

引用格式:Chai Dong, Xu Suhui, Luo Chang, *et al.* Object Accurate Localization of Remote Sensing Image based on Bayesian Optimization[J]. Remote Sensing Technology and Application, 2020, 35(6): 1377-1385. [柴栋, 许凤晖, 罗畅, 等. 使用贝叶斯优化对遥感影像目标进行精确定位[J]. 遥感技术与应用, 2020, 35(6): 1377-1385.]
doi:10.11873/j.issn.1004-0323.2020.6.1377

使用贝叶斯优化对遥感影像目标进行精确定位

柴 栋¹, 许凤晖^{1,2}, 罗 畅³, 鲁彦辰⁴

(1. 空军研究院, 北京 100085;

2. 中国人民解放军78102部队, 四川 成都 610031;

3. 中国人民解放军78092部队, 四川 成都 610031;

4. 中国人民解放军96946部队, 北京 102202)

摘要:针对大尺寸遥感影像目标检测中检测边框不精确的问题, 提出使用高斯过程贝叶斯优化对遥感影像中的目标进行精确检测与定位。研究分为两个阶段, 第一阶段使用基于边缘信息的EdgeBoxes算法对大尺寸遥感影像进行目标候选区域的选取, 用分类器得到初始检测结果; 为了得到更加准确的边框, 在第二阶段, 基于高斯过程的贝叶斯优化对每个目标的边框进行微调: ①以目标初始边框为基准, 在其周围选取与初始边框相交的边框集合, 并得到一个高斯过程分布; ②使用贝叶斯优化估计出下一个边框, 并将其加入边框集; ③求分类器对所有边框的得分, 得分最高的边框作为下次迭代的基准边框; ④重复若干次贝叶斯优化后得到最终的边框。实验结果表明: EdgeBoxes方法以较少的候选框可以得到较大的召回率, 使用高斯过程的贝叶斯优化可以明显地提高检测边框的精度。

关键词:遥感影像; 目标检测; 目标精确定位; 区域候选框; 高斯过程; 贝叶斯优化

中图分类号: TP751 **文献标志码:** A **文章编号:** 1004-0323(2020)06-1377-09

1 引言

目标检测是遥感影像理解的重要内容^[1]。遥感影像目标检测需要完成两个任务: 一是判断这幅图像中是否含有目标; 二是确定目标在图像中的位置。在目标检测评价体系中, 通常使用交叠率 (Intersection over Union, IOU) 来判定单个目标是否被正确检测, IOU 为检测框与真实目标标记窗口的交叠率。当检测框与真实标记窗口的交叠率大于一定值 (通常为 0.5) 时, 则判定该检测框检测出了目标。然而当目标检测框与目标真实边框的 IOU 较小时, 检测效果如图 1 左所示的飞机目标, 蓝色边框为检测边框。该边框包括了飞机的主体, 但是两边机翼的一部分被排除在外。因此本研究的目标是

对初始检测结果进行进一步优化, 使检测框更精确包围住整个目标, 效果如图 1 右红色边框所示。

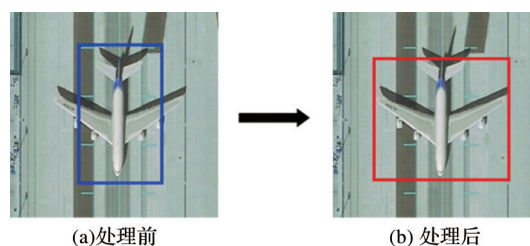


图 1 本研究解决的问题

Fig.1 The issue to be solved in this study

提出了基于高斯过程贝叶斯优化的遥感影像目标精确检测方法: 首先使用 EdgeBoxes 产生目标候选区域, 之后使用线性分类器进行目标初步检测; 以

收稿日期: 2019-12-09; 修订日期: 2020-10-15

基金项目: 军内某科研项目。

作者简介: 柴 栋 (1988—), 男, 河南博爱人, 博士, 助理研究员, 主要从事飞行器总体、人工智能研究。E-mail: chaibaodong@126.com

通讯作者: 许凤晖 (1989—), 女, 河南博爱人, 博士, 工程师, 主要从事遥感影像智能识别研究。E-mail: xu_suhui@163.com

此为基础,使用基于高斯过程的贝叶斯优化估计目标的精确位置。此外还进行了飞机目标的检测实验,并加入主流的区域建议方法进行对比,以验证 EdgeBoxes 方法与基于贝叶斯优化的遥感影像目标定位方法在提高检测精度上的有效性。

2 相关工作

2.1 遥感影像的目标检测研究现状

Cheng 等^[2]将遥感影像目标检测的方法分为 4 类:基于模板的方法,基于特定知识的方法,基于面向对象的方法,基于机器学习的方法。近年来更多的学者利用机器学习的方法来解决遥感影像目标检测问题,这类方法将目标检测问题转化为分类问题,并取得了更好的检测效果。

使用深度学习方法对遥感影像特定目标进行检测可以获得图像更好的特征,因此相对于传统特征^[3-5],如方向梯度直方图特征(Histogram of Oriented Gradient, HOG)、词袋模型特征(Bag of Words, BoW)及深度学习方法可以取得更高的检测精度,然而其主要瓶颈是对正样本数量要求高。因此主流的方法都沿着两个方向进行:①尽可能地提取出更多的训练样本,如使用弱监督方法^[6-7]从含有目标的遥感图像中提取出可靠的图像块进行训练;②使用深度学习方法提取出目标更好的特征。Han 等^[6]使用 BoW 特征与玻尔兹曼机学习目标影像块的深度特征。Long 等^[8]选择了深度卷积神经网络来提取遥感影像特征。Tang 等^[9]则使用了自动编码器(Autoencoder)对图像进行了特征学习。Yang 等^[10]使用全卷积网络实现了飞机目标的检测。以上方法使用了复杂的深度学习方法,提高了检测精度,然而目标检测框并不精确,很少有文献关注如何对检测框进行后续处理。

