

融合特征的深度学习遥感图像目标检测模型

张永福,宋海林,班越,汪西莉

(陕西师范大学 计算机科学学院,陕西 西安 710119)

摘要:为了提高目标检测模型对遥感图像中排列密集、尺度不一的目标,特别是小目标的检测性能,提出了融合特征的深度学习遥感图像目标检测模型和方法。模型采用小规模的网络结构,以应对标记样本较少的情况,并提出了融合多级特征的策略获取更为有效的特征,使模型在不增加检测时间的同时,提高遥感图像中较为密集且大小不一的目标的检测精度。模型中提出了一种新的后处理算法——分组融合剔除检测框算法,在剔除冗余检测框的同时微调检测框位置,使检测框对目标定位更精确,进一步提升检测精度。实验结果表明,所提模型在 UCAS-AOD 和 RSOD-Dataset 数据集上检测飞机,精度比 Faster R-CNN 的结果提高了 4.2% 和 7.3%,漏检率和误检率均有降低。在 UCAS-AOD 数据集上检测更小的汽车目标,所提模型比 Faster R-CNN 检测精度提高了 7.9%,漏检率下降了 5.91%,误检率下降了 2.06%。和 Faster R-CNN 相比,所提的融合处理和检测框后处理算法使得模型针对复杂场景中多尺度密集目标和小目标取得了更高的检测性能。

关键词:遥感图像;目标检测;卷积神经网络;特征融合;后处理

中图分类号:TP391

文献标识码:A

文章编号:1673-629X(2021)09-0048-07

doi:10.3969/j.issn.1673-629X.2021.09.009

Fusion Features Based Deep Learning Remote Sensing Image Target Detection Model

ZHANG Yong-fu, SONG Hai-lin, BAN Yue, WANG Xi-li

(School of Computer Science, Shaanxi Normal University, Xi'an 710119, China)

Abstract: In order to improve the performance of the object detection model for dense, different sizes targets, especially for small targets, in remote sensing images, a fusion features based deep learning remote sensing image target detection model is proposed. The model adopts small-scale network structure aiming at the small labeled sample data sets. The strategy of multi-level feature fusion is proposed to obtain more effective features, so that the model can improve the detection accuracy for dense and different sizes targets in remote sensing images without increasing the detection time. A new post-processing algorithm, namely packet fusion reject detection bounding boxes algorithm, is proposed, which can remove the redundant detection bounding box and fine-tune the position of the detection box, thus improving the box's locating accuracy and further improving the detection accuracy. The experimental results on the UCAS-AOD and RSOD-Dataset remote sensing datasets for detecting aircraft and vehicle targets show that the proposed model achieves high detection performance for multi-scale and dense targets on small sample datasets. Compared with Faster R-CNN, it improves the detection accuracy by 4.2% and 7.3% for plane target, and by 7.9% for car targets, and reduces the missing ratio by 5.91% and false alarm rate by 2.06%. Compared with the existing models for dense and different sizes targets detection in complex remote sensing image, the proposed model achieves higher performance due to the feature fusion processing and the detection post-processing algorithm.

Key words: remote sensing image; object detection; convolutional neural network; feature fusion; post processing

0 引言

目标检测是计算机视觉的研究热点之一,随着遥感卫星以及无人机的增多,遥感图像数据的大数据特点日益凸显。遥感图像目标检测,即在遥感图像中获取特定目标的位置和类别,在公共安全、车辆监控、智

慧城市等诸多方面具有重要的应用价值。遥感图像目标检测和普通光学图像目标检测相比有很多不同,如遥感图像尺寸较大,而目标相对所占比例小,目标可能非常密集,且目标的大小、姿态等可能会有很大的不同,遥感图像中的复杂背景以及遮挡、阴影等也会使目

收稿日期:2020-10-22

修回日期:2021-02-25

基金项目:科技部第二次青藏高原综合科学考察研究项目(2019QZKK0405)

作者简介:张永福(1973-),男,实验师,研究方向为智能信息处理。

标检测难度增加。针对这些挑战及标记样本较少的实际情况,提出了一种融合特征的深度学习遥感图像目标检测模型,贡献主要有:小规模的网络结构应对标记样本少的问题,多级特征融合策略应对不同尺度目标检测难题,新的后处理算法进一步提升检测精度。所提模型和 Faster R-CNN 相比在提高检测精度的同时降低了漏检率和误检率。

