

# 基于 Faster-RCNN 的遥感图像飞机检测算法

张中宝,王洪元,张 继,杨 薇

(常州大学信息科学与工程学院,江苏 常州 213164)

[摘要] CCCV2017 发布遥感图像飞机数据集,用于评测飞机检测算法. 针对该遥感图像数据集中的飞机朝向不确定、图像覆盖范围广、图像背景复杂度高,导致飞机检测难度大、检测算法准确率和算法泛化能力低等问题,提出了基于 Faster-RCNN 的飞机检测改进算法. 首先,通过对图像采用翻转以及角度旋转等方式对数据集进行合理的扩增;然后,在扩增后的数据集上,使用深度残差网络对图像进行特征提取,针对数据集中飞机目标的长宽比特点优化区域建议网络;同时,为了防止训练集中正负样本不平衡,采用在线困难样本挖掘方法对数据进行训练. 在 CCCV2017 数据集上评估表明,改进后的 Faster-RCNN 算法极大提高了初始的 Faster-RCNN 算法性能,在测试集上 mAP 达到了 89.93%. 在 NWPUVHR-10、NWPU-RESISC45、UCAS-AOD 遥感图像飞机数据集测试表明,该改进模型同样具有良好的性能,从而验证了该模型具有良好的鲁棒性和泛化能力.

[关键词] 遥感图像,飞机检测,Faster-RCNN,残差网络,区域建议网络,在线困难样本挖掘

[中图分类号] TP391 [文献标志码] A [文章编号] 1001-4616(2018)04-0079-08

## Airplane Detection in Remote Sensing Image Based on Faster-RCNN Algorithm

Zhang Zhongbao, Wang Hongyuan, Zhang Ji, Yang Wei

(School of Information Science and Engineering, Changzhou University, Changzhou 213164, China)

**Abstract:** CCCV2017 releases remote sensing image airplane dataset for evaluating airplane detection algorithm. Due to the uncertainty of the orientation of airplanes in remote sensing images and the images with a wide coverage and high background complexity, airplane detection is difficult, the precision and the generalization ability of the model are low. This paper proposes an improved airplane detection algorithm based on Faster-RCNN. First of all, the dataset is reasonably augmented by flipping and rotating the images; then, on the augmented dataset, the residual network is used to extract features from the images and the region proposal network is optimized based on the characteristics of the aspect ratio of airplanes; at the same time, in order to prevent imbalance between positive and negative samples in the training set, the online hard example mining method is used to train the data. The evaluation on the CCCV2017 dataset shows that the improved Faster-RCNN algorithm greatly improved the performance of the initial Faster-RCNN algorithm. In the test set, the mAP (mean Average Precision, mAP) has reached 89.93%. Tests on NWPU VHR-10, NWPU VHR-45, and UCAS-AOD remote sensing image datasets show that the improved model also has good performance, which verifies that the model has good robustness and generalization ability.

**Key words:** remote sensing images, airplane detection, Faster-RCNN, residual network, region proposal network, online hard example mining

遥感图像技术自发展以来,无论在民用领域还是军事领域都扮演着极其重要的角色,有着极高的应用价值. 在民用领域,遥感图像被广泛用于国土规划、农业生产、天气预报、交通导航,为国民生产活动提供重要信息指导. 在军事领域,由于遥感技术可以在不直接接触物体的情况下获取目标信息,且观察范围广、观察周期短,因此被广泛用于战略侦察、地形勘探、海洋监测等方面. 由于遥感图像中飞机目标对于国

收稿日期:2018-08-15.

基金项目:国家自然科学基金(61572085).

通讯联系人:王洪元,博士,教授,CCF 高级会员,研究方向:计算机视觉. E-mail:hywang@cczu.edu.cn

防建设以及国民经济活动有着极其重要的意义,因此本文以飞机为代表,探讨快速准确的目标检测算法。

近年来,有众多的理论研究集中到目标检测领域<sup>[1]</sup>,涌现出不少优秀的算法和解决方案。根据提取图像特征方法的不同,可以将其分为基于传统图像处理与机器学习算法相结合的目标检测算法和基于深度卷积神经网络的目标检测算法两种类型。

传统的图像处理与机器学习算法,主要对图像梯度<sup>[2-3]</sup>、小波<sup>[4]</sup>、纹理、形状等特征进行提取,然后将提取的特征输入到支持向量机(Support Vector Machine, SVM)<sup>[5]</sup>、贝叶斯分类器<sup>[6]</sup>中学习分类规则,最后利用训练完成的分类器实现对目标的检测。然而,这些特征都是人工设计的特征,特征提取过程繁琐,依赖大量的先验知识,另外,提取的特征属于浅层特征,特征表达能力弱,对于背景复杂多变的遥感图像目标检测效果欠佳。