本研究侧重于目标检测边框的优化,因此并没有选择相对复杂的深度学习框架,而是选择了经典的目标检测框架:基于 HOG 和支持向量机(Support Vector Machine, SVM)的目标检测框架^[11]。该框架分为两个阶段:训练阶段,构造正样本集和负样本集,提取所有样本的 HOG 特征,并将正负样本标记为 1 和 0,将正负样本的特征和标签都输入到 SVM 中进行训练;测试阶段,在测试图像上使用不同大小的滑动窗口来提取大量候选框,对候选框进行 HOG 特征提取,使用之前训练好的 SVM 模型进行预测。预测标签为 1 的候选框再进行非极大值抑

制(Non Maximum Suppression, NMS),形成最终的检测结果。遥感影像尺寸往往较大,例如本文使用的测试图像,尺寸为 $4\,800 \times 2\,589$,即使使用较大的步长如 5,尺寸为 100×100 滑动窗口产生候选框,则候选框数目将大于 4×10^6 。若使用更小的滑动窗口或者更小的步长,候选框数目则会更多,将导致检测效率十分低下^[12]。

2.2 目标检测候选框选取研究现状

为解决滑动窗口效率低下的问题,近年来研究者们提出了各种产生目标候选框的方法。具有代表性的方法有选择性搜索算法 Selective Search^[13]、Bing^[14]、Multibox^[15]以及 EdgeBoxes^[16]等。与普通图像中的目标不同,遥感影像中的目标主要表现为边缘特征,相对于整个遥感影像,目标往往较小,内部纹理信息单一;目标的规模已知,目标视觉效果随着观察角度也是低变化的。

因此本文使用 EdgeBoxes 产生目标候选框,EdgeBoxes 是由 Zitnick 等^[16]提出的基于边缘的产生目标候选框方法,该方法的原理是用包含在候选框中的等高线数表示框中含有对象的可能性。实验证明在遥感影像目标检测中,控制 EdgeBoxes 产生较少的候选框数量,就可以得到较高的目标召回率。

3 遥感影像目标初步检测

目标初步检测框架分为两个阶段,如图 2 所示:①训练阶段(Training stage),负样本(Negative image)为收集到的不含有目标的遥感影像。使用 EdgeBoxes 方法产生候选框,产生大量负样本,将其统一为相同大小正方形,形成负样本块(Negative patches),部分样本块如图中 Negative patches 上方所示。正样本块(Positive patches)为目标图像块;②使用预训练过的 VGG-16 网络^[17](蓝色虚线框)进行训练,将网络的最后一层的输出节点数改为 2,使用正负样本进行网络参数微调。微调后的网络作为接下来的特征提取网络。将待处理图像输入到 VGG-16 网络中,取倒数第二层的 4 096 维输出作为表征图像的特征向量;③使用正样本和初始负样本训练线性 SVM 分类器;④由于非目标数量可能远远大于目标数量,这样训练出来的 SVM 效果并不好,所以使用难样本挖掘(Hard samples mining)方法,从负样本中选取一些有代表性的负样本,使得分类器的训练结果更好。具体做法是用训练好的 SVM 分类器预测剩余的负样本块,如果样

本块的预测结果为正样本,则将其加入负样本库中,当所有负样本进行该步骤之后,对SVM进行第二次训练,得到最终的分类器模型。检测阶段(De-

tection stage)中,对测试图像使用EdgeBoxes产生候选框,使用训练阶段产生的分类器进行目标检测,得到初步目标检测结果。

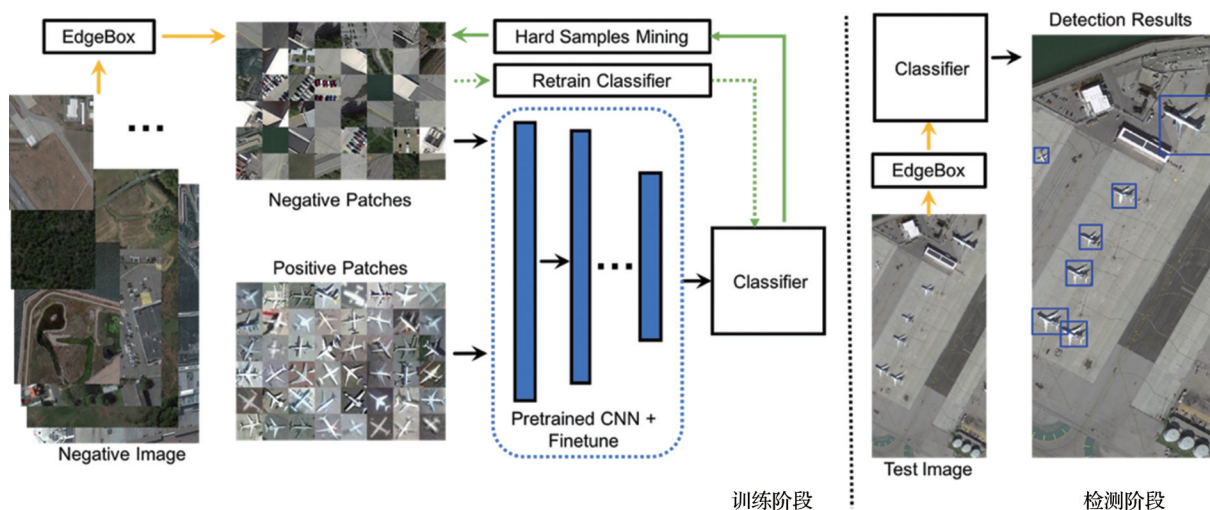


图2 遥感影像目标初步检测框架

Fig.2 The initial remote sensing image object detection framework

4 基于高斯过程贝叶斯优化的目标边框微调

遥感影像目标初步检测存在两个主要问题: ①若EdgeBoxes没有将影像上的目标选中,这就导致该目标不可能被正确检测。这是因为对于大幅遥感影像来说,如尺寸为 $4\,800 \times 2\,589$,目标相对于该幅影像来说尺寸较小,如尺寸为 100×100 ,这就需要设置EdgeBoxes参数让其产生更多的候选框。而候选框的数量过多(通常是数万个),检测效率很低。②若使EdgeBoxes产生适当数量的候选框,可能导致目标被一个不合适的框选中,若这个框与真实框之间的IOU过小,此目标就不能被检测出来。因此本文提出在目标检测步骤之后,使用贝叶斯优化进行边框的微调,提高候选框与真实标记的IOU。

