

全卷积网络多层特征融合的飞机快速检测

辛鹏, 许悦雷, 唐红, 马时平, 李帅, 吕超

空军工程大学航空航天工程学院, 陕西 西安 710038

摘要 针对传统飞机检测方法准确率低、虚警率高、速度慢等问题, 提出一种全卷积神经网络多层特征融合的飞机快速检测方法。将浅层和深层的特征经过采样后在同一尺度进行融合, 以缓解由于深层特征图维度过低造成的对小目标表达不足的问题; 修改区域提取时的选框尺寸以适应实际图像中飞机的尺寸特征; 用卷积层代替全连接层以减少网络参数并适应不同大小的输入图像; 复用区域提取网络和检测网络的卷积层和学习的特征参数以保证检测的高效性。仿真结果表明, 与典型的飞机检测方法相比, 所提方法在测试集上取得了更高的准确率和更低的虚警率, 同时大大加快了检测速度。

关键词 机器视觉; 飞机检测; 全卷积神经网络; 浅层和深层特征; 特征融合

中图分类号 TN957.52; TP183

文献标识码 A

doi: 10.3788/AOS201838.0315003

Fast Airplane Detection Based on Multi-Layer Feature Fusion of Fully Convolutional Networks

Xin Peng, Xu Yuelei, Tang Hong, Ma Shiping, Li Shuai, Lü Chao

*Aeronautics and Astronautics Engineering College, Air Force Engineering University,
Xi'an, Shaanxi 710038, China*

Abstract In order to solve the problems of traditional airplane detection methods, such as low accuracy, high false alarm rate, and low speed, we propose a fast airplane detection method based on multi-layer feature fusion in a fully convolutional neural network. Firstly, we sample the shallow and deep features separately and fuse them at the same scale, which can alleviate the problem that the deep features are too sparse to express the small-size objects. Secondly, we redesign the size of the reference boxes to adjust to the practical size of the airplane in the input image. Thirdly, we replace the fully connected layers by convolutional layers to reduce the network parameters and adapt to input images with different sizes. Fourthly, we multiplex the convolutional layers and the learning-feature parameters of the proposal network and the detection network to improve the detection efficiency. The simulation results show that compared with typical airplane detection methods, the proposed method achieves higher accuracy and lower false alarm rate and greatly accelerates the detection speed.

Key words machine vision; airplane detection; fully convolutional neural networks; shallow and deep features; feature fusion

OCIS codes 150.1135; 100.3008; 100.4996; 330.4270

1 引 言

飞机检测可用于战场监测、对地打击等, 被广泛应用于军事领域。待检测的飞机对象可粗略分为空中飞行的飞机和地面停泊的飞机两种, 从图像上看, 二者主要的区别是背景的复杂度。空中的飞机背景

单一, 检测相对容易; 地面的飞机多位于机场、港口等复杂的区域, 检测难度大、应用价值高。因此, 将研究对象定为处于如机场内的复杂背景下的飞机。所提方法也可以应用于空中或其他状态的飞机检测。地面飞机在图像中处于复杂的背景下, 尺寸很小且并不显眼, 还具有多尺寸和多色特征, 大大增加

收稿日期: 2017-09-07; **收到修改稿日期:** 2017-10-15

基金项目: 国家自然科学基金(61372167, 61473309, 61379104)

作者简介: 辛鹏(1993—), 男, 硕士研究生, 主要从事模式识别与人工智能方面的研究。E-mail: wszxxmx@163.com

导师简介: 许悦雷(1975—), 男, 博士, 教授, 博士生导师, 主要从事图像处理、模式识别方面的研究。

E-mail: yueleixu@126.com(通信联系人)

