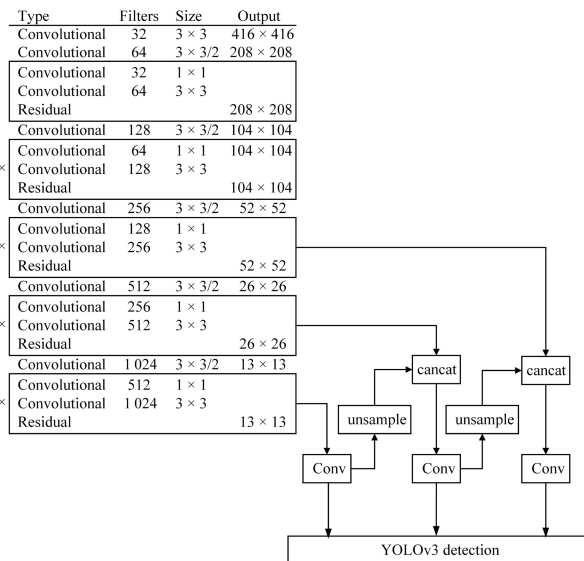


图 1 YOLOv3 网络结构

Fig. 1 YOLOv3 network structure

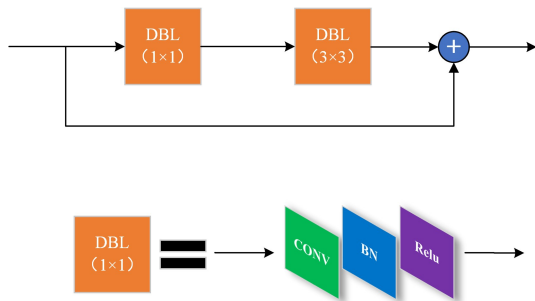


图 2 残差单元结构和 DBL 模块

Fig. 2 Residual unit structure and DBL module

YOLOv3 网络的检测层由 13×13 , 26×26 , 52×52 , 3 个尺度组成,这 3 种尺度分别输出 3 种不同尺度的特征图到检测层中,其中大尺度特征图包含的小目标信息较多,负责检测小目标,小尺度特征图对大目标较为敏感,容易检测出大目标。因此,YOLOv3 在检测大目标和小目标上都有不错的表现。

3 改进的 YOLOv3 网络模型

如前文所述,由于遥感图像成像距离远,图像中不仅包含十字路口、桥梁这样的大目标,也包含许多如汽车这样的小目标,而 YOLOv3 模型虽然比 YOLOv2 对小目标的检测效果好,但仍然不能很好地检测遥感小目标。同时,由于本文所采

用的 HRRSD 遥感数据集^[18]中每幅图像中大小目标分布极不均匀。因此,本文首先采用 Stitcher 数据增强方法缓解大小目标样本不平衡的问题,然后改进 DarkNet-53 基础网络以提高网络检测精度,最后采用分布式排序损失计算分类损失,缓解正负样本不平衡造成的精度下降。

3.1 Stitcher 数据增强

Stitcher 数据增强由 Chen 等^[19]提出。由于数据集中小目标物体分布不均匀,训练中网络对小目标学习不充分,导致训练后的模型对小目标的检测精度差。因此,Stitcher 把每个训练批次内的 4 张图都缩小到同样大小,之后拼成一张与原始图像相同大小的图用作训练,以此均衡不同尺度物体在训练过程中的分布。

遥感数据集中不仅包含桥梁、十字路口等大目标,也有车辆、储罐等小目标,而 HRRSD 遥感数据集中每幅图像中的大小目标分布极不均匀,一幅遥感图像中通常只包含一类目标,在训练神经网络时,极易出现在同一批次甚至几个批次内网络未学习到某一类目标的情况,导致网络对这一类目标的学习不充分,从而影响网络的检测效果。为了解决这一问题,本文参考了 Stitcher 数据增强方法(见图 3),将每个训练批次内的 4 张图拼接成一张图进行训练,在训练过程中监控各类目标损失的反馈情况,如果这一批次内某一类目标的损失下降值占总损失下降值的比例较低,则在该批次内加入含有该类目标的图片进行训练,以此来解决目标分布不平衡的问题,从而提高网络的检测精度。



