

# 改进区域卷积神经网络的机场检测方法

朱明明, 许悦雷, 马时平, 唐红, 辛鹏, 马红强

空军工程大学研究生院, 陕西 西安 710038

**摘要** 提出了一种结合级联的区域建议网络和检测网络的遥感图像机场检测方法。通过改进区域建议网络, 以获得高质量的机场建议框; 通过改进检测网络的损失函数, 以提高机场检测的准确性; 使用交叉优化策略, 实现了两个网络的卷积层共享, 机场检测时间大幅缩减。结果表明, 所提方法在复杂背景下能准确地检测出不同类型的机场, 检测率高, 虚警率低, 平均运行时间短。

**关键词** 遥感; 遥感图像; 机场检测; 卷积神经网络; 交叉优化

**中图分类号** TP183; TP751.1

**文献标识码** A

**doi:** 10.3788/AOS201838.0728001

## Airport Detection Method with Improved Region-Based Convolutional Neural Network

Zhu Mingming, Xu Yuelei, Ma Shiping, Tang Hong, Xin Peng, Ma Hongqiang

Graduate School, Air Force Engineering University, Xi'an, Shaanxi 710038, China

**Abstract** An airport detection method based on remote sensing images which combines the cascaded regional proposal network with the detection network is proposed. The regional proposal network is improved to get the airport proposal boxes with a high quality, and the loss function of the detection network is improved to increase the accuracy of the airport detection. In addition, the alternating optimization strategy is adopted to share the convolution layers between the two networks, and thus the airport detection time is greatly shortened. The results show that this proposed method can be used to accurately detect different types of airports under complex backgrounds with a high detection rate, a low false-alarm rate and short average processing time.

**Key words** remote sensing; remote sensing image; airport detection; convolutional neural network; alternating optimization

**OCIS codes** 280.4788; 100.2000; 100.4996

## 1 引 言

遥感图像中典型目标检测是当前图像处理领域的研究热点之一<sup>[1]</sup>, 机场作为军事和民用基础设施, 在飞机起降、交通运输、能量补给等方面发挥着重要作用。然而, 机场检测也面临着许多问题和挑战, 比如机场所处背景十分复杂, 机场的形状和大小也不尽相同等。因此, 遥感图像机场检测方法的研究具有十分重要的应用意义。

迄今为止, 国内外学者提出并研究了多种机

场检测方法, 主要可以分为基于边缘和基于区域分割两类方法<sup>[2]</sup>。前者专注于边缘直线特征, 通过检测跑道来实现机场检测。这种方法速度快, 复杂度低, 但易受其他具备长直线特征的非机场目标干扰。后者专注于机场的显著性结构特征, 但由于滑动窗口重叠问题, 该方法效率低。上述方法存在的主要问题表现在: 一是区域选择中采用的滑动窗口策略无针对性, 时间复杂度高且窗口冗余; 二是对于机场多样性的变化, 手工设计的特征稳健性不强。

**收稿日期:** 2018-01-03; **收到修改稿日期:** 2018-01-23

**基金项目:** 国家自然科学基金(61372167, 61379104)、航空科学基金(20175896022)

**作者简介:** 朱明明(1992—), 男, 硕士研究生, 主要从事图像处理和计算机视觉方面的研究。

E-mail: ming\_paper@163.com

**导师简介:** 许悦雷(1975—), 男, 博士, 教授, 硕士生导师, 主要从事图像处理和视觉神经计算方面的研究。

E-mail: 649014294@qq.com

近来,深度学习的兴起引领了工业界和学术界的潮流<sup>[3]</sup>,也为机场检测提供了一种新的思路。Zhang 等<sup>[4]</sup>基于平行线特征,利用卷积神经网络(CNN)的迁移学习能力来识别机场跑道,随后又利用 CNN 的迁移学习能力直接识别机场<sup>[5]</sup>。Xiao 等<sup>[6]</sup>利用 CNN 来提取图像的多尺度深度融合特征,再用支持向量机进行分类。辛鹏等<sup>[7]</sup>借鉴深度 CNN 架构,提出了一种机场快速检测方法。然而,这些方法大多只是利用 CNN 强大的分类识别能力,机场建议框的提取还是基于边缘或区域分割的手工方法,仍然存在局限性。因此,本文基于区域的

