

DOI: 10.7652/xjtuxb201806002

一种高分辨率遥感图像视感知目标检测算法

李策^{1,2}, 张亚超¹, 蓝天¹, 杜少毅²

(1. 兰州理工大学电气工程与信息工程学院, 730050, 兰州; 2. 西安交通大学电子与信息工程学院, 710049, 西安)

摘要: 针对大幅面高分辨率光学遥感图像目标检测尚存在着检测精度和效率低的问题,提出了一种高分辨率遥感图像视感知目标检测算法。该算法首先通过显著区域有选择性的引导获取场景中的子区域,将计算资源转移到可能包含目标的区域中,以降低计算复杂度;然后,利用基于单次检测器(YOLO)卷积神经网络目标检测模型获取预选目标;最后,提出目标语义关联抑制对获取的预选目标进行筛选得到有效目标,能够减少虚假目标的干扰,降低虚警率。所提算法在公开 NWPU_VHR-10 数据集上的平均检测精度为 0.865,高于对比算法,在包含更多高分辨率的 LUT_VHR-VOC-2 数据集上,比 YOLO 的检测效果更好。实验结果表明,所提算法提高了大幅面高分辨率遥感图像的目标检测精度。

关键词: 高分辨率遥感图像;目标检测;目标语义关联抑制;卷积神经网络

中图分类号: TP751.1 **文献标志码:** A **文章编号:** 0253-987X(2018)06-0009-08

An Object Detection Algorithm with Visual Perception for High-Resolution Remote Sensing Images

LI Ce^{1,2}, ZHANG Yachao¹, LAN Tian¹, DU Shaoyi²

(1. College of Electrical and Information Engineering, Lanzhou University of Technology, Lanzhou 730050, China;

2. School of Electronic and Information Engineering, Xi'an Jiaotong University, Xi'an 710049, China)

Abstract: An effective object detection approach with visual perception for high-resolution remote sensing images is proposed to address the problem that the accuracy and speed of existing object detection algorithms of remote sensing images are low, especially in large-scale and high-resolution remote sensing images. Firstly, some sub-regions are selected in the scene using a saliency map, and then transfer computing resources to the area that may contain objects to reduce the computational complexity. Then, pre-selected objects are obtained by a fast learning model YOLO (you only look once). An object semantic association suppress is proposed to screen the pre-selected objects for effective objects. It reduces the interference of false objects for reducing the false alarm probability. Experimental results on NWPU_VHR-10 dataset show that the proposed algorithm is the best, and the mean average precision (mAP) is 86.53%. The results of the proposed algorithm are much better than those of YOLO on LUT_VHRVOC-2 dataset which contains more large-scale and high-resolution remote sensing images. It is concluded that the performance of the high-resolution remote sensing image is improved by the

收稿日期: 2017-09-20。 作者简介: 李策(1974—),男,教授;杜少毅(通信作者),男,副教授。 基金项目: 国家自然科学基金资助项目(61365003,61573274);甘肃省基础研究创新群体计划资助项目(1506RJIA031);西安交通大学基本科研业务费综合交叉类资助项目(xjj2017005)。

网络出版时间: 2018-03-15 网络出版地址: <http://kns.cnki.net/kcms/detail/61.1069.T.20180315.2049.006.html>

proposed algorithm.

Keywords: high-resolution remote sensing image; object detection; object semantic correlation suppression; convolutional neural network

大幅面高分辨率的遥感图像具有目标清晰、视野范围广的特点,在为用户提供更多感兴趣目标信息的同时也带来了更复杂的背景信息。因此,如何高效准确的从大幅面高分辨率的遥感图像中检测出所需要的目标,已成为遥感图像应用于实际工程中的关键问题之一。遥感图像目标检测不仅要对图像中目标的类别进行判断,还要给出目标的位置信息。因图像会受不同光照和观察点变化的影响,目标检测任务是一种具有挑战性的视觉任务。大幅面高分辨率遥感图像更加具有挑战性的3个原因:①同一幅图像中目标的尺寸差异较大,且目标的位置和姿态都无法确定;②白天和夜晚的遥感图像颜色空间差异较大;③高分辨率的遥感图像目标背景信息通常较为复杂。

遥感图像目标检测是传统目标检测问题中一个极具应用背景的分支。常见的遥感图像目标检测方法主要有利用模板匹配的方法^[1,2]、背景建模的方法^[3]以及基于浅层学习的方法^[4-6]等。近年来深度学习的蓬勃发展,亦使得引入深度学习方法后的目标检测精度不断提升,主要方法有区域建议卷积神经网络(region convolutional neural network, R-CNN)^[7]、面向实时的区域建议卷积神经网络(faster region convolutional neural network, Faster R-CNN)^[8]、单次检测器(you only look once, YOLO)^[9]、单网多尺度检测器(single shot multibox detector, SSD)^[10]等。同时,诸多学者从深度学习框架出发,亦设计了多种算法拟提升遥感图像中目标检测精度和速度,如旋转不变卷积神经网络(rotation-invariant convolutional neural networks, RICNN)^[11]、转化的神经网络(transferred convolutional neural networks, T-CNN)^[12]、分层模型^[13]等。