1 相关工作

随着深度学习的兴起,卷积神经网络(convolutional neural networks, CNN)被用于图像目标检测, CNN 通过大量的训练数据自主学习到图像的多层次特征,相对于传统的手工设计特征,展现了其自动和有效性^[1]。但是卷积网络由于卷积和池化操作在信息传递时会导致信息丢失,进而对多尺度目标检测效果不理想。通常采用两种方式解决此问题,一种是改进卷积网络结构,如 He 等^[2]提出了残差网络模型(residual neural network, ResNet),该模型建立前后层之间的“短路连接”,直接将输入信息与输出特征结合使特征更加完整,并有助于训练过程中梯度的反向传播,从而训练出更深的 CNN 网络,提高目标检测精度。Kuang 等^[3]提出了 DenseNet 网络,其每一层的输入都是前面所有层输出的并集,而该层所学习的特征图也会被直接传给其后面所有层作为输入,通过这些操作达到保留特征的目的。

另一种解决思路是融合多级特征图。Sean 等^[4]提出了 inside-outside net (ION), ION 在最后一层卷积输出的特征图上分别在上、下、左、右四个方向独立地使用 RNN,并将它们的输出连接成一个特征输出,将两次相同操作得到的特征作为上下文信息,再与前面不同的卷积层得到的特征连接起来作为一个特征输出。这样得到的特征既包括上下文信息,又包含多尺度信息,有利于对不同尺度的目标进行检测,使检测结果更加精确。Lin 等^[5]提出了特征金字塔网络 feature pyramid network (FPN),利用特征图金字塔,通过 bottom-up 和 top-down 的方法创建了一个具有横向连接的自顶向下架构用在多尺度上构建高级语义特征图。Kong 等^[6]提出一种端到端的全卷积检测方法(RON),通过反向连接块将相邻的特征图融合起来,这样浅层的特征图融合了高层的特征图信息,然后分别对融合后的多个特征图做预测,预测得到的检测框整合在一起输出就得到整个网络的输出结果。Cai 等^[7]通过级联几个不同阈值训练得到的网络,前一个检测模型的输出作为后一个检测模型的输入,达到不断优化预测结果的目的。Singh 等^[8]提出了图像金字塔尺度归一化(SNIP),用图像金字塔来处理数据集中

不同大小的数据。Zhang 等^[9]通过 scale-depend pooling (SDP)来处理目标尺度多样的问题,根据目标大小选择合适的卷积层的特征作为分类器和边框回归器的输入特征。

上述文献研究基于深度学习的图像目标检测仍然存在一些不足。首先,GoogLeNet、ResNet 等大型网络在训练时需要大量的标记样本^[10],但遥感图像有标记样本往往有限,对于大规模的深度网络,少量训练样本会造成训练不充分。其次,级联不同网络会造成参数的增加,使时间成本增加。最后,现有特征融合方式未高效和有效地将浅层网络与深层网络信息融合用于预测,对遥感图像中小目标的检测帮助不大^[11]。

本研究以 Faster R-CNN 目标检测框架为基础,提出了一种融合特征的深度学习遥感图像目标检测模型(fusion features based deep learning remote sensing image target detection model, FF-DLM),主要工作和创新包括:一,模型用小型卷积深度网络作为特征提取的主干网络,适用于训练样本较少的情况;二,在特征提取主干网络中通过融合特征图提高对不同尺度目标的检测率,在融合后的特征图上进行预测,降低漏检率,并采用 RPN 网络提取候选框,加快处理速度;三,提出一种新的检测框后处理算法——分组融合剔除检测框算法(packet fusion reject detection bounding boxes, PFR-DBB),在减少冗余检测框的同时微调检测框位置,使检测框对目标定位更精确。模型在有限的训练样本和相对较短的时间内,对遥感图像中较小且密集的目标取得了较高的检测正确率,较低的漏检率、误检率,且对目标定位更为准确。