图 3 Stitcher 图像拼接
Fig. 3 Stitcher image stitching

3.2 VOVDarkNet-53 基础网络

由于 DenseNet 任何两层之间都有直接连接,且每个层都会聚合前面层的特征,造成了特征冗余,并且从模型权重的 L1 范数会发现,中间层对最后的分类层贡献较少,这是因为后面的特征已经学习到了这些中间层的核心信息。这种信息冗余是可以优化的方向,因此 Lee 等^[20]提出 CenterMask 网络,CenterMask 网络的基础网络为 VOVNet,是由多个 OSA (One-Shot Aggregation) 模块组成,如图 4 所示。OSA 模块中每个卷积层以改进的密集连接方式连接,增强特征的复用,使网络充分利用不同层级的特征图,更多地利用低层特征图,从而提升网络的检测效果,尤其是对小目标的检测效果。因此,本文采用 CenterMask 网络中的 VOVNet 密集连接方式改进 DarkNet-53 基础网络。

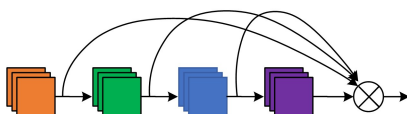


图 4 OSA 模块
Fig. 4 OSA module

参考 VOVNet 的密集连接方式,本文提出 VOVDarkNet-53 基础网络。首先,为了更加充分利用浅层小目标信息,将 DarkNet-53 基础网络每一次下采样之后直到下一次下采样之前的所有残差模块作为一个 DOSA (DarkNet-one-shot-aggregation) 模块,如图 5 所示,在 DOSA 模块中,每一个残差模块产生两种连接,一种是通过卷积层和下一个残差模块连接,产生感受野更大的特征图,另一种是与最后的输出层相连,以聚合足够好的特征。其次,为了减少网络计算量,提升网络检测速度,将网络第 4 次下采样后的 8 个残差模块减少为 4 个残差模块。最后,基于 DOSA 模块构建 VOVDarkNet-53 基础网络,如图 6 所示,图中一个蓝色立方体代表一个残差模块。将 DarkNet-53 基础网络第 2 次下采样后的 2 个残差模块的输出特征图和最后的输出层相连,将 DarkNet-53 基础网络 3 次、4 次和 5 次下采样后的 4 个残差模块的输出特征图和每次下采样最后的输出层相连,在每次下采样之后采用改进的密集连接方式使网络能利用更多的浅层小目标信息,以此来增强 DarkNet-53 基础网络对特征的复用,更多地利用低层特征图,从而提升网络的检测效果。

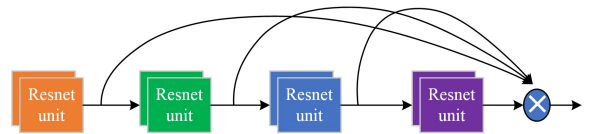


图 5 DOSA 模块
Fig. 5 DOSA model

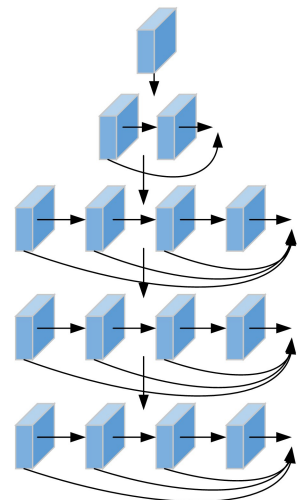


图 6 VOVDarkNet-53 基础网络(电子版为彩色)
Fig. 6 VOVDarkNet-53 backbone

3.3 分布式排序损失

为了保证检测速度,单阶段目标检测没有区域建议(region proposal)阶段,直接从所生成的大量候选对象中识别出前景对象。因此,背景候选框的数量远远超过前景候选框的数量,尤其是对于 YOLOv3 网络来说,其每个预测输出的背景候选框有几百甚至上千个,然而真正的正确目标却只有几个,这会导致网络预测的正负样本不平衡,影响网络的检测精度,而分布式排序损失(Distributional Ranking Loss, DR Loss)^[21]可以减轻类不平衡问题的影响。因此,本文参考 DR