本文提出的基于高斯过程贝叶斯优化的边框微调方法如图3所示:I为目标检测初步检测的结果,为了更清楚地演示,截取1个目标放大,如II所示,可以看出EdgeBoxes方法产生的飞机边框范围相对于目标的真实轮廓较大。以目标初始边框为基准,在之前EdgeBoxes选中的边框中,按照相应规则(将在4.1节中介绍)搜索近邻边框的集合(III中所示的黄色边框)。利用边框集合及各自SVM得分,使用基于高斯过程的贝叶斯优化估计出下一个

边框,并将其加入边框集中,求所有边框中SVM得分,得分最高的边框为下一个基准边框(IV中的蓝色边框),接着从图V开始重复III和IV的步骤,迭代终止后得到最终的边框(VI中蓝色边框)。

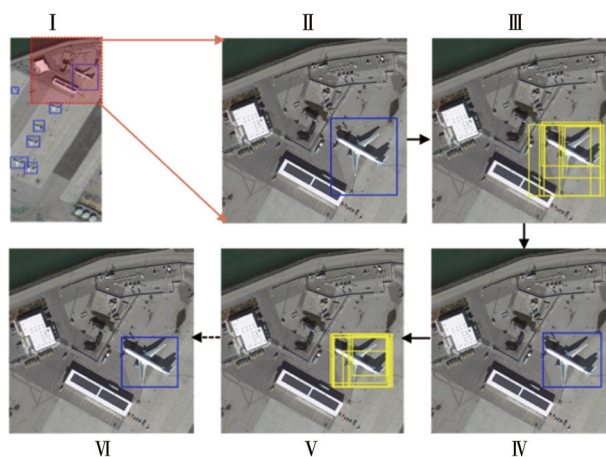


图3 基于高斯过程贝叶斯优化的目标边框微调

Fig.3 Fine-tuning bounding box based on gaussian process and bayesian optimization

4.1 近邻框的选择

图3中II到III或IV到V的步骤为近邻框的选择。将进行检测之前EdgeBoxes产生的边框集合表示 $\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_M\}$,其中M是EdgeBoxes生成候选框的数量, $\mathbf{x}_i = (\alpha_i, \beta_i, w_i, h_i)$, $i = 1, 2, \dots, M$, α_i, β_i 为中心点的坐标, w_i, h_i 为该边框的宽和高。将

当前步骤中的局部最优框(如图3中II和IV中的蓝色边框)设为 $\mathbf{x}_0 = (\alpha_0, \beta_0, w_0, h_0)$ 。将选中的框的集合表示为 \mathbb{X}^* 。对于 \mathbb{X}^* 中任意元素 \mathbf{x}_i ,如果其满足与 \mathbf{x}_0 的交叠率IOU大于一定值(本文实验中取0.3),则将其加入 \mathbb{X}^* 。该过程可以描述为算法1,其中 $\text{calculateIOU}(\mathbf{x}_i, \mathbf{x}_0)$ 为计算边框 \mathbf{x}_0 与 \mathbf{x}_i 之间交叠率的函数, $\mathbb{X}^* \leftarrow \mathbf{x}_i$ 表示将 \mathbf{x}_i 加入集合 \mathbb{X}^* 。

算法1近邻框的选择策略

输入: $\mathbb{X} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_M\}, \mathbf{x}_0 = (\alpha_0, \beta_0, w_0, h_0), \mathbb{X}^* = \emptyset, \text{threshold}$
 输出: \mathbb{X}^*

1. for \mathbf{x}_i in \mathbb{X}
2. $\text{iou} = \text{calculateIOU}(\mathbf{x}_i, \mathbf{x}_0)$
3. if $\text{iou} > \text{threshold}$
4. $\mathbb{X}^* \leftarrow \mathbf{x}_i$
5. end if
6. end for

4.2 高斯过程

贝叶斯优化需要解决两个问题:①为函数 $f(\mathbf{x})$ 选择一个合适的先验分布;②选择一个合适的获得函数,这个函数在后验概率中找到下一个采样点。对于本文的具体问题,假设当前有 N 个观察值,每个样本记为 $\mathbf{x}_i = (\alpha_i, \beta_i, w_i, h_i), i = 1, 2, \dots, N$ 。 $f(\mathbf{x}_i)$ 为目标函数在观察点 \mathbf{x}_i 处的观察值。使用高斯过程用来描述边框坐标与其得分的分布。高斯分布可以由均值函数 m 和协方差函数 k 描述:

$$f(\mathbf{x}) \sim_{\zeta} (m(\mathbf{x}), k(\mathbf{x}, \mathbf{x}')) \quad (1)$$

为方便通常设置 $m(\mathbf{x}) = 0$,协方差函数 k 为基于自动相关性确定(Automatic Relevance Determination, ARD)距离的平方指数协方差函数:

$$k(\mathbf{x}_i, \mathbf{x}_j) = \lambda^2 \exp \left(-\frac{1}{2} (\mathbf{x}_i - \mathbf{x}_j)^T \text{inv}(\mathbf{A}) (\mathbf{x}_i - \mathbf{x}_j) \right) \quad (2)$$