CNN 架构,抛弃以往的滑动窗口固定模式,对网络模型进行改进和优化,提出了一种结合级联区域建议网络和检测网络的机场检测方法。

## 2 改进区域 CNN 的机场检测方法

区域 CNN 检测方法通常由两个子任务组成<sup>[8]</sup>,这里将两个子任务进行进一步优化和改进,以实现机场的快速准确检测。所提方法主要由级联的区域建议网络(RPN)<sup>[9]</sup>和检测网络组成,两个网络间共享卷积层,整体结构如图 1 所示,其中 RoI 表示感兴趣区域。

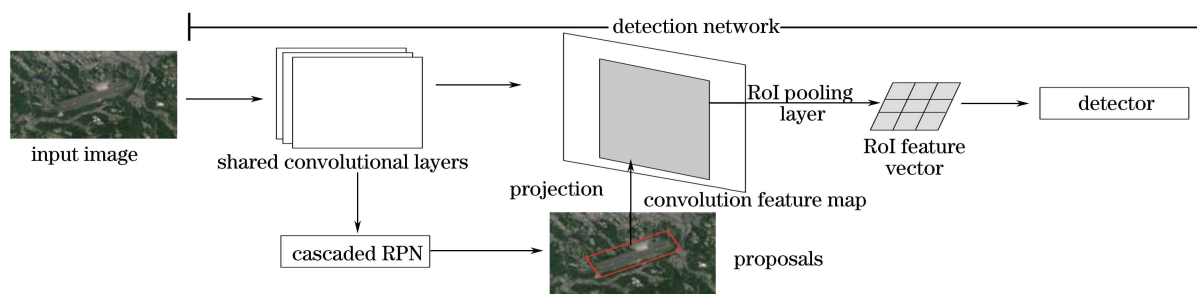


图 1 整体结构示意图

Fig. 1 Schematic of overall architecture

### 2.1 级联的区域建议网络

RPN 将建议框的生成整合到网络中,并与后续的检测网络共享卷积层,大大降低了计算成本且建议框的质量更高,建议框是指可能含有目标的区域

包围框。

#### 2.1.1 区域建议网络

经共享卷积层输出的卷积特征图由 RPN 处理后,输出图 2 所示的矩形建议框集合。

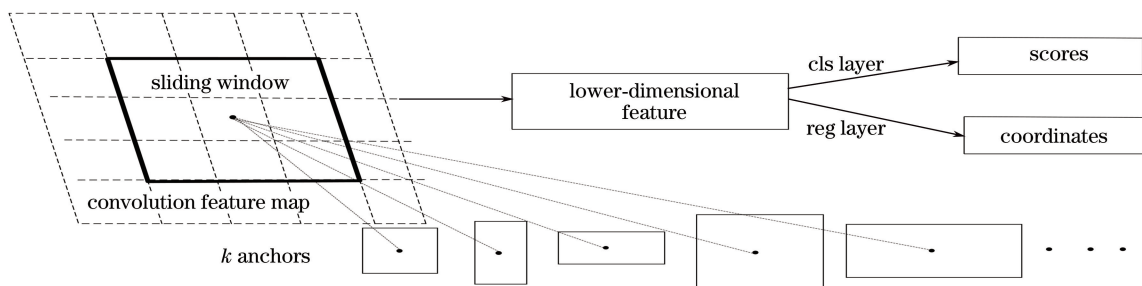


图 2 RPN 结构示意图

Fig. 2 Schematic of RPN structure

RPN 使用一个  $3 \times 3$  的滑动窗口在特征图上进行卷积运算,每个滑动窗口映射到一个低维向量。经过  $1 \times 1$  的卷积核卷积后输入到分类层(cls layer)和回归层(reg layer),同时进行目标分类和建议框的定位回归。特别需要指出的是,这里的分类只是判断建议框中是不是目标,并不是具体的类别。建议框的参数化表示称为 anchor,即每个 anchor 以滑动窗口的中心为中心对应一组尺寸大小和长宽比。对于卷积特征图上的每个位置,这里考虑三种尺寸