Loss,首先将分类问题转化为排序问题,从而优化正负样本的排序。每一对样本分别由一个前景候选和一个背景候选组成,所以正负样本间是平衡的。同时,考虑到背景候选框和前景候选框的不平衡问题,本文引入分布式排序损失,将前景候选框的约束分布排序在背景候选框的约束分布之上。通过对选项重新加权,得到对应于最坏情况的损失分布,使损失可以集中在前景和背景分布之间的决策边界上。此外,将分布的期望值进行排序以取代原来的样本排序,从而减少排序中的对数,提高效率。同时,每次对样本排序时考虑图像中的样本对(包含一个正样本和一个负样本),然后根据前景候选框和背景候选框的置信度作决策,最后对样本置信度分布进行优化,以减轻类不平衡问题。损失计算公式如式(1)所示:

$$\min \sum_i^N \tau l_{DR} + l_{YL} \quad (1)$$

其中, l_{DR} 是通过分布式排序损失函数计算出的损失; l_{YL} 是YOLOv3中的原始回归损失; τ 是用于平衡分类和回归的权重,设置为4。

4 实验结果及分析

4.1 数据集和实验平台

本文采用HRRSD(TGRS-HRRSD-Dataset)数据集,数据集提供了21761幅尺寸大小不一的遥感图像,共计4万个目标对象。数据集中有13类对象,分别为飞机、棒球场、篮球场、桥梁、十字路口、地面运动场、海港、停车场、船、储罐、T 字路口、网球场、车辆。按照原数据集设定,将其中5401幅图像作为训练集,10943幅图像作为测试集。

实验平台如下:操作系统为ubuntu16.04,深度学习框架为pytorch1.4,CPU为AMD-R2700,内存为32GB,训练GPU为NVIDIA GeForce 1080ti * 2,显存为22GB,测试GPU为NVIDIA RTX 2070,显存为8GB。

4.2 损失函数对比

本节对原网络和使用DRLoss作为分类损失的改进后网络的损失曲线进行对比,结果如图7所示。可以看出,两个网络的分类损失均降到一定水平且趋于平衡,表明网络在分类上皆可达到一定效果。但在使用DRLoss作为分类损失后,网络的分类损失下降得更快,同时损失下降到更低位置才趋于平衡,表明改进后的网络能更好地解决正负样本不平衡的问题,降低分类损失。

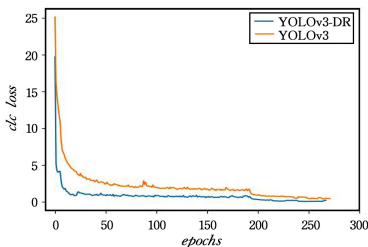


图7 改进前后网络分类损失的对比

Fig. 7 Comparison of classification loss of network before and after improvement

4.3 检测精度及P-R曲线对比

本文以平均检测精度均值(Mean Average Precisions,

mAP)、各类检测精度(Average Precisions, AP)、平均单张图片处理时间和准确率-召回率(P-R)曲线等作为评价指标。首先测试训练后的两个模型,计算其精确度(Precision, P)和召回率(Recall, R),计算精确度和召回率的公式分别为:

$$P = \frac{TP}{TP + FP} \quad (2)$$

$$R = \frac{TP}{TP + FN} \quad (3)$$

其中, TP 代表将正例正确预测为正类数, FP 代表将负类错误预测为正类数, FN 代表将正类错误预测为负类数。

设定固定阈值,按照置信度分数对检测器预测结果降序排列,逐个将样本作为正例输出进行预测,计算其 P, R 值并绘制 $P-R$ 曲线。如图8所示,改进后的YOLOv3网络的准确率和召回率均比YOLOv3网络有一定程度的提升(由于单个改进点的 $P-R$ 曲线变化不明显,这里只对融合了3个改进点的YOLOv3-D进行对比)。