2 方法原理

两阶段目标检测深度学习模型包括定位和识别两个阶段,其中 region with convolutional neural network (R-CNN)^[12]目标检测方法重复提取所有的候选区域特征而造成时间花费巨大,效率低下。为此, Fast R-CNN^[13]使用 ROI-Pooling 将提取的候选区映射到最后一个卷积层的特征图上,只需要提取一次特征,大幅度提高了检测速度,但它使用 Selective Search 进行候选区域的提取耗时过多^[14],为此 Faster R-CNN^[15]提出了区域建议网络(region proposal network, RPN),自动提取候选区域,这样可以使目标检测在一个网络中进行,真正形成端到端的训练。因此 Faster R-CNN 在时间、效率上比之前的模型有巨大的优势。遥感图像目标检测比普通光学图像目标的检测难度更高,文中以 Faster R-CNN 框架为基础,提出 FF-DLM 模型应对遥感图像训练样本较少、目标较小且密集等对目标检测带来的困难。

2.1 融合特征的目标检测模型

FF-DLM 和 Faster R-CNN 目标检测模型具有相同的框架,模型包括 3 个部分:特征提取、区域提议和检测部分。首先利用一个小型的卷积网络提取特征;其次,融合多层特征图,将融合的特征图送到区域提取

网络产生候选区域,再将候选区域送到目标检测网络中得到目标检测框和目标的类别;最后,使用分组融合剔除检测框 (PFR-DBB) 算法去除冗余的检测框,并调整检测框位置使其对目标定位更精确。

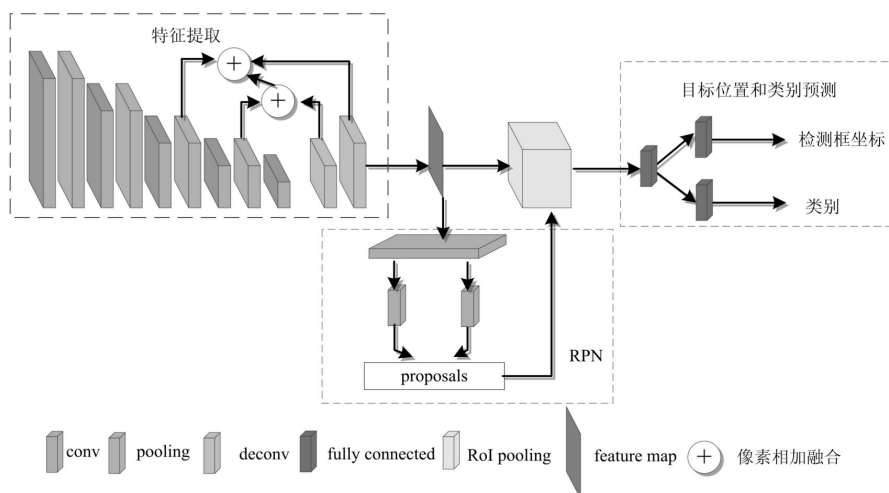


图 1 FF-DLM 模型结构

FF-DLM 模型结构如图 1 所示。其中特征提取卷积神经网络有 7 层,前 5 层是卷积层,后 2 层是反卷积层。网络中的卷积核均采用较小的尺寸。为了减少参数个数,降低计算量,设置第一个卷积层卷积核的尺寸为 7×7 ,第二层卷积核的尺寸为 5×5 ,剩下的卷积层的卷积核尺寸为 3×3 。5 个卷积层的卷积核个数分别为 96、256、384、384、256。卷积核作用于前一层卷积操作后输出的所有特征图上,同一卷积核对于前一层不同特征图的权重不同以提取不同的特征。前四个卷积层包括卷积运算、Relu 非线性激活处理、数据归一化和池化操作。采用最大池化方式,池化时对特征图进行 $\text{Padding} = 1$ 的扩充。最后一个卷积层不进行池化操作。反卷积阶段使用 3×3 的卷积核进行反卷积操作恢复 conv4、conv5 的特征图。这里通过卷积操作提取输入的遥感图像特征,池化操作保留卷积后的主要特征,减少干扰信息。

鉴于越深层次的特征语义信息越丰富,特征越来越抽象,分辨率越来越低;层次较浅的特征具有较高的分辨率,适于小目标的检测,故提出将多级特征融合以保留不同尺度目标的特征来适应不同尺度目标的检测任务^[16]。为了进行融合操作,对 conv4、conv5 通过反卷积操作得到增大的特征图,在高分辨率特征图上保留语义信息。然后将 conv3、conv4、conv5 的特征图进行逐像素相加的融合,融合结果既包含浅层次的细节特征,又包含深层次的语义特征。接着,用 3×3 的卷积核在卷积神经网络融合后的特征图上滑动,每个滑动窗口产生 9 个不同比例 $\{1:1, 1:2, 2:1\}$ 、不同尺寸 $\{322, 642, 1282\}$ 的锚框 (anchor box),之后用两个