大小和三种长宽比的 anchor,以应对遥感图像中的多尺度机场目标。

#### 2.1.2 多任务损失函数

为了训练 RPN,每个 anchor 都被赋予一个二进制标签。正标签被赋予两类 anchors:与某一个人工标定框的交集并集之比(IoU)最高的 anchors;与任意人工标定框的 IoU 大于 0.7 的 anchors。负标签被赋予与所有人工标定框的 IoU 都小于 0.3 的 anchors。

由于 RPN 要完成目标分类和建议框的回归两

个任务,因此遵循多任务损失原则<sup>[10]</sup>,定义 RPN 损失函数为

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

式中  $i$  是一个小批量样本中 anchor 的索引; $p_i$  表示第  $i$  个 anchor 中目标的预测概率; $p_i^*$  表示第  $i$  个 anchor 中目标的人工标定标签,如果 anchor 的真实标签为正,则  $p_i^*=1$ ,如果 anchor 的真实标签为负,则  $p_i^*=0$ ; $t_i$  为建议框的回归值, $t_i^*$  为人工标定框的坐标; $N_{\text{cls}}、N_{\text{reg}}$  为归一化常数,分别为小批量样本数量和 anchor 的数量; $\lambda$  为平衡权重; $L_{\text{cls}}$  和  $L_{\text{reg}}$  分别表示分类和回归损失,定义如下:

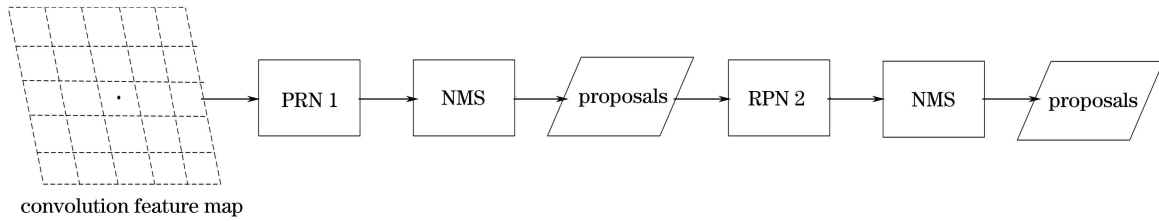


图 3 级联 RPN 结构示意图

Fig. 3 Schematic of cascaded RPN structure

采用两个标准的 RPN 进行前后连接。第一个 RPN 依然利用滑动窗口和 anchor 的对应关系来生成建议框,第二个 RPN 利用输入的建议框与 anchor 在特征图上的对应关系来生成新的建议框。根据建议框的分类得分,在每个 RPN 的后面使用 NMS 进一步减少冗余。NMS 可以通过设置 IoU 阈值来进一步减少建议框的数量,但不会影响最终检测准确率,具体细节可参考文献[13]。所提方法简单,额外增加的计算成本非常小,能够在提升机场建议框质量的基础上,实现机场检测性能的提高。

## 2.2 建议框检测网络

### 2.2.1 感兴趣区域池化层

同一幅图像生成的几千个建议框之间的重叠率很高,如果重复提取建议框特征,将导致网络的计算成本很高。由于建议框对应特征图与图像完整特征图间的固定映射关系,因此,利用感兴趣区域池化层<sup>[10]</sup>,可以从完整特征图上直接得到不同建议框对应的特征矢量,从而实现建议框的特征提取共享,如图 1 所示。