其中: \mathbf{A} 是对角矩阵,对角元素为 $\{\theta_1^2, \theta_2^2, \dots, \theta_D^2\}$, inv 为取逆函数, D 表示输入空间的维度,本文中 \mathbf{x} 有4个坐标,因此 $D = 4$ 。 λ 为信号方差(Signal Variance)参数,因此函数 k 有5个超参数 $\{\theta_1, \theta_2, \theta_3, \theta_4, \lambda\}$ 。从先验分布中进行抽样时,样本值可以通过多元正态分布 $\mathcal{N}(0, \mathbf{K})$ 得到,其中:

$$\mathbf{K} = \begin{bmatrix} k(\mathbf{x}_1, \mathbf{x}_1) & \cdots & k(\mathbf{x}_1, \mathbf{x}_N) \\ \vdots & \ddots & \vdots \\ k(\mathbf{x}_N, \mathbf{x}_1) & \cdots & k(\mathbf{x}_N, \mathbf{x}_N) \end{bmatrix} \quad (3)$$

现有观察到的样本及其观察值 $\mathbb{D}_{1:N} = \{\mathbf{x}_{1:N}, f(\mathbf{x}_{1:N})\}$,接下来要找到下一个点 \mathbf{x}_{N+1} ,其对应的函数值为 $f(\mathbf{x}_{N+1})$ 。因此, $f(\mathbf{x}_{1:N})$ 与 $f(\mathbf{x}_{N+1})$ 满足联合分布:

$$\begin{bmatrix} f(\mathbf{x}_{1:N}) \\ f(\mathbf{x}_{N+1}) \end{bmatrix} \sim \mathcal{N} \left(0, \begin{bmatrix} \mathbf{K} & \mathbf{k} \\ \mathbf{k}^T & k(\mathbf{x}_{N+1}, \mathbf{x}_{N+1}) \end{bmatrix} \right) \quad (4)$$

$$P(f_{N+1} | \mathbf{D}_{1:N}, \mathbf{x}_{N+1}) = \mathcal{N}(\mu_N(\mathbf{x}_{N+1}), \sigma_N^2(\mathbf{x}_{N+1})) \quad (5)$$

$$\mu_N(\mathbf{x}_{N+1}) = \mathbf{k}^T \mathbf{K}^{-1} f(\mathbf{x}_{1:N}) \quad (6)$$

$$\sigma_N^2(\mathbf{x}_{N+1}) = k(\mathbf{x}_{N+1}, \mathbf{x}_{N+1}) - \mathbf{k}^T \mathbf{K}^{-1} \mathbf{k} \quad (7)$$

4.3 获得函数

获得函数是用来引导贝叶斯优化方向的,最大化获得函数就是采样下一个样本点的过程。在 $\text{argmax}_{\mathbf{x}} \mathcal{A}(\mathbf{x} | \mathbf{D})$ 处对 $f(\mathbf{x})$ 进行采样, $\mathcal{A}(\cdot)$ 表示获得函数。本文使用EI(Expected Improvement)获得函数^[18],EI函数能够兼顾全局搜索和局部优化,其形式如下:

$\text{EI}(\mathbf{x}) =$

$$\begin{cases} (\mu(\mathbf{x}) - f(\mathbf{x}^+) - \xi) \Phi(Z) + \sigma(\mathbf{x}) \phi(Z), & \text{if } \sigma(\mathbf{x}) > 0 \\ 0, & \text{if } \sigma(\mathbf{x}) = 0 \end{cases} \quad (8)$$

$$Z = \begin{cases} \frac{\mu(\mathbf{x}) - f(\mathbf{x}^+) - \xi}{\sigma(\mathbf{x})}, & \text{if } \sigma(\mathbf{x}) > 0 \\ 0, & \text{if } \sigma(\mathbf{x}) = 0 \end{cases} \quad (9)$$

其中: $\Phi(\cdot)$ 和 $\phi(\cdot)$ 分别为标准正态分布的分布函数和概率密度函数, $\mathbf{x}^+ = \text{argmax}_{\mathbf{x}_i \in \mathbf{x}_{1:N}} f(\mathbf{x}_i)$, ξ 为参数。

4.4 算法实现

本章使用贝叶斯优化进行边框微调的过程如下,设 $\mathbb{B} = \{\mathbf{bbs}_1, \mathbf{bbs}_2, \dots, \mathbf{bbs}_{\text{NUM}}\}$ 为原始检测框, $\text{search}()$ 表示4.1节中近邻框的选择过程, $\text{SvmScore}()$ 是计算SVM分类器得分的函数, gp.fit 表示高斯过程的参数优化过程, θ 代表高斯过程的超参数。

5 实 验

5.1 数据集介绍

本节进行了飞机目标的检测实验,正样本来自文献[8]中的RSOD数据集。实验选用了其中446幅飞机类别的遥感影像,每张图像都含有若干个飞机目标,该数据集提供了每个目标坐标信息。本实验中使用坐标信息截取了446个图片中共5061个飞机图像块,作为训练集中飞机正样本。如图4所示,图4(a)中箭头左侧为446幅中部分影像,右侧为裁剪后飞机正样本块。在Google地图上截取不含有飞机目标的影像若干张,作为整个训练的负样本集,其中的2幅影像如图4(b)所示。对于测试集,从

算法 2 基于高斯过程贝叶斯优化的边框微调算法

输入: 图像初步检测结果 $\mathbb{B}=\{bbs_1, bbs_2, \dots, bbs_{NUM}\}$, $i=j=0$, 初始化参数 θ , 最大优化次数 numstep。