模板匹配方法是通过手动设计特征或者从训练集中学习得到模板,计算待检测的图像和模板之间的相似性来找到最佳匹配。基于浅层学习的检测方法主要依据特征提取、特征融合和分类器训练3个步骤来实现遥感图像目标检测。其中,特征提取通常使用方向梯度直方图特征(histogram of oriented gridients, HOG)、词袋特征(bag of words, BoW)、Gabor特征、基于稀疏表达的特征等。分类器主要使用支持向量机、Adaboost、K近邻算法、条件随机

场等机器学习算法。

基于深度学习的遥感图像目标检测有较好的检测结果,根本原因是深度卷积神经网络(convolutional neural networks, CNN)通过大量的有监督训练,能直接从图像像素级获取更具有表达力的特征。尽管CNN的深度结构能够提取更加鲁棒的特征,但是这些在自然图像中提取的特征并不能直接应用到垂直于地面拍摄的遥感图像的目标检测任务中。基于感兴趣区域的CNN目标检测算法存在效率低的问题,例如RICNN算法^[11]是将检测任务分为生成感兴趣区域、CNN训练优化、区域分类3个模块。为保证检测精度,RICNN算法需要多个感兴趣区域,这些感兴趣区域之间存在较大的区域重叠,会被重复性的CNN计算,会造成计算资源的浪费。另外,若所获取的感兴趣区域不好,将直接影响目标检测精度,但高质量感兴趣区域的获取也需要极大的计算资源,因此RICNN算法的检测效率较低,YOLO算法^[9]是将图像只进行一次CNN计算,相对减少了大量重复计算。由于YOLO算法的第一步是将图像归一化到一个固定的尺度,对于大幅面高分辨率的遥感图像,目标范围占图像大小的比例较小,归一化后会导致目标特征丢失,若直接使用原图像,消耗的计算资源亦巨大。为解决大幅面高分辨率遥感图像的目标检测问题,考虑到遥感图像目标检测应用对检测效率的要求,本文提出了一种高分辨率遥感图像视感知目标检测算法。

1 本文所提算法

一般定义满足目标面积与图像面积之比小于0.5%,且图像分辨率大于 800×800 像素即为大幅面高分辨率遥感图像。本文提出了一种高分辨率遥感图像视感知目标检测算法,主要包括3个部分:基于视觉注意机制的子区域提取、子区域目标检测、目标语义关联抑制,算法框架如图1所示。本文所提算法首先利用基于视觉注意机制的子区域提取方法,获取可能包含目标的子区域。对待检测的大幅面高分辨率的图像,进行切片,利用模拟人类的视觉注意机制的显著图,去除部分背景信息,将计算资源转移到包含目标的子区域中,降低了计算复杂度。然后,对子区域进行目标检测,利用基于YOLO学

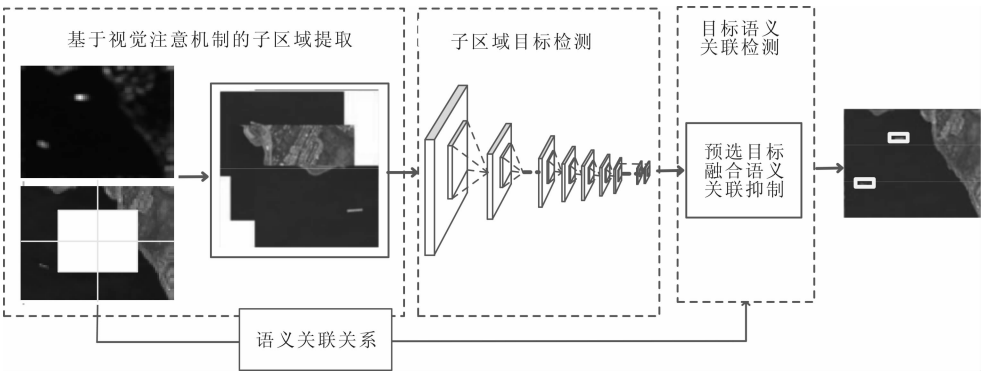


图 1 本文所提算法框架

习网络的目标检测模型得到预选目标。最后,将子区域的预选目标融合,利用所提目标语义关联抑制对预选目标进行筛选,最终获取检测结果。目标语义关联抑制通过学习目标之间的关联性,来抑制图像中不可能出现的目标,可降低虚警率,本文所提算法详述如下。