随着计算机性能的提升,尤其是借助 GPU(Graphics Processing Unit, GPU)的高性能计算能力,基于卷积神经网络(Convolution Neural Network, CNN)算法<sup>[7]</sup>在通用目标检测领域的表现远远超过传统算法。由于 CNN 可以对图像中的特征进行自动学习并提取具有判别力的特征,因此模型具有较强的泛化能力和鲁棒性,受到了众多研究者的青睐。2013 年, Ross Girshick 率先提出 Region-CNN(RCNN)<sup>[8]</sup>算法,通过选择性搜索算法<sup>[9]</sup>(Selective Search, SS)提取图像中目标候选区域,然后使用 CNN 对候选区域进行特征提取,并使用 SVM 分类器对提取的特征进行分类,最后根据边界框回归算法重新定位目标边界框。然而,由于该方法需要对每一个候选区域计算卷积特征,因此算法时间复杂度高,空间复杂度大。针对以上这些问题,2015 年, Ross Girshick 提出 Fast-RCNN 模型<sup>[10]</sup>,通过将候选区域映射到最后一层卷积特征图上,然后在特征图上直接提取相应区域的卷积特征,避免重叠候选区域的卷积计算,以此降低时间和空间开销。为了将候选区域生成、深度特征提取、以及目标识别和检测融入到同一个神经网络模型中,2015 年, Ross Girshick 团队再次提出 Faster-RCNN<sup>[11]</sup>算法,将 SS 算法替换成区域建议网络(Region Proposal Networks, RPN)来完成目标候选区域的提取,实现了端到端(end-to-end)的训练和检测。同年, Joseph Redmond 等人受到以上论文的启发,提出 YOLO(You Only Look Once)<sup>[12]</sup>算法,该算法省去了候选区域提取的过程,直接将输入图片分成  $s \times s$  个网格,然后将图片输入到 CNN,预测每一个网格是否存在目标、目标的边界框以及类别,最后通过非极大抑制算法筛选出最好的边界框。

上述方法在通用目标检测 Pascal VOC、MS COCO 等数据集上取得不错的效果,有力地推动了目标检测技术的发展。但是针对背景复杂多变的遥感图像飞机检测任务效果欠佳,主要原因在于通用目标检测数据集中目标长宽比分布和遥感图像中飞机目标长宽比分布有所不同,因此本文针对遥感图像飞机数据集的特点,提出了基于 Faster-RCNN 的目标检测改进算法,主要工作包含以下 3 个方面:

(1)数据扩增阶段,在已经标注完成的飞机数据集上,采用左右、上下翻转图像,以及角度旋转等方式对数据集进行扩充,以此获得足够多的训练数据。

(2)数据训练阶段,本文统计训练数据集中目标长宽比分布,根据此分布,增加 RPN 网络中生成 anchors 所对应的尺度以及长宽比,最终最后一层卷积特征图上每个点可以产生大小和比例不同的 16 个 anchors,同时为了提取更具判别力的特征,本文采用残差网络(Residual Network, ResNet)<sup>[13]</sup>提取图像特征。另外,为了保持训练数据正负样本比例平衡,本文结合在线困难样本挖掘方法<sup>[14]</sup>对数据进行训练。

(3)模型评估阶段,本文从 mAP 以及假阳率、漏检率 3 个指标评估本文算法,并对比主流算法 R-FCN(Object Detection via Region based Fully Convolutional Networks, R-FCN)<sup>[15]</sup>,同时在不同的飞机数据集上评估模型的泛化能力,说明改进后的 Faster-RCNN 算法对于提升遥感图像飞机检测效果的有效性。

## 1 遥感图像数据集建立

基于 Faster-RCNN 遥感图像目标检测训练需要大量数据,更多的训练数据不仅可以提高模型的准确率,同时还可以提高模型的泛化性能,增强模型的鲁棒性。因此,遥感图像数据集的建立是训练 Faster-RCNN 算法非常重要的一环。

### 1.1 数据标注

CCCV2017 遥感图像目标识别技术挑战赛发布的飞机数据集总共 1 200 张图像,其中训练集 1 000 张,测试集 200 张,图像主要以低分辨率彩色融合图像为主。由于基于 Faster-RCNN 遥感图像目标检测是

一种有监督的深度学习算法,因此需要手动标注 CCCV2017 遥感图像数据,标注方式参照 PASCAL VOC 2007 数据集的格式对 1 200 张图片中飞机的位置、大小、属性进行标定,标定的结果保存为格式 xml 文件. 根据对 xml 文件中标定的目标进行统计,总共标记飞机目标 3 559 个.