了飞机检测的难度。

飞机检测方法包括候选区域的定位和识别。不同的定位和识别方法以及将二者结合方法的不同会产生不同的检测效果。Liu 等^[1]提出一种基于边缘的由粗到精的方法来识别飞机; Li 等^[2]用显著性和对称检测的方法进行飞机检测; Yildiz 等^[3]将 Gabor 滤波器和支持向量机结合起来进行飞机检测。以上方法都是对如边缘、对称性等的低级特征进行检测, 虚警率较高、计算效率低, 难以满足快速、准确检测的需求。近年来快速发展的深度学习方法能自动学习目标特征, 获得稀疏、表征能力强的深层特征, 相比于传统方法, 其识别能力大大提高。刘峰等^[4]把深度卷积网络和手工提取特征相结合, 提出一种基于全局模型和局部搜索的显著性检测方法。Chen 等^[5]将深度置信网络与目标定位方法相结合, 用于飞机检测, 提高了检测性能, 但其定位方法基于滑动窗口搜索, 耗费时间长, 因此检测速度较慢; Wu 等^[6]采用区域提取方法并结合卷积神经网络进行飞机检测, 进一步提高了检测性能, 同时缩短了检测时间, 但在该方法的区域提取过程中, 每张图像会产生 2000~3000 个候选区域, 若送入同一检测网络, 会造成计算冗余, 降低检测速度。

为提高检测效率, Girshick 等^[7]提出了一种基于空间金字塔池化方法的检测网络, 将检测效率提高了几百倍。为缩短生成候选区域的时间, Ren 等^[8]提出了区域提取网络, 通过共享检测网络卷积层提取的特征生成候选区域, 生成候选区域时几乎不消耗多余时间。但多个卷积和池化层使最终特征图维度很低, 导致网络对小目标的检测效果不理想。Long 等^[9]提出的全卷积网络将网络后几层的特征

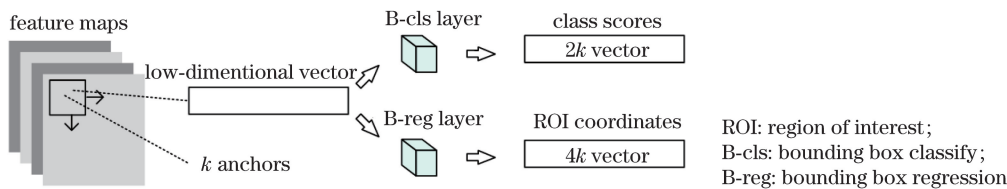


图 1 区域提取网络结构示意图

Fig. 1 Structural schematic of regional proposal network

检测网络是在文献^[11]提出的在空间金字塔池化的基础上, 借鉴自适应池化窗口大小和步长的方法, 使不同尺寸的特征图产生固定尺寸的特征图输出, 以适应不同尺寸的候选区域输入图。为减少计算冗余, 将来自同一输入图的候选区域和对应的特征图进行空间位置的匹配, 不需要对有大量重叠部

图融合在一起, 丰富了特征信息, 得到了逐像素的语义分割图。王鑫等^[10]提出一种多层卷积特征融合的目标尺度自适应稳健跟踪方法, 取得了良好的跟踪成功率和精度。基于该方案, 提出一种卷积神经网络多层特征融合的飞机检测方法, 采用区域提取网络和自适应池化网络作为基本框架, 通过上、下采样等方法融合不同层的特征图代替最后的特征图进行区域生成和识别。用卷积层代替全连接层, 训练和测试过程都使用图形处理单元(GPU), 以提高检测速度。

2 卷积神经网络浅层和深层特征的融合

2.1 经典的目标检测网络及其局限性

经典的目标检测网络一般分为区域提取和目标识别两部分, 目前常用的区域提取方法采用区域提取网络^[8], 它可以通过复用检测网络卷积层提取的特征图进行区域提取, 极大地提高了区域提取的效率。但实验发现, 区域提取网络对小目标的检测效果不佳, 且存在定位不精确的问题, 下面对该问题进行讨论分析。

2.1.1 网络结构和特点

区域提取网络的特点是: 在卷积网络最后一层卷积层输出的特征图上, 设定一个滑动窗口(尺寸一般指定为 3×3), 并以此窗口的中心所对应的输入图像位置生成候选边界框(每个滑窗位置可产生多个不同尺度、不同长宽比的边界框)。每个滑动窗口位置产生一个低维的向量, 将其送入两个同级的全连接层——边界框分类层和边界框回归层, 对区域是否包含目标进行估计, 并对预测的边界框的位置信息进行调整。区域提取网络结构示意图如图 1 所示, k 为特征图每个点对应产生的候选边界框的个数。

