

引用格式: Lin Na, Feng Lirong, Zhang Xiaoqing. Aircraft Detection in Remote Sensing Image based on Optimized Faster-RCNN[J]. Remote Sensing Technology and Application, 2021, 36(2): 275-284. [林娜, 冯丽蓉, 张小青. 基于优化Faster-RCNN的遥感影像飞机检测[J]. 遥感技术与应用, 2021, 36(2): 275-284.]
doi: 10.11873/j.issn.1004-0323.2021.2.0275

基于优化 Faster-RCNN 的遥感影像飞机检测

林娜^{1,2}, 冯丽蓉¹, 张小青¹

(1. 重庆交通大学 土木工程学院, 重庆 400074;
2. 重庆市地理信息和遥感应用中心, 重庆 401147)

摘要: 针对传统飞机检测算法特征学习能力较弱, 在背景复杂、目标密集、成像质量较差的遥感影像上检测精度较低的问题, 提出了一种基于 Faster-RCNN (Faster-Regions with Convolutional Neural Network) 框架的遥感影像飞机检测优化算法。以 ResNet50 为基础特征提取网络, 引入空洞残差块进行多层特征融合, 构建新的特征提取网络, 提高算法的特征提取能力。首先在 UCAS-AOD 数据集上采用交叉验证训练方法验证模型在不同训练集与测试集上的稳定性, 同时比较不同算法的检测性能; 然后在 NWPU VHR-10 数据集上进行飞机检测对比实验, 验证模型泛化性。实验结果表明: 在 UCAS-AOD 数据集上优化算法平均精度为 97.1%, 在 NWPU VHR-10 数据集上优化算法平均精度为 96.2%。该优化算法能够提升遥感影像中飞机的检测精度, 且泛化性更强, 对实现遥感影像飞机快速检测具有一定的参考意义。

关键词: 深度学习; 遥感影像; 目标检测; 特征融合; 空洞残差块

中图分类号: TP75 **文献标志码:** A **文章编号:** 1004-0323(2021)02-0275-10

1 引言

随着遥感卫星、无人机等技术的进步, 能够获取到的遥感影像数据量呈爆炸式地增长。为了能在海量的遥感影像中快速获得有用信息, 目标检测发挥了极大的作用, 具有极高的应用价值。在遥感影像目标检测中, 飞机因其在军事与民用领域中的巨大价值, 一直是重要的研究目标^[1]。通过遥感影像实现飞机快速检测对提高军队作战效率、民用航线规划、安全搜救等具有重要意义^[2]。

近年来, 在飞机检测领域, 传统的飞机检测算法, 如圆周频率滤波法^[3]、边缘提取法^[4]、稀疏表示法^[5]等, 主要停留在对浅层、中层特征的利用, 算法鲁棒性差且检测过程繁琐, 检测结果极易受人工设计的特征质量影响, 不适用于遥感影像的飞机目标检测。目前, 深度学习发展迅速, 其中, 卷积神经网络

(Convolutional Neural Network, CNN) 因其强大的特征提取能力在影像分类、影像分割、语音识别等领域均超越了传统的机器学习算法。2014 年, Girshick 等^[6]首次将深度学习应用到目标检测领域, 提出了区域卷积神经网络 (Regions with CNN, R-CNN) 检测框架, 该框架首先利用选择性搜索算法生成约 2 000 个候选区域, 然后利用 CNN 提取候选区域特征, 最后根据 SVM (Support Vector Machines) 分类器、回归器进行目标分类、候选区域位置回归。R-CNN 极大地提高了自然场景影像的目标检测精度, 开创了基于候选区域的目标检测算法。为了提高 R-CNN 目标检测速度, 减少冗余计算, Girshick 等^[7]提出了快速区域卷积神经网络 (Fast R-CNN) 框架, 该框架将整张影像进行特征提取生成特征图, 再将候选框映射到特征图上, 通

收稿日期: 2020-04-21; 修订日期: 2021-03-11

基金项目: 重庆市教委科技项目 (KJQN201800747), 重庆交通大学研究生教育创新基金项目 (2020S0001)。

作者简介: 林娜 (1981—), 女, 湖北襄阳人, 博士, 副教授, 主要从事深度学习在遥感图像处理中的应用研究。E-mail: linnawb@126.com

通讯作者: 冯丽蓉 (1996—), 女, 四川绵阳人, 硕士研究生, 主要从事深度学习在遥感图像处理中的应用研究。E-mail: 862456630@qq.com

过感兴趣区域池化层(Region of Interest Pooling, ROI Pooling)得到固定尺寸的特征图,实现了特征层共享。为了提高Fast R-CNN中候选框生成效率, Ren等^[8]提出了更快速区域卷积神经网络(Faster-RCNN)框架,利用区域生成网络(Region Proposal Network, RPN)取代了选择性搜索算法,减少了候选框的数量,同时将特征提取、候选区域生成、目标分类和边框回归这4个步骤整合为一个检测框架,实现了端到端(end to end)的检测。与前述基于候选区域的目标检测算法不同, Redmon等^[9]提出了一种基于回归的目标检测算法:YOLO(You Only Look Once),该框架舍弃了预设候选框策略,将目标检测任务视为回归问题,直接在输出层进行边框回归和类别判定,能够达到实时的检测要求,但检测精度较低。为了兼顾检测速度与精度, Liu等^[10]提出了SSD(Single Shot MultiBox Detector)检测框架,融合了预设候选框与YOLO思想,通过特征金字塔同时进行多层检测来提高检测精度,但小目标检测效果不理想。