1×1 的卷积核判断其是否为目标并进行位置回归,将所有锚框按目标置信度排序,取前 300 个作为候选框。最后将候选框送到检测网络再次对候选区域进行分类,并回归得到目标边界框的坐标。由于多级尺度特征的融合使提取的目标特征更为有效,模型不仅能检测出特征明显的目标,对于尺寸较小的目标、干扰较多的大的目标也提升了检测率,降低了背景对目标的干扰造成的误检。

2.2 分组融合剔除检测框算法 (PFR-DBB)

目标检测网络最终输出的检测结果中一个目标可能会被多个检测框检出,且对目标定位可能不够准确。经典的非极大值抑制算法 (non-maximum suppression, NMS)^[17] 将所有的检测框按照置信度从小到大排序,保留得分最高的检测框,计算该检测框与其他检测框的 IOU,当 IOU 大于设定的阈值,则删除该检测框。NMS 算法仅去除多余的检测框,没有对检测框位置进行调整,最终剩下的检测框对目标的定位不够准确,且仍然存在多个目标被框在一个检测框中或一个目标被多个检测框包围的情况。

为了同时减少漏检、误检,并使检测框位置更为准确,文中提出了一种新的后处理方法,分组融合剔除检测框算法 (PFR-DBB)。这种方法和 NMS 类方法的思想完全不同,通过分组处理的方式,在去除冗余框的同时对检测框进行位置的微调,使其对目标定位更精准。PFR-DBB 先将检测框分成不同的组,认为每组检测框框住的是同一个目标,这样做是为了防止两个位置较近的物体被框到一个检测框中的情况;然后对每组中的检测框再进行处理,剔除多余检测框,同时调

整检测框的大小和位置,使检测框对目标的定位更为精确。

算法包括两个部分,首先是分组过程,对应于算法步骤中的(1)~(4),其次是在每个分组中去除冗余的检测框,同时微调检测框位置的过程,对应于步骤(5)、(6)。后处理算法过程如下:

(1) 将所有的检测框按照置信度递增排序,得到 m 个检测框 $\{R_1, R_2, \dots, R_m\}$, 放入队列 Q 中, 令 $i = 1$;

(2) 将 R_i 放入组 G_i 中, 并将其从 Q 中移除, 将 $i + 1 \rightarrow j$;

(3) 如果 $\text{IOU}(R_i, R_j) > 0.6$, 则将 R_j 放入组 G_i 中并从 Q 中移除, $j + 1 \rightarrow j$ 。若 $j > m$ 转到(4), 否则转到(3);

(4) $i + 1 \geq i$, 如果 Q 不为空, 转到(2); 若 Q 为空, 分组过程结束, 这时把检测框分成了多个组, 每个待检测目标对应一组检测框;

(5) 将组内的所有检测框按照面积递增顺序加入到一个队列中;

(6) 对于每个组对应的队列, 从第一个检测框开始依次计算它与队列中其他检测框间的 IOU 值, 如果 IOU 值大于 0.7, 则计算重叠区域的中心, 以此为新检测框的中心, 以两个框的长之和的一半为长, 宽之和的一半为宽, 重新构造一个检测框并将该框加入队尾, 删除这两个检测框; 否则和队列中的下一个检测框对比; 重复本步骤直到队列中检测框间的 IOU 均小于 0.7, 得到最终的检测框。

3 实验设置与评价指标

实验使用的是 Ubuntu14.04 系统, 硬件配置为 Intel(R) Xeon(R) CPU E5-2690 2.60 GHz 处理器、256 GB 内存和 4 TB 硬盘, 采用 Caffe 深度学习框架。训练过程中参数设置如下: 初始学习率为 0.01, 迭代次数每增加 10 000 次, 学习率降低 0.1 倍, gamma 为