1.1 基于视觉注意机制的子区域提取

大幅面高分辨率的遥感图像中小目标是非常普遍的,且大多图像中目标是稀疏的。小目标的检测在目标检测任务中是一种困难的任务。像大多基于深度学习的目标检测算法一样,在利用卷积神经网络模型提取特征之前,将图像大小归一化到一定尺寸,图像中的小目标特征会丢失,如图 2 所示。若使用整幅图像作为网络的输入,则高分辨率图像会显著增加计算量。研究发现视觉注意力是将注意力引导到空间中的一个位置。视空间注意力是人类通过在视觉区域内的一个区域的优先级来选择性地处理视觉信息的机制。因此,本文可将这种机制应用到大幅面高分辨率的遥感图像目标检测任务中,有选择性的获取子区域,使得该区域内的信息得到进一步处理,来解决因图像归一化目标特征丢失严重的问题,具体步骤如下。

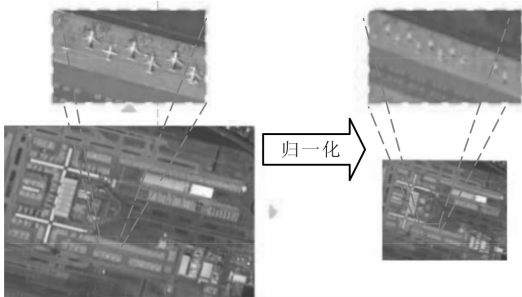
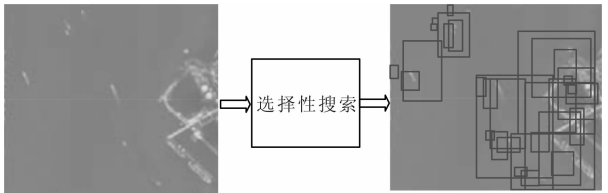


图 2 大幅面图像归一化目标特征丢失示例

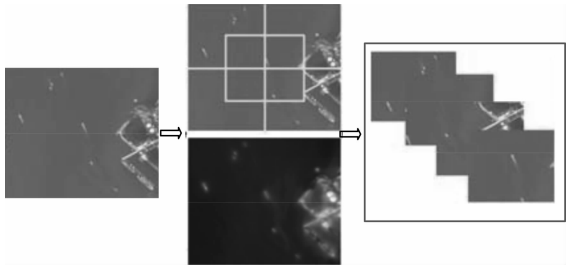
(1)计算视觉注意机制的显著图。视觉注意机制的显著图计算方法,为不失一般性,本文选用的是

Itti 等人提出的基于视觉注意机制的显著性检测算法^[14]。选择该算法的原因是本文并不关心目标的精确边界,只需要找到视觉注意的显著区域,给出大致范围。

(2)子区域提取。首先,对遥感图像进行 5 切片操作,判断每个切片是否包含视觉注意机制的显著区域,将包含显著区域的切片保留,舍弃不包含显著区域的切片,以将计算资源转移到含有显著区域的切片上。另外,对于目标稀疏的图像可以有效去除背景信息,以提高检测效率。不同于 R-CNN 模型^[7]中利用选择性搜索算法^[15]提取的感兴趣区域,期望感兴趣区域最小包围于目标物体,两者差异如图 3 所示。本文的方法是找到最可能包含目标的子区域,该区域可能会包含一个或者多个目标,且不过分依赖于区域的选择。



(a)R-CNN 模型^[7]中感兴趣区域提取



(b)本文方法子区域提取

图 3 本文方法与文献[7]方法感兴趣区域提取对比

1.2 子区域目标检测

遥感图像目标检测在军事安防、地面监测等应

用中要求较高的检测效率。YOLO 算法^[9]是一种具有快速性特点的深度学习目标检测模型。由于遥感图像中目标尺度差别较大,直接使用 YOLO 模型检测效果并不好。本文使用网络模型,其结构如图 4 所示,包含有 22 个卷积层,用来提取目标物体的特征,5 个最大池化层。网络低层特征包含更清晰的轮廓信息,高层信息更能表达目标的语义。图 4 中 16 层和 24 层之间的连接是为了能够更好的对小目标进行准确定位。将 16 层的输出信息与 24 层输出信息结合作为下一层的输入,兼顾大目标和小目标检测。