分的所有候选区域重复进行卷积操作。其余结构参照区域提取网络, 区别是分类层为多目标分类。检测网络结构示意图如图 2 所示。

2.1.2 网络的局限性

随着卷积神经网络层数的增加, 特征图的维度越来越小, 特征越来越抽象, 语义特征越来越明显,

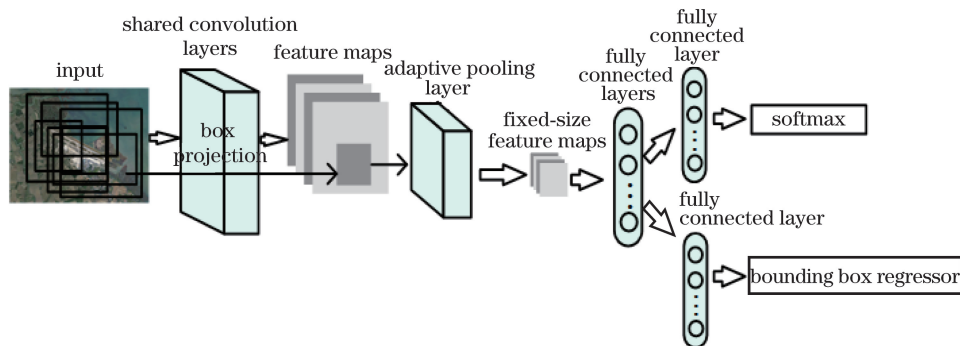


图 2 检测网络结构示意图

Fig. 2 Structural schematic of detection network

换言之,位置信息越来越模糊,因此必然导致定位不精确。虽然边界框回归层能在一定程度上缓解这个问题,但边界框回归层对于不同的输入图像仍然存在差异,有时修正效果很差。此外,多个卷积和池化操作交替运行,使得尺寸较小的目标在特征图上的对应单元尺寸很小,特征表达不足,导致网络对小目标的检测性能较差。

2.2 卷积神经网络浅层和深层特征的融合

基于以上分析,得出两个结论:1)用于区域提取和检测的特征图应该包含丰富的信息,包括语义信息及位置信息;2)特征图应该有合适的尺寸,尺寸太小会导致对特征表达的不足,尺寸太大则会影响计算效率。

卷积网络的深层特征抽象能力强,语义特征强,但位置信息不足;浅层特征能很好地定位目标,但语义特征不明显,特征表达能力不足。因此,一个好的检测网络应该兼顾浅层和深层的特征,以特定方式对浅层和深层特征进行融合,以折中语义信息和位置信息,从而获得合适尺寸的特征图。Kong 等^[12]通过池化、反卷积、加权求和等方式将多层特征图进行融合,获得了更加精确的检测网络。

基于这一思路,对网络的多层特征采用不同的采样方法,将各层化为同一尺度,如采用对浅层特征进行下采样、对深层特征进行上采样、对中层特征不做处理的方法,对各层分别进行一次卷积操作后再进行融合,生成合适大小的特征图,用该特征图代

替最后一层卷积层的输出特征图,用来提取和识别候选区域,卷积神经网络多层融合结构示意图如图 3 所示。

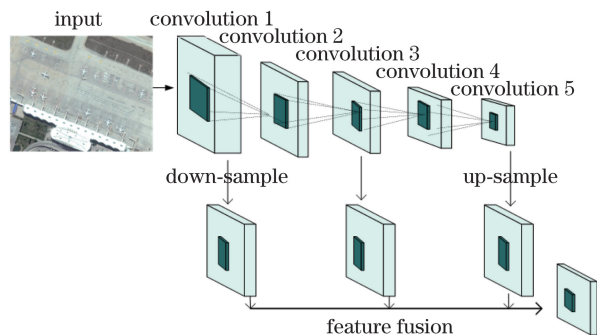


图 3 卷积神经网络多层融合结构示意图

Fig. 3 Structural diagram of multi-layer fusion in convolutional neural network

3 全卷积神经网络多层特征融合的飞机检测

3.1 检测网络的整体框架和技术细节

整个网络的结构参考经典的基于区域的卷积神经网络,为了使网络和飞机检测任务相匹配,需对一些细节进行修改:

1) 考虑到飞机在图像中尺寸很小,形状接近方形,在生成候选区域时选择 3 种较小的尺度和 3 种长宽比,生成的 9 个候选区域边界框的尺寸如表 1 所示。

表 1 9 个候选区域边界框的尺寸

Table 1 Bounding box sizes of nine candidate regions

Set	30×30	60×60	110×110	30×30	60×60	110×110	30×30	60×60	110×110
Proportion	1:1	1:1	1:1	4:5	4:5	4:5	5:4	5:4	5:4
Size	30×30	60×60	110×110	27×33	53×67	98×122	33×27	67×53	122×98

2) 为适应不同尺寸的输入图像、减小网络的规模,将所有的全连接层都替换为相对应的卷积层。为了实现不同层特征图的融合,选择最大池化的方法作为该飞机检测任务的浅层下采样方法,并采用反卷积方法作为深层上采样方法,不同层的特征图在融合前再经过一次卷积,以将特征图进一步调整至同一尺寸,同时增加特征的语义性。

3) 多层融合会降低整个检测系统的速度,为了提速,在用多层融合的特征图生成候选区域前加入一个卷积层,对特征图进行降维,这样能降低用于区域生成的特征图的维度,同时也使得之后的分类过程更加简洁和高效,但是这会以牺牲检测精度为代价。检测过程中,使用的特征图是在该卷积操作之前生成的。测试过程中也用相同的操作来加速。

4) 在训练阶段,区域提取网络会产生数以万计的候选区域,其中有许多区域是冗余的或相似的,为提高候选区域的质量,用非极大值抑制的方法对候选区域集进行筛选。筛选后会保留约 1000 个质量较高的候选区域,选择排位靠前的 200~300 个区域作为训练集。给候选框分配一个二进制的标签(标注是不是飞机),并将其作为训练样本。正样本条件为:与任意人工标定边界框有大于 0.75 的交集并集比(IoU),与某个人工标定的边界框有最高的 IoU(可能小于 0.75)。负样本条件为:与所有人工标定边界框的 IoU 都低于 0.3。一个人工标定的边界框可能对应多个正样本。除正、负样本外,其余的样本对训练不起作用。为了保持训练过程中正、负样本的均衡,随机选取正、负样本各 100 个组成一个小批次。

5) 多任务损失的联合训练可以使多个任务的信息互相补充,从而更好地提取公用特征,因此采用文献[7]中的多任务损失对分类和边界框回归两个任务进行联合训练。定义损失函数为

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

式中 i 为训练时某个候选区域的索引; p_i 为预测区域 i 含有目标的概率,若该区域为正样本,则标签 $p_i^* = 1$,若为负样本,则 $p_i^* = 0$; 分类层和边界框回归层的输出分别为 $\{p_i\}$ 和 $\{t_i\}$; N_{cls} 和 N_{reg} 为归一化因子,一般 N_{cls} 为训练时小批量样本的数量, N_{reg} 为候选区域的数量;为使两个任务的损失权重大致相同,引入平衡权重 λ 。

分类损失 $L_{\text{cls}}(p, u)$ 和边界框回归损失 $L_{\text{reg}}(t_i, t_i^*)$ 分别为

$$L_{\text{cls}}(p, u) = -\lg p_u, \quad (2)$$

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

式中 u 为分类类别的索引; $R(x)$ 参考文献[7]中的定义,可表示为

$$R(x) = \begin{cases} 0.5x^2, & |x| < 1 \\ |x| - 0.5, & \text{else} \end{cases}. \quad (4)$$

6) 检测网络的训练依赖于区域提取网络提取的区域,因此两个网络不能同时训练,而是进行先后训练。为使两个网络复用卷积层提取的特征,参考文献[8]提出的交叉优化训练策略,将训练步骤安排如下:

① 选择预训练的卷积网络,丢弃卷积层后的部分,其余部分作为区域提取和检测网络的初始值,并根据网络结构创建完整的网络模型;

② 训练区域提取网络部分,微调网络权值,用多层特征融合后的特征图产生候选区域,经过非极大值抑制操作舍弃冗余的区域,最后选择每张输入图像分类得分前 200 的区域作为训练样本;

③ 用步骤②产生的训练样本训练检测网络,微调网络权值,注意此时检测网络的初始化方式需与步骤①相同,而不可使用步骤②微调过的网络权值;

④ 将步骤③训练好的检测网络的卷积层权重作为初始化的卷积层并保持其固定不变,重新训练区域提取网络部分,产生新的候选区域,将与步骤②相同的方法挑选的候选区域作为训练样本;

⑤ 保持步骤④的卷积层不变,用新生成的训练样本重新训练检测网络,此时两个网络的卷积层是共用的。这时整个网络训练完毕,并且可以应用于飞机检测任务。

3.2 网络模型的选择和初始化

训练一个任务特异的全新网络需要大量的标注图像数据集,而这种标注数据集的制作十分耗费人力,且标注越多,越容易出现遗漏和偏差等,导致反复检查的工作量大大增加。研究发现,不同类别的图像目标提取的卷积网络的浅层特征是很相似的,可以把在通用数据集上训练的网络权值迁移到特异的检测任务中,这就是常用的迁移学习方法。Zhang 等^[13]将迁移学习方法应用于机场检测中,取得了不错的效果。用大型数据集 ImageNet^[14]上预训练的网络作为模型的初始值,经过较小的标注数据集的再训练以微调网络权值。具体到飞机检测任

务时,为避免过拟合,选择预训练的小型网络 ZF nets (Zeiler and Fergus nets)^[15] 作为卷积层的初始值,其余网络结构用均值为 0、方差为 0.0001 的高斯函数随机进行初始化。

4 仿真和分析

4.1 仿真平台和机场数据库

以 Matlab 2014b 作为仿真平台,用 Visual Studio 2013 编译的 caffe 作为神经网络的框架,电脑配置的 CPU 为 I7-7700 3.6 GHz、16 G 内存、NVIDIA GTX 1060 GPU。选取 200 张飞机图像作为训练数据集,其尺寸不固定,大致范围从 800 × 600 到 1200 × 800,为克服飞机的旋转问题并防止过拟合,对训练数据集的图像进行 90°、180°、270° 旋转和水平翻转,以进行数据扩张增强。

本实验中的飞机检测只关注图像中是否含飞机,属于二分类问题。飞机在图像中所占区域很小,一张图像中约有 10~20 个飞机。训练数据集中约有 3000 个飞机,足以训练一个标准的卷积神经网络。

4.2 训练和测试的仿真

为探索不同层的特征图融合对检测结果的影响,用相同的训练方法完成了多组对比仿真,分别融合 1,2,3 层、3,4,5 层和 1,3,5 层。选取每幅训练图像的前 300 个提取区域,计算查准率和查全率,得到融合不同层特征的仿真结果如表 2 所示。为了更加直观地展示不用融合方法的性能,绘制各自的 P-R (precision-recall) 曲线。融合不同层特征得到的网络 P-R 曲线如图 4 所示。

表 2 融合不同层特征的仿真结果

Table 2 Simulation results of feature maps by fusing different layers

Layer	5	1,2,3	1,3,5	3,4,5
Precision / %	78.7	88.4	94.6	91.8
Recall / %	74.5	85.3	89.5	87.6

由表 2 和图 4 的结果可得,融合不同层的特征会对检测结果产生较大的影响。不融合特征而只用第 5 层的特征进行仿真时,得到的网络对飞机的检测性能不够好,这是因为飞机目标很小,在输入图像中所占区域太小,经过多个卷积和池化后得到的特征也很短,表征能力不足。融合多层特征后能显著提高检测网络对飞机的检测性能,这说明多层特征图融合能加强特征对飞机的表征能力。融合不同的层,如 1,2,3 层、3,4,5 层和 1,3,5 层,得到的结果

