

基于深度学习的高分辨率遥感影像飞机掩体检测方法

史姝姝^{1,2}, 陈永强^{1*}, 王樱洁¹, 王春乐¹

¹中国科学院空天信息创新研究院, 北京 100190;

²中国科学院大学电子电气与通信工程学院, 北京 100049

摘要 飞机掩体是关键的飞机防护工事, 利用遥感影像实现飞机掩体的快速准确检测有重要意义。为探究遥感影像飞机掩体检测方法, 收集了 60 个包含飞机掩体的机场信息及 Google Earth 影像, 构建了一个飞机掩体高分辨率遥感影像数据集。对比 Faster R-CNN、SSD、RetinaNet、YOLOv3 和 YOLOX 等 5 个深度学习目标检测模型的综合性能, 结果表明, 在飞机掩体影像数据集上 YOLOX 模型表现更佳, 平均精度可达 97.7%, 但水平框的检测结果无法获得飞机掩体的精确边界和朝向。为此, 对 YOLOX 模型进行改进, 提出针对不同朝向下的飞机掩体检测新方法 R-YOLOX, 实现对飞机掩体的旋转检测, 旋转预测框更加贴合目标轮廓, 采用 KL 散度损失改进后的模型精度显著提升, 准确率提升了 7.24 个百分点, 对飞机掩体具有更好的检测效果。从水平框和旋转框这 2 个角度都能实现飞机掩体的准确检测, 为高分辨率遥感影像中飞机掩体的准确识别提供了新思路。

关键词 遥感; 目标检测; 深度学习; 遥感影像; 旋转框

中图分类号 TP701

文献标志码 A

DOI: 10.3788/LOP22287

Aircraft-Bunker Detection Method Based on Deep Learning in High-Resolution Remote-Sensing Images

Shi Shushu^{1,2}, Chen Yongqiang^{1*}, Wang Yingjie¹, Wang Chunle¹

¹Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100190, China;

²School of Electronic, Electrical and Communication Engineering, University of Chinese Academy of Sciences, Beijing 100049, China

Abstract Aircraft bunkers are the key aircraft protection fortifications. Therefore, the use of remote sensing images to achieve rapid and accurate detection of aircraft bunkers is of great significance. To develop a method for detecting aircraft bunkers through remote sensing images, we collected information and Google Earth images of 60 airfields with aircraft bunkers and constructed a high-resolution remote-sensing-image dataset of aircraft bunkers. Then, we compared the comprehensive performance of five deep-learning target-detection models, namely, Faster R-CNN, SSD, RetinaNet, YOLOv3, and YOLOX. The research results show that the YOLOX model performs better on the aircraft-bunkers-image dataset with an average precision of 97.7%. However, the results of the horizontal frame cannot obtain a precise boundary and orientation of the aircraft bunkers. Therefore, we propose a new method R-YOLOX, which is an improved version of the YOLOX model, for detecting aircraft bunkers under different orientations. Our method achieves the rotational detection of aircraft bunkers. Compared with the YOLOX model, the rotational prediction frame of our method fits the target contour more closely, and the model accuracy with respect to Kullback-Leibler divergence loss is significantly improved, with an increase of 7.24 percentage points, showing a better detection effect on aircraft bunkers. Further, the accurate detection of aircraft bunkers is achieved from the perspective of horizontal and rotating frames, thereby providing a new idea for the accurate identification of aircraft bunkers in remote sensing images.

Key words remote sensing; target detection; deep learning; remote sensing image; rotating frame

收稿日期: 2022-10-19; 修回日期: 2022-11-23; 录用日期: 2022-12-15; 网络首发日期: 2023-01-05

基金项目: 国家自然科学基金(61901445)

通信作者: *clyq@163.com

1 引言

飞机掩体是机场中保护和隐蔽飞机的关键设施,其快速检测和定位能为制导装置实现精准打击提供关键支撑。飞机掩体形态多样,在方向、尺寸及表面覆盖物等方面具有较大差异,且部分机场附近环境复杂,有些背景地物与目标形态特征较为相似,多样性和复杂性给目标检测任务带来挑战。遥感影像覆盖范围广、更新时间快、受地面条件限制少,可以为飞机掩体检测提供数据支持。因此,研究如何利用遥感影像快速准确检测飞机掩体值得关注,在战场态势评估和敌方战备分析等方面具有十分重要的战略意义。

遥感影像目标检测就是确认影像中是否存在感兴趣的目标并找到目标位置^[1],能够快速有效地获取影像信息。针对飞机掩体这一目标,孙军领等^[2]基于光学遥感影像提取掩体的位置特征和形状特征进行融合检测,徐万朋等^[3]利用基于几何活动轮廓模型分割方法对合成孔径雷达图像中的掩体目标进行检测。这些都是通过传统检测方法进行飞机掩体检测,依赖于目标的纹理、几何和边缘等人工设计的底层特征,检测效率和精度低。而随着高分辨率遥感卫星的快速发展,影像包含更多复杂的局部信息,仅依靠传统的遥感影像检测方法和人工目视解译不能满足当前数据处理精度和效率的要求。

如今以深度学习为主的人工智能方法在计算机视觉领域成就显著^[4],深层卷积神经网络的非线性变换特性使其具有自学习能力强、泛化性高、鲁棒性好和检测精度高等优点^[5-6],为遥感影像信息的快速获取创造了有利条件^[7],在车辆^[8]、飞机^[9]、舰船^[10-11]、建筑^[12]、云^[13]和油罐^[14]等遥感影像目标检测中得到了广泛应用。深度学习目标检测模型根据实现过程可分为双阶段模型和单阶段模型两大类。双阶段检测模型是指基于候选区域的目标检测方法,Faster R-CNN^[15]是经典的双阶段检测模型,在精度上表现更为突出,在遥感影像目标检测中被广泛应用。Yan 等^[16]利用添加注意力机制的 Faster R-CNN 模型对高分辨率遥感影像中的尾矿池进行检测,提高了检测精度。单阶段检测模型是基于回归的目标检测方法,包括 SSD^[17]、RetinaNet^[18]和 YOLO 系列^[19]。单阶段检测模型在速度方面更具优势,不同学者对模型网络结构等继续进行优化,使其在遥感影像目标检测中保持速度的同时具有更高的检测精度。史文旭等^[20]指出对 SSD 模型中的金字塔结构特征进行特征融合后,可以有效增强网络对不同尺度遥感影像的适应能力。Yin 等^[21]利用改进的 RetinaNet 模型来检测具有复杂分布特性的火力发电厂。牛浩青等^[22]通过添加门控通道注意力机制和自适应上采样模块改进 YOLOv3 模型,以提高模型对小目标的检测能力。以上都是基于锚框的检测,近些年,又陆续出现无锚框的检测方法,如 CenterNet^[23]、

ExtremeNet^[24]和 CornerNet^[25]等,消除了复杂的锚框设计过程,通过预测关键点来取代锚框,为目标检测开辟了新方向,YOLOX 模型^[26]也借鉴了无锚框的检测方法。

一般来说,深度学习目标检测领域常采用水平框来标注目标,但遥感影像是俯视成像,具有目标方向多样性的特点。水平框包围区域会出现大量的空白和重叠交叉区域,影响检测效率且无法提供准确的角度信息,针对这一特殊性,探索采用旋转框对目标进行检测。Yang 等^[27-29]利用公开的遥感影像数据集,连续提出 R3Det、SCRDet、SCRDet++ 等旋转目标检测方法。Liu 等^[30]提出旋转感兴趣区域池化层(RROI pooling)来提取旋转区域的特征,利用基于旋转区域的卷积神经网络(RRCNN)实现遥感影像中船舶的检测。Tang 等^[31]在 RetinaNet 基础上,结合多特征融合模块和旋转锚框生成机制,对具有任意方向的飞机进行了有效检测。