### 2.2.2 改进的损失函数

建议框对应的特征图被输入到全连接层,最终从两个同级输出层输出。其中一个输出为对背景和  $K$  类目标的预测概率集合即  $\mathbf{p}=(p_0,p_1,\dots,p_K)$ ,另

$$L_{\text{cls}}(x,y) = -\text{lb } x_y, \quad (2)$$

$$L_{\text{reg}}(t_i,t_i^*) = \begin{cases} 0.5(t_i-t_i^*)^2, & |t_i-t_i^*| < 1 \\ |t_i-t_i^*|-0.5, & \text{otherwise} \end{cases}. \quad (3)$$

### 2.1.3 级联的区域建议网络

理想的区域建议方法是尽可能生成少的建议框,同时能覆盖图像中的每个目标。但是,RPN 生成的机场建议框中依旧存在大量的背景区域。Yan 等<sup>[11]</sup>通过减少由聚类产生的候选框数量,实现了性能的提高;Kuo 等<sup>[12]</sup>用 CNN 重新排序候选框,提高了候选框的质量;辛鹏等<sup>[7]</sup>引入一个二分类器,筛除掉许多较差的候选框。受此启发,提出图 3 所示的改进的级联 RPN 结构,其中 NMS 表示非极大值抑制。

一个为  $K$  类目标的检测框回归值  $\mathbf{t}^K=(t_x^K,t_y^K,t_w^K,t_h^K)$ ,其中  $x、y$  为检测框的中心坐标, $w$  为宽, $h$  为高。通常采用损失函数  $L$  来联合训练:

$$L(\mathbf{p},\mathbf{u},\mathbf{t}^u,\mathbf{v}) = L_{\text{cls}}(\mathbf{p},\mathbf{u}) + [u \geq 1]L_{\text{loc}}(\mathbf{t}^u,\mathbf{v}), \quad (4)$$

$$[u \geq 1] = \begin{cases} 1, & u \geq 1 \\ 0, & u = 0 \end{cases}, \quad (5)$$

$$L_{\text{loc}}(\mathbf{t}^u,\mathbf{v}) = \sum_{i \in \{x,y,w,h\}} R(t_i^u,v_i), \quad (6)$$

式中  $L_{\text{cls}}(\mathbf{p},\mathbf{u})$  表示分类损失; $L_{\text{loc}}(\mathbf{t}^u,\mathbf{v})$  表示检测框的回归损失; $\mathbf{v}=(v_x,v_y,v_w,v_h)$  为检测框回归目标值; $\mathbf{u}(0 \leq u \leq K)$  为类别预测目标值; $R$  为稳健的损失函数,详见文献[8]。损失函数  $L$  训练建议框时,如果建议框与某个人工标定框的 IoU 大于阈值 0.5,则建议框被赋予该人工标定框的类别  $u$ ,否则  $u=0$  即背景类。

(4)式所示的分类损失简单地将所有检测框分成两类,无区别地处理 IoU 大于阈值的所有检测框,导致机场最终定位准确性不高。实际上,IoU 较大的建议框应该优于 IoU 较小的建议框。机场检测不单单追求高的识别率,机场定位的准确性同样重要,且定位准确性的提升也能够提高最终的机场检测率。因此,这里对分类损失进行改进,将分类损失定义为  $L_{\text{cls}}$  在 IoU 阈值  $g$  下的积分,为了计算简

单,利用矩形法近似计算定积分,即

$$\int_{50}^{100} L_{\text{cls}}(p_g, u_g) dg \approx 50 \sum_g L_{\text{cls}}(p_g, u_g) / n, \quad (7)$$

式中  $u_g$  表示阈值为  $g$  时的类别预测目标值,  $p_g$  表示阈值为  $g$  时的预测概率,  $n$  为定积分区间等分个数。  $L_{\text{cls}}(p_g, u_g)$  计算公式为(2)式,则最终的损失函数可改写为

$$L(\mathbf{p}, \mathbf{u}, \mathbf{t}^u, \mathbf{v}) = \sum_g \{L_{\text{cls}}(p_g, u_g) + [u_g \geq 1]L_{\text{loc}}(\mathbf{t}^u, \mathbf{v})\} / n, \quad (8)$$