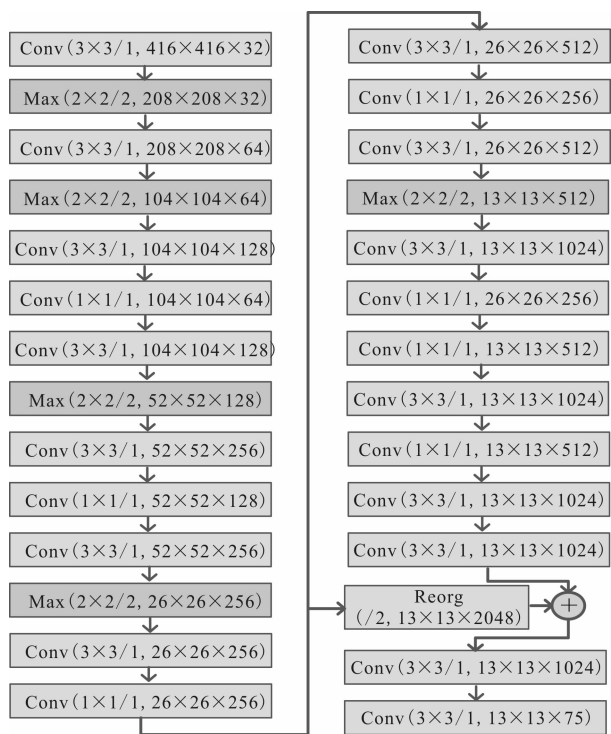


图 4 本文网络结构模型图

在训练过程中本文使用的损失函数可以分为 3 个部分:坐标误差、IoU(Intersection over Union)误差和分类误差,其损失函数可以表示为

$$L = L_1 + L_2 + L_3 \quad (1)$$

式中: L 为本文使用的损失函数; L_1 、 L_2 、 L_3 分别是坐标误差、IoU 误差和分类误差。因此,在检测阶段得到包含有位置信息和类别信息的特征图。

坐标误差 L_1 表示为

$$L_1 = \lambda_{\text{coord}} \sum_{i=0}^{S^2} \sum_{j=0}^B I_{ij}^{\text{obj}} [(x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2] + \lambda_{\text{coord}} \sum_{i=0}^{S^2} \sum_{j=0}^B I_{ij}^{\text{obj}} [(\sqrt{w_i} - \sqrt{\hat{w}_i})^2] + (\sqrt{h_i} - \sqrt{\hat{h}_i})^2 \quad (2)$$

式中: x_i 、 y_i 、 w_i 、 h_i 和 $\hat{x}_i$ 、 $\hat{y}_i$ 、 $\hat{w}_i$ 、 $\hat{h}_i$ 分别是预测框和真实标注的坐标值与尺度; $\lambda_{\text{coord}}=0.5$; i 为网格数; j 为目标框数; I_{ij}^{obj} 表示目标在第 i 个网格上的第 j 个预测框上存在目标;利用 w 、 h 的平方根来减少大、小目标的 w 、 h 对 L_1 影响的差异。

IoU 误差中,IoU 是指选择区域与标注区域交集与并集的比值。 I_{ij}^{noobj} 表示目标在第 i 个网络上第 j 个预测框不存在目标, $\lambda_{\text{noobj}}=0.5$, C_i 是当前区域与第 i 个网格的 IoU 的值, $\hat{C}_i$ 是落在第 i 个网格内真实标注图与第 i 个格的 IoU 的值,则 IoU 误差可表示为

$$L_2 = \sum_{i=0}^{S^2} \sum_{j=0}^B I_{ij}^{\text{obj}} (C_i - \hat{C}_i)^2 + \lambda_{\text{noobj}} \sum_{i=0}^{S^2} \sum_{j=0}^B I_{ij}^{\text{noobj}} (C_i - \hat{C}_i)^2 \quad (3)$$

分类误差 L_3 是指所有目标在网格上的概率的差异, $p_i(c)$ 表示的某个类别在第 i 个网格上的概率值。分类误差表示为

$$L_3 = \sum_{i=0}^{S^2} I_{ij}^{\text{obj}} \sum_{c \in C} (p_i(c) - p_i(\hat{c}))^2 \quad (4)$$

式中: c 是 C 的元素,表示具体类别, C 是所有类别的集合。使用平方差误差是因为其易于优化,另外为了缓解过拟合,针对遥感图像特殊性,在随机裁剪、缩放、旋转的基础之上增加了增强对比度和加入高斯白噪声两种数据增广方法。

在检测阶段,图像经过网络模型得到特征图,在特征图上的每个位置采用不同长宽比和不同大小的预测框进行目标属性和位置判别。预测框是通过对训练集中真实标注目标区域进行聚类选取,选取聚类中心的维度为预测框的维度。

聚类类别的增加会得到较好的检测结果,在权衡计算复杂度和检测精度后选取 5 类,如图 5 所示,在 13×13 像素的特征图上进行预测,每个网格上取 5 个预测框,预测坐标值和属性值,得到目标的预选区域。

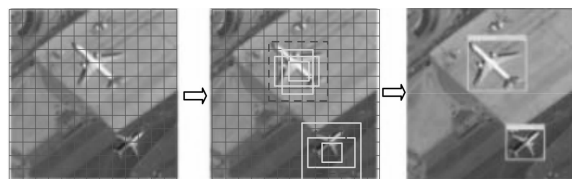


图 5 飞机类目标预测过程示意图(不同网格上选取 5 个候选框,然后再对每个候选框进行属性和位置回归)

1.3 目标语义关联抑制

子区域的合并会造成检测窗口的冗余,子区域

检测的窗口本身是存在冗余的,导致主目标可能会被重复框选,影响检测结果。为了能够选取一个较好的目标区域,非极大值抑制算法(non-maximum suppression, NMS)是解决该问题的常用方法。本文在实验中发现,目标检测出来的目标框是存在虚警的,例如在飞机场中给出了舰船目标,利用目标语义关联抑制去除了舰船的虚警,该过程如图 6 所示。