尽管深度学习技术的发展为目标检测领域带来了新的突破,但相比自然场景下的目标检测,遥感影像尺寸大、覆盖区域广、背景复杂且具有特殊的视角。遥感影像飞机掩体检测会面临目标尺度变化大、多角度、小目标以及类不平衡等问题。高分辨率影像对地表信息刻画得更为精细,也使地物视觉特征和影像的空间模式更加复杂。飞机掩体在高分辨率影像中呈现了多种色调、形态且具有不同的朝向,为检测任务带来诸多挑战。为此,本文分别从水平框和旋转框两个方面进行飞机掩体检测算法研究。首先,对飞机掩体的特征进行详细分析,构建高分辨率遥感影像飞机掩体样本数据集。然后,对多个卷积神经网络目标检测模型进行综合性能对比分析,探究适用于遥感影像多尺度、小目标检测的算法,用于飞机掩体检测。针对遥感影像中飞机掩体具有任意朝向的分布特点,提出一种旋转目标检测方法 R-YOLOX,通过添加对角度参量的回归,扩展模型输出维度并采用 KL 散度(KLD)损失作为检测框的回归损失函数,提高旋转目标检测精度,解决了水平框和旋转目标轮廓不能贴合的问题,实现了对高分辨率遥感影像中飞机掩体的快速高效检测。

2 飞机掩体特征分析及数据集构建

2.1 飞机掩体特征分析

当前研究多利用公开发布的数据集进行检测,获取到的目标信息大多是船舶、机场、飞机和油罐等,而目前公开发布的数据集中尚没有飞机掩体这一类别的信息。基于此需求,对飞机掩体的特征进行相应分析,构建包含飞机掩体单一类别的数据集。早期的飞机掩体多为露天工事,由一块修整过的平地 and 周围用土或砖坯构筑的防护墙组成,俗称“机窝”,但没有顶盖,抗打击能力非常弱,随着技术的提升,逐渐被取缔。现在常用的多为由钢筋水泥以及波纹钢板等材料构建的

带有顶盖的飞机掩体,用来增强飞机在地面的防弹能力、提高机场的战时生存能力。根据其构筑情况不同

又可细分为飞机洞库和飞机掩蔽库,不同类型飞机掩体图像如图 1 所示。

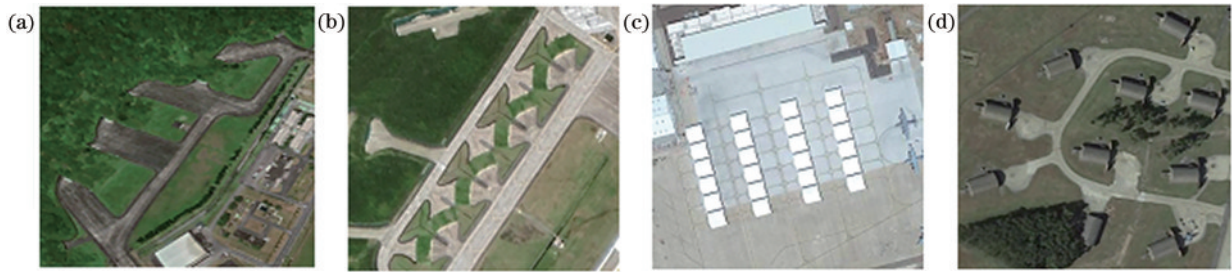


图 1 不同类型飞机掩体 Google Earth 图像。(a)坑道式机库;(b)开掘式机库;(c)机棚;(d)机堡

Fig. 1 Google Earth images of different types of aircraft bunkers. (a) Tunnel hangar; (b) excavation hangar; (c) hangar; (d) revetment

飞机洞库包括坑道式和开掘式两种。进行过伪装的飞机洞库在遥感影像中较难分辨,洞库表面有植被和裸土等覆盖时表现出植被和裸土的光谱特征,与背景地表特征相似,但会有滑行道与机场跑道相连,可根据上下文语义信息进行识别,在检测时容易造成误检和漏检。飞机掩蔽库通常位于停机坪或滑行道的一侧,呈拱形结构、带顶盖或是利用地形构筑的一种抗力较强的单机掩蔽工事,被广泛采用,包括机棚和机堡两种。机棚是一种直壁式机库,有拱形洞式和平顶式,部分会对顶棚进行伪装,在集体停机坪上联排修筑。钢结构材料顶棚反射性强,在光学遥感影像中呈高亮度,与机场不透水地面对比较为明显,几何形状为一对边是直线、另一对边为弧形的类长方形,主要在靠近机场一侧的大块不透水面上集中分布。机堡主要为半圆形的落地拱券式,结构坚固、抗打击能力强,一般位于机场滑行道两侧,通过滑行道与跑道相连,多为树杈形分散分布。在光学遥感影像中呈现白色、灰色及绿色等多种色调,部分与机场周围的建筑屋顶色调较为相似,增加了检测难度。在遥感影像中可以看到较为明显的拱形结构,飞机掩体具有一定高度,在一定方向上可以看到明显的阴影特征,可作为判别依据。存放不同类型的飞机的掩体大小不一致,具有尺寸多样化的特点。

基于上述分析,主要利用飞机掩体的位置、形状、色调及阴影等典型特征作为解译标志,选择全球分布的多个机场的多种飞机洞库和飞机掩蔽库构建飞机掩体数据集,保证数据集的准确性与多样性。

2.2 飞机掩体数据集

选取 60 个建有飞机掩体的机场作为目标区域,选择影像分辨率高、时间跨度大及数据量大的 Google Earth 影像作为数据源,下载 300 张 17~19 级的高分辨率影像数据,空间分辨率最高达 0.42 m,数据尺寸大小为 640 pixel×640 pixel。选取不同背景、不同季节及不同方向的飞机掩体,并尽可能包含主要类型,保证目标的分辨率、时间分辨率、方向和尺度的多样性,构建包含复杂背景、多目标、小目标及不同尺度下飞机掩体的遥感影像数据集。数据集按照 VOC 2007

格式制作,旋转框利用 rolLabelImg 程序进行标注,标注结果通过专家审核,保证了标注的准确性。取其最小外接水平矩形,微调后作为水平框标注,并通过水平翻转、垂直翻转、旋转以及增加高斯噪声等方法进行 4 次数据增强,经过筛选后最终得到包含 1056 张图片的数据集。按照 8:1:1 的比例将数据集划分为训练集、验证集和测试集,数据集部分样本如图 2 所示,数据集中每张图像的目标分布情况如图 3 所示,保证数据的多样性。

3 研究方法

3.1 深度学习目标检测模型

凭借卷积神经网络在特征提取中的显著优势,深度学习为目标检测领域提供了许多出色的模型。为探究深度学习目标检测模型在遥感影像飞机掩体检测中的适用性,选择 Faster R-CNN、SSD、RetinaNet、YOLOv3 以及 YOLOX 模型进行实验,涵盖双阶段、单阶段、有锚框和无锚框的不同检测模型,它们的性能可以代表最先进的深度学习目标检测模型检测飞机掩体的能力。

Faster R-CNN^[15]是经典的双阶段检测模型,网络结构如图 4 所示。通过特征提取网络获取图像特征后,利用区域生成网络(RPN)生成一系列候选区域,RPN 训练过程就是寻找最接近真实框的候选锚框,感兴趣区域池化层(ROI pooling)主要根据 RPN 生成的特征图和特征提取网络生成的特征图进行最大池化,将其变为固定尺寸的特征图。最后进行目标分类和锚框精确回归定位,获取目标所属类别和检测框的精确位置。