## 1.2 数据增强

为了增加训练集数据量,提高模型的泛化能力,本文主要采用翻转以及旋转两种方式增强数据. 对于飞机而言,由于实际停放位置不同,在遥感图像中呈现的姿态以及角度也各不相同. CCCV2017 发布的飞机数据集中,仅包含 1200 张图片,因此不可能包含飞机朝向的所有角度. 为了呈现更加自然的情况,就需要对图像进行角度的旋转以及进行镜像操作,增强模型对飞机朝向的敏感度. 在 1.1 节标注数据中,目标包围盒与图像的坐标轴平行,因此在实际处理中,为了方便计算,本文仅对图像进行 180 度的旋转,以及左右翻转、上下翻转. 经过本节采用的两种数据增强方式,最终将 CCCV2017 遥感图像飞机数据集扩充到 4200 张图片,并将 4000 张图片用于训练,扩充前的测试集 200 张图片仍作为扩充后的测试集.

## 2 基于 Faster-RCNN 飞机检测算法

本文采用 Faster-RCNN 算法实现对遥感图像中飞机目标的检测,具体结构如图 1 所示. 首先对输入任意尺寸大小的图片(飞机)进行卷积特征提取,然后通过区域建议网络中滑动窗口机制得到区域建议以及区域得分,对于区域得分采用非极大抑制算法输出前  $N$  个得分的区域建议,然后将其输入到 RoI 池化层,通过该层得到区域建议特征,最后将其输入到全连接层后输出区域的分类得分以及区域位置.

为了训练区域建议网络, Faster-RCNN 算法设计了如式(1)所示的多任务损失函数. 其中,  $i$  是一个小批量数据中选取的候选框(anchor)的索引,  $p_i$  表示为预测候选框  $i$  为目标的概率,  $p_i^*$  表示为选取的候选框中对应的样本所属标签. 当候选框中的目标为正样本时,  $p_i^*$  取值为 1, 否则  $p_i^*$  为 0.  $t_i$  表示预测的包围盒(bounding box)的 4 个参数化坐标向量,  $t_i^*$  表示真实区域(ground truth)包围盒的坐标向量.  $L_{cls}$  表示为二值分类器(目标/非目标)的分类损失,具体定义如式(2)所示.

$$L(\{p_i\}, \{t_i\}) = \frac{1}{N_{cls}} \sum_i L_{cls}(p_i, p_i^*) + \lambda \frac{1}{N_{reg}} \sum_i p_i^* L_{reg}(t_i, t_i^*), \quad (1)$$

$$L_{cls}(p_i, p_i^*) = -\log[p_i p_i^* + (1-p_i)(1-p_i^*)]. \quad (2)$$

$L_{reg}$  表示为回归损失,定义为:

$$L_{reg}(t_i, t_i^*) = \sum_{i \in \{x, y, w, h\}} smooth_{L_1}(t_i, t_i^*), \quad (3)$$

式中,  $smooth_{L_1}(x)$  为:

$$smooth_{L_1}(x) = \begin{cases} 0.5x^2, & |x| < 1, \\ |x| - 0.5, & |x| \geq 1. \end{cases} \quad (4)$$

对于回归,使用如下 4 个坐标:

$$\begin{aligned} t_x &= (x - x_a) / w_a, & t_y &= (y - y_a) / h_a, \\ t_w &= \log(w / w_a), & t_h &= \log(h / h_a), \\ t_x^* &= (x^* - x_a) / w_a, & t_y^* &= (y^* - y_a) / h_a, \\ t_w^* &= \log(w^* / w_a), & t_h^* &= \log(h^* / h_a), \end{aligned} \quad (5)$$

式中,  $[x, y, w, h]$  表示预测框坐标,  $[x_a, y_a, w_a, h_a]$  表示候选框(anchor)的坐标,  $[x^*, y^*, w^*, h^*]$  表示真实区域(ground truth)的坐标. 另外,  $\lambda$  为权重参数,  $N_{cls}$  和  $N_{reg}$  为归一化参数,用于标准化两个损失.