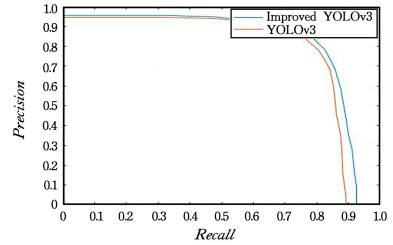


图8 改进前后网络P-R曲线的对比

Fig. 8 Comparison of P-R curve of network before and after improvement

为了进一步证明3种改进方法的效果,我们整理出5个不同的模型进行对比实验,分别是YOLOv3-A(只将YOLOv3的分类损失改进为DR Loss)、YOLOv3-B(只对YOLOv3网络进行Stitcher数据增强)、YOLOv3-C(只改进YOLOv3的基础网络为VOVDarkNet-53)、YOLOv3-D(融合了以上3种改进的模型)以及YOLOv3网络模型。同时,为了检验模型在低计算能力的计算机上的处理速度,在GPU为Nvidia RTX2070的台式计算机上用测试集中的10943幅图像对5个模型进行测试,输出5个模型的平均检测精度均值(mAP)、各类检测精度(AP)、平均每张图片处理时间,结果如表1所列。由于数据集中共包含13类目标,因此根据目标大小分布,将飞机、车辆、储罐合并为小目标,将棒球场、篮球场、网球场等合并为中目标,将十字路口、T字路口和桥梁合并为大目标,分别计算出小、中、大3种目标各类检测精度。

表1 改进前后网络检测精度和速度对比

Table 1 Comparison of network detection accuracy and speed before and after improvement

| Model | mAP | AP (Small) | AP (Medium) | AP (Large) | Speed/ms |
|----------|-------|------------|-------------|------------|----------|
| YOLOv3 | 0.787 | 0.685 | 0.822 | 0.808 | 31.6 |
| YOLOv3-A | 0.791 | 0.688 | 0.822 | 0.811 | 31.5 |
| YOLOv3-B | 0.799 | 0.696 | 0.825 | 0.809 | 32.0 |
| YOLOv3-C | 0.819 | 0.744 | 0.837 | 0.787 | 32.5 |
| YOLOv3-D | 0.828 | 0.757 | 0.843 | 0.798 | 32.8 |

从实验结果可以看出, YOLOv3-A 的效果最不明显, 但其可以使分类损失下降得更快, 使网络的收敛速度更快; YOLOv3-B 的总体效果和 YOLOv3-A 较接近, 因为其本质上都是使 YOLOv3 网络的损失下降得更快, 但由于 YOLOv3-B 减轻了小目标样本分布不均衡的问题, 其对小目标的检测精度有所提升; YOLOv3-C 对小目标的检测结果有明显提升, 证明了 VOVDdarkNet-53 基础网络的有效性; 融合了改进方法的 YOLOv3-D 表现最好, 改进后的网络对于车辆、储罐等小目标的检测精度提升了 7.2%, 同时对篮球场、足球场等中型目标的检测精度也提升了 2.1%, 但对大型目标的检测精度有所下降, 但不明显, 总体的平均检测精度均值上升了 4.1%。同时, 由于删除了网络中多余的层数, 网络平均每张图片的处理时间变化不大, 可以达到 32.8 ms, 具有较好的实时性。