SSD^[18]是经典的单阶段检测模型,无需进行候选区域提取,能对图像的全局信息进行预测,并将分类和识别进行统一的回归处理,有效提升检测速度的同时保证了较高的检测精度,网络结构如图 5 所示。RetinaNet 在 SSD 模型的基础上进行优化,采用 Focal 损失^[19]代替标准交叉熵损失函数进行模型训练,训练时能够自适应调整损失权重,网络结构如图 6 所示。



图 2 数据集示例图像。(a)不同遮盖物的目标;(b)不同背景下多目标;(c)复杂背景下目标

Fig. 2 Dataset sample images. (a) Objectives of different coverings; (b) multiple objectives under different backgrounds; (c) objectives under complex background

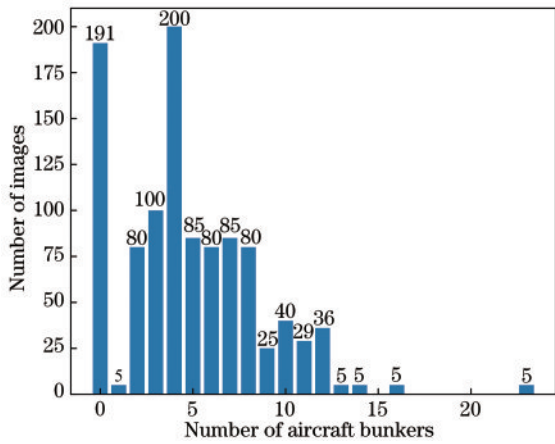


图 3 数据集每张图像中目标分布情况

Fig. 3 Objectives distribution in each image of dataset

YOLO 系列是最早提出的单阶段检测模型。YOLOv3^[32]用更深结构的 DarkNet53 基础特征提取网络替换之前的 DarkNet19^[33],并引入特征金字塔网络进行多尺度预测,结构清晰且实时性好,在实际工程

中应用较多。2021 年,旷视提出的 YOLOX^[26]在 YOLO 系列模型基础上进行多处改进,采用无锚框的方法,模型更加灵活有效^[34]。YOLOX 主干特征提取网络采用基于 YOLOv3 的 DarkNet53,并在其基础上添加 SPP 组件,如图 7 所示,通过主干网络获取 3 个特征层进行下一步网络的构建。特征融合部分采用 FPN 结构,自顶向下通过上采样的方式将高层次的特征信息进行传递;并结合 YOLOv4 中用到的 PANet 结构^[35],再次进行下采样实现特征融合,增强对多尺度目标的特征表达能力。在输出层,从解耦头、无锚框、标签分配、损失函数计算等角度进行改进:解耦头的结构使网络具有更高的精度和更快的收敛速度;无锚框的方式有效减少了预测框数量,缓解了正负样本不均衡的问题。YOLOX 保留了 YOLOv5 中将网络结构做成可选择配置的方式,根据主干网络结构的宽度和高度不同分为 YOLOX-s、YOLOX-m、YOLOX-l、YOLOX-x 等版本,本研究主要采用 YOLOX-s 模型。

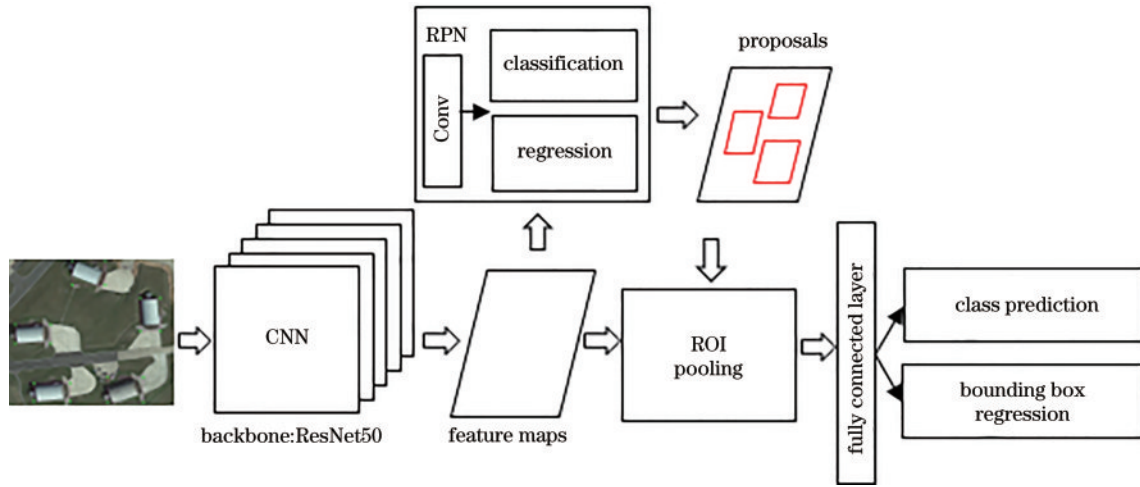


图 4 Faster R-CNN 结构
Fig. 4 Faster R-CNN framework

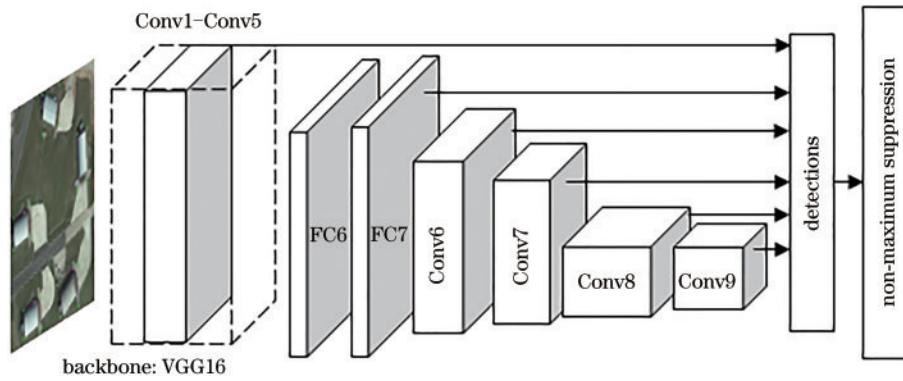


图 5 SSD 结构
Fig. 5 SSD framework

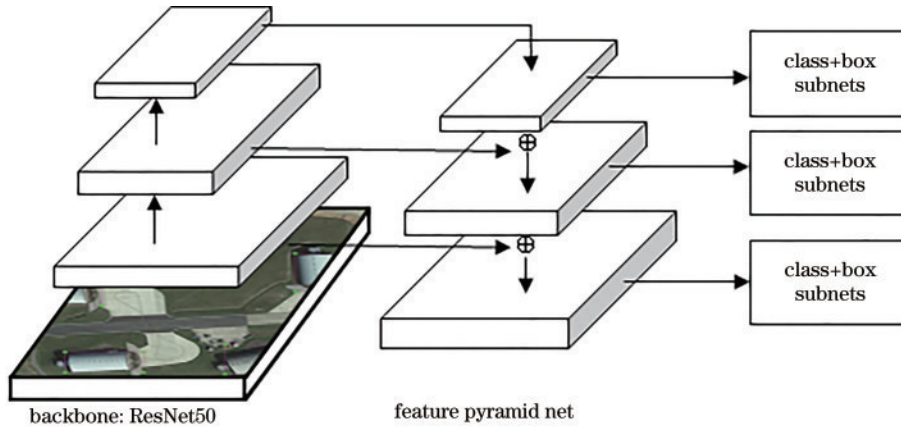


图 6 RetinaNet 结构
Fig. 6 RetinaNet framework

3.2 飞机掩体旋转检测改进方法

针对具有多方向性的飞机掩体目标,为实现更高效的检测,在YOLOX模型基础上进行改进,提出一种旋转目标检测方法R-YOLOX。通过添加对角度参数的预测,采用KLD损失作为Reg分支的损失函数,以减少角度的微小误差对精度的影响,提高边界框的回归精度,进一步实现具有任意朝向的飞机掩体检测。