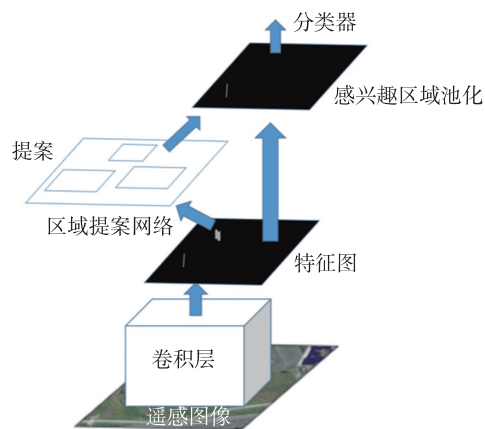


图 1 Faster-RCNN 的网络结构

Fig. 1 Faster-RCNN network structure

2.1 Faster-RCNN 模型的改进

2.1.1 残差网络提取特征

Faster-RCNN 目标检测算法中,卷积操作分别使用经典的 ZF<sup>[16]</sup> 和 VGG16<sup>[17]</sup> 网络对图像进行特征提取. 为了提取遥感图像样本中更加本质的特征,本文采用更深层次的网络. 然而深层次的神经网络训练十分困难,随着网络层数的加深,随之而来的一个问题就是梯度弥散,但是使用残差网络可以在增加网络层数的同时减轻深层网络训练的负担. 它的核心在于残差学习模块,如图 2 所示,假设该部分神经网络的输入为  $x$ ,期望输出是  $H(x)$ ,可以定义另外一个残差映射  $F(x)$  为  $H(x)-x$ ,则原始的函数映射  $H(x)$  可以表示为  $F(x)+x$ . He<sup>[13]</sup> 通过实验证明,优化残差映射  $F(x)$  比优化原始映射  $H(x)$  容易得多. 由于学习目标改变,不再是学习一个完整的输出  $H(x)$ ,只是输出和输入的差别  $H(x)-x$ ,即残差. 这种前馈的短连接形式既没有引进新的参数,也没有增加计算复杂度,因而简化了学习目标的难度,从而避免随着网络深度的不断增加,分类精度逐步达到饱和问题.

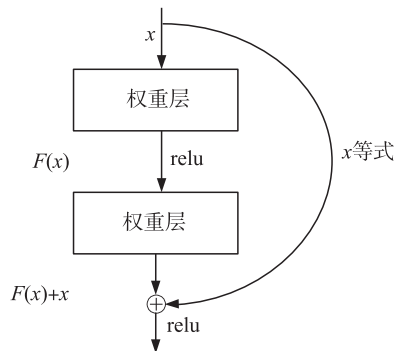


图 2 残差网络学习模块

Fig. 2 The learning module of ResNet

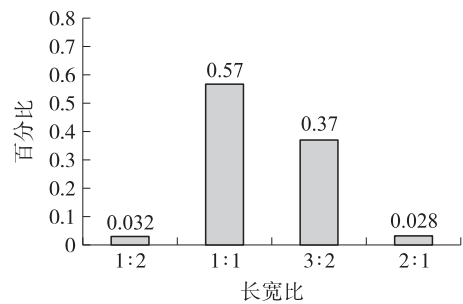


图 3 飞机长宽比条形图

Fig. 3 The aspect ratio bar chart of the airplanes

2.1.2 优化区域建议网络

Faster-RCNN 目标检测算法中,RPN 连接在卷积提取特征之后,以最后一层卷积产生的特征图上的每个点为中心滑动窗口,然后生成区域建议框. Faster-RCNN 预设了 9 种 anchors,分别对应 3 种尺度( $128^2$ 、 $256^2$ 、 $512^2$ )和 3 种长宽比(1:1、1:2、2:1),这种设置方式可以覆盖 Pascal VOC、MS COCO 数据集中大多数目标,但是针对 CCCV2017 飞机数据集,由于数据来源于遥感卫星,图像覆盖范围广,数据集中包含部分小目标,同时由于飞机机身呈瘦长型特点,Faster-RCNN 算法中预设的 3 种尺度和 3 种长宽比对于 CCCV2017 飞机检测任务并不合理. 本文从生成 anchor 的尺度和长宽比出发,统计遥感图像数据集中飞机长宽比特点,根据统计得到的飞机长宽比条形图(如图 3 所示),本文增加一组尺度和一组长宽比,此时 anchor 对应尺度为( $64^2$ 、 $128^2$ 、 $256^2$ 、 $512^2$ ),长宽比为(1:2,1:1,3:2,2:1).