输出: 最终 $\mathbb{B}=\{bbs_1, bbs_2, \dots, bbs_{NUM}\}$

```

1. for  $bbs_i$  in  $\mathbb{B}$ 
2.    $\mathbb{X}^* = \text{search}(bbs_i)$ 
3.    $f = \emptyset$ 
4.   for  $x$  in  $\mathbb{X}^*$ 
5.      $y = \text{SvmScore}(x)$ 
6.      $f \leftarrow y$ 
7.   end for
8.    $\mathbb{D} = \{\mathbb{X}^*, f\}$ 
9.    $\theta = \text{gp.fit}(\mathbb{D})$ 
10.   $y_{\max} = \max(f)$ 
11.  for  $j$  in numstep
12.     $\hat{x}_{\max} = \text{argmax}(\mathcal{A}(y_{\max}, \theta, \mathbb{D}))$ 
13.     $\mathbb{X}^* \leftarrow \hat{x}_{\max}$ 
14.     $\hat{f}_{\max} = \text{SvmScore}(\hat{x}_{\max})$ 
15.     $f \leftarrow \hat{f}_{\max}$ 
16.     $\mathbb{D} = \{\mathbb{X}^*, f\}$ 
17.     $\theta = \text{gp.fit}(\mathbb{D})$ 
18.    if  $\hat{f}_{\max} > y_{\max}$ 
19.       $y_{\max} = \hat{f}_{\max}$ 
20.       $x_{\max} = \hat{x}_{\max}$ 
21.    end if
22.     $bbs_i = x_{\max}$ 
23.  end for
24. end for

```

Google 地图上截取 SanFrancisco 机场、Denver 机场、LosAngeles 机场区域的影像, 作为飞机目标检测的测试图像, Denver 机场、SanFrancisco 机场各自选中 1 幅示于图 4(c) 中, 表 1 列出了各个样本集的张数、目标数目及尺寸信息。

5.2 对照实验设置

为检验本文算法在不同目标候选框生成方法上的效果, 本文除了使用 EdgeBoxes 方法产生候选

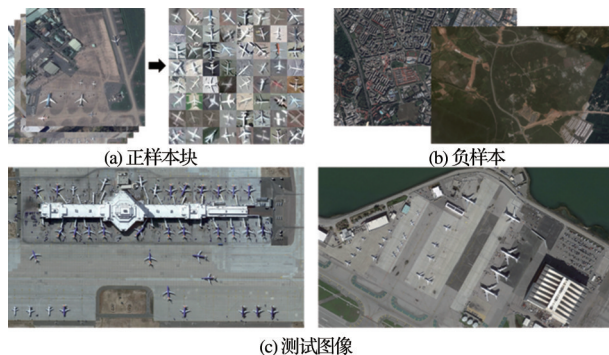


图 4 数据集样例

Fig.4 Samples of the datasets

表 1 数据集信息

Table 1 Details of the datasets

数据集	用途	张数	目标数	尺寸
RSOD-airplane	训练正样本	446	5 061	1 044×915
SanFrancisco 机场	测试图	4	77	4 800×2 589
Denver 机场	测试图	4	103	4 800×2 589
LosAngeles 机场	测试图	2	39	4 800×2 589
其他	训练负样本	30		4 800×2 589

框, 还使用了另外一种比较常用的方法 Selective Search(SS)。对本实验中的所有待检测图像中, 设置 EdgeBoxes 产生的候选框数量都为 20 000。SS 对不同的图像产生的候选框数量不同, 实验中控制其参数, 使其产生两种候选框数量, 分别大于和小于 20 000。因此对于每幅测试图像, 共有 3 个不同的候选框数量。对于每一批候选框, 记录其初始检测结果后, 再进行边框优化。这样可以对比 EdgeBoxes 和 SS 方法在遥感影像检测中的召回率, 也可以使用 IOU 指标来检验本文候选框优化方法在不同生成目标候选框方法上的优化效果。

5.3 实验结果

5.3.1 目标检测结果

首先将区域建议方法固定为 EdgeBoxes, 检验基于贝叶斯优化的边框微调方法的效果, 实验设置的 EdgeBoxes 参数为如表 2 所示: α 控制窗口的大小, β 为非极大抑制 NMS 的阈值, EdgeBoxes 对每个边框计算出来的得分小于 MinScore 时, 该边框不计入结果, 边框面积小于 MinBoxArea 时也将被排除, MaxBoxes 是 EdgeBoxes 生成的边框的最大的数量。分类器为 Liblinear 工具包提供的 SVM 分类和测试函数。对于贝叶斯优化参数, 设置每个边框迭代优化次数 numstep 为 25, 高斯过程中的核函数信号方差参数 λ 设为 1, 获得函数的参数 ξ 设为 0.005。

表 2 EdgeBoxes 参数设置

Table 2 Parameters of the EdgeBoxes algorithm

参数名称	α	β	MinScore	MinBoxArea	MaxBoxes
设定值	0.01	0.75	0.01	2 000	20 000

图 5 上部分为 10 幅测试图像中的 2 幅测试图像, 影像上方的文字为测试图像名称。图中黄色边框为目标的真实边框, 蓝色边框为初始检测边框, 其上方蓝色背景数字为该边框与真实边框的 IOU, 相应地, 红色边框为进行高斯过程贝叶斯优化后的边框, 其下方红色背景数字为该边框与真实边框的 IOU。为了更清楚地显示, 选择其中 12 个目标, 分