3.2.1 旋转框表示方法

常规目标检测方法是水平检测,通常用 x 、 y 、 w 、 h 这4个参数表示水平边界框,分别代表边界框的中心点坐标、宽度和高度,但水平框不能很好地覆盖到不同方向的目标^[36],并存在信息冗余,具有一定局限性。旋转目标检测在获取目标的位置信息基础上还可得到旋转角度的预测信息,能够感知飞机掩体更完整全面的

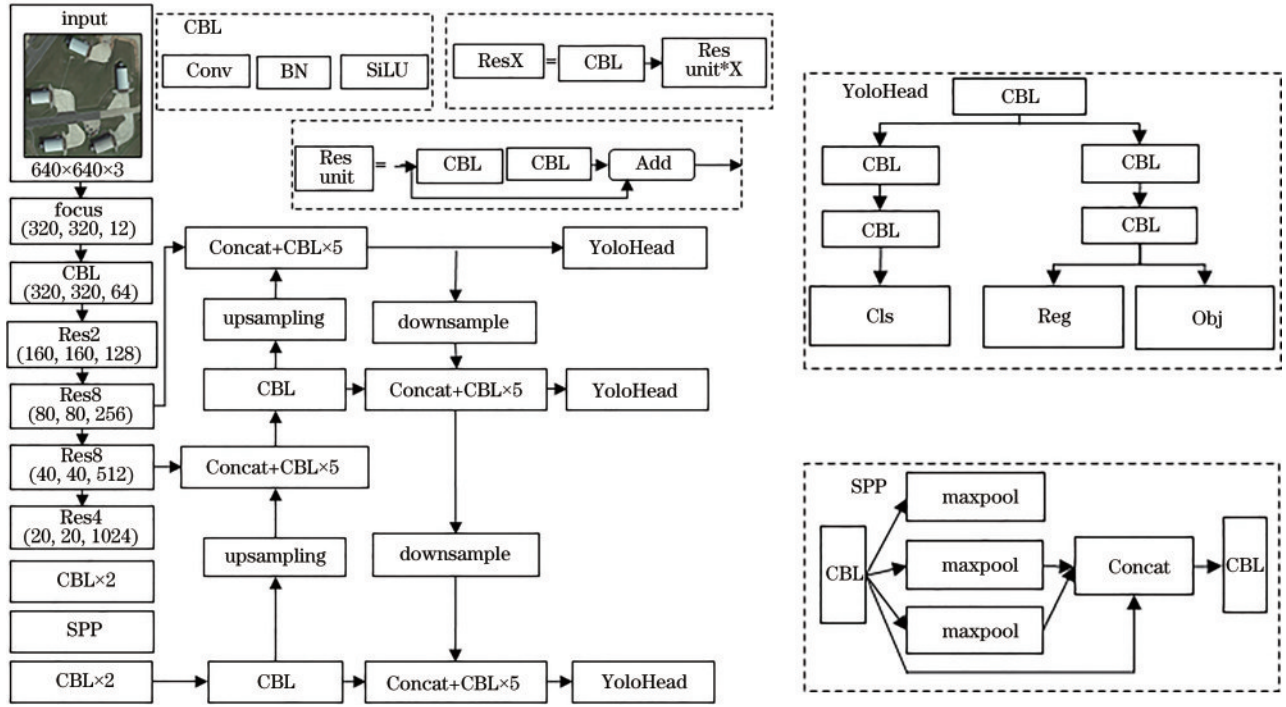


图 7 YOLOX 结构

Fig. 7 YOLOX framework

信息。具体而言,就是要引入目标的角度信息维度和相应的损失回归函数。图 8 中,通过 x 、 y 、 w 、 h 和 θ 这 5 个参数来确定任意方向上的边界框,其中, x 、 y 分别表示边界框中心点的横纵坐标, w 、 h 分别表示边界框的

长边和短边, θ 由 x 轴正方向和长边 w 的夹角决定^[37], 取值范围在 $[-90^\circ, 90^\circ)$, 通过 θ 将目标的角度信息维度引入模型中。

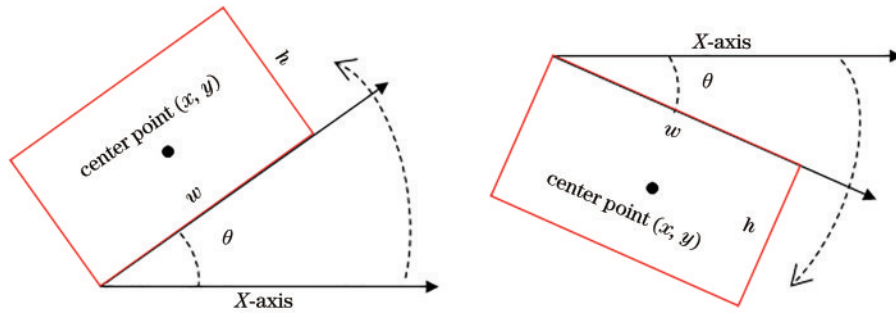


图 8 旋转框表示方法

Fig. 8 Definition of rotated object representation

3.2.2 损失函数

YOLOX 模型的损失函数分为 3 部分:类别损失 (L_{cls})、边界框回归损失 (L_{reg}) 以及置信度损失 (L_{obj}):

$$L = L_{reg} + L_{cls} + L_{obj} \quad (1)$$

类别损失和置信度损失采用的是二元交叉熵损失函数:

$$L_{BCE} = -\frac{1}{N} \sum_{i=1}^N y_i \cdot \log[\sigma(x_i)] + (1 - y_i) \cdot \log[1 - \sigma(x_i)] \quad (2)$$

式中: $\sigma(x_i)$ 为 Sigmoid 函数, 能把 x 映射到 $(0, 1)$ 的区间:

$$\sigma(x) = \frac{1}{1 + \exp(-x)} \quad (3)$$

对增添的角度参数进行回归时, 旋转边界框的回归会引起旋转敏感度误差问题, 导致模型训练不稳定从而影响检测精度。借鉴 Yang 等^[29] 的研究结果, 采用 KLD 损失函数进行边界框回归损失计算。其原理是将旋转边界框转换为二维高斯分布, 然后计算高斯分布之间的相对熵即 KLD:

$$D_{kl}(\mathcal{N}_p \| \mathcal{N}_t) = \frac{1}{2} (\mathbf{u}_p - \mathbf{u}_t)^T \sum_t^{-1} (\mathbf{u}_p - \mathbf{u}_t) + \frac{1}{2} \text{Tr} \left(\sum_t^{-1} \sum_p \right) + \frac{1}{2} \ln \frac{|\sum_t|}{|\sum_p|} - 1 \quad (4)$$

$$(\mathbf{u}_p - \mathbf{u}_t)^T \sum_t^{-1} (\mathbf{u}_p - \mathbf{u}_t) = \frac{4(\Delta x \cos \theta_t + \Delta y \sin \theta_t)^2}{w_t^2} + \frac{4(\Delta y \cos \theta_t - \Delta x \sin \theta_t)^2}{h_t^2}, \quad (5)$$

$$\text{Tr} \left(\sum_t^{-1} \sum_p \right) = \frac{h_p^2}{w_t^2} \sin^2 \Delta \theta + \frac{w_p^2}{h_t^2} \sin^2 \Delta \theta + \frac{h_p^2}{h_t^2} \cos^2 \Delta \theta + \frac{w_p^2}{w_t^2} \cos^2 \Delta \theta, \quad (6)$$

$$\ln \left| \frac{\sum_i}{\sum_p} \right| = \ln \frac{h_i^2}{h_p^2} + \ln \frac{w_i^2}{w_p^2}, \quad (7)$$