0.1, 权重衰减为 0.000 5, 动量为 0.9。

评价指标包括 IOU (intersection over union)、AP (average precision)、误检率 FAR (false alarm rate) 和漏检率 MR (missing ratio)。IOU 指检测得到的结果和标注结果重合的程度; AP 是查准率的平均值, 用来衡量检测效果; FAR 计算的是误检为目标的数量的数量占目标总数的比例; MR 计算的是没有被检测出的目标占目标总数的比例。将召回率分成 n 级, 则 IOU、AP、FAR、MR 分别用下式计算:

$$\text{IOU} = \frac{\text{area}(B \cap B_r)}{\text{area}(B \cup B_r)} \quad (1)$$

$$\text{AP} = \frac{1}{n} \sum_{i=1}^n \max_{r \in [i-1/n, i/n]} p(r) \quad (2)$$

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

$$\text{MR} = \frac{\text{FN}}{\text{TP} + \text{FN}} \quad (4)$$

$$\text{FAR} = \frac{\text{TN}}{\text{TP} + \text{TN}} \quad (5)$$

其中, B 是检测得到的检测框, B_r 是 Ground Truth, r 是查全率, TP 是真正例, TN 是真反例, FP 是假正例, FN 假反例。

4 实验结果和分析

4.1 飞机目标检测实验

遥感数据集中的样本相对较少, 为了提高检测精度, 利用迁移学习的思想, 先对模型使用 Pascal VOC 2007 数据集进行预训练。预训练结束后, 分别使用 UCAS-AOD^[18]、RSOD-Dataset^[19] 遥感数据集中的飞机目标样本对模型进行微调训练, 使网络中的权值学习更充分。UCAS-AOD 和 RSOD-Dataset 中的飞机数据集在实验中被划分成训练集和验证集 (70%) 以及测试集 (30%), 具体信息如表 1 所示。

表 1 UCAS-AOD、RSOD-Dataset 飞机数据集

数据集	图像尺寸	图像数量	训练集	测试集	验证集
UCAS-AOD	1 280×659	1 000	489	300	211
RSOD-Dataset	1 044×915	446	218	134	94

图 2 给出了 Faster R-CNN 和 FF-DLM 的一些测试样本的检测结果。由结果可见, 在不同的场景以及光照条件下, FF-DLM 的检测结果与 Faster R-CNN 的检测结果相比目标检出率均有提高。FF-DLM 对分布密集的小目标 (即使背景和目标的光谱特征相似)、不同尺度的目标检出情况较好。图 2 中 (1)~(4), 由于场景中目标和背景特征相似或目标较小等情况, Faster R-CNN 都有漏检的目标, 而 FF-DLM 较好地检

出了这些目标。另外图像中的被漏检绿色的小飞机被检测出, 说明 FF-DLM 的特征融合策略有利于复杂场景不同尺度目标的检出。小飞机和一些背景在形状、颜色等方面和飞机较为相似, 容易将背景误检为目标。图 2 中的 (5)、(6) 显示了 Faster R-CNN 误将背景检测为目标, 而 FF-DLM 减少误检的情况。实验结果表明 FF-DLM 模型可以在小样本数据集上充分学习, 和 Faster R-CNN 比较, 在提高检测率的同时减少了误