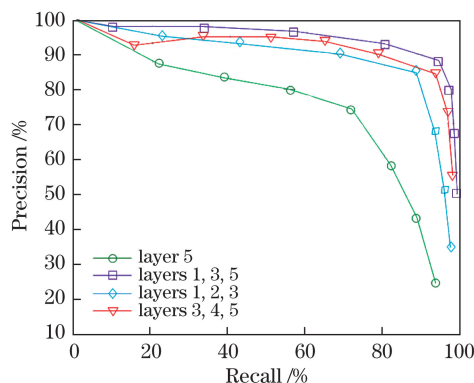


图 4 融合不同层特征得到网络的 P-R 曲线

Fig. 4 P-R curves of networks obtained from feature maps by fusing different layers

差别不大,但融合 1,3,5 层得到的结果最好,这是因为相邻层的相关性较好,融合特征不够丰富,该结果也证明了融合浅层和深层特征确实能对检测结果产生有益影响。

图 5 为所提方法的飞机检测效果图,均采用 1,3,5 层特征图融合得到的网络。由图 5 可以看出,所提方法对飞机这一类小目标的检测效果很好,除一些模糊的和形状比较特别的飞机之外,基本能检测出图像中所有的飞机。



图 5 所提方法的飞机检测效果图

Fig. 5 Airplane detection results of proposed method

4.3 与其他方法的对比仿真

为检验所提方法的性能,另选取 100 张飞机图像作为测试数据集,将两种典型的飞机检测方法作为对比方法,统计最终的检测率、虚警率和平均检测时间 3 项指标,并绘制 3 种方法的 P-R 曲线,如图 6 所示。Location-DBN 指文献[5]中所提出的基于定位和深度置信网络的飞机检测方法;BING-CNN 指文献[6]中提出的基于区域提取方法 BING 和卷积神经网络的飞机检测方法;Fast RCNN 指在文

献[7]的基础上修改的方法,主要是对部分检测网络用飞机数据集进行再训练,得到新的模型;Faster RCNN指在文献[8]的基础上进行修改所得的方法,主要修改区域提取网络的候选边界框尺寸,再对飞机数据集进行再训练,得到新的模型。为复现这几种方法,对网络上下载的相关公开代

码进行修改,对每种方法的结果取多次仿真的平均值。由于测试数据库和复现方法的差异,以及研究者的主观操作的偏差,数据可能存在一定偶然性,为了方便比较,数据均保留整数位或小数点后一位。所提方法与4种典型方法的测试结果对比如表3所示。

表3 所提方法和4种典型方法的测试结果对比

Table 3 Comparison of detection results of proposed method and four typical methods

Method	Location-DBN	BING-CNN	Fast RCNN	Faster RCNN	Proposed method
Detection rate /%	84	86	80	79	93
False alarm rate /%	36	19	22	23	8
Average time /s	>100	6.1	2.9	0.2	0.3

表3中,检测率和虚警率的计算公式分别为

$$R_D = \frac{N_{\text{correct}}}{N_{\text{all}}} \times 100\%, \quad (5)$$

$$R_F = \frac{N_{\text{wrong}}}{N_{\text{all}}^*} \times 100\%, \quad (6)$$

式中 N_{correct} 为检测正确的飞机数目, N_{wrong} 为检测错误的飞机数目, N_{all} 为图像中飞机总数, N_{all}^* 为检测到的飞机的数目。

以上对比仿真中,所提方法采用1,3,5层融合,区域提取阶段选取排名靠前的300个候选区域。需要注意的是,表3中后3列的方法均采用GPU加速,当不用GPU时,检测速度下降了10%。当不用GPU时,表3中后两列方法虽然仍优于前两种对比方法,但速度优势不再明显,而第3列Fast RCNN方法的检测时间甚至会比BING-CNN方法略长。从表3中可以看出,所提方法在检测率、虚警率和检测时间上均超过了前3列的方法,虽然检测时间略慢于Faster RCNN方法,但检测率和虚警率更高,总体而言,所提方法的综合性能是最优的,这得益于卷积神经网络强大的特征表征能力,并且多层特征融合将浅层和深层特征融合在一起,加强了网络对小目标的检测能力。