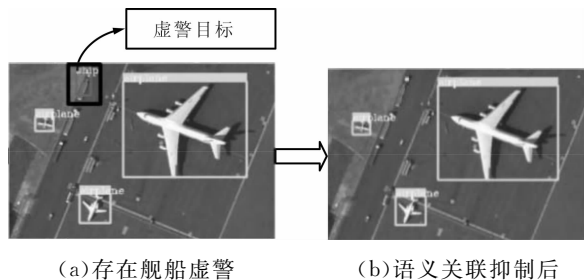


图 6 目标语义关联抑制检测过程

遥感图像是对真实场景的拍摄,而在机场中舰船的出现并不符合客观事实,原因在于这种方法忽略了目标高层语义之间的关联关系。在自然场景中一些目标是关联存在的,例如,船舶与港口、棒球场与篮球场等,而船舶与棒球场,船舶与田径场等几乎不可能出现在同一场景中。为此本文提出了一种目标语义关联抑制方法,该方法通过对样本的统计学习得到目标语义的关联关系 R_{ij} ,然后 R_{ij} 对图像中的预选目标进行选择性地抑制,结合非极大值抑制算法得到最终的检测结果,达到降低目标检测虚警率的目的。本文定义了两个目标之间的关联关系 R_{ij} ,即统计数据集中包含两个目标图像的交集与并集之比

$$R_{ij} = \frac{\text{num}(I_i \cap I_j)}{\text{num}(I_i \cup I_j)}, \quad i, j \in M, I_{i,j} \subseteq S \quad (5)$$

式中: I_i 是包含目标 i 的图像的集合; $\text{num}(\cdot)$ 是统计个数的函数; M 是所有目标类别的集合; S 为训练数据集。显然,当 $i=j$ 时, R_{ij} 的值为 1。目标语义关联抑制方法的具体步骤如下。

(1) 选取主目标。子区域在基于 YOLO 的学习网络模型进行目标检测,得到多个目标预测框,一幅图像中总有些目标是容易被检测到的,表现为预测框具有较高的属性概率值,选取所有预测框中属性概率最大者作为选取的主目标。

(2) 抑制非关联目标。基于 R_{ij} 的计算的基础之上,查询主目标和一幅图像中可能其他目标之间语义关联关系,本文选取 $t=0.1$ 为语义关联阈值,大于等于 t 的认为是有关联的目标,反之是没有关联的目标,通过设定一个抑制比 f ,抑制比乘以目标的

属性概率来抑制不具有关联关系的目标。抑制后的目标属性概率表示为

$$C' = \begin{cases} C, & R_{ij} \geq t \\ Cf, & R_{ij} < t \end{cases} \quad (6)$$

式中: C 、 C' 分别是抑制前、后的属性概率值。

(3) 非极大值抑制。NMS 算法是通过设定固定 IoU 阈值,选取满足 IoU 阈值的所有预测框的目标属性的概率值,选择概率值最大预测框,作为该目标的检测结果,对于不满足 IoU 阈值的预测框就直接保留下来,作为检测结果。

对于区域检测结果合并后,经过步骤 1 和步骤 2,抑制与主目标不关联目标预选框,再利用步骤 3 非极大值抑制去除重复的目标框,得到最终的检测结果。

2 实验结果与分析

为了验证本文算法的有效性,在公开数据集 NWPU_VHR-10 上与 BoW^[5]、FDDL^[6]、局部融合检测器(collection of part detectors, COPD)^[16]、T-CNN^[12]、RICNN^[11] 算法进行定性和定量的实验,验证了所提算法对于遥感图像目标检测是有效的。此外,为了验证本文算法针对大幅面高分辨率的遥感图像改进的有效性,本文利用包含大幅面高分辨率的遥感图像目标检测数据集 LUT_VHROVOC-2 与原 YOLO 模型进行对比实验,实验结果表明本文算法对大幅面高分辨率的遥感图像目标检测有有效提升。本文测试平台 CPU 为 Intel Core i7 6700, GPU 为英伟达 GTX 1070 8 GB 显存,使用 Ubuntu x64 操作系统,实验中设定 f 值为 0.6, IoU 阈值为 0.65。网络初始学习率设定为 0.003,在训练迭代次数达到 100 和 25 000 时下降均降低 10 倍,使用的激活函数为 Leaky。目标检测单目标和整体的性能用检测精度和平均精度来评价。

2.1 NWPU_VHR-10 数据集检测结果对比分析

NWPU_VHR-10^[16] 数据集是一个公开的卫星图像目标检测数据集,数据集包含以下 10 个类别目标:飞机、舰船、储油罐、棒球场、网球场、篮球场、田径场、港口、桥梁和车辆。

2.1.1 主观结果分析 采用本文算法在 NWPU_VHR-10 数据集上对 12 幅图像进行主观检测,检测结果如图 7 所示。从图 7 可以看出,对目标较大的田径场,或者目标较小的储油罐,本文算法都能够检测到目标,并且给出的框能够合理包围于目标物体,检测结果与预检测目标相符。