2.1.3 在线困难样本挖掘

Faster-RCNN 目标检测算法中,IOU(Intersection Over Union)表示预测的目标包围盒与真实区域的重叠度. 如果 IOU 超过 0.7,则被认为是正样本,如果 IOU 大于 0.1 小于 0.3,则被认为是负样本. 为了保持正负样本比例平衡,初始 Faster-RCNN 算法从所有负样本中随机选取部分负样本,使得正负样本比例保持在 1:3 左右. 但是,遥感图像中目标的区域相比背景区域要小得多,数据集中样本经过标注后,负样本数量通常远大于正样本数量,正负样本分布的不均衡可能导致训练不稳定,模型最终可能倾向将所有样本分为负样本,影响模型的性能. 为了平衡模型训练时正负样本比例,本文引入在线困难样本挖掘方法,对于给定输入的图像,经过残差网络提取图像特征图后,RPN 网络在最后一层卷积特征图提取感兴趣区域(Region of Interest, RoI),此时对提取到的感兴趣区域边框计算损失,根据计算的损失值按从高到低进行排序,选取前  $k$  个 RoI 作为正负样本进行训练,从而保持正负样本分布平衡.

3 实验与分析

3.1 数据集与评价指标

本文采用 4 种公开数据集进行训练、测试以及评估,如图 4 所示,分别展示了不同数据集的部分样



本. 其中,CCCV2017 数据集在 1.1、1.2 节中进行了详细介绍,该数据集主要用于模型的训练以及测试. 另外,本文分别选取了 NWPU VHR-10<sup>[18-20]</sup>、NWPU-RESISC45<sup>[21]</sup> 以及 UCAS-AOD<sup>[22]</sup> 这 3 种数据集用于评估模型的泛化能力. NWPU VHR-10 是一个包含 10 个类别的目标检测数据集,数据主要从 Google Earth 和 Vaihingen 中裁剪得到,本文选取其中 87 张飞机图片. NWPU-RESISC45 是由西北工业大学创建的、涵盖 45 个类别、其中每个类别 700 张图片的遥感图像场景分类数据集,本文选取其中 700 张飞机图片. UCAS-AOD 为中科院大学使用 Google Earth 软件在全球部分区域截取图像创建的数据集,本文选取其中 1000 张飞机图像,图像中总共包含 7 482 个样本.



图 4 不同数据集样本

Fig. 4 Samples in different datasets

对于遥感图像飞机目标检测,本文采用  $mAP$  (mean Average Precision)、假阳率 (False Positive Rate, FPR)、漏检率 (Miss Rate, MR) 3 个评价指标用于评估模型的性能.  $mAP$  根据精确率 (Precision) 以及召回率 (Recall) 计算得到. 精确率和召回率的计算如式 (6) 和 (7) 所示. 其中,  $t_p$  表示正确检测的飞机框个数,  $f_p$  表示错误检测的个数,  $f_n$  表示未被检测出来的飞机框的个数,  $t_n$  表示正确检测不是飞机框的个数.

$$Precision = t_p / (t_p + f_p), \quad (6)$$

$$Recall = t_p / (t_p + f_n). \quad (7)$$

根据式 (6) 和 (7) 计算得到的精确率以及召回率,以召回率为横坐标,精确率为纵坐标画出  $P-R$  曲线,此时  $AveP$  (average precision, AveP) 即为精确度和召回率曲线下的面积,数学意义上表示为  $P-R$  曲线的积分,如式 (8) 所示. 为了求解式 (8) 中的积分,将召回率划分为  $n$  块,  $[0, 1/n, \dots, (n-1)/n, 1]$ , 则  $AveP$  此时可以表示为式 (9). 最后得到所有类的平均精度,如式 (10) 所示,其中  $N$  为测试集中目标种类的个数. 由于本文中只检测飞机一类目标,因此  $N=1$ . 检测的假阳率和漏检率计算方式如式 (11) 和 (12) 所示.

$$AveP = \int_0^1 p(r) dr, \quad (8)$$

$$AveP = \frac{1}{n} \sum_{i=1}^n \max_{r \in [(i-1)/n, i/n]} p(r), \quad (9)$$

$$mAP = \frac{\sum_{n=1}^N AveP(n)}{N},$$

(10)

$$FPR = f_p / (f_p + t_n),$$

(11)

$$MR = f_n / (t_p + f_n).$$

(12)

3.2 实验环境及参数设置

本文实验使用的仿真平台为 Ubuntu16.04, 显卡采用 NVIDIA 公司生产的 TiTan XP 12G 独立显卡, 另外安装了 GPU 开发包 CUDA8.0 以及深度学习加速库 CUDNN5.1, 深度学习框架为 Caffe, 代码运行环境为 Python 2.7.12. 训练阶段, 采用在 ImageNet 数据集上预训练得到的模型进行权重初始化, 超参数设置: 学习率为 0.001; momentum 为 0.9; decay 为 0.000 5; batchsize 为 64, iterations 为 80 000. 模型训练得到的 loss/iterations 分析如图 5 所示, 其中 loss-iterations 表示模型损失随训练次数产生的变化, 网络的平均训练速度为 0.497 s/iter.