5 结 论

提出一种将卷积神经网络浅层和深层的特征经过采样后在同一尺度融合的飞机检测方法。将浅层特征位置信息与深层特征语义信息进行融合,以提高特征的表征能力,同时缓解深层特征维度太低对小目标表达不足的问题。采用9个尺度的候选框以适应实际图像中不同尺度的飞机;用卷积层代替全连接层,以减少网络参数并适应不同尺寸的输入图像;保留区域提取网络复用检测网络结构和学习的

特征参数,以保证检测的高效性。所提方法对飞机的高效、实时检测方案进行了有效探索,有很强的理论和实用价值。在下一步工作中,将把所提方法应用到多类目标的检测任务中,并进一步提高网络的检测性能。

参 考 文 献

- [1] Liu G, Sun X, Fu K, *et al.* Aircraft recognition in high-resolution satellite images using coarse-to-fine shape prior[J]. IEEE Geoscience and Remote Sensing Letters, 2013, 10(3): 573-577.
- [2] Li W, Xiang S M, Wang H B, *et al.* Robust airplane detection in satellite images[C]. 18th IEEE International Conference on Image Processing (ICIP), 2011: 12504922.
- [3] Yildiz C, Polat E. Detection of stationary aircrafts from satellite images[C]. IEEE Conference on Signal Processing and Communications Applications (SIU), 2011: 12075218.
- [4] Liu F, Shen T S, Lou S L, *et al.* Deep saliency detection based on global model and local optimization[J]. Acta Optica Sinica, 2017, 37(12): 1215005.
刘峰, 沈同圣, 娄树理, 等. 全局模型和局部优化的深度网络显著性检测[J]. 光学学报, 2017, 37(12): 1215005.
- [5] Chen X Y, Xiang S M, Liu C L, *et al.* Aircraft detection by deep belief nets[C]. IAPR Asian Conference on Pattern Recognition (ACPR), 2013: 14198071.
- [6] Wu H, Zhang H, Zhang J F, *et al.* Fast aircraft detection in satellite images based on convolutional neural networks[C]. IEEE International Conference on Image Processing (ICIP), 2015: 15654338.
- [7] Girshick R. Fast R-CNN[C]. IEEE International Conference on Computer Vision, 2015: 1440-1448.
- [8] Ren S Q, He K M, Girshick R, *et al.* Faster R-

- CNN: Towards real-time object detection with region proposal networks[C]. *Advances in Neural Information Processing Systems*, 2015: 91-99.
- [9] Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation[C]. *IEEE Conference on Computer Vision and Pattern Recognition*, 2015: 3431-3440.
- [10] Wang X, Hou Z Q, Yu W S, *et al.* Target scale adaptive robust tracking based on fusion of multilayer convolutional features[J]. *Acta Optica Sinica*, 2017, 37(11): 1115005.
王鑫, 侯志强, 余旺盛, 等. 基于多层卷积特征融合的目标尺度自适应稳健跟踪[J]. *光学学报*, 2017, 37(11): 1115005.
- [11] He K M, Zhang X Y, Ren S Q, *et al.* Spatial pyramid pooling in deep convolutional networks for visual recognition[C]. *European Conference on Computer Vision*, 2014: 346-361.
- [12] Kong T, Yao A B, Chen Y R, *et al.* Hypernet: Towards accurate region proposal generation and joint object detection[C]. *IEEE Conference on Computer Vision and Pattern Recognition*, 2016: 845-853.
- [13] Zhang P, Niu X, Dou Y, *et al.* Airport detection from remote sensing images using transferable convolutional neural networks[C]. *International Joint Conference on Neural Networks (IJCNN)*, 2016: 16430277.
- [14] Deng J, Dong W, Socher R, *et al.* Imagenet: A large-scale hierarchical image database[C]. *IEEE Conference on Computer Vision and Pattern Recognition*, 2009: 10836047.
- [15] Zeiler M D, Fergus R. Visualizing and understanding convolutional networks[C]. *European Conference on Computer Vision*, 2014: 818-833.