4.4 与其他目标检测算法的对比

为了进一步验证改进方法的有效性, 将改进后的 YOLOv3 算法 (YOLOv3-D) 与 FCOS, RetinaNet, CornerNet 等目标检测算法在数据集上作对比实验, 实验结果如表 2 所列。

表 2 改进后网络与其他网络的对比

Table 2 Comparison between improved network and other networks

| Model | mAP | Speed/ms |
|----------------|-------|----------|
| YOLOv3 | 0.787 | 31.6 |
| RetinaNet-101 | 0.807 | 96.2 |
| FCOS-ResNet101 | 0.835 | 85.5 |
| CornerNet | 0.813 | 33.2 |
| YOLOv3-D | 0.828 | 32.8 |

从实验结果可以看出, 仅有 FCOS 在精度上超过了 YOLOv3-D, 但 FCOS 的速度较慢; CornerNet 在速度与 YOLOv3-D 很接近, 但其 mAP 比 YOLOv3-D 低 1.5%。综上所述, 改进后的 YOLOv3 能更好地检测遥感图像中的目标。

4.5 真实检测结果

图 9 是 YOLOv3 和改进 YOLOv3 网络模型对遥感图像的检测结果。我们随机从数据集中选择了 4 张包含小目标的图像, 用来测试改进后网络对小目标检测的性能。图 9 中的 (a), (c), (e), (g) 表示用原 YOLOv3 网络生成的检测结果, 图 9 中的 (b), (d), (f), (h) 表示用改进后的 YOLOv3 网络生成的检测结果。图 9(a) 中 YOLOv3 漏检了位于图像右上方边缘的一辆汽车, 而图 9(b) 中改进的 YOLOv3 检测出了这一车辆目标; 图 9(c) 中有许多储罐未被检测出来, 而图 9(d) 中改进的 YOLOv3 检测出了几乎全部的储罐; 图 9(e) 中 YOLOv3 将路边杂物和建筑物错检为车辆目标, 而图 9(f) 中, 改进的 YOLOv3 可以检测出所有正确的车辆目标, 并且没有将其他目标误检为车辆目标; 同样地, 图 9(g) 中 YOLOv3 将机场杂物检测为飞机, 而图 9(h) 中改进的 YOLOv3 可以正确地检测出所有目标。这些对比结果表明改进后的 YOLOv3 网络能更好地检测出图像中的小目标。



(a) YOLOv3

(b) Improved YOLOv3



(c) YOLOv3

(d) Improved YOLOv3



(e) YOLOv3

(f) Improved YOLOv3



(g) YOLOv3

(h) Improved YOLOv3

图 9 YOLOv3 和改进的 YOLOv3 网络检测效果的对比

Fig. 9 Comparison of detection effect between improved YOLOv3 and YOLOv3 network

结束语 本文提出一种改进的 YOLOv3 目标检测算法, 通过修改其基础网络结构和分类损失, 在检测速度不下降的情况下, 提升网络对遥感小目标的检测精度, 从而提高网络的平均检测精度均值、召回率和准确率。在网络检测精确度较高的情况下提升检测速度, 将是下一步的研究方向。

参考文献

- [1] HUANG G Q. Research on Satellite Image Target Detection and recognition based on Deep Learning[D]. Hangzhou: Zhejiang University, 2019.
- [2] PANG J, LI C, SHI J, et al. R²-CNN: Fast Tiny Object Detec-