式中: $\Delta x = x_p - x_t$, $\Delta y = y_p - y_t$, $\Delta \theta = \theta_p - \theta_t$ 分别是预测框和真实框之间的中心点坐标和角度差值。KLD 损失函数如式 (8) 所示: 如果 KL 散度 $D_{\text{KL}}(\mathcal{N}_p \parallel \mathcal{N}_t)$ 等于 0, 那么损失函数也为 0; 如果 KL 散度很大, 则损失函数为 1, 在 KL 散度上利用 \ln 函数来平滑损失函数。

$$L_{\text{KLD}}(\mathcal{N}_p \parallel \mathcal{N}_t) = 1 - \frac{1}{1 + \ln[1 + D_{\text{KL}}(\mathcal{N}_p \parallel \mathcal{N}_t)]}. \quad (8)$$

KLD 损失函数可以根据目标的尺度动态调整权重, 减少微小的角度误差对精度的影响、实现高精度的旋转检测^[38]。

3.2.3 性能指标

目标检测领域常用的精度评价指标包括准确率 (P)、召回率 (R)、平均精度 (AP) 和 F1 得分 (s_{F1})。准确率表示在所有检测出的目标中检测正确的概率, 即“找的对”的比例; 召回率是指所有正样本中正确识别的概率, 即“找的全”的比例; 平均精度是对不同召回率点上的准确率进行平均, 在 PR 曲线图 (即 P 为纵坐标、 R 为横坐标的曲线图) 上表现为 PR 曲线下方的面积。F1 得分是准确率和召回率的调和平均数。具体公式如下:

$$P = \frac{N_{\text{TP}}}{N_{\text{TP}} + N_{\text{FP}}}, \quad (9)$$

表 1 模型训练参数设置

Table 1 Parameter settings in model training

| Model | Backbone | Batch size | Input image size | Initial learning rate | Iterations |
|--------------|---------------|------------|------------------|-----------------------|------------|
| Faster R-CNN | ResNet50 | 4 | 600×600 | 0.0001 | 150 |
| SSD | VGG16 | 8 | 300×300 | 0.0020 | 150 |
| RetinaNet | ResNet50 | 8 | 600×600 | 0.0001 | 150 |
| YOLOv3 | DarkNet53 | 8 | 640×640 | 0.0010 | 150 |
| YOLOX-s | SPP-DarkNet53 | 8 | 640×640 | 0.0001 | 150 |

4.2 结果分析

为了更好地测试模型性能, 在原先划分的测试集基础上, 新加入包含 115 张图像与原数据集制备方法相

$$R = \frac{N_{\text{TP}}}{N_{\text{TP}} + N_{\text{FN}}}, \quad (10)$$

$$R_{\text{AP}} = \int_0^1 P(R) dR, \quad (11)$$

$$s_{F1} = 2 \times \frac{P \times R}{P + R}, \quad (12)$$

式中: N_{TP} 表示正确识别的飞机掩体数量; N_{FP} 表示错误提取的飞机掩体数量; N_{FN} 表示遗漏提取的飞机掩体数量。

训练时间 T_{train} 表示在模型训练过程中, 平均迭代一次所花费的时间, 公式为

$$T_{\text{train}} = \frac{1}{n} \sum_i t_i, \quad (13)$$

式中: n 为模型训练至收敛的迭代次数; t_i 表示训练第 i 次所需的时间。

用每秒帧数 (S_{FPS}) 评估模型的检测速度, 即检测一张图像所需时间 (T_{FPS}), 公式为

$$S_{\text{FPS}} = \frac{1}{T_{\text{FPS}}}. \quad (14)$$

4 实验及结果分析

4.1 实验平台及参数设置

实验处理器为 7 核 Intel(R) Xeon(R) CPU E5-2680v4 @2.40GHz, GPU 为 NVIDIA GeForce RTX 3060 (12G 显存), 使用 PyTorch 开发框架。

选用 Faster R-CNN 这一双阶段模型和 SSD、RetinaNet、YOLOv3 和 YOLOX 这 4 个单阶段检测模型进行实验, 模型训练参数设置如表 1 所示。Faster R-CNN 和 RetinaNet 采用 ResNet50 作为基础特征提取网络; SSD 采用 VGG16 作为特征提取网络; YOLOv3 以 DarkNet53 作为基础特征提取网络, 并采用 K-means 聚类方法获取锚框预设尺寸: (39, 40)、(44, 60)、(60, 48)、(59, 68)、(77, 71)、(80, 102)、(97, 89)、(117, 112)、(151, 149); YOLOX 以改进的 SPP-DarkNet53 作为特征提取网络。所有模型都对应使用在 ImageNet 数据集预训练得到的权重, 采用“Step”学习率下降方式和 Adam 优化器。

同的独立数据集, 整体作为新测试集进行模型性能的评估, 结果如表 2 所示, 表内加粗内容为该项的最佳值。

表 2 不同模型性能对比

Table 2 Comparison of detection accuracy of different model

| Model | Precision / % | Recall / % | AP ₅₀ / % | F1 | FPS | Size / MB | Time / s |
|--------------|---------------|--------------|----------------------|-------------|--------------|--------------|-------------|
| Faster R-CNN | 78.28 | 96.00 | 95.58 | 0.86 | 12.14 | 108.17 | 0.61 |
| SSD | 90.49 | 92.90 | 95.46 | 0.92 | 87.80 | 90.61 | 0.19 |
| RetinaNet | 92.94 | 95.01 | 97.31 | 0.94 | 28.60 | 138.91 | 0.62 |
| YOLOv3 | 94.64 | 96.00 | 96.95 | 0.95 | 29.54 | 235.04 | 0.68 |
| YOLOX | 96.85 | 95.45 | 97.70 | 0.96 | 42.97 | 34.30 | 0.32 |

在交并比(IoU)阈值为 0.5 时, AP 值均在 95% 以上, 所选模型的检测精度都能满足检测要求。其中, YOLOX 模型检测精度较高, 准确率、AP 值和 F1 得分这 3 项指标都为最佳值, AP 值达 97.7%。Faster R-CNN 和 YOLOv3 模型取得了最高的召回率, 为 96%; Faster R-CNN 模型准确率稍低, AP 值为 95.58%, YOLOv3 的 AP 值为 96.95%, F1 得分为 0.95, 仅次于 YOLOX 模型。RetinaNet 模型 AP 值较高, 为 97.31%, 略低于 YOLOX 模型。SSD 模型准确率和召回率相对较好, AP 值为 95.46%。此外, SSD 模型因输入数据尺寸较小而具有最快的训练时间和检测速度; YOLOX 在输入尺寸较大的情况下同样具备较快的训练时间和检测速度, 还拥有最小的模型权重。基于深度学习的目标检测模型在飞机掩体数据集上获得优异表现, 都能满足对飞机掩体的自动识别任务, YOLOX 模型性能更佳。

选取具有代表性的测试集数据的检测结果进行可视化, 效果如图 9~14 所示。针对有多目标分布的图 9(a); Faster R-CNN 和 YOLOv3 实现了准确检测; SSD 将一个带有阴影的树丛误检为目标; RetinaNet 和