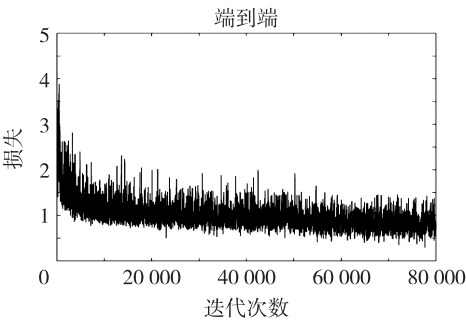


图 5 训练损失变化  
Fig. 5 The changes of training loss

3.3 对比实验

3.3.1 数据增强实验对比

本文实验数据集来源于遥感卫星, 图像采用俯拍方式获得, 因此可以通过旋转以及翻转两种数据增强方式扩充训练数据集. 从表 1 实验可知, 训练集从 1 000 张图像扩充到 4 000 张图像后, 在 CCCV2017 飞机测试集上分别使用 ZF 和 VGG16 两种特征提取网络, 遥感图像飞机检测的 mAP 分别提高了 0.57%、0.75%, 假阳率分别下降 0.06%、0.12%, 漏检率分别下降了 4.25%、1.35%. 由此可见, 合理的扩充训练数据集, 有利于网络充分学习遥感图像的特征, 提升模型的检测效果. 之后的实验则是建立在数据集扩充的基础之上.

3.3.2 不同特征提取网络实验对比

一般情况下, 不同深度的特征提取网络, 特征提取能力不同, 更深层次的卷积网络可以有效提取具有判别力的特征. 因此本文在 Faster-RCNN 算法特征提取阶段, 通过引入 ResNet101 残差网络, 同时采取在线困难样本挖掘策略对训练数据集进行训练. 通过表 2 的实验数据可以发现, ResNet101-OHEM 网络模型相较于 ZF 以及 VGG16, mAP 分别提高了 9.98%、8.85%, 假阳率以及漏检率也均有所下降.

表 1 数据增强实验对比

Table 1 Experiment comparison of data augmentation

数据集大小	特征提取网络	mAP/%	FPR/%	MR/%
1 000	ZF	78.17	1.26	18.46
	VGG16	79.12	0.90	8.61
4 000	ZF	78.74	1.20	14.21
	VGG16	79.87	0.78	7.26

表 2 不同特征提取网络实验对比

Table 2 Experiment comparison of different feature extraction networks

特征提取网络	mAP/%	FPR/%	MR/%
ZF	78.74	1.20	14.21
VGG16	79.87	0.78	7.26
ResNet101-OHEM	88.72	0.60	7.01

3.3.3 增加 anchor 数目实验对比

通过对数据集中 4 200 张图片飞机目标的长宽比分布统计, 有针对性地修改生成 anchor 的尺度以及长宽比, 最后将 anchor 数量从 9 增加到 16, 增加了候选目标框的数目. 从表 3 可以看出, 增加生成 anchors 数目, 虽然漏检率有所上升, 但是 mAP 和假阳率指标有更好的改善. 可见, 合理地增加生成 anchor 数目, 可以有效提升目标检测的效果. 最终, 模型在 CCCV2017 遥感图像飞机测试集上 mAP 达到了 89.93%, 假阳率降低到 0.50%, 漏检率降低到 7.28%.

3.3.4 不同算法的比较

实验最终将改进后的算法同初始 Faster-RCNN 算法以及基于全卷积神经网络的 R-FCN 算法进行对比. 其中, 初始 Faster-RCNN 算法特征提取网络采用 VGG16, R-FCN 算法特征提取网络采用 ResNet101, 并且引用在线困难样本挖掘策略对数据进行训练, 实验对比如表 4 以及图 6 所示.