检、漏检的情况。

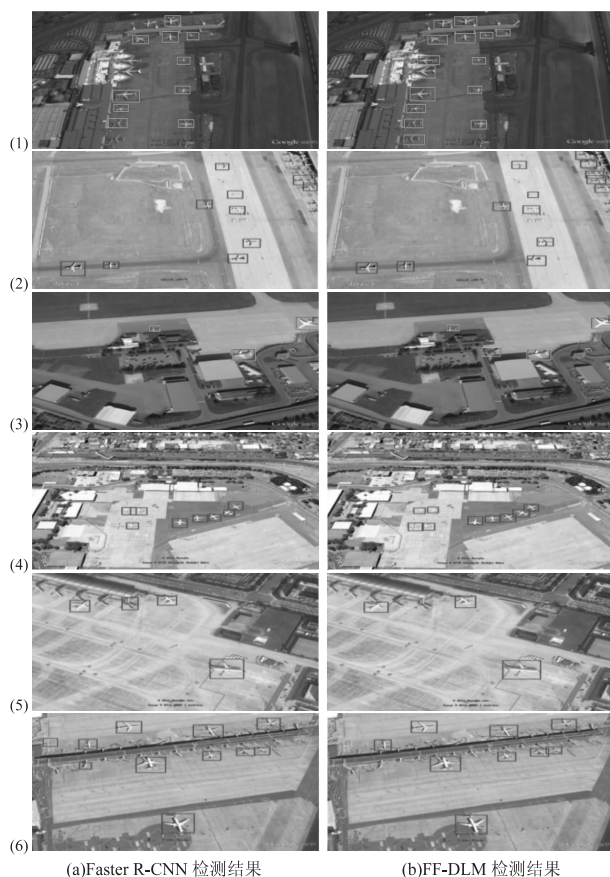


图 2 不同深度目标检测模型飞机检测结果

图 3 显示了 FF-DLM 采用 PFR-DBB 算法和 NMS 算法对检测框进行后处理的结果。图 3 的(1)图中(a)子图是用 NMS 算法后处理的结果,其中有两个小飞机被检测到一个框里面,(2)、(3)图中有的小飞机被多个检测框框住。而采用 PFR-DBB 算法有效去除了多余检测框,且没有将两个目标框到一个检测框中的情况,同时对一些检测框的位置和大小进行了调整,使目标定位更加精确,从而提升了检测精度。

表 2 不同数据集上不同模型的结果比较

模型	UCAS-AOD 检测精度 AP/%	RSOD-Dataset 检测精度 AP/%	每张图测试时间/s
Fast R-CNN	70.1	58.1	4.80
Faster R-CNN	88.5	63.2	0.15
FF-DLM	92.7	70.5	0.14

表 3 展示了飞机数据测试集在不同的模型上的漏检率和误检率。可见,采用 PFR-DBB 后处理算法的

表 3 Faster R-CNN 和 FF-DLM 模型在 UCAS-AOD 测试集上的漏检率和误检率

模型	飞机总个数	检测到的飞机个数	被误检为飞机个数	漏检率 MR/%	误检率 FAR/%
Faster R-CNN	2 138	1 929	59	9.78	2.76
FF-DLM(NMS)	2 138	1 992	42	6.82	1.96
FF-DLM(PFR-DBB)	2 138	2 024	38	5.33	1.78

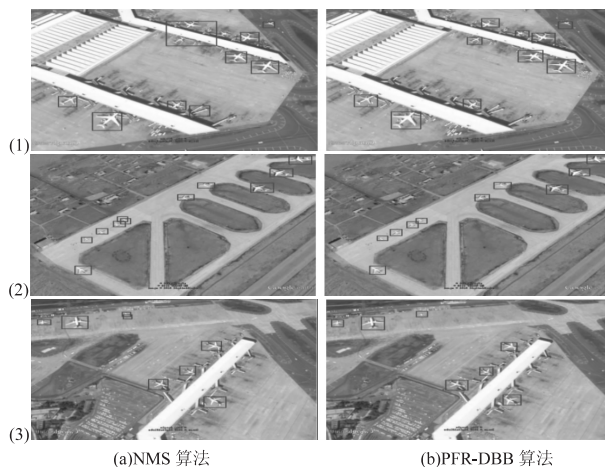


图 3 检测框后处理结果

采用平均准确率 AP 对检测结果进行定量评价, Fast R-CNN、Faster R-CNN 模型和 FF-DLM 的检测结果和所用时间如表 2 所示。其中 Fast R-CNN 和 Faster R-CNN 采用 ZF 网络作为特征提取主干网络,它和 FF-DLM 模型都采用训练 Pascal VOC 数据集得到的参数为初始参数进行训练,测试时 IOU 设置为 0.7。Fast R-CNN 和 Faster R-CNN 未进行特征融合,后处理方法和 FF-DLM 模型不同。从表 2 可见,FF-DLM 在检测精度、检测时间上都具有明显的优势。FF-DLM 使用 RPN 提取候选区域比 Fast R-CNN 使用 Select search 穷举搜索提取候选区域减少了大量的时间,并且实现了端到端的训练使精度提高。和 Faster R-CNN 相比,FF-DLM 进行多级特征的融合,使模型可以更好地提取有效的特征,提高了检测精度,降低了误检率。还因为 FF-DLM 使用了 PFR-DBB 后处理算法,比 Faster R-CNN 使用 NMS 算法对目标的定位更加精确,也为检测精度的提高做出了贡献。FF-DLM 使用小型网络作为主干网络,在样本不是很多的情况下可以使学习更加充分,得到更好的检测结果。