式中  $n=4$ , 阈值  $g \in \{50, 60, 70, 80\}$ 。网络模型最终输出的预测概率为不同阈值下  $p_g$  的平均值。

### 3 交叉优化策略

为了解决小样本难以训练甚至过拟合等问题,共享卷积层的权重用预训练网络 VGG16 进行初始化,其余层的权重用均值为 0、标准差为 0.01 的高斯分布函数进行随机初始化。网络的基本学习率为 0.001,动量为 0.9,权重衰减为 0.0005。为了实现级联的 RPN 和检测网络共享卷积层,采用交替优化策略来训练整个网络,具体分为四步。

第一步,训练级联的 RPN 和输出机场建议框集合。利用预训练网络 VGG16 初始化建议框生成网络权重后,进行端到端微调训练,训练结束后输出机场建议框集合。

第二步,训练建议框检测网络。利用 VGG16 网络初始化建议框检测网络权重后,利用第一步生成的机场建议框来训练检测网络,此时两个网络没有共享卷积层。

第三步,再次微调级联的 RPN 和输出机场建议框集合。用第二步检测网络的最终权值初始化级联的 RPN,固定共享卷积层权重,仅微调训练级联的 RPN 独有的层,训练结束后再次输出机场建议框集合。这个过程中两个网络共享卷积层。

第四步,再次微调训练网络模型。保持共享的卷积层和级联的 RPN 权重参数固定,用第三步输出的机场建议框微调检测网络的全连接层,最终得到卷积层共享的机场检测网络。

### 4 实验与结果分析

所提机场检测方法在 1000 pixel  $\times$  600 pixel 遥感图像数据集上进行仿真实验,实验数据均来自 Google Earth 软件。数据集总量为 600 张,其中有 400 张存在机场区域的图像,剩余图像主要为铁路、

公路、桥梁、建筑物等背景场景。随机选取 240 张机场图像和 120 张非机场图像组成训练集,其余图像组成测试集。为了防止过拟合,训练时图像以 0.5 的概率水平翻转。图像测试时 IoU 阈值设定为 0.5,采用检测率(DR)、虚警率(FAR)和平均运行时间作为实验评价指标,具体计算公式参照文献[2]。实验中的所有数据结果均取多次随机实验的平均值。

#### 4.1 参数分析

为了确定(1)式中  $\lambda$  的值,采用控制变量法在测试集上进行了一组对比实验,结果见表 1。

表 1 不同  $\lambda$  值下的机场检测率

Table 1 Airport detection rates under different  $\lambda$

$\lambda$	0.1	1	10	100
DR / %	91.25	93.75	96.25	94.4

实验中采用检测率作为评价指标,结果表明: $\lambda$  取不同的值时,检测率也随之改变;当  $\lambda=10$  时,检测率最高为 96.25%。究其原因,本文  $N_{\text{cls}}$  和  $N_{\text{reg}}$  的取值分别为 256 和 2400,只有当  $\lambda=10$  时,才能保证(1)式中两个任务损失权重几乎相等。此时,级联的 RPN 输出建议框的质量最优,所提机场检测方法效果最好。

#### 4.2 结果分析

为了证明级联 RPN 和改进损失函数的有效性,采用控制变量法进行对比实验,结果见表 2。

表 2 区域建议方法的对比

Table 2 Comparison among regional proposal methods

Method	RPN	Cascaded RPN
DR / %	93.75	96.25
Average processing time / s	0.2	0.2

区域建议方法的对比实验采用检测率和每张测试图像平均运行时间为评价指标。表 2 结果表明:级联的 RPN 在不增加额外时间成本的前提下,能够较为明显地提高约 3% 的检测率,表明级联的 RPN 通过提高建议框的质量可实现机场检测方法性能的提高。