表 3 增加 anchor 数目实验对比

Table 3 Experiment comparison of increasing the number of anchors

候选框数目	<i>mAP</i> /%	<i>FPR</i> /%	<i>MR</i> /%
9	88.72	0.60	7.01
16	89.93	0.50	7.28

通过表 4 第二行以及第四行数据对比,在经过数据增强、采用 ResNet101 特征提取网络和结合在线困难样本挖掘方式训练数据、以及增加 anchors 数目等改进后,最终本文算法比初始 Faster-RCNN 算法在 *mAP* 上提高了 10.81%,假阳率降低了 0.40%,漏检率降低了 1.33%。对比表 4 第三行以及第四行数据可知,同样基于 ResNet101,以及采用在线困难样本挖掘方式训练数据,本文算法在 *mAP*、假阳率、漏检率 3 个指标上优于 R-FCN。从图 6 中 P-R 曲线也可以直观显示出本文算法的优越性以及对于提高遥感图像飞机检测效果的可行性。

3.3.5 不同数据集实验对比

为了测试模型的泛化能力,实验最后分别选取了 NWPU VHR-10、NWPU-RESISC45、UCAS-AOD 数据集中飞机数据集进行测试,测试效果示意图如图 7 所示。

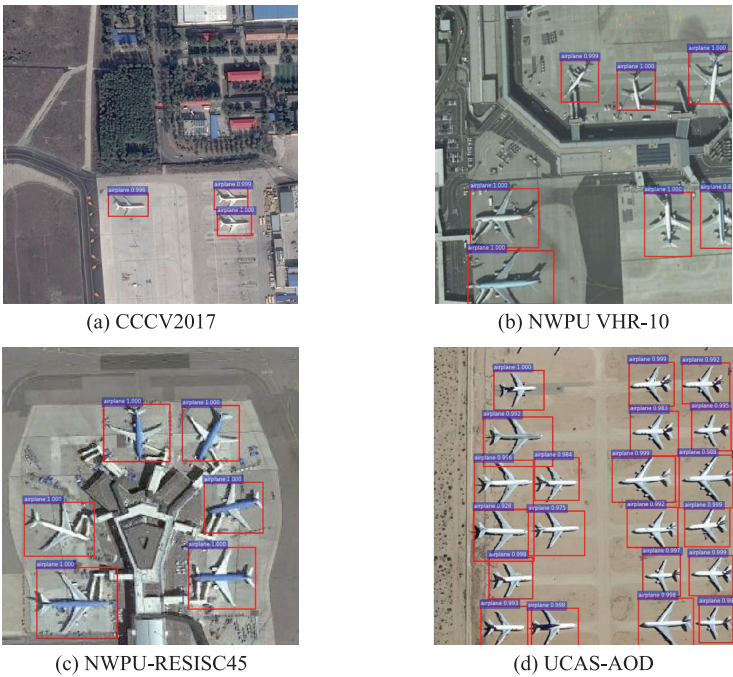


图 7 不同数据集检测结果

Fig. 7 The detection results in different datasets

从图 7 中,可以看出本文算法可以准确检测出不同数据集中飞机目标的位置。结合表 5 实验数据可知,本文算法在 CCCV2017 遥感图像飞机数据集上获得的性能,在其他数据集上同样具有良好的性能,表明本文算法的鲁棒性好、泛化能力强。

4 结论

本文针对遥感图像飞机检测问题进行深入的研究,在初始 Faster-RCNN 目标检测算法的基础上,通过

表 4 不同算法的比较

Table 4 Comparison of different algorithms

算法	<i>mAP</i> /%	<i>FPR</i> /%	<i>MR</i> /%
Faster-RCNN (VGG16)	79.12	0.90	8.61
R-FCN	88.71	0.80	9.23
Ours	89.93	0.50	7.28

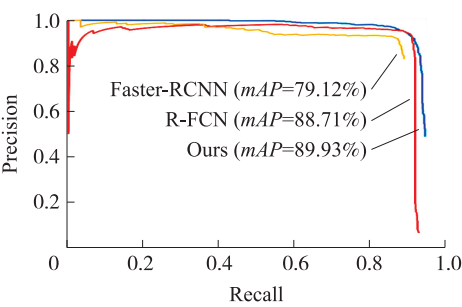


图 6 Precision-recall 曲线

Fig. 6 The curve of precision-recall

表 5 不同数据集实验对比

Table 5 Experimental comparison of different datasets

数据集	<i>mAP</i> /%	<i>FPR</i> /%	<i>MR</i> /%
CCCV2017	89.93	0.50	7.28
NWPU VHR-10	88.80	1.49	6.31
NWPU-RESISC45	80.00	4.36	14.88
UCAS-AOD	88.20	1.90	17.56

采用常用的数据增强方法扩充数据集,在特征提取阶段,引入深度残差网络对遥感图像进行特征提取,在模型训练阶段,根据遥感图像数据集中飞机目标的特点,有针对性地优化区域建议网络.另外,为了防止训练阶段正负样本不平衡导致训练不稳定,本文采取在线困难样本挖掘策略来提高模型训练的稳定性,使得训练的模型更加鲁棒,泛化能力更强.模型最终在 CCCV2017 测试集上 mAP 达到了 89.93%,假阳率降低到 0.50%,漏检率降低到 7.28%,优于初始的 Faster-RCNN 以及 R-FCN 算法,具有很强的实用价值.在 NWPUVHR-10、NWPU-RESISC45、UCAS-AOD 遥感图像飞机数据集测试表明,该改进模型同样具有良好的性能,从而验证了该模型具有良好的鲁棒性和泛化能力.但是该模型针对遥感图像中飞机目标较小情况仍有漏检,未来将进行更深一步的探讨,提高模型对于小目标的检测率.