YOLOX 存在漏检情况, 都将图中偏左的一个目标漏检, 该目标附近遮蔽物较多, 影响了目标的准确识别。且目标入口面向太阳照射方向, 影像中目标地面没有明显阴影时易出现漏检情况。图 9(b) 中有多个与目标特征较为相似的屋顶目标, 易与目标混淆, YOLOv3 和 YOLOX 实现了准确识别, 其他模型均有误检情况出现。图 9(c) 中的目标尺寸稍小, 具有多个分布方向, 仅 YOLOX 实现了准确检测, 其他模型漏检较多, YOLOX 无需预先设定锚框, 对小目标检测效果较好。针对图 9(d) 目标表面有雪覆盖、背景较为模糊复杂的情况, SSD 和 YOLOv3 漏检了两个目标, RetinaNet 和 YOLOX 均有一个目标漏检, 只有 Faster R-CNN 全部检出。图 9(e) 为并排分布的双联飞机掩体, YOLOX 和 Faster R-CNN 实现了准确检测, RetinaNet 漏检了一个目标, SSD 和 YOLOv3 将一个建筑屋顶误检为目标。并排密集分布的目标易造成检测框混叠, 在进行非极大抑制时会删除掉某一检测框, 从而造成目标的漏检。综上所述: 在背景较为模糊复杂时, 容易造成模型的漏检误检; 且部分方向的目标没有阴影作为特征, 易被漏检; 就综合性能而言, YOLOX 模型表现更优。

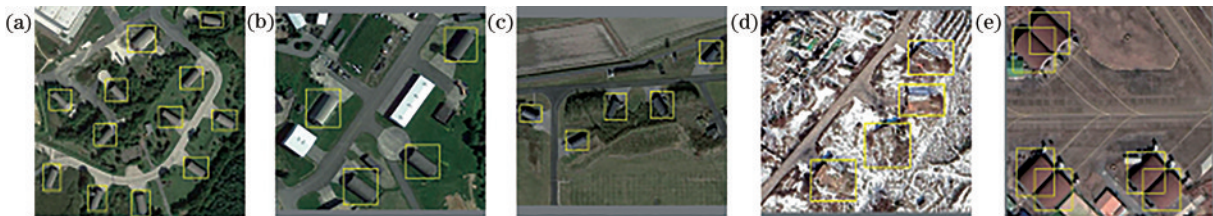


图 9 真实框。(a)多目标分布;(b)包含易混淆目标;(c)多方向小目标;(d)雪覆盖下目标;(e)双联飞机掩体

Fig. 9 Ground truth. (a) Multi-objective distribution; (b) with confusing objectives; (c) multi-directional small objectives; (d) objectives under snow cover; (e) duplex aircraft bunkers



图 10 Faster R-CNN 检测结果。(a)多目标分布;(b)包含易混淆目标;(c)多方向小目标;(d)雪覆盖下目标;(e)双联飞机掩体

Fig. 10 Results of Faster R-CNN. (a) Multi-objective distribution; (b) with confusing objectives; (c) multi-directional small objectives; (d) objectives under snow cover; (e) duplex aircraft bunkers

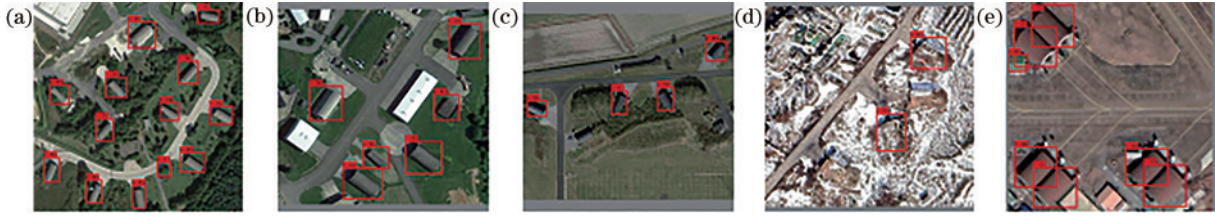


图 11 SSD 检测结果。(a)多目标分布;(b)包含易混淆目标;(c)多方向小目标;(d)雪覆盖下目标;(e)双联飞机掩体
Fig. 11 Results of SSD. (a) Multi-objective distribution; (b) with confusing objectives; (c) multi-directional small objectives; (d) objectives under snow cover; (e) duplex aircraft bunkers



图 12 RetinaNet 检测结果。(a)多目标分布;(b)包含易混淆目标;(c)多方向小目标;(d)雪覆盖下目标;(e)双联飞机掩体
Fig. 12 Results of RetinaNet. (a) Multi-objective distribution; (b) with confusing objectives; (c) multi-directional small objectives; (d) objectives under snow cover; (e) duplex aircraft bunkers



图 13 YOLOv3 检测结果。(a)多目标分布;(b)包含易混淆目标;(c)多方向小目标;(d)雪覆盖下目标;(e)双联飞机掩体
Fig. 13 Results of YOLOv3. (a) Multi-objective distribution; (b) with confusing objectives; (c) multi-directional small objectives; (d) objectives under snow cover; (e) duplex aircraft bunkers



图 14 YOLOX 检测结果。(a)多目标分布;(b)包含易混淆目标;(c)多方向小目标;(d)雪覆盖下目标;(e)双联飞机掩体
Fig. 14 Results of YOLOX. (a) Multi-objective distribution; (b) with confusing objectives; (c) multi-directional small objectives; (d) objectives under snow cover; (e) duplex aircraft bunkers

用所提 R-YOLOX 模型进行对比实验,精度评价结果如表 3 所示。YOLOX 模型由水平框改进为旋转框检测后,AP 值由 97.7% 下降为 97.25%,整体精度稍有下降。因为多增加了一个参数的回归,旋转框对

表 3 旋转检测精度对比

Table 3 The comparison of rotation detection accuracy

| Model | L_{Reg} | Precision / % | Recall / % | AP ₅₀ / % | F1 |
|---------|-----------|---------------|------------|----------------------|----------|
| YOLOX | Smooth | 86.60 | 91.69 | 92.06 | 0.8 |
| | L1 loss | | | | 9 |
| R-YOLOX | KLD loss | 93.84 | 94.57 | 97.25 | 0.9 4 |

目标的标注更为细致,旋转框检测比水平框检测更难,对检测精确度的要求也更高。且旋转框和水平框的交并比的计算方式不同,旋转框的交并比结果受角度影响比较大,角度的微小变化会引起精度的较大改变,从而表征出略低于水平框检测模型的精度。用 YOLOX 和 R-YOLOX 进行预测的可视化结果如图 15、16 所示。从图中可以看到,旋转框检测效果更好,水平框预测结果有大量的重叠和嵌套,而旋转框预测结果更加贴合飞机掩体的轮廓,标注边界清晰,对密集分布的目标检测效果更好,对飞机掩体的旋转检测可以获得更加准确的位置和尺寸信息。

改进前采用 Smooth L1 损失的网路收敛较慢,模

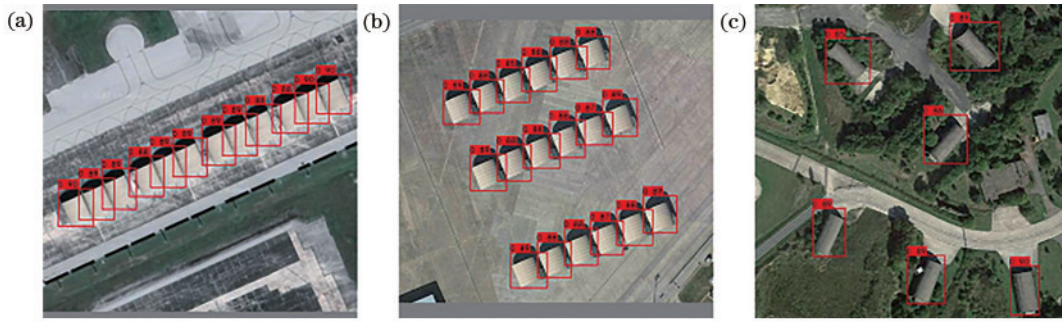


图 15 YOLOX 检测结果。(a) 并排分布目标；(b) 集中分布多目标；(c) 多方向目标

Fig. 15 Results of YOLOX. (a) Side-by-side distribution of objectives; (b) centrally distributed multiple objectives; (c) multi-directional objectives