遥感影像的成像特点不同于常用的由数码相机捕获的自然场景影像,遥感影像主要在高空拍摄,覆盖的地物范围较广、影像背景复杂;目标分布密集、尺寸较小;且成像质量不如数码相机捕获的影像好。直接将现有的深度学习检测框架应用到遥感影像上进行目标检测,不能取得同自然场景影像下相同的检测精度。目前,大量学者对深度学习检测框架进行改进,将其运用于遥感影像中的飞机检测。殷文斌等^[11]在R-CNN模型基础上结合级联AdaBoost算法,提出了一种由粗到精的飞机识别算法,在减少计算量的同时提高了高分遥感影像的目标检测准确率。朱明明等^[12]基于Faster-RCNN,提出了一种特征融合结合软判决的飞机检测算法,提高了小目标的检测精度。Xu等^[13]通过融合浅层特征与上采样后的深层特征,克服了深层特征丢失小目标信息的缺陷,提高了航空遥感影像飞机快速检测精度。张中宝等^[14]针对飞机目标的长宽比特点对RPN进行优化,同时,采用在线困难样本挖掘方法平衡正负样本,提高了模型算法性能。戴媛等^[15]针对遥感影像中目标方向的不确定性,将旋转区域网络融入Faster-RCNN,同时,在网络的全连接层前引入卷积层,降低特征图维数的同时提高分类器性能,获得了更好的检测效果。李淑敏等^[16]基于YOLO V2检测框架,引入迁移学习策略,提高了小

样本下的高分遥感影像飞机检测精度,但小目标检测精度不高。裴伟等^[17]基于SSD检测框架,利用ResNet50替换前置网络VGG16,在航拍数据集上进行目标检测,在飞机类别数据集上效果提升明显。姚群力等^[18]针对复杂场景下小目标检测精度较低的问题,提出了轻量级多尺度特征融合检测框架MultDet,该框架提高了对小尺度飞机目标的表征能力,实现了飞机目标的高精度检测。

考虑到实验数据集样本较少、目标尺寸偏小,因此本文选用基于候选区域的目标检测方法,在经典检测框架Faster-RCNN的基础上进行改进,以ResNet50^[19]为基础特征提取层,然后引入空洞卷积残差块融合多层特征提高遥感影像中飞机的检测精度,减少错分、漏检、定位不准确的情况发生。将改进后的算法应用于高分遥感影像飞机检测中,验证本文方法的有效性。

2 基于优化Faster-RCNN的目标检测算法

2.1 Faster-RCNN基本原理

2.1.1 检测流程

Faster-RCNN是一种基于区域的卷积神经网络框架,主要由特征提取层、RPN层、ROI Pooling层、检测子网络4部分组成,如图1所示。特征提取层(通常是ZF^[20]、VGG^[21]网络)通过卷积神经网络提取输入影像的特征;候选区域生成网络在得到的影像特征图上进行卷积操作,生成可能包含目标物体的候选框(anchor);感兴趣区域池化层将不同候选框对应的区域特征图转换为尺寸一致的特征图;检测网络对感兴趣区域进行目标识别及候选框位置校正,得到最后的检测结果。

Faster-RCNN检测流程:①首先将训练集影像进行缩放后输入网络,通过特征提取层提取影像特征图;②接着将特征图传入RPN层,RPN层根据设定的交并比(Intersection over Union, IOU)阈值生成候选框;③然后将RPN层产生的候选框与步骤①产生的特征图传入ROI Pooling层,得到固定尺度(7×7)的候选框的映射特征图;④最后将其输入检测子网络,回归经过ROI Pooling层之后的候选框,得到最终边框位置,同时,进行目标类别识别,生成分类结果。

2.1.2 RPN工作原理

RPN是一个全卷积网络,减少了检测过程中候

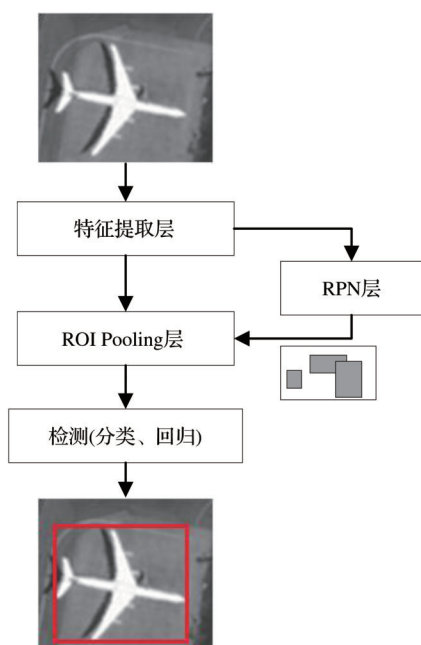


图1 Faster-RCNN框架结构

Fig.1 The structure of the Faster-RCNN framework

选框的数量,提高了Faster-RCNN检测效率。由图2可以看出,RPN网络主要有两个分支:一个是分类层,用于分类正负样本,另一个是回归层,用于回归正样本候选框位置。RPN网络实现流程:在特征图(假设大小为 $N \times 60 \times 60$)上使用 3×3 的滑动窗口进行滑动,遍历特征图每个像元,产生低维特征图($256 \times 60 \times 60$),并且在特征图上每个像元位置生成 k 个预先定义的候选框;然后对低维特征图分别进行两个 1×1 卷积操作,在每个像元处得到 $2k$ 个概率值和 $4k$ 个候选框偏移值;最后,结合预先定义的候选框,进行跨边界裁剪、小候选框移除、非极大值抑制(Non-Maximum Suppression, NMS)等后处理操作,得到感兴趣候选框。

RPN网络的训练采用的是多任务损失函数,如式(1)所示:

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

其中: i 为anchor的索引值; p_i, t_i 分别为第 i 个anchor包含目标的概率值与anchor对应的4个坐标值(中心点坐标与anchor宽高); N_{cls}, N_{reg} 为标准化常数; λ 是平衡系数;当anchor为正样本时, p_i^* 为1,否则 p_i^* 为0; t_i^* 为正样本对应的真实区域的坐标值(中心点坐标与宽高); L_{cls}, L_{reg} 分别为分类损失函数与回归损失函数,表达式如下:

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

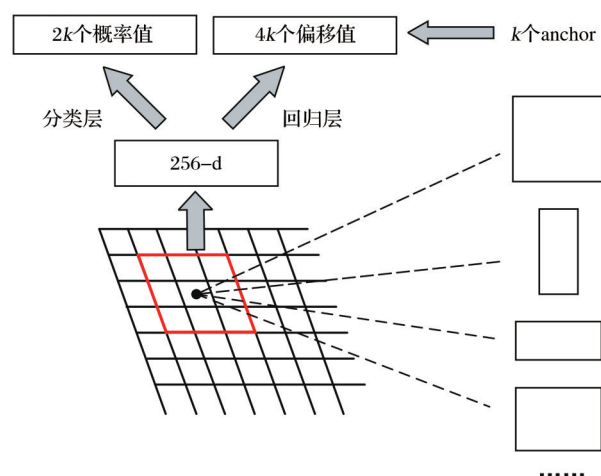


图2 候选区域生成网络(RPN)

Fig.2 Region Proposal Network

$$L_{reg}(t_i, t_i^*) = R(t_i - t_i^*) \quad (3)$$

其中: R 是Smooth L1函数,为反映鲁棒性的函数,定义如下:

$$R(x) = \begin{cases} 0.5x^2 & |x| \leq 1 \\ |x| - 0.5 & |x| > 1 \end{cases} \quad (4)$$

2.2 特征提取网络选取

与原始特征提取层ZF、VGG相比,ResNet50具有更深的网络层次,能够提取更深层、更抽象的目标特征,而且ResNet50中的残差结构有利于解决网络深度加深时的梯度弥散、梯度爆炸以及网络性能提升不明显的问题。选用ResNet50作为基础特征提取网络。Resnet50网络结构如表1所示。在表1中,两个数字相乘,表示卷积核的大小,后面的数字表示卷积核的个数。以conv1为例, 7×7 表示卷积核大小为 7×7 ,64表示卷积核个数。矩阵表示残差块,矩阵内的数字表示残差块的构成情况。以conv3为例,conv3由4个残差块构成,每个残差块由128个 1×1 大小的卷积核,128个 3×3 大小的卷积核,512个 1×1 大小的卷积核组成,其中第一个残差块中的 3×3 大小的卷积核步长为2,其余都为1。

然而,直接将conv1-conv5作为特征提取层,应用于Faster-RCNN结构中,并没有明显提高模型检测精度。为了解决特征提取层加深而目标检测精度提升不明显的问题,Ren等^[22]采用特征图卷积映射网络(Networks on Convolutional Feature Maps, NoCs)结构,将ResNet50中conv1-conv4作为特征提取层,conv5和全连接层作为最后的检测子网络,具体结构如图3所示。将conv5移至检测网络可以提高分类器性能,从而提高整体检测精度。

表1 ResNet50网络结构

Table1 The structure of ResNet50

阶段	结构
conv1	$7 \times 7, 64, \text{步长}=2$ $3 \times 3, \text{最大池化}, \text{步长}=2$
conv2	$\begin{bmatrix} 1 \times 1 & 64 \\ 3 \times 3 & 64 \\ 1 \times 1 & 256 \end{bmatrix} \times 3, \text{步长}=1$
conv3	$\begin{bmatrix} 1 \times 1 & 128 \\ 3 \times 3 & 128 \\ 1 \times 1 & 512 \end{bmatrix} \times 4, \text{步长}=2$
conv4	$\begin{bmatrix} 1 \times 1 & 256 \\ 3 \times 3 & 256 \\ 1 \times 1 & 1024 \end{bmatrix} \times 6, \text{步长}=2$
conv5	$\begin{bmatrix} 1 \times 1 & 512 \\ 3 \times 3 & 512 \\ 1 \times 1 & 2048 \end{bmatrix} \times 3, \text{步长}=2$
$14 \times 14, \text{平均池化}, 1000\text{-d fc}, \text{SoftMax}$	

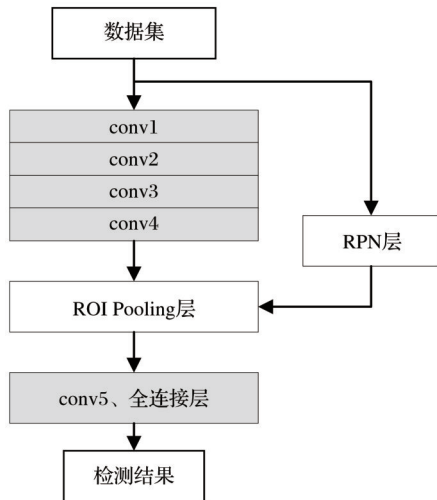


图3 Faster-RCNN与ResNet50结合

Fig.3 A combination of Faster-RCNN and ResNet50

2.3 特征提取网络优化

图4(a)是Resnet50中的原始残差块,图4(b)和4(c)中的空洞残差块^[23]是在Resnet50中原始残差块的基础上引入扩张率为2的空洞卷积而形成的结构。空洞残差块在保证空间分辨率不变的情况下,能扩大深层网络感受野,有利于提取目标的深层语义信息。实验利用空洞残差块设计一种特征融合模型,具体结构如图4(d)所示。

首先,由于原有的ResNet50中的conv5移到了检测网络中,导致特征提取层只有前4层,网络深度较浅,所以在conv4之后,使用一次带有 1×1 卷积层映射的空洞残差块(图4(b)),然后再使用两次空洞残差块(图4(c))作为特征提取网络中新的第5层。空洞残差块的组合保证了第5层与conv4空间分辨