#### [参考文献]

- [1] 尹宏鹏,陈波,柴毅,等.基于视觉的目标检测与跟踪综述[J].自动化学报,2016,42(10):1466-1489.
- [2] LOWE D G. Distinctive image features from scale-invariant keypoints[J]. International journal of computer vision,2004,60(2):91-110.
- [3] DALAL N,TRIGGS B. Histograms of oriented gradients for human detection[C]//Proceedings of the IEEE Computer Society Conference on Computer Vision & Pattern Recognition. San Diego,2005,1(12):886-893.
- [4] MALLAT S. A wavelet tour of signal processing[M]. Beijing:China Machine Press,2010.
- [5] VAPNIK V,CORTES C. Support vector networks[J]. Machine learning,1995,20(3):273-297.
- [6] CHENG J,GREINER R. Comparing bayesian network classifiers[J]. IEEE transactions on vehicular technology,2013,63(5):2002-2012.
- [7] LECUN Y,BOTTOU L,BENGIO Y,et al. Gradient-based learning applied to document recognition[J]. Proceedings of the IEEE,1998,86(11):2278-2324.
- [8] UIJLINGS J R,SANDE K E,GEVERS T,et al. Selective search for object recognition[J]. International journal of computer vision,2013,104(2):154-171.
- [9] GIRSHICK R,DONAHUE J,DARRELL T. Rich feature hierarchies for accurate object detection and semantic segmentation[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Columbus,2014:580-587.
- [10] GIRSHICK R. Fast R-CNN[DB/OL]. [2018-10-22]. <https://arxiv.org/pdf/1504.08083.pdf>.
- [11] REN S Q,HE K M,GIRSHICK R,et al. Faster R-CNN:towards real-time object detection with region proposal networks[J]. IEEE transactions on pattern analysis & machine intelligence,2017,39(6):1137-1149.
- [12] REDMON J,DIVVALA S,GIRSHICK R,et al. You only look once:unified,real-time object detection[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas,2016:779-788.
- [13] HE K M,ZHANG X Y,REN S Q,et al. Deep residual learning for image recognition[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas,2016:770-778.
- [14] SHRIVASTAVA A,GUPTA A,GIRSHICK R. Training region-based object detectors with online hard example mining[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas,2016:761-769.
- [15] DAI J,LI Y,HE K,et al. R-FCN:object detection via region-based fully convolutional networks[DB/OL]. [2018-10-22]. <https://arxiv.org/pdf/1605.06409.pdf>.
- [16] ZEILER M D,FERGUS R. Visualizing and understanding convolutional networks[DB/OL]. [2018-10-22]. <https://arxiv.org/pdf/1311.2901.pdf>.
- [17] SIMONYAN K,ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[DB/OL]. [2018-10-22]. <https://arxiv.org/pdf/1409.1556.pdf>.
- [18] CHENG G,HAN J W,ZHOU P C,et al. Multi-class geospatial object detection and geographic image classification based on collection of part detectors[J]. Isprs journal of photogrammetry & remote sensing,2014,98(1):119-132.
- [19] CHENG G,HAN J W. A survey on object detection in optical remote sensing images[J]. Isprs journal of photogrammetry & remote sensing,2016,117:11-28.
- [20] CHENG G,ZHOU P C,HAN J W. Learning rotation-invariant convolutional neural networks for object detection in VHR optical remote sensing images[J]. IEEE transactions on geoscience & remote sensing,2016,54(12):7405-7415.
- [21] CHENG G,HAN J W,LU X Q. Remote sensing image scene classification:benchmark and state of the art[J]. Proceedings of the IEEE,2017,105(10):1865-1883.
- [22] ZHU H G,CHEN X G,DAI W Q,et al. Orientation robust object detection in aerial images using deep convolutional neural network[C]//Proceedings of the IEEE International Conference on Image Processing. Quebec,2015:3735-3739.

[责任编辑:顾晓天]