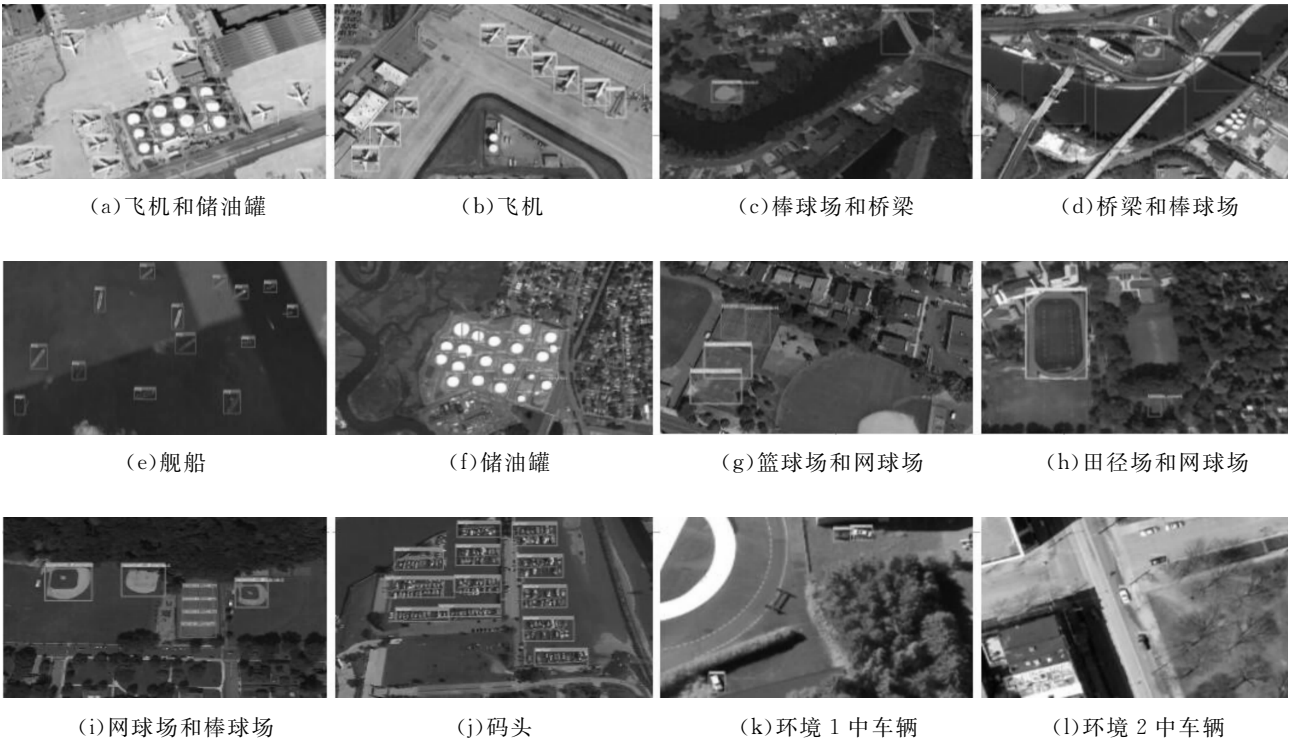


图 7 本文算法对 NWPU_VHR-10 数据集 中 12 幅图像的主观检测结果示例

2.1.2 客观结果对比 表 1 是 6 种检测方法在 NWPU_VHR-10 数据集上的检测精度对比结果, 对比了 3 种传统目标检测算法: BoW 特征检测算法^[5]、基于稀疏编码特征(FDDL)的检测算法^[6]、

COPD 模型检测算法^[16]。另外, 对比了两种基于 CNN 的算法 T-CNN^[12]和 RICNN^[11]。从客观数据平均精度值的对比可以看出, 本文算法的检测效果明显好于传统的检测模型。

表 1 采用 6 种算法在 NWPU_VHR-10 数据集上的检测精度对比

图像	检测精度					
	BoW 算法 ^[5]	FDDL 算法 ^[6]	COPD 算法 ^[16]	T-CNN 算法 ^[12]	RICNN 算法 ^[11]	OURS 算法
飞机	0.250	0.292	0.623	0.661	0.883	0.974
舰船	0.585	0.376	0.689	0.569	0.773	0.864
储油罐	0.632	0.770	0.637	0.843	0.853	0.894
棒球场	0.090	0.258	0.833	0.816	0.881	0.975
网球场	0.047	0.028	0.321	0.350	0.408	0.760
篮球场	0.032	0.037	0.363	0.459	0.584	0.726
田径场	0.078	0.201	0.853	0.800	0.867	0.981
码头	0.530	0.254	0.553	0.620	0.686	0.886
桥梁	0.122	0.215	0.148	0.423	0.615	0.948
车辆	0.091	0.045	0.440	0.429	0.711	0.645
平均精度	0.246	0.248	0.546	0.597	0.726	0.865