损失函数的实验对比结果如图 4 所示,其中  $g=50$  表示(5)式所示损失函数中 IoU 阈值恒定为 0.5。由图 4 可知,所提方法的检测率始终高于  $g=50$  对应的模型方法,且两者差值随 IoU 阈值的增加而逐渐变大。这说明改进的损失函数使定位更加准确,模型能得到更高的检测率,且随 IoU 阈值的增加,定位优势越来越明显。

图 5 显示了部分机场的检测结果,其中机场背景包含道路、建筑物、河流、山脉等多种类型,第二行

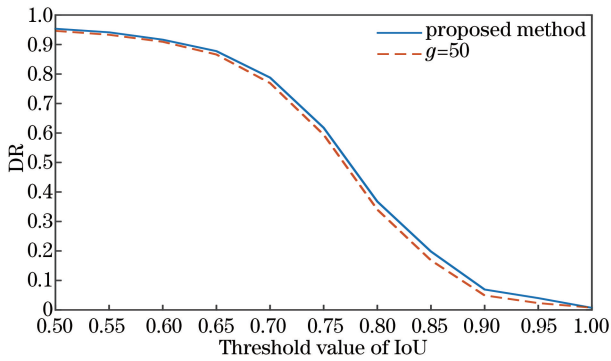


图 4 损失函数的对比

Fig. 4 Comparison among loss functions



图 5 各种机场的检测结果

Fig. 5 Various airport detection results

图像显示了一些形状较为独特的机场,所提方法均能准确地识别出机场。表明所提方法能在复杂背景下准确地检测出不同类型的机场。

### 4.3 方法比较

为了证明所提方法的优越性,选取已有的机场检测方法进行对比实验。根据对比方法的网上公开代码,在本文实验环境下进行仿真实验,将每种方法在测试集上的检测率、虚警率和平均运行时间进行记录和比较,结果见表 3。文献[14]是基于边缘的机场检测方法,文献[15]是基于区域分割的机场检测方法,文献[5]是基于卷积神经网络的机场检测方法,Faster R-CNN 是指基于区域的 CNN 目标检测方法。

表 3 不同机场检测方法的对比

Table 3 Comparison among different airport detection methods

Method	DR /%	FAR /%	Average processing time /s
Ref. [14]	65	22.5	2.46
Ref. [15]	71.88	27.5	>100
Ref. [5]	83.13	35	21.37
Faster R-CNN	90.63	15	0.16
Proposed method	96.25	7.5	0.20

由表 3 结果可知,所提方法的检测率、虚警率和

平均运行时间均明显优于以往的机场检测方法<sup>[5,14-15]</sup>,这是因为所提方法抛弃了以往机场检测方法中采用的利用滑动窗口加手工设计特征的方式,将候选区域生成、特征提取、目标分类和边界回归整合到一个深度卷积神经网络框架中,充分利用 CNN 强大的特征表达能力,实现了端到端的机场检测,并且区域建议网络和检测网络间共享卷积层,从而大幅提高了机场检测的准确性和速度。虽然所提方法的速度略慢于 Faster R-CNN 的,但检测率和虚警率要优于 Faster R-CNN 的,总体来说,所提方法的综合性能是最优的。本文的级联区域建议网络通过筛除大量较差的建议框来提高建议框的质量,从而提升了机场检测的检测率和虚警率性能。虽然这导致检测时间相对增加,但影响较小,且检测时间仍明显优于现有的其他机场检测方法的。损失函数中充分考虑了多个 IoU 阈值的影响,通过定位准确性的提高提升了检测率和虚警率的性能。