率一致,特征信息损失更小。为了减少计算量与内存消耗,残差块的通道数设为512。其次,由于遥感影像中飞机尺寸偏小,在经过4层下采样后,空间分辨率变小,小目标的语义信息丢失严重,基于此,将conv1、conv3与新的第5层生成的特征进行融合,充分利用低空间分辨率下的深层语义信息与高空间分辨率下的浅层几何、纹理信息等。然而,浅层特征虽保留了高空间分辨率与丰富的小目标信息,但浅层信息过于浅显,直接将浅层特征与深层特征进行融合不能明显提高飞机的检测精度。因此,在特征融合之前,于conv1、conv3层特征后引入与第5层一致的结构,残差块通道数设为256,目的是充分学习高空间分辨率下的目标信息,提高大目标的定位精度与小目标的识别精度,增强网络模型对目标的检测能力。最后,对浅层特征进行下采样、深层特征进行上采样进行特征融合,得到改进后的特征提取网络。

3 实验与分析

实验在Python3.5上进行,所用深度学习框架为TensorFlow与Keras,计算机硬件配置为i5-9300H CPU @ 2.40GHz, 8.00GB内存,显卡为NVIDIA GTX1660Ti, 6GB显存。操作系统为Windows。

3.1 实验数据

实验数据采用UCAS-AOD数据集^[23]和NWPU VHR-10数据集^[25-27]中的飞机样本影像,数据集示例影像如图5所示。UCAS-AOD数据集由中科院大学提供,影像来源于Google Earth,包含飞机、汽车两种样本,其中飞机影像1 000张,共7 482个飞机样本。NWPU VHR-10是西北工业大学提供的航天遥感目标检测数据集,共800张影像,其中88张影像包含飞机样本,共标注飞机样本757个。由于数据集较小,为了防止实验过程中出现过拟合,训练时对数据集进行增强处理。增强方式包括:水平翻转,竖直翻转,旋转 90° 、 180° 、 270° 。

3.2 实验方法

实验采用3种不同的飞机检测算法:以VGG16为特征提取层的原始Faster-RCNN方法,称为Faster-RCNN+VGG16;以ResNet50为特征提取层采用NoCs连接方式的Faster-RCNN方法,称为Faster-RCNN+ResNet50;在Faster-RCNN+ResNet50方法基础上,结合空洞残差块进行特征提取网络优化的算法称为优化Faster-RCNN。通过统

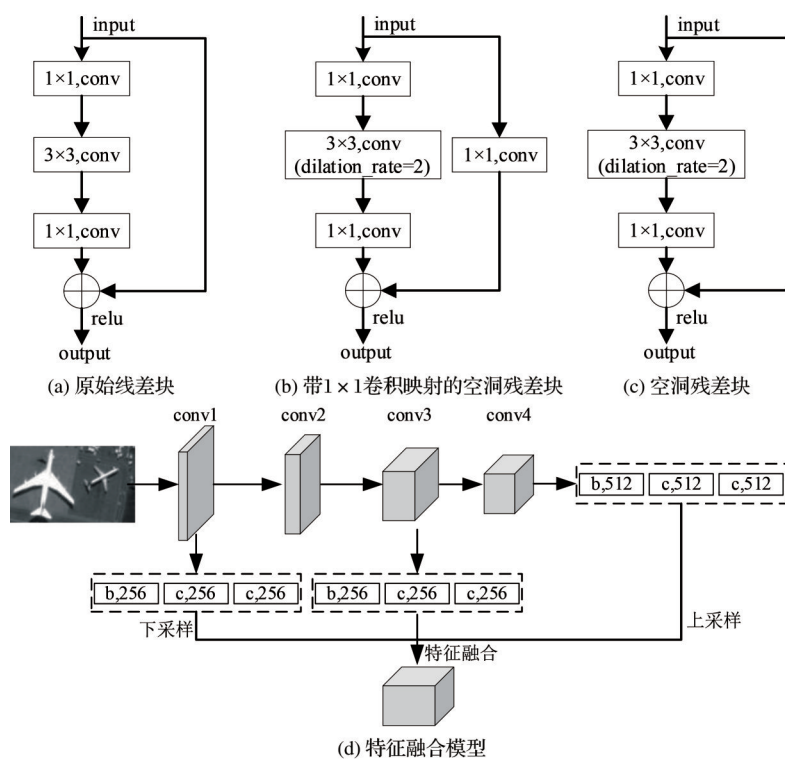


图 4 改进后的特征提取网络

Fig.4 The improved feature extraction network



(b) NWPU VHR-10数据集

图 5 实验数据示例图

Fig.5 The sample of experimental data sets

计数据集中飞机真实外包框的长、宽、纵横比,设置3种算法候选框的尺寸为(32、64、128、256),长宽比为(1:1、1:2、2:1)。

交叉验证是机器学习中一种评估模型泛化性

和稳定性的常用方法。在数据集样本较少的情况下,利用交叉验证有助于避免因数据集划分的不确定性而引起的模型不稳定问题。基于此,实验的模型训练方式采用交叉验证思想:①将 UCAS-AOD

数据集划分为5份等量子集;②取其中一份作为测试集,其余4份作为训练集;③训练模型,得到测试结果;④轮换测试集,重复步骤②、③,直到5份全部测试一遍为止,得到5个模型测试结果。训练过程使用Adam优化方法,依据经验初始学习率设为0.000 01,迭代次数设为40 000次。