FF-DLM 模型比 Faster R-CNN 和采用 NMS 后处理算法的 FF-DLM 模型的效果要好。由于融合不同层

特征使模型对于遥感图像中小目标的特征能够更好的学习与定位,提高对小目标的检测性能。同时因为对特征学习的更加充分,可以降低误检率。此外,采用所提 PFR-DBB 后处理算法对于提升检测精度也带来了贡献。

4.2 汽车目标检测实验

采用 UCAS-AOD 遥感数据集汽车目标进行实验。图像大小为 $1\ 280 \times 659$,共用了 310 张图像,其中

表4 Faster R-CNN 和 FF-DLM 模型在汽车数据集上的检测精度、漏检率、误检率

方法	检测精度 AP/%	漏检率 MR/%	误检率 FAR/%	每张图所用时间/s
Faster R-CNN	64.8	15.3	5.18	0.16
FF-DLM(NMS)	70.9	10.8	3.26	0.15
FF-DLM(PFR-DBB)	72.7	9.39	3.12	0.15

遥感图像中的汽车目标与飞机目标相比,所占像素更少,汽车的方向、大小、颜色差别大,所处场景多变,检测起来更加困难。从图4可见,FF-DLM 模型对不同场景中的汽车的检测结果比 Faster R-CNN 的检测效果都好。图4中的(1)图显示 FF-DLM 不仅减少了漏检目标个数,而且检测框对目标的定位更加精确;(2)图中 FF-DLM 模型将较小以及颜色与背景相似的汽车均能检测出来,与 Faster R-CNN 相比误检汽车数大幅度减少;(3)图中汽车尺寸更小,背景与汽车的颜

训练集和验证集的图像 268 张,测试集图像 42 张。使用不同的模型进行检测的结果如表4所示。FF-DLM 模型和 Faster R-CNN 的结果进行比较,得到了较低的误检率、漏检率和较高的精度。对测试集在检测精度上 FF-DLM 模型比 Faster R-CNN 提高了 7.9%。同时,后处理算法 PFR-DBB 比 NMS 在提高检测精度以及降低漏检率和误检率上表现更好。

色更为相似,可以明显看出 FF-DLM 的检测结果更好。

Faster R-CNN 仅使用较深层的特征图,对于小目标的检测效果并不好,漏检汽车的数量较多,而且检测框对目标的框取不够精确,且将很多阴影部分错检为汽车;FF-DLM 检出的目标更多,没有将阴影错检为汽车。可见所提模型对于汽车目标同样取得了较好的检测效果,表明模型在少量样本下对于多尺度的目标检测具有优势。

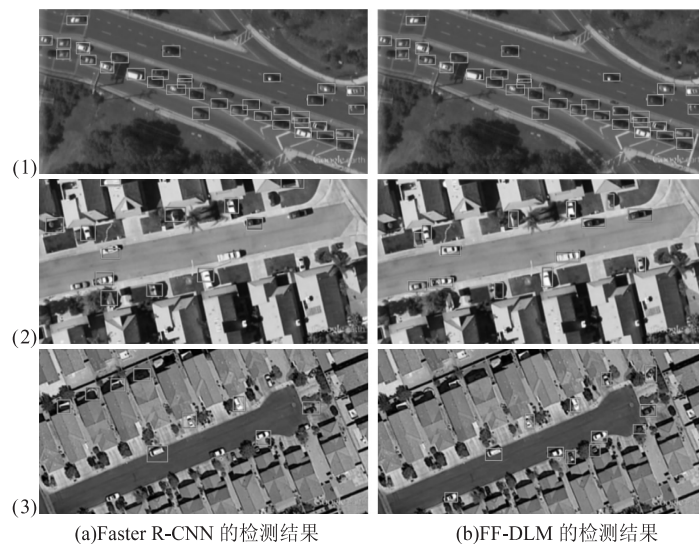


图4 不同检测模型对遥感图像中汽车的检测

5 结束语

提出了一种深度学习遥感图像目标检测模型,针对复杂场景遥感图像中密集的、大小不一、姿态各异的目标,以及训练样本有限的情况,提出融合多级特征的小型卷积神经网络提取有效的目标特征,并提出 PFR-DBB 后处理算法在有效去除冗余检测框的同时微调检测框位置。所提模型在两个遥感图像数据集上获得