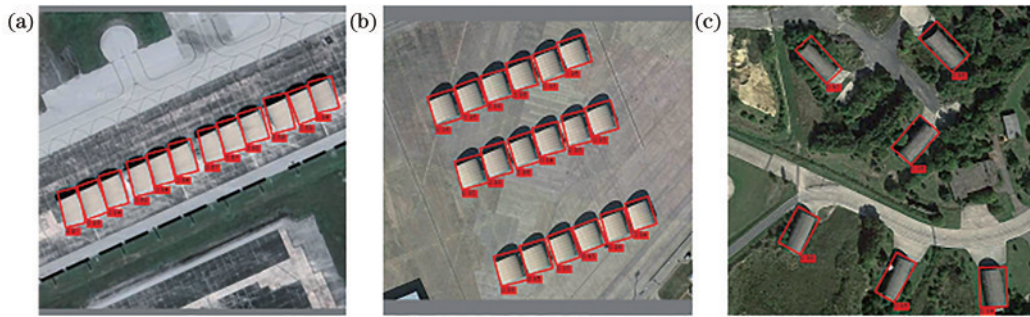


图 16 R-YOLOX + KLD 损失检测结果。(a) 并排分布目标；(b) 集中分布多目标；(c) 多方向目标

Fig. 16 Results of R-YOLOX with KLD loss. (a) Side-by-side distribution of objectives; (b) centrally distributed multiple objectives; (c) multi-directional objectives

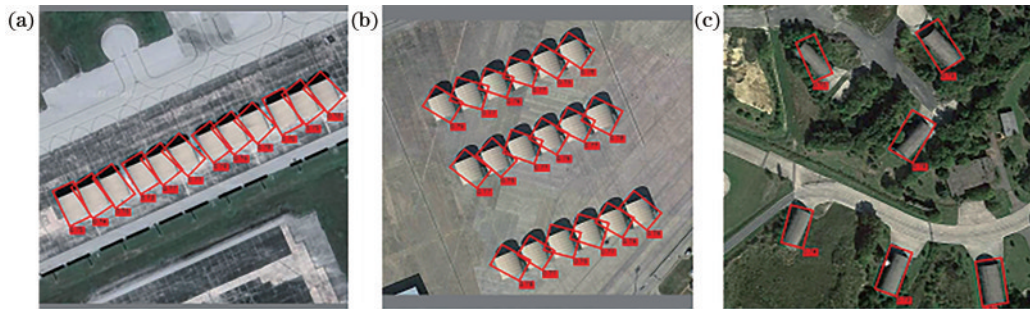


图 17 YOLOX + Smooth L1 损失检测结果。(a) 并排分布目标；(b) 集中分布多目标；(c) 多方向目标

Fig. 17 Results of YOLOX with Smooth L1 loss. (a) Side-by-side distribution of objectives; (b) centrally distributed multiple objectives; (c) multi-directional objectives

型整体精度不高,预测结果如图 17 所示,部分目标角度预测偏差严重。采用 KLD 损失进行改进后,避免了角度微小变化引起精度的下降,模型精度显著提升,平均精度提高了 5.19 个百分点,达到 97.25%。模型的准确率和召回率分别提高 7.24 和 2.88 个百分点,改进后的模型实现了飞机掩体的旋转检测。

总体而言,飞机掩体的旋转检测显著减少了检测框的重叠,能够更加贴合目标边界轮廓。采用 KLD 损失改进后的模型 R-YOLOX 实现了检测精度的显著提升,旋转框的检测结果更能准确显示飞机掩体所在位置。

选择某机场大幅 Google Earth 影像进行飞机掩体自动检测性能验证,进行切片预测后再拼接,最终识别结果如图 18 所示。该区域影像图幅大小为

3200 pixel \times 3200 pixel,空间分辨率为 0.77 m,影像中共包含 74 个飞机掩体目标,所提算法 R-YOLOX 将其全部准确识别,证明了该模型的有效性,在大幅遥感影像中仍然表现良好。

5 结 论

飞机掩体是重要的飞机防护工事,对其进行快速高效检测具有十分重要的战略意义。现有的飞机掩体类型多样、形态各异,在遥感影像中存在多尺度、小目标分布等问题,传统方法检测精度差、效率低,难以满足当下需求。利用先进的深度学习目标检测模型,从水平框和旋转框两个角度出发,针对飞机掩体特征对模型进行优化改进,使其更适用于遥感影像中飞机掩



图 18 某机场飞机掩体检测结果

Fig. 18 Detection results of aircraft bunkers at an airport

体的检测,通过模型训练和检测验证,得出以下结论:1)经过多个深度学习目标检测模型对比分析,在飞机掩体数据集上,YOLOX 的检测精度更高,AP 值达 97.7%,且模型尺寸较小、所需训练时间短、检测速度快,可实现飞机掩体高效检测。2)所提旋转目标检测方法 R-YOLOX,在采用 KLD 损失作为损失函数的改进后,模型检测精度显著提升,AP 提升了 5.19 个百分点,有效实现了飞机掩体旋转检测。本研究从水平框和预测框两个角度实现了飞机掩体的识别,获取了目标更完整的信息。与水平框检测效果对比,旋转框检测能够获得更加准确的规模和位置信息,预测框更贴近目标轮廓。

在光学遥感影像飞机掩体检测中,所用模型整体上具备了快速准确检测的能力,但在部分复杂背景下的检测中仍存在误检和漏检的情况。为实现更加精准的飞机掩体识别,未来还要进一步思考如何利用多源遥感影像进行检测。