3.3 评价指标

采用准确率(Precision)、召回率(Recall)以及不同IOU阈值下的平均精度(Average Precision, AP)值作为评价指标。准确率为模型正确检测出的飞机数量与测试集中被识别为飞机目标的总数的比值;召回率为模型正确检测出的飞机数量与测试集中飞机总样本数量的比值;AP值为综合评价指标,由Precision与Recall绘制的P-R曲线下的面积确定。各指标计算公式如下:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (6)$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (7)$$

$$\text{AP} = \int_0^1 P(r) dr \quad (8)$$

其中:FN表示被误判为背景的飞机样本,TP表示为正确识别出的飞机样本,FP表示被误判为飞机的背景样本, r 表示Recall值, $P(r)$ 表示 r 值所对应的Precision值。

3.4 结果分析

3.4.1 交叉验证对比实验

为了验证本文所提出的优化Faster-RCNN算法在遥感影像飞机检测中的有效性,本文分别采用3种不同的飞机检测算法在UCAS-AOD数据集上进行五折交叉验证,得到3种实验方法在UCAS-AOD数据集上各测试指标(IOU阈值为0.5时的准确率与召回率,和IOU分别为0.5、0.75、0.85时的AP值)的均值和方差,如表2所示。

表2 UCAS-AOD数据集上交叉验证测试结果对比

Table2 Comparison of cross-validation detection results on UCAS-AOD data set

实验方法		准确率	召回率	AP(0.50)	AP(0.75)	AP(0.85)
Faster-RCNN+VGG	均值	88.3	92.5	89.2	46.7	23.2
	方差	0.020	0.052	0.020	0.076	0.068
Faster-RCNN+ResNet50	均值	91.7	94.6	92.8	52.3	25.4
	方差	0.004	0.028	0.008	0.028	0.032
优化Faster-RCNN	均值	97.6	98.7	97.1	58.9	30.4
	方差	0.032	0.028	0.038	0.020	0.044

由表2可以看出各指标的方差均小于0.1,说明轮换数据集对模型影响较小,模型精度具有良好的稳定性。在IOU阈值为0.5时,优化Faster-RCNN的AP值分别为97.2%,比Faster-RCNN+VGG、Faster-RCNN+ResNet50分别提高了7.9%、4.3%。实验选用深层网络ResNet50替换VGG能提取更深层的语义信息,同时利用空洞残差块进行多层特征融合进一步提升了算法检测精度。IOU阈值越大,对算法定位的准确性要求越高。从表中可以看出在IOU阈值增大的情况下,3种方法的AP均值排序依旧是优化Faster-RCNN>Faster-RCNN+ResNet50>Faster-RCNN+VGG,表明本文方法的整体分类与定位精度更高。

为了直观展示3种实验方法的检测效果,抽取交叉验证中3种方法在同一测试集上的检测图片进行展示,如图6所示。

图6展示了3种不同场景影像下的部分测试结果图,图中红色框为模型正确检测结果;绿色框为

模型错分结果;黄色框为模型漏检结果。图6(a)测试影像中包含建筑物、植被、汽车、道路等目标,干扰信息较多,背景复杂,从检测结果可以看出本文方法无错分、漏检情况。图6(b)测试影像主要集中于停机坪附近,飞机目标分布较密集,从检测结果可以看出改进前的两种方法在目标密集区域存在漏检的情况,且均有将登机桥误判为飞机的情况,而本文方法能够正确检测出所有飞机目标,在密集区域检测精度较高。图6(c)测试影像亮度过强,成像质量较差,且目标尺寸变化较大,从检测结果可以看出本文方法在影像质量较差区域仍有较高的检测精度。对比3组影像图可以看出,采用优化Faster-RCNN算法进行飞机检测,检测精度更高,抗干扰能力更强,具有良好的自适应性,更有利于飞机目标的检测。