了明显优于比较方法的检测结果,在提高检测正确率的同时降低了漏检率和误检率。在更小的汽车目标上同样取得了明显优于对比方法的结果,表明所提模型在样本有限的情况下对不同尺度的遥感目标检测具有优势。

由于所提模型在规模上、采取的特征提取手段上都是简单、适合标记样本不是很多的情况的,如果结合注意力、特征金字塔等更多特征提取阶段的措施,性能

将会进一步提升,当然训练和测试时间相应也会增加。

参考文献:

- [1] 许必霄,宫婧,孙知信.基于卷积神经网络的目标检测模型综述[J].计算机技术与发展,2019,29(12):87-92.
- [2] HE K M,ZHANG X Y,REN S Q,et al. Deep residual learning for image recognition[C]//IEEE conference on computer vision and pattern recognition. Las Vegas:IEEE,2016:770-778.
- [3] KUANG P,MA T,CHEN Z,et al. Image super-resolution with densely connected convolutional networks[J]. Applied Intelligence,2019,49(12):125-136.
- [4] BELL S,ZITNICK C L,BALA K,et al. Inside-outside net: detecting objects in context with skip pooling and recurrent neural networks[C]//IEEE conference on computer vision and pattern recognition. Las Vegas:IEEE,2016:2874-2883.
- [5] LIN T Y,DOLLAR P,GIRSHICK R,et al. Feature pyramid networks for object detection[C]//IEEE conference on computer vision and pattern recognition. Honolulu:IEEE,2017:2117-2125.
- [6] KONG T,SUN F C,YAO A B,et al. RON:reverse connection with objectness prior networks for object detection[C]//IEEE conference on computer vision and pattern recognition. Honolulu:IEEE,2017:5936-5944.
- [7] CAI Z W,VASCONCELOS N. Cascade R-CNN:delving into high quality object detection[C]//IEEE conference on computer vision and pattern recognition. Salt Lake City:IEEE,2018:6154-6162.
- [8] SINGH B,DAVIS L S. An analysis of scale invariance in object detection SNIP[C]//IEEE conference on computer vision and pattern recognition. Salt Lake City:IEEE,2018:3578-3587.
- [9] ZHANG Z S,QIAO S Y,XIE C H,et al. Single-shot object detection with enriched semantics[C]//IEEE conference on computer vision and pattern recognition. Salt Lake City:IEEE,2018:5813-5821.
- [10] 施泽浩,赵启军.基于全卷积网络的目标检测算法[J].计算机技术与发展,2018,28(5):55-58.
- [11] 姚群力,胡显,雷宏.基于多尺度卷积神经网络的遥感目标检测研究[J].光学学报,2019,39(11):1128002.
- [12] GIRSHICK R,DONAHUE J,DARRELL T,et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]//IEEE conference on computer vision and pattern recognition. Columbus:IEEE,2014:580-587.
- [13] GIRSHICK R. Fast R-CNN[C]//IEEE international conference on computer vision. Santiago:IEEE,2015:1440-1448.
- [14] 方梦梁,黄刚.一种光学遥感图像船舶目标检测技术[J].计算机技术与发展,2019,29(8):136-141.
- [15] REN HE K,GIRSHICK R,SUN J. Faster R-CNN:towards real-time object detection with region proposal network[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,2017,39(6):1137-1149.
- [16] 张小娟,汪西莉.完全残差连接与多尺度特征融合遥感图像分割[J].遥感学报,2020,24(9):1120-1133.
- [17] ZHU H,CHEN X,DAI W,et al. Orientation robust object detection in aerial images using deep convolutional neural network[C]//IEEE international conference on image processing. Quebec City:IEEE,2015:3735-3739.
- [18] XIAO Z,LIU Q,TANG G,et al. Elliptic Fourier transformation-based histograms of oriented gradients for rotationally invariant object detection in remote-sensing images[J]. International Journal of Remote Sensing,2015,36(2):618-644.
- [19] ZEILER M D,FERGUS R. Visualizing and understanding convolutional neural networks[C]//European conference on computer vision. [s.l.]:Springer,2014:818-833.