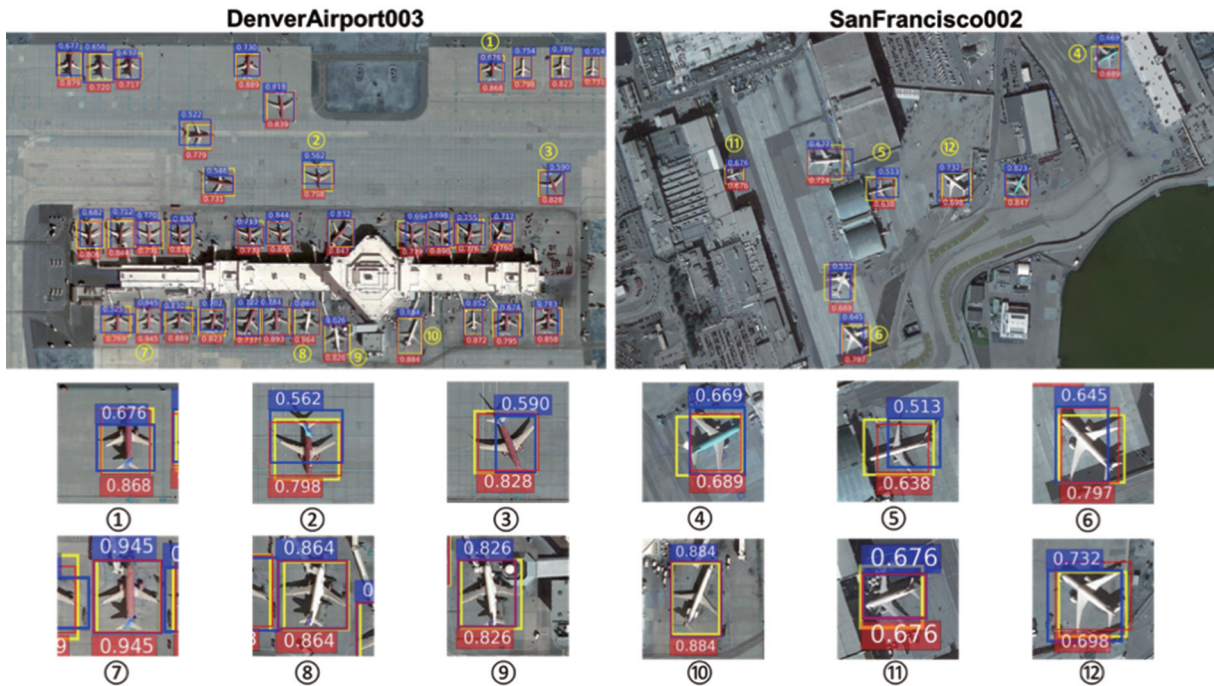


图5 两幅测试图像进行优化前后的检测结果与真实目标框对比

Fig.5 Comparison of bounding box before and after optimization for two test images

别编号为1~12,将其裁剪放大,示于图5下部分:1~6中,原始的蓝框并没有很好地包围目标,优化后的边框提高了与真实边框的IOU;7~11在优化过程中并没有找到更好的解,黄色边框与蓝色边框重合,IOU没有变化;12为在优化后虽然找到了更大的分数,但是新的边框的IOU反而小于蓝色边框,这与目标分类器的性能有关,表明在此目标上更高的SVM得分并不一定对应对合适的边框。综上,在大多数情况下,基于高斯过程的贝叶斯优化方法能够有效提高目标检测的IOU。

为了观察算法的优化过程,以图5中的标号为1、2、5、6的4个目标为对象,每个目标迭代优化25次。记录每次迭代后最佳候选框的SVM得分,及其与真实边框的IOU,绘于图6,4幅曲线的左侧纵轴是边框得分,右侧纵轴为IOU。从图中可以看

出,得分与IOU随着迭代次数增加均呈增大趋势,表明基于高斯过程的贝叶斯优化方法的有效性。

5.3.2 与其他方法对比

召回率表示测试集中的正例有多少被正确检测出^[19]。不同的IOU标准,召回率结果不同,IOU值越大,则检测标准越苛刻。本文统计了IOU分别为0.5与0.7时各方法的召回率,并计算每幅图像上所有的检测边框与对应真实边框的IOU的平均值。

选取其中8幅图像的结果展示于图7,每幅图上方的文字为对应的测试图像的名称。每幅测试图像使用了6种方法:①SS-I:为使用SS方法选取小于20 000个数目的候选框,再使用SVM进行目标初步检测;②SSGP-I是在SS-I的基础上,再进行本章提出的边框优化;③SS-II是使用SS方法选取了大于20 000个数目的候选框,再使用SVM进行目

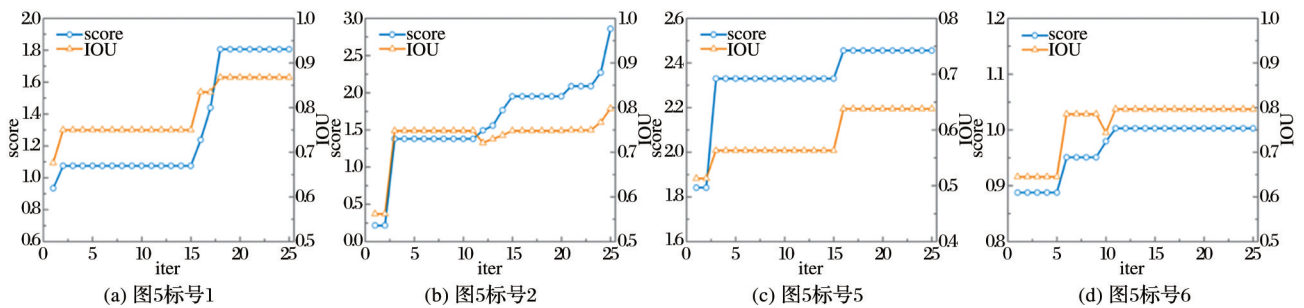


图6 4个目标的得分和IOU与迭代次数的关系曲线

Fig.6 The score and IOU curves with the number of iteration for four objects

标初步检测;④SSGP-II是在SS-II的基础上,再进行本章提出的边框优化;⑤Edge为使用EdgeBoxes方法选取20 000个数目的候选框,再使用SVM进行初步检测;⑥EdgeGP在Edge的基础上,再进行本章提出的边框优化。

对于每一种方法,统计其在IOU标准分别为0.5与0.7下的召回率,以及IOU的平均值。从图7可以观察到:①对比IOU=0.5时召回率,EdgeGP与Edge表现最好,其次是SS-II与SSGP-II,最后是SS-I与SSGP-I。EdgeBoxes方法产生框个数介于SS-I与SS-II之间,但其召回率却优于后两者。这说明当IOU=0.5时,EdgeBoxes表现效果要好于SS;分别对比EdgeGP与Edge,SS-II与SSGP-II和