## 5 结 论

抛弃了以往的滑动窗口模式,以区域卷积神经网络为基本框架,提出了一种结合级联的区域建议网络和检测网络的遥感图像机场检测方法。级联的区域建议网络通过进一步筛选和优化建议框,提高了建议框的质量;在充分考虑 IoU 训练阈值对定位精度的影响下,对检测网络的损失函数进行了改进,最终提高了机场检测率;两个网络间共享卷积层,大大提高了机场检测的效率。实验结果表明,所提方法在复杂背景下能准确地检测出不同类型的机场,优于现有的其他机场检测方法的,但处理速度距实时处理还有一定差距,这也是下一步研究的方向。

## 参 考 文 献

- [1] Song M Z, Qu H S, Jin G. Weak ship target detection of noisy optical remote sensing image on sea surface [J]. Acta Optica Sinica, 2017, 37 (10): 1011004.  
宋明珠, 曲宏松, 金光. 含噪光学遥感图像海面弱小舰船目标检测 [J]. 光学学报, 2017, 37 (10): 1011004.
- [2] Zhu D, Wang B, Zhang L, et al. Airport target detection in remote sensing images: A new method based on two-way saliency [J]. IEEE Geoscience and Remote Sensing Letters, 2015, 12(5): 1096-1100.
- [3] Qu L, Wang K R, Chen L L, et al. Fast road detection based on RGBD images and convolutional neural network [J]. Acta Optica Sinica, 2017,

- 37(10): 1010003.  
 曲磊, 王康如, 陈利利, 等. 基于 RGBD 图像和卷积神经网络的快速道路检测 [J]. 光学学报, 2017, 37(10): 1010003.
- [4] 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 Network, 2016: 2590-2595.
- [5] Zhang P, Niu X, Dou Y, *et al.* Airport detection on optical satellite images using deep convolutional neural networks [J]. IEEE Geoscience and Remote Sensing Letters, 2017, 14(8): 1183-1187.
- [6] Xiao Z F, Gong Y P, Long Y, *et al.* Airport detection based on a multiscale fusion feature for optical remote sensing images [J]. IEEE Geoscience and Remote Sensing Letters, 2017, 14(9): 1469-1473.
- [7] Xin P, Xu Y L, Ma S P, *et al.* Airport detection combining region proposal networks and adaptive pooling networks [J]. Journal of Xidian University, 2018(3): 121-127.  
 辛鹏, 许悦雷, 马时平, 等. 区域提取网络结合自适应池化网络的机场检测 [J]. 西安电子科技大学学报, 2018(3): 121-127.
- [8] Ye G L, Sun S Y, Gao K J, *et al.* Nighttime pedestrian detection based on faster region convolution neural network [J]. Laser & Optoelectronics Progress, 2017, 54(8): 081003.  
 叶国林, 孙韶媛, 高凯珺, 等. 基于加速区域卷积神经网络的夜间行人检测研究 [J]. 激光与光电子学进展, 2017, 54(8): 081003.
- [9] 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, 2017, 39(6): 1137-1149.
- [10] Girshick R B. Fast R-CNN [C]. International Conference on Computer Vision, 2015: 1440-1448.
- [11] Yan J, Yu Y, Zhu X, *et al.* Object detection by labeling superpixels [C]. Computer Vision and Pattern Recognition, 2015: 5107-5116.
- [12] Kuo W, Hariharan B, Malik J, *et al.* DeepBox: Learning objectness with convolutional network [C]. International Conference On Computer Vision, 2015: 2479-2487.
- [13] Rothe R, Guillaumin M, Van Gool L, *et al.* Non-maximum suppression for object detection by passing messages between windows [C]. Asian Conference on Computer Vision, 2014: 290-306.
- [14] Qu Y, Li C, Zheng N, *et al.* Airport detection base on support vector machine from a single image [C]. International Conference on Signal Processing, 2005: 546-549.
- [15] Tao C, Tan Y, Cai H, *et al.* Airport detection from large IKONOS images using clustered SIFT keypoints and region information [J]. IEEE Geoscience and Remote Sensing Letters, 2011, 8(1): 128-132.