由表 1 可见, 在兼顾检测速度与精度前提下, 本文算法对车辆目标的检测精度值略低于 RICNN 算法的 0.06, 但是要高于 T-CNN 算法的 0.22, 对其余目标的检测精度均高于对比算法, 尤其是网球场、篮球场、桥梁目标的检测精度值均有约 0.3 的提升,

说明本文算法具有良好的检测能力。
对于一幅图像而言, 从计算复杂度上来看, 由于 RICNN 算法^[11]采用了选择性搜索的算法获取了约 2 000 个感兴趣区域, 这种选择性搜索算法^[15]本身是消耗较多的计算资源的, 并且选取的这 2 000 个

感兴趣区域都要经过归一化送入深度神经网络中计算。相当于原图像经过了约 2 000 次深度神经网络计算,而本文是将图像进行 5 切片操作,并对切块进行了筛选,每一幅图使用了小于 5 个子区域图像经过了深度神经网络计算。由于深度神经网络的计算量是巨大的,本文算法显著降低了利用深度神经网络计算的次数,能够减少计算量。

2.2 在 LUT_VHRVOC-2 数据集上的检测结果

NWPU_VHR-10 数据集中大幅面高分辨率的图像约占 28%,为证明本文算法在大幅面高分辨率遥感图像上有较好的检测效果,本文研究团队多名专业人员通过大量标注,创建了大幅面高分辨率图像约占 46%的 LUT_VHRVOC-2 数据集,包含飞机和舰船两个类别目标、图像分辨率从 200×200 像素到8 000×8 000 像素不等的 3 039 幅遥感图像。遥感图像数据一部分是由北京航天宏图信息技术股份有限公司提供(该公司是中国计算机视觉大会 2017 遥感图像目标检测竞赛承办方),另一部分是通过百度地图采集。

2.2.1 主观结果对比 从图 8 和图 9 中的对比可以看出:对于高分辨率大幅面的遥感图像,本文所提

算法对于目标的检测效果要好于 YOLO^[9] 算法;无论是飞机还是舰船,YOLO 算法给出的标记框存在多个不准确标记框,同时还存在目标的漏检,本文所提算法能够检测到 YOLO 算法中未检测到和检测不准的目标,给出的目标区域框更能够最小包围于目标物体。对比结果表明,本文算法在大幅面高分辨率的遥感图像目标检测任务上,效果好于 YOLO 算法。

2.2.2 客观结果对比 在 LUT_VHRVOC-2 数据集上进一步验证了本文算法,表 2 是本文算法与 YOLO 算法在 LUT_VHRVOC-2 数据集上检测精度值的对比结果。

表 2 2 种算法在 LUT_VHRVOC-2 数据集上的检测精度对比

检测目标	检测精度	
	YOLO 算法 ^[9]	本文算法
飞机	0.903	0.928
舰船	0.566	0.704

从表 2 可见,本文算法在 LUT_VHRVOC-2 数据集上有较高的检测精度。相比 YOLO 算法,本文所提算法在大幅面高分辨率的遥感图像舰船目标检测任务上提高较大,因为舰船目标在图像中尺寸更小,本文所提算法有效改善了小目标的检测精度。图 10 给出了查全率曲线的对比图,可以看出本文所提算法在保证较高检测精度的同时,有较高的召回率。

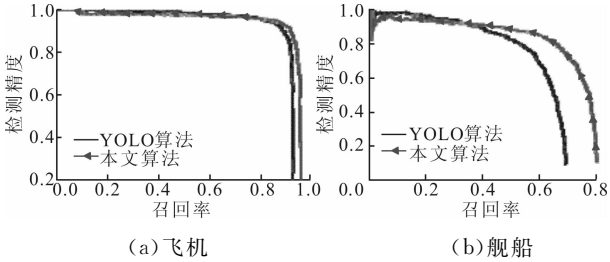
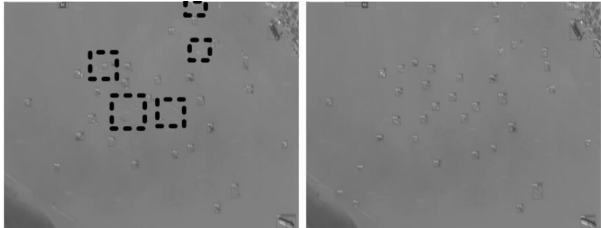