SS-I与SSGP-I,在大多数情况下,使用高斯过程优化后的召回率得到提升;②IOU=0.7时的召回率相对于IOU=0.5I时均下降。对比此时6种方法的召回率表现,①中结论同样适用;③对比各方法的IOU,SSGP-I,SSGP-II和EdgeGP分别高于SS-I, SS-II和Edge,这说明使用高斯过程优化过程能够有效地提高检测框的精度。表3统计了上述评价指标在全部测试图像上的平均值。表中召回率分子上的219为10幅测试图像的目标总数量,分子为在相应IOU标准下检测到的目标数。表3同样表明EdgeBoxes方法使用了较少数量的候选框,取得了更高的召回率,并且本文提出的高斯过程优化方法能够有效提高边框精度。

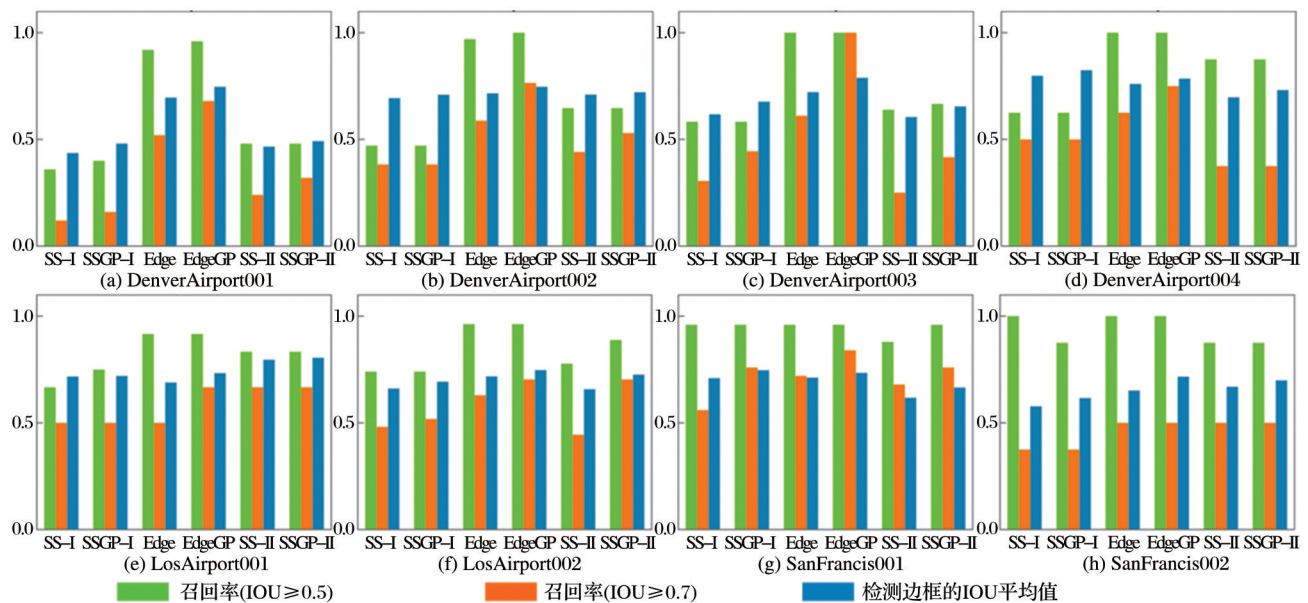


图7 6种方法对8幅测试图像检测结果评价

Fig.7 8 test image detection results evaluation with 6 methods

表3 6种方法对测试图像检测结果评价

Table 3 Test image detection results evaluation with 6 methods

	SS-I	SSGP-I	SS-II	SSGP-II	Edge	EdgeGP
召回率($\text{IOU} \geq 0.5$)	150/219	160/219	144/219	146/219	210/219	214/219
召回率($\text{IOU} \geq 0.7$)	78/219	107/219	80/219	92/219	125/219	161/219
平均值 IOU	0.620 2	0.653 3	0.621 2	0.652 2	0.700 3	0.743 2

6 结 论

针对大尺寸遥感影像目标检测中存在的检测边框不精确的问题,本文提出使用高斯过程贝叶斯优化对遥感影像中的目标进行精确检测定位。首先采用基于边缘信息的EdgeBox算法对大尺寸遥感影像选取候选边框,并用SVM得出初始检测结

果;然后以初始检测结果为基准选取近邻边框,以此为样本得出高斯过程分布;最后利用贝叶斯优化估计出新的边框,迭代优化得到最终的检测结果。实验结果表明,基于边缘信息的EdgeBox算法可以用较少的候选框取得较高的召回率,使用高斯过程的贝叶斯优化明显提高了检测边框的精度。实验中少数候选框的检测精度在优化后并没有得到改

善,这与分类器的性能有关,需要进一步优化分类模型。

参考文献 (References):