3.4.2 模型泛化性测试

上述交叉验证对比实验中测试子集与训练集样本相似度较高,因此模型检测精度偏高。为了验

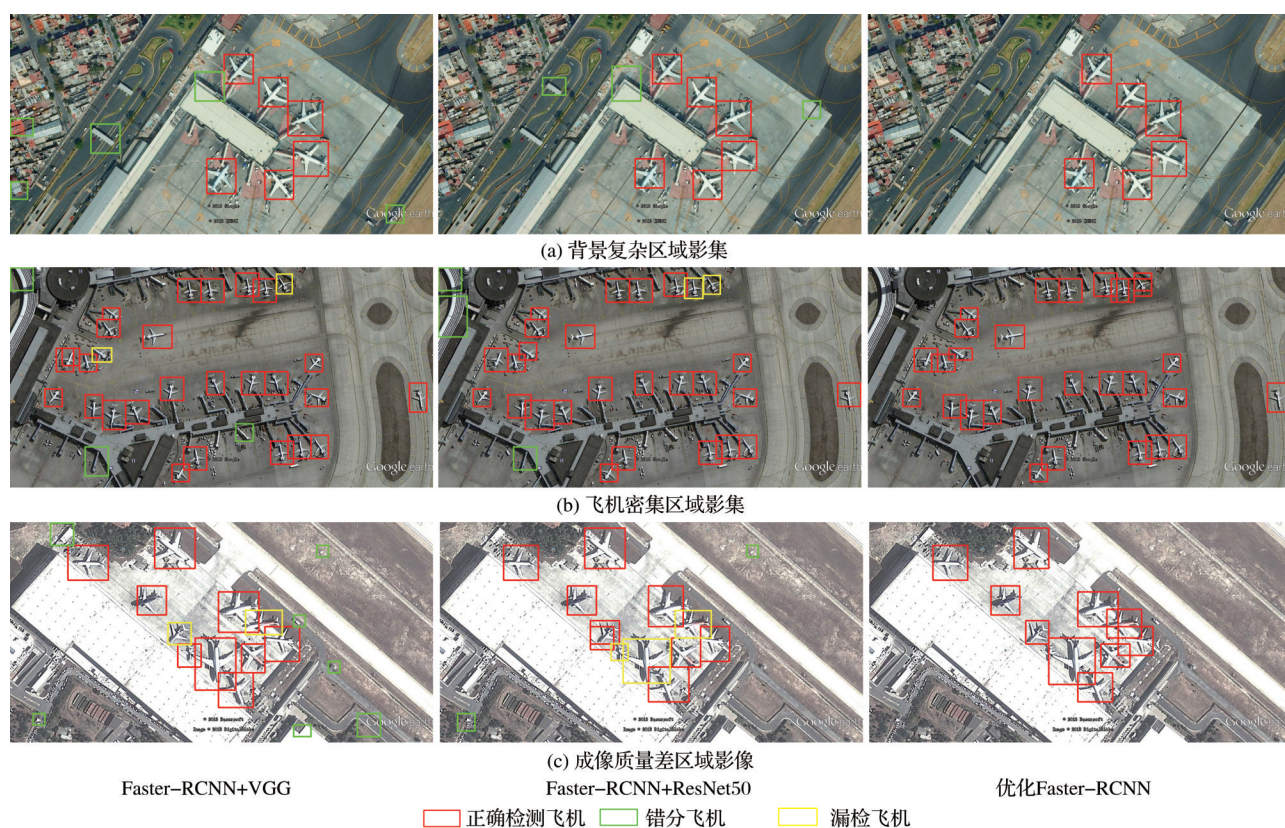


图 6 UCAS-AOD 数据集上测试结果对比图

Fig.6 The comparison of detection results on UCAS-AOD data set

证本文优化算法所得模型在其他飞机遥感影像数据集上的泛化性,利用交叉验证建立的模型对 NWPU VHR-10 数据集上的 88 张飞机影像进行检测,通过对比检测结果评估模型泛化性能力。检测结果如表 3 所示。

表 3 飞机检测结果对比

Table3 Comparison of aircraft detection results

实验方法	准确率	召回率	AP (0.50)
Faster-RCNN+VGG	92.8	94.1	89.3
Faster-RCNN+ ResNet50	93.2	94.9	92.6
优化 Faster-RCNN	96.8	97.6	96.2

从表 3 中可知,优化 Faster-RCNN 算法在检测 NWPU VHR-10 数据集时:准确率较 Faster-RCNN+VGG、Faster-RCNN+ResNet50 算法分别高 4.0%、3.6%,表明优化算法错分较少,检测较精确;召回率较 Faster-RCNN+VGG、Faster-RCNN+ResNet50 算法分别高 3.5%、2.7%,表明优化算法漏检较少,正确检测出数据集中绝大多数飞机目标;IOU 阈值为 0.5 时的平均精度 AP 值较 Faster-RCNN+VGG、Faster-RCNN+ResNet50 算法分别高 6.9%、3.6%,表明优化 Faster-RCNN

算法具有良好的检测性能。结合图 7 在 NWPU VHR-10 数据集上的检测结果图,优化算法整体检测效果较好,定位较准确,模型鲁棒性较强,具有较好的泛化效果。

4 讨 论

基于 Faster-RCNN 框架实现了遥感影像飞机目标的高精度检测,具有一定参考意义。为了深入理解实验细节,为相关研究提供参考建议,特对研究方法与不足之处进行探讨。

从特征提取网络选取来看,选用深层卷积神经网络 ResNet50 作为特征提取网络,在 UCAS-AOD 数据集和 NWPU VHR-10 数据集上的实验结果表明 Faster-RCNN+ResNet50 的算法结果优于以 VGG 为特征提取网络的原始算法,采用 ResNet50 作为特征提取主干网络更适用于本文所用遥感影像飞机检测数据集,实验结果与裴伟等^[17]在 SSD 框架上的研究结果一致。在特征网络选取中,数据集与网络复杂度是两个重要的参考指标。不同的数据集适用于不同的特征提取网络,如在杂草数据集上,选用 VGG 比 ResNet50 检测效果更好^[28];网络复杂度的大小决定了训练速度的快慢和内存占用大



图 7 NWPU VHR-10数据集上检测结果图

Fig.7 The detection results on NWPU VHR-10 data set

小。深层网络虽能提取更丰富的语义信息但通常训练时间更长、内存占用更大,选用 ResNet50 而未选择更深层的 ResNet101 正是基于此番考量。

从特征融合优化方法来看,首先通过引入空洞残差块结构新增了第 5 层来加深网络深度,与姚群力等^[18]选用的方法相似,增添新的卷积层可以提取更深层的语义信息;然后利用空洞残差块结构提取浅层特征信息,并将提取后的浅层特征与第 5 层深层特征进行融合,与 Xu 等^[13]采用的融合方式相似。目标检测的任务是进行类别判定与目标定位。浅层特征空间分辨率大,信息丢失少,有利于目标定位;深层特征空间分辨率小,语义信息丰富,有利于目标分类。综合利用浅层与深层特征更有利于提高影像目标检测精度。

针对上述实验,仍存在不足之处:①针对基于候选区域的目标检测算法,候选框的合理选取是模型训练的关键。虽然对影像数据集外边框进行统计,选取合适的候选框尺寸、比例,但没有像有关研究^[29]一样进行候选框对比实验,未对候选框的影响进行验证,在后续研究中,应设置对比实验进行深入讨论。②此外,实验未对空洞残差块结构效果进行细化验证,可以通过轮换去掉 conv1、conv3 后的空洞残差块结构增加对比实验,验证有效性。③实验所用飞机检测遥感影像主要采集于全球各地的飞机场,场景范围存在局限性。在未来的研究中,应扩大数据集来源范围,建立样本丰富的飞机检测数据集。