图 10 LUT_VHRVOC-2 数据集上 2 种算法的查全率曲线

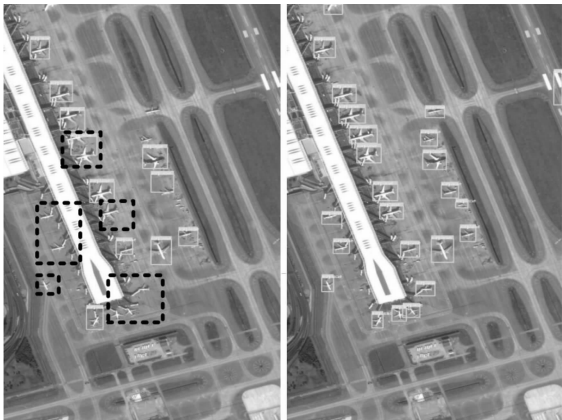
3 结 论

本文在分析现有算法对于大幅面高分辨率遥感图像目标检测局限性的基础上,结合大幅面高分辨率遥感图像的特点,本文提出了一种高分辨率遥感图像视觉感知目标检测算法。所提算法利用视觉注意力机制和目标语义关联关系,结合 YOLO 算法快速性的特点,能够实现大幅面高分辨率的遥感图像目标检测。对比实验取得了较好的检测结果。在未来的工作中将借鉴深监督学习思想改进目标检测模



(a)YOLO 算法 (b)本文算法

图 8 LUT_VHRVOC-2 数据集中舰船检测主观结果对比 (图像虚线框的目标是未检测到目标)



(a)YOLO 算法 (b)本文算法

图 9 LUT_VHRVOC-2 数据集中飞机检测主观结果对比 (图像虚线框的目标是未检测到目标)

型,另外,可将视觉注意机制的子区域提取加入到并行运算中,以提高检测效率。

致谢 感谢西北工业大学韩军伟教授提供的 NW-PU_VHR-10 数据集,以及北京航天宏图信息技术股份有限公司提供的遥感数据。

参考文献:

- [1] FELZENSZWALB P, MCALLESTER D, RAMANAN D. A discriminatively trained, multiscale, deformable part model [C]//Proceedings of the IEEE Computer Vision and Pattern Recognition. Piscataway, NJ, USA: IEEE, 2008: 1-8.
- [2] AHMADI S, ZOEJ M J V, EBADI H, et al. Automatic urban building boundary extraction from high resolution aerial images using an innovative model of active contours [J]. International Journal of Applied Earth Observation and Geoinformation, 2010, 12(3): 150-157.
- [3] 李亚超,周瑞雨,全英汇,等. 采用自适应背景窗的舰船目标检测算法 [J]. 西安交通大学学报, 2013, 47(6): 25-30.
LI Yachao, ZHOU Ruiyu, CHUAN Yinghui, et al. A ship detection algorithm of adaptive background window [J]. Journal of Xi'an Jiaotong University, 2013, 47(6): 25-30.
- [4] 杜春,孙即祥,李智勇,等. 光学遥感舰船目标识别方法 [J]. 中国图象图形学报, 2012, 17(4): 589-595.
DU Chun, SUN Jixiang, LI Zhiyong, et al. Ship target recognition method for optical remote sensing [J]. Journal of Image and Graphics, 2012, 17(4): 589-595.
- [5] XU Sheng, FANG Tao, LI Deren, et al. Object classification of aerial images with bag-of-visual words [J]. IEEE Geoscience and Remote Sensing Letters, 2010, 7(2): 366-370.
- [6] CHEN Yi, NASRABADI N M, TRAN T D. Sparse representation for target detection in hyperspectral imagery [J]. IEEE Selected Topics in Signal Processing, 2011, 5(3): 629-640.
- [7] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation [C]// Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Piscataway, NJ, USA: IEEE, 2014: 580-587.
- [8] REN Shaoqing, HE Kaiming, GIRSHICK R, et al. Faster R-CNN: towards real-time object detection with region proposal networks [J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2017, 39(6): 1137-1149.
- [9] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection [C]//Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Piscataway, NJ, USA: IEEE, 2016: 779-788.
- [10] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single shot multibox detector [C]//Proceedings of the 14th European Conference on Computer Vision. Berlin, Germany: Springer Verlag, 2016: 21-37.
- [11] CHENG Gong, ZHOU Peicheng, HAN Junwei. Learning rotation-invariant convolutional neural networks for object detection in VHR optical remote sensing images [J]. IEEE Geoscience and Remote Sensing, 2016, 54(12): 7405-7415.
- [12] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks [J]. Communications of the ACM, 2012, 60(6): 84-90.
- [13] 高常鑫,桑农. 基于深度学习的高分辨率遥感影像目标检测 [J]. 测绘通报, 2014(s1): 108-111.
GAO Changxin, SANG Nong. Object detection of high resolution remote sensing image based on the deep study [J]. Bulletin of Surveying and Mapping, 2014(s1): 108-111.
- [14] ITTI L, KOCH C, NIEBUR E. A model of saliency-based visual attention for rapid scene analysis [J]. IEEE Pattern Analysis and Machine Intelligence, 1998, 20(11): 1254-1259.
- [15] UIJLINGS J R R, SANDE K E A, GEVERS T, et al. Selective search for object recognition [J]. International Journal of Computer Vision, 2013, 104(2): 154-171.
- [16] CHENG Gong, HAN Junwei, ZHOU Peicheng, 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: 119-132.

(编辑 刘杨 荆树蓉)