参 考 文 献

- [1] 张欣. 基于 YOLO 算法的遥感图像飞机目标检测技术研究[D]. 长春: 中国科学院长春光学精密机械与物理研究所, 2021.
Zhang X. Research on aircraft target detection technology in remote sensing image based on YOLO algorithm[D]. Changchun: Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, 2021.
- [2] 孙军领, 陈天泽, 粟毅. 基于特征融合的遥感图像机场掩体目标检测方法研究[J]. 遥感技术与应用, 2009, 24(3): 379-384.
Sun J L, Chen T Z, Su Y. A method of airport shelter detection based on feature fusion in remote sensing images [J]. Remote Sensing Technology and Application, 2009, 24(3): 379-384.
- [3] 徐万朋, 陈向宁, 徐春宇, 等. SAR 图像机场区域的掩体目标检测方法[J]. 测绘科学技术学报, 2013, 30(5): 494-499.
Xu W P, Chen X N, Xu C Y, et al. SAR image of airport areas cover target detection method[J]. Journal of Geomatics Science and Technology, 2013, 30(5): 494-499.
- [4] LeCun Y, Bengio Y, Hinton G. Deep learning[J]. Nature, 2015, 521(7553): 436-444.
- [5] Cheng G, Yang C Y, Yao X W, et al. When deep learning meets metric learning: remote sensing image scene classification via learning discriminative CNNs[J]. IEEE Transactions on Geoscience and Remote Sensing, 2018, 56(5): 2811-2821.
- [6] 周飞燕, 金林鹏, 董军. 卷积神经网络研究综述[J]. 计算机学报, 2017, 40(6): 1229-1251.
Zhou F Y, Jin L P, Dong J. Review of convolutional neural network[J]. Chinese Journal of Computers, 2017, 40(6): 1229-1251.
- [7] Zhu X X, Tuia D, Mou L C, et al. Deep learning in remote sensing: a comprehensive review and list of resources[J]. IEEE Geoscience and Remote Sensing Magazine, 2017, 5(4): 8-36.
- [8] Tan Q L, Ling J, Hu J, et al. Vehicle detection in high resolution satellite remote sensing images based on deep learning[J]. IEEE Access, 2020, 8: 153394-153402.
- [9] Ge J, Wang C, Zhang B, et al. Azimuth-sensitive object detection of high-resolution SAR images in complex scenes by using a spatial orientation attention enhancement network[J]. Remote Sensing, 2022, 14(9): 2198.
- [10] Wang Z Q, Zhou Y, Wang F T, et al. SDGH-net: ship detection in optical remote sensing images based on Gaussian heatmap regression[J]. Remote Sensing, 2021, 13(3): 499.
- [11] Wu J X, Pan Z X, Lei B, et al. LR-TSDet: towards tiny ship detection in low-resolution remote sensing images[J]. Remote Sensing, 2021, 13(19): 3890.
- [12] Zheng L J, Ai P, Wu Y. Building recognition of UAV remote sensing images by deep learning[C]//IGARSS 2020-2020 IEEE International Geoscience and Remote Sensing Symposium, September 26-October 2, 2020, Waikoloa, HI, USA. New York: IEEE Press, 2020: 1185-1188.
- [13] Lu J Y, Wang Y Z, Zhu Y Q, et al. P_Segnet and NP_Segnet: new neural network architectures for cloud recognition of remote sensing images[J]. IEEE Access, 2019, 7: 87323-87333.
- [14] 郭海涛, 李宸尧, 崔君, 等. 多尺度并联卷积神经网络遥感影像油罐目标检测算法[J]. 测绘科学技术学报, 2021, 38(2): 148-154.
Guo H T, Li C Y, Cui J, et al. Oil tank detection algorithm on remote sensing image using multi-scale parallel convolutional neural networks[J]. Journal of Geomatics Science and Technology, 2021, 38(2): 148-154.
- [15] Girshick R. Fast R-CNN[C]//2015 IEEE International Conference on Computer Vision, December 7-13, 2015, Santiago, Chile. New York: IEEE Press, 2015: 1440-1448.
- [16] Yan D C, Li G Q, Li X Q, et al. An improved faster R-CNN method to detect tailings ponds from high-resolution remote sensing images[J]. Remote Sensing,

- 2021, 13(11): 2052.
- [17] Liu W, Anguelov D, Erhan D, et al. SSD: single shot MultiBox detector[M]//Leibe B, Matas J, Sebe N, et al. Computer vision-ECCV 2016. Lecture notes in computer science. Cham: Springer, 2016, 9905: 21-37.
- [18] Lin T Y, Goyal P, Girshick R, et al. Focal loss for dense object detection[C]//2017 IEEE International Conference on Computer Vision, October 22-29, 2017, Venice, Italy. New York: IEEE Press, 2017: 2999-3007.
- [19] Redmon J, Divvala S, Girshick R, et al. You only look once: unified, real-time object detection[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition, June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 779-788.
- [20] 史文旭, 鲍佳慧, 姚宇. 基于深度学习的遥感图像目标检测与识别[J]. 计算机应用, 2020, 40(12): 3558-3562.
Shi W X, Bao J H, Yao Y. Remote sensing image target detection and identification based on deep learning[J]. Journal of Computer Applications, 2020, 40(12): 3558-3562.
- [21] Yin W X, Diao W H, Wang P J, et al. PCAN: part-based context attention network for thermal power plant detection in remote sensing imagery[J]. Remote Sensing, 2021, 13(7): 1243.
- [22] 牛浩青, 欧鸥, 饶姗姗, 等. 改进 YOLOv3 的遥感影像小目标检测方法[J]. 计算机工程与应用, 2022, 58(13): 241-248.
Niu H Q, Ou O, Rao S S, et al. Small object detection method based on improved YOLOv3 in remote sensing image[J]. Computer Engineering and Applications, 2022, 58(13): 241-248.
- [23] Duan K W, Bai S, Xie L X, et al. CenterNet: keypoint triplets for object detection[C]//2019 IEEE/CVF International Conference on Computer Vision (ICCV), October 27-November 2, 2019, Seoul, Republic of Korea. New York: IEEE Press, 2019: 6568-6577.
- [24] Zhou X Y, Zhuo J C, Krähenbühl P. Bottom-up object detection by grouping extreme and center points[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 15-20, 2019, Long Beach, CA, USA. New York: IEEE Press, 2019: 850-859.
- [25] Law H, Deng J. CornerNet: detecting objects as paired keypoints[J]. International Journal of Computer Vision, 2020, 128(3): 642-656.
- [26] Li S, Fu X J, Dong J. Improved ship detection algorithm based on YOLOX for SAR outline enhancement image [J]. Remote Sensing, 2022, 14(16): 4070.
- [27] Yang X, Yang J R, Yan J C, et al. SCRDet: towards more robust detection for small, cluttered and rotated objects[C]//2019 IEEE/CVF International Conference on Computer Vision (ICCV), October 27-November 2, 2019, Seoul, Republic of Korea. New York: IEEE Press, 2019: 8231-8240.
- [28] Yang X, Yan J C, Ming Q, et al. Rethinking rotated object detection with gaussian wasserstein distance loss [C]//Proceedings of International Conference on Machine Learning (ICML), July 18-24, 2021, Virtual Event. Copenhagen: MLR Press, 2021: 11830-11841.
- [29] Yang X, Yang X J, Yang J R, et al. Learning high-precision bounding box for rotated object detection via Kullback-Leibler Divergence[C]//Proceedings of Neural Information Processing Systems (NeurIPS), December 6-14, 2021, Virtual. Copenhagen: MLR Press, 2021: 18381-18394.
- [30] Liu Z K, Hu J G, Weng L B, et al. Rotated region based CNN for ship detection[C]//2017 IEEE International Conference on Image Processing, September 17-20, 2017, Beijing, China. New York: IEEE Press, 2017: 900-904.
- [31] Tang F F, Wang W, Li J, et al. Aircraft rotation detection in remote sensing image based on multi-feature fusion and rotation-aware anchor[J]. Applied Sciences, 2022, 12(3): 1291.
- [32] Redmon J, Farhadi A. YOLOv3: an incremental improvement[EB/OL]. (2018-04-08)[2022-08-06]. <https://arxiv.org/abs/1804.02767>.
- [33] Redmon J, Farhadi A. YOLO9000: better, faster, stronger[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition, July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 6517-6525.
- [34] Ge Z, Liu S T, Wang F, et al. YOLOX: exceeding YOLO series in 2021[EB/OL]. (2021-07-18)[2022-08-06]. <https://arxiv.org/abs/2107.08430>.
- [35] Bochkovskiy A, Wang C Y, Liao H. YOLOv4: optimal speed and accuracy of object detection[EB/OL]. (2020-04-23)[2022-08-09]. <https://arxiv.org/abs/2004.10934>.
- [36] Yu Y, Yang X, Li J, et al. A cascade rotated anchor-aided detector for ship detection in remote sensing images [J]. IEEE Transactions on Geoscience and Remote Sensing, 2022, 60: 1-14.
- [37] Chan S X, Zheng J C, Wang L N, et al. Rotating object detection in remote-sensing environment[J]. Soft Computing, 2022, 26(16): 8037-8045.
- [38] Zakaria Y, Mokhtar S A, Baraka H, et al. Improving small and cluttered object detection by incorporating instance level denoising into single-shot alignment network for remote sensing imagery[J]. IEEE Access, 2022, 10: 51176-51190.