- tion in Large-Scale Remote Sensing Images[J]. IEEE Transactions on Geoscience and Remote Sensing, 2019, 57(8): 5512-5524.
- [3] SHERMEYER J, VAN ET TEN A. The Effects of Super-Resolution on Object Detection Performance in Satellite Imagery [C]//IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops. 2019:1432-1441.
- [4] RUSSAKOVSKY O, DENG J, SU H, et al. ImageNet Large Scale Visual Recognition Challenge[J]. International Journal of Computer Vision, 2015, 115(3): 211-252.
- [5] EVERINGHAM M, GOOL L V, WILLIAMS C K I, et al. The Pascal Visual Object Classes (VOC) Challenge[J]. International Journal of Computer Vision, 2010, 88(2): 303-338.
- [6] LIN T, MAIRE M, BELONGIE S, et al. Microsoft COCO: Common Objects in Context [C]// European Conference on Computer Vision. 2014:740-755.
- [7] LIU Y, WANG Y, WANG S, et al. CBNNet: A Novel Composite Backbone Network Architecture for Object Detection [C]// National Conference on Artificial Intelligence. 2020.
- [8] TAN M, PANG R, LE Q V, et al. EfficientDet: Scalable and Efficient Object Detection [EB/OL]. (2019-11-09) [2020-08-01]. <https://arxiv.org/abs/1911.09070>. pdf.
- [9] SONG G, LIU Y, WANG X, et al. Revisiting the Sibling Head in Object Detector [EB/OL]. (2020-03-07) [2020-08-01]. <https://arxiv.org/abs/2003.07540>. pdf.
- [10] GHIASI G, TAN M, JIN P, et al. SpineNet: Learning Scale-Permuted Backbone for Recognition and Localization [EB/OL]. (2019-12-05) [2020-08-01]. <https://arxiv.org/abs/1912.05027>.
- [11] LIN T, GOYAL P, GIRSHICK R, et al. Focal Loss for Dense Object Detection [C]// International Conference on Computer Vision. 2017:2999-3007.
- [12] REDMON J, FARHADI A. YOLOv3: An Incremental Improvement [EB/OL]. (2018-04-08) [2020-07-01]. <https://arxiv.org/abs/1804.02767>.
- [13] TIAN Z, SHEN C, CHEN H, et al. FCOS: Fully Convolutional One-Stage Object Detection [C]// International Conference on Computer Vision. 2019:9627-9636.
- [14] LAW H, DENG J. CornerNet: Detecting Objects as Paired Key-points [C]// European Conference on Computer Vision. 2018: 765-781.
- [15] ZHU W T, XIE B R, WANG M, et al. A review of aircraft target detection in optical remote sensing images [J]. Computer Science, 2020, 47(S2): 165-171.
- [16] BELL S, ZITNICK C L, BALA K, et al. Inside-Outside Net: Detecting Objects in Context with Skip Pooling and Recurrent Neural Networks [C]// Computer Vision and Pattern Recognition. 2016:2874-2883.
- [17] ZHANG M, LI J, DING R L, et al. Research on remote sensing image target detection technology based on improved YOLO-V2 algorithm [J]. Computer Science, 2020, 547(S1): 176-180.
- [18] LONG Y, GONG Y, XIAO Z, et al. Accurate Object Localization in Remote Sensing Images Based on Convolutional Neural Networks [J]. IEEE Transactions on Geoscience & Remote Sensing, 2017, 55(5): 2486-2498.
- [19] CHEN Y, ZHANG P, LI Z, et al. Stitcher: Feedback-driven Data Provider for Object Detection [EB/OL]. (2020-04-12) [2020-08-01]. <https://arxiv.org/abs/2004.12432>.
- [20] LEE Y, PARK J. CenterMask: Real-Time Anchor-Free Instance Segmentation [EB/OL]. (2019-11-06) [2020-08-01]. <https://arxiv.org/abs/1911.06667>. pdf.
- [21] QIAN Q, CHEN L, LI H, et al. DR Loss: Improving Object Detection by Distributional Ranking [EB/OL]. (2019-07-10) [2020-08-01]. <https://arxiv.org/abs/1907.10156>.



YUAN Lei, born in 1995, postgraduate. His main research interests include deep learning and object detection.



LIU Zi-yan, born in 1974, associate professor, postgraduate supervisor, is a member of China Computer Federation. Her main research interests include wireless communication system, mobile robot and big data mining analysis.