- [1] 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.]
- [2] Cheng G, Han J. A Survey on Object Detection in Optical Remote Sensing Images[J]. ISPRS Journal of Photogrammetry & Remote Sensing, 2016, 117:11-28. doi: 10.1109/TGRS.2014.2374218.
- [3] Yang Bisheng, Zong Zeliang, Chen Chi, *et al.* Real Time Approach for Underground Objects Detection from Vehicle-borne Ground Penetrating Radar[J]. Acta Geodaetica et Cartographica Sinica, 2020, 49(7):874-883.[杨必胜, 宗泽亮, 陈驰, 等. 车载探地雷达地下目标实时探测法[J]. 测绘学报, 2020, 49(7):874-882.]
- [4] Yang Rui, Qi Yuan, Su Yang. U-Net Neural Networks and Its Application in High Resolution Satellite Image Classification. Remote Sensing Technology and Application[J], 2020, 35(4): 767-774.[杨瑞, 祁元, 苏阳. 深度学习 U-Net 方法及其在高分辨卫星影像分类中的应用. 遥感技术与应用[J], 2020, 35(4): 767-774.]
- [5] Zheng Xin, Pan Bin, Zhang Jian. Power Tower Detection in Remote Sensing Imagery based on Deformable Network and Transfer Learning[J]. Acta Geodaetica et Cartographica Sinica, 2020, 49(8):1042-1050.[郑鑫, 潘斌, 张健. 可变形网络与迁移学习相结合的电力塔遥感影像目标检测法[J]. 测绘学报, 2020, 49(8):1042-1050.]
- [6] Han J, Zhang D, Cheng G, *et al.* Object Detection in Optical Remote Sensing Images based on Weakly Supervised Learning and High-level Feature Learning[J]. IEEE Transactions on Geoscience and Remote Sensing, 2015, 53(6): 3325-3337. doi: 10.1109/TGRS.2014.2374218.
- [7] Zhang F, Du B, Zhang L, *et al.* Weakly Supervised Learning based on Coupled Convolutional Neural Networks for Aircraft Detection[J]. IEEE Transactions on Geoscience and Remote Sensing, 2016, 54(9): 5553. doi: 10.1109/TGRS.2016.2569141.
- [8] Long Y, Gong Y, Xiao Z, *et al.* Accurate Object Localization in Remote Sensing Images based on Convolutional Neural Networks [J]. IEEE Transactions on Geoscience and Remote Sensing, 2017, 55(5): 2486-2498. doi: 10.1109/TGRS.2016.2645610.
- [9] Tang J, Deng C, Huang G B, *et al.* Compressed-Domain Ship Detection on Spaceborne Optical Image Using Deep Neural Network and Extreme Learning Machine[J]. IEEE Transactions on Geoscience and Remote Sensing, 2014, 53(3): 1174-1185. doi: 10.1109/TGRS.2014.2335751.
- [10] Yang Y, Zhuang Y, Bi F, *et al.* M-FCN: Effective Fully Convolutional Network-based Airplane Detection Framework [J]. IEEE Geoscience and Remote Sensing Letters, 2017, 14(8): 1293-1297. doi: 10.1109/LGRS.2017.2708722.
- [11] Dalal N, Triggs B. Histograms of Oriented Gradients for Human Detection[C]// Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on. IEEE, 2005, 1: 886-893.
- [12] Hosang J, Benenson R, Dollár P, *et al.* What Makes for Effective Detection Proposals? [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2016, 38(4): 814-830. doi: 10.1109/TPAMI.2015.2465908.
- [13] Uijlings J R R, Sande K E A V D, Gevers T, *et al.* Selective Search for Object Recognition[J]. International Journal of Computer Vision, 2013, 104(2): 154-171. doi: 10.1007/s11263-013-0620-5.
- [14] Cheng M M, Zhang Z, Lin W Y, *et al.* BING: Binarized Normed Gradients for Objectness Estimation at 300fps[C]// IEEE Conference on Computer Vision and Pattern Recognition. IEEE Computer Society, 2014: 3286-3293. doi: 10.1109/CVPR.2014.414.
- [15] Erhan D, Szegedy C, Toshev A, *et al.* Scalable Object Detection Using Deep Neural Networks[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2014: 2147-2154. doi: 10.1109/CVPR.2014.276.
- [16] Zitnick C L, Dollár P. Edge Boxes: Locating Object Proposals from edges[C]// European Conference on Computer Vision. Springer, Cham, 2014: 391-405. doi: 10.1007/978-3-319-10602-1_26.
- [17] Simonyan K, Zisserman A. Very Deep Convolutional Networks for Large-scale Image Recognition[J]. Computerence, 2015. doi:arxiv.org/abs/1409.1556.
- [18] Lizotte D J. Practical Bayesian Optimization[M]. Canadian: University of Alberta, 2008.
- [19] Lu Pengjie, Xu Dalu, Ren Fu, *et al.* Auto-detection and Hiding of Sensitive Targets in Emergency Mapping based on Remote Sensing Data[J]. Geomatics and Information Science of Wuhan University, 2020, 45(8):1263-1272.[鲁鹏杰, 许大璐, 任福, 等. 应急遥感制图中敏感目标自动检测与隐藏方法[J]. 武汉大学学报(信息科学版), 2020, 45(8):1263-1272.]

Object Accurate Localization of Remote Sensing Image based on Bayesian Optimization

Chai Dong¹, Xu Suhui^{1,2}, Luo Chang³, Lu Yanchen⁴

(1. *Beijing Aviation Engineering Technology Research Center, Beijing 100076, China;*

2. *The Chinese People's Liberation Army(78102), Chengdu, Sichuan 610031, China;*

3. *The Chinese People's Liberation Army(78092), Chengdu, Sichuan 610031, China;*

4. *The Chinese People's Liberation Army(96946), Beijing 102202, China)*

Abstract: To solve the issues of inaccurate bounding box in large-scale remote sensing image object detection, an accurate object detection and localization approach of remote sensing image based on Bayesian Optimization is proposed. The method consists of two stages: In the first stage, the EdgeBoxes which is based on edges information is adopted to generate object proposals. The classifier is applied to get initial object detection result. To obtain more accurate bounding box, a bayesian optimization based on gaussian process is applied to fine-tune the bounding box around each object in the second stage. Firstly, a set of boxes that intersect the initial bounding box around each initial box is selected to form a gaussian process. Secondly, a new bounding box is estimated through bayesian optimization and added to the set of boxes. Thirdly, the score of each box is calculated by the classifier, and the box with the highest score is set as the base box in the next iteration. At last, the bayesian optimization process is repeated and final bounding boxes is obtained. Experiments demonstrate the EdgeBoxes method can achieve a better recall evaluation with less number of proposals. The bayesian optimization based on gaussian process can significantly improve the localization accuracy of the detection bounding box.

Key words: Remote sensing image; Object detection; Object accurate localization; Region proposal; Gaussian process; Bayesian optimization