5 结 语

为提高遥感影像飞机检测精度,在原有 Faster-RCNN 框架基础上进行优化,以 ResNet50 为基础特

征提取网络,引入空洞残差块结构,进行不同尺度下的多层特征融合,有利于小目标的识别与大目标的定位。对比不同实验方法在 UCAS-AOD 数据集上的交叉验证结果,本文所提出的优化 Faster-RCNN 算法平均精度 AP 为 97.1%,检测性能更好,同时,比较 3 种算法在影像背景复杂区域、飞机密集区域、成像质量差区域时的检测结果图,可以看出优化 Faster-RCNN 算法抗干扰能力更强,具有良好的鲁棒性。在 NWPU VHR-10 数据集上的模型泛化性能对比测试中,优化 Faster-RCNN 算法精确率、召回率、平均精度均优于原始方法,表明优化 Faster-RCNN 算法泛化能力更强。如何利用深度学习技术实现大范围遥感影像目标快速检测仍是未来研究重点。

参考文献 (References):

- [1] Yu Donghang, Guo Haitao, Zhang Baoming, *et al.* Aircraft Detection in Remote Sensing Images Using Cascade Convolutional Neural Networks[J]. *Acta Geodaetica et Cartographica Sinica*, 2019, 48(8): 1046-1058. [余东行, 郭海涛, 张保明, 等. 级联卷积神经网络的遥感影像飞机目标检测[J]. *测绘学报*, 2019, 48(8): 1046-1058.]
- [2] Li Wenbin, He Ran. Aircraft Detection Models of Remote Sensing Images based on FCN and CNN[J]. *Computer Engineering*, 2020-03-11. doi: 10.19678/j.issn.1000-3428.057070. [李文斌, 何冉. 基于 FCN 与 CNN 的遥感影像飞机目标检测方法[J]. *计算机工程*, 2020-03-11. doi: 10.19678/j.issn.1000-3428.0057070.]
- [3] Cai Hongping, Geng Zhenwei, Su Yi. A New Method to Detect Airplanes in Remote Sensing Image: Circle-Frequency Filter[J]. *Signal Processing*, 2007, 33(4): 539-543. [蔡红苹, 耿振伟, 粟毅. 遥感图像飞机检测新方法——圆周频率滤波法[J]. *信号处理*, 2007(4): 539-543.]
- [4] Qiu Jianbin, Li Shijin, Wang Wei. A New Approach to De-

- tect Aircrafts in Remote Sensing Images based on Corner and Edge Information Fusion[J]. *Microelectronics and Computer*, 2011, 28(9):214-216.[仇建斌, 李士进, 王玮. 角点与边缘信息相结合的遥感图像飞机检测新方法[J]. *微电子学与计算机*, 2011, 28(9):214-216.]
- [5] Lin Yudong, He Hongjie, Yin Zhongke, *et al.* Air-plane Detection in Optical Remote Sensing Image based on Sparse-Representation[J]. *Acta Photonica Sinica*, 2014, 43(9):1-6. [林煜东, 和红杰, 尹忠科, 等. 基于稀疏表示的可见光遥感图像飞机检测算法[J]. *光子学报*, 2014, 43(9):1-6.]
- [6] Girshick R, Donahue J, Darrell T, *et al.* Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation[C]// *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, 2014:580-587. doi: 10.1109/CVPR.2014.81.
- [7] Girshick R. Fast R-CNN[C]// *Proceedings of 2015 IEEE International Conference on Computer Vision*. Santiago: IEEE, 2015:1440-1448. doi:10.1109/ICCV.2015.169.
- [8] 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 and Machine Intelligence*, 2015, 39(6):1137-1149. doi: 10.1109/TPAMI.2016.2577031.
- [9] Redmon J, Divvala S, Girshick R, *et al.* You only Look Once: Unified, Real-time Object Detection[C]// *IEEE Conference on Computer Vision and Pattern Recognition*. IEEE Computer Society, 2016: 779-788. doi: 10.1109/CVPR.2016.91.
- [10] Liu Wei, Anguelov D, Erhan D, *et al.* SSD: Single Shot MultiBox Detector [M]. *Computer Vision-ECCV 2016*. Springer International Publishing, 2016:21-37. doi: 10.1007/978-3-319-46448-0_2.
- [11] Yin Wenbin, Wang Chengbo, Yuan Cui, *et al.* A Method of Aircraft Recognition in Remote Sensing Images[J]. *Bulletin of Surveying and Mapping*, 2017(3):34-37.[殷文斌, 王成波, 袁翠, 等. 一种飞机目标的遥感识别方法[J]. *测绘通报*, 2017(3):34-37.]
- [12] Zhu Mingming, Xu Yuelei, Ma Shiping, *et al.* Airplane Detection based on Feature Fusion and Soft Decision in Remote Sensing Images[J]. *Acta Optica Sinica*, 2019, 39(2):71-77. [朱明明, 许悦雷, 马时平, 等. 基于特征融合与软判决的遥感图像飞机检测[J]. *光学学报*, 2019, 39(2):71-77.]
- [13] Xu Y L, Zhu M M, Peng X, *et al.* Rapid Airplane Detection in Remote Sensing Images based on Multilayer Feature Fusion in Fully Convolutional Neural Networks [J]. *Sensors*, 2018, 18(7):2335-2347. doi:10.3390/s18072335.
- [14] Zhang Zhongbao, Wang Hongyuan, Zhang Ji, *et al.* Remote Sensing Image Aircraft Detection Algorithm based on Faster RCNN[J]. *Journal of Nanjing Normal University(Natural Science Edition)*, 2018, 41(4):79-86.[张中宝, 王洪元, 张继, 等. 基于Faster-RCNN的遥感图像飞机检测算法[J]. *南京师大学报(自然科学版)*, 2018, 41(4):79-86.]
- [15] Dai Yuan, Yi Benshun, Xiao Jinsheng, *et al.* Object Detection of Remote Sensing Image based on Improved Rotation Region Proposal Network[J]. *Acta Optica Sinica*, 2020, 40(1):270-280.[戴媛, 易本顺, 肖进胜, 等. 基于改进旋转区域生成网络的遥感图像目标检测[J]. *光学学报*, 2020, 40(1):270-280.]
- [16] Li Shumin, Feng Quanlong, Liang Qichun, *et al.* Aircraft Auto-detection in Domestic High Resolution Remote Sensing Images Using Deep Learning [J]. *Remote Sensing Technology and Application*, 2018, 33(6):1095-1102.[李淑敏, 冯权洸, 梁其椿, 等. 基于深度学习的国产高分遥感影像飞机目标自动检测[J]. *遥感技术与应用*, 2018, 33(6):1095-1102.]
- [17] Pei Wei, Xu Yanming, Zhu Yongying, *et al.* The Target Detection Method of Aerial Photography Images with Improved SSD[J]. *Journal of Software*, 2019, 30(3):738-758.[裴伟, 许晏铭, 朱永英, 等. 改进的SSD航拍目标检测方法. *软件学报*, 2019, 30(3):738-758.]
- [18] Yao Qunli, Hu Xian, Lei Hong. Aircraft Detection in Remote Sensing Imagery with Multi-scale Feature Fusion Convolutional Neural Networks[J]. *Acta Geodaetica et Cartographica Sinica*, 2019, 48(10):1266-1274.[姚群力, 胡显, 雷宏. 基于多尺度融合特征卷积神经网络的遥感图像飞机目标检测[J]. *测绘学报*, 2019, 48(10):1266-1274.]
- [19] He K M, Zhang X Y, Ren S Q, *et al.* Deep Residual Learning for Image Recognition [C]// *Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition*. Las Vegas, NV, United States: IEEE, 2016: 770-778. doi: 10.1109/CVPR.2016.90.
- [20] Zeiler M D, Fergus R. Visualizing and Understanding Convolutional Networks [C]// *European Conference on Computer Vision*. Zurich: Springer, 2014:818-833.
- [21] Simonyan K, Zisserman A. Very Deep Convolutional Networks for Large-scale Image Recognition [C]// *International Conference on Learning Representations*. San Diego: IEEE, 2015:1-14.
- [22] Ren S Q, He K M, Girshick R, *et al.* Object Detection Networks on Convolutional Feature Maps[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2016, 39(7):1476-1481. doi: 10.1109/TPAMI.2016.2601099.
- [23] Li Z M, Chao P, Yu G, *et al.* DetNet: A Backbone Network for Object Detection [J/OL]. (2018-04-19) [2019-04-26]. <https://arxiv.org/abs/1804.06215>.
- [24] 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.
- [25] 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 and Remote Sensing*, 2014, 98(1):119-132. doi: 10.1016/j.isprsjprs.2014.10.002.
- [26] Cheng G, Han J W. A Survey on Object Detection in Optical

- Remote Sensing Images[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2016, 117: 11-28. doi: 10.1016/j.isprsjprs.2016.03.014.
- [27] 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 and Remote Sensing, 2016, 54(12): 7405-7415. doi: 10.1109/TGRS.2016.2601622.
- [28] Zhang Le, Jin Xiu, Fu Leiyang, *et al.* Recognition Method for Weeds in Rapeseed Field based on Faster R-CNN Deep Network[J]. Laser & Optoelectronics Progress, 2020, 57(2): 304-312.[张乐,金秀,傅雷扬,等.基于Faster R-CNN深度网络的油菜田间杂草识别方法[J].激光与光电子学进展, 2020, 57(2): 304-312.]
- [29] Wang Yingjie, Zhang Qiao, Zhang Yanmei, *et al.* Oil Tank Detection from Remote Sensing Images based on Deep Convolutional Neural Network[J]. Remote Sensing Technology and Application, 2019, 34(4): 727-735.[王颖洁,张荞,张艳梅,等.基于深度卷积神经网络的油罐目标检测研究[J].遥感技术与应用, 2019, 34(4): 727-735.]

Aircraft Detection in Remote Sensing Image based on Optimized Faster-RCNN

Lin Na^{1,2}, Feng Lirong¹, Zhang Xiaoqing¹

(1.School of Civil Engineering, Chongqing Jiaotong University, Chongqing 400074, China;

2.Chongqing Geomatics and Remote Sensing Application Center, Chongqing 401147, China)

Abstract: To address the problem that traditional aircraft detection methods have low detection accuracy on remote sensing images with complex backgrounds and dense targets, an improved remote sensing image aircraft target detection algorithm based on Faster-RCNN (Faster-Regions with Convolutional Neural Network) is proposed. ResNet50 is used as the basic feature extraction network of the algorithm, and the dilated bottlenecks are introduced for multi-layer feature fusion to construct a new feature extraction network, which improve the feature extraction capability of the algorithm. First, the cross-validation training method is used on the UCAS-AOD data set to verify the stability of the model on different training sets and test sets, and compare the detection performance of different algorithms. Then, comparative experiment is conducted on the NWPU VHR-10 data set to verify the generalization of the model. Experimental results showed that: The average precision of the proposed algorithm is 97.1% on the UCAS-AOD data set and 96.2% on the NWPU VHR-10 data set. The study indicated that the proposed algorithm in this paper can not only improve the detection accuracy of aircraft in remote sensing images, but also have a stronger generalization, which has certain reference significance to the rapid detection of aircraft in remote sensing images.

Key words: Deep learning; Remote sensing image; Object detection; Feature fusion; Dilated bottleneck