

doi:10.3969/j.issn.1673-9833.2022.03.007

# 基于YOLOv5的遥感图像目标检测

董丽君, 曾志高, 易胜秋, 文志强, 孟辰

(湖南工业大学 计算机学院, 湖南 株洲 412007;  
湖南省智能信息感知与处理技术重点实验室, 湖南 株洲 412007)

**摘要:** 为了解决在遥感图像目标检测任务中目标背景繁杂难以识别且目标尺寸复杂的问题, 提出一种基于YOLOv5的遥感图像检测优化模型。首先, 对输入数据进行马赛克增强, 增加样本多样性, 同时采用自适应锚框计算, 寻求最优初值锚框; 然后, 把通过主干网络提取到的特征层进行特征融合得到最优特征层, 再对定位损失进行优化, 采用CIoU loss作为定位损失函数, Focal loss作为分类损失函数; 最后, 在测试时对输入图片采用自适应图片缩放, 以减少信息冗余, 加快模型检测速率。该模型能有效捕捉图像特征, 实现快速精准的目标定位。对公开10类地理空间物体检测数据集(NWPU-VHR 10)和RSOD数据集进行了训练测试, 对比试验表明, 优化模型mAP达到0.9896, 比优化前的模型mAP提升了2.31%, 与使用相同数据集的其他模型的最优值进行比较, 其mAP提升了8.19%, 该方法能有效提高遥感图像检测精度。

**关键词:** 遥感图像检测; YOLOv5算法; CIoU loss; Focal loss; 马赛克数据增强; 自适应方法

**中图分类号:** TP391.4      **文献标志码:** A      **文章编号:** 1673-9833(2022)03-0044-07

**引文格式:** 董丽君, 曾志高, 易胜秋, 等. 基于YOLOv5的遥感图像目标检测[J]. 湖南工业大学学报, 2022, 36(3): 44-50.

## Research on a YOLOv5-Baed Remote Sensing Image Target Detection

DONG Lijun, ZENG Zhigao, YI Shengqiu, WEN Zhiqiang, MENG Chen

(College of Computer Science, Hunan University of Technology, Zhuzhou Hunan 412007, China;  
Intelligent Information Perception and Processing Technology Hunan  
Province Key Laboratory, Zhuzhou Hunan 412007, China)

**Abstract:** In order to solve the problem of a complex target background which is hard to identify, as well as a complex target size in remote sensing image target detection task, a remote sensing image detection optimization model based on YOLOv5 has thus been proposed. Firstly, a mosaic enhancement is carried out on the input data to increase the diversity of samples; meanwhile, an adaptive anchor frame calculation is used to obtain the optimal initial value of the anchor frame. Then the feature layer extracted through the backbone network is to be fused, thus obtaining the optimal feature layer, to be followed by an optimization of the positioning loss, with CIoU loss used as the positioning loss function and Focal loss as the classification loss. Finally, adaptive image scaling is used for a reduction in information redundancy and speeding up of model detection. The optimized model can capture image features effectively and

收稿日期: 2021-09-26

基金项目: 国家重点研发计划基金资助项目(2018AAA0100400); 湖南省自然科学基金资助项目(2020JJ6089, 2020JJ6088); 湖南省教育厅科研基金资助项目(21A0350, 21C0439, 19A133)

作者简介: 董丽君(1996-), 女, 湖南衡南人, 湖南工业大学硕士生, 主要研究方向为计算机视觉和模式识别,  
E-mail: dlj15115359886@163.com

通信作者: 易胜秋(1974-), 女, 湖南醴陵人, 湖南工业大学教师, 主要研究方向为机器学习与图像处理,  
E-mail: 447374124@qq.com

achieve a fast and accurate target location. Training tests are carried out on 10 types of geospatial object detection data sets (NWPU-VHR 10) as well as RSOD data sets. Comparative experiments show that, the mAP of the optimized model reaches 0.989 6, which is 2.31% higher than the model prior to the optimization. Compared with the optimization of other models using the same data sets, the mAP is 8.19% higher, verifying the fact that the optimized model can effectively improve the detection accuracy of remote sensing images.

**Keywords:** remote sensing image detection; YOLOv5 algorithm; CIoU loss; Focal loss; mosaic enhancement; adaptive method

## 1 研究背景

遥感技术<sup>[1]</sup>在航空航海动态监控、生态资源环境监测、军事目标检测等众多领域发挥着重要作用。而遥感图像目标检测是遥感图像分析技术的重要一环。近年来,人工智能快速崛起,基于深度学习的目标检测算法<sup>[2]</sup>层出不穷,但是随着工业应用要求的不断严格,提高目标检测的精度仍然是研究的重点方向。传统的目标检测算法利用滑动窗口遍历原始图像,易造成候选区域选择效率低下、窗口冗余等弊端。遥感图像背景复杂难以检测<sup>[3-4]</sup>,而传统目标检测算法只适用于背景简单、特征显著的图像,实际应用有很大的局限性。随着深度学习技术<sup>[5]</sup>的飞速发展,基于深度学习的目标检测算法凭借其结构灵活、自动提取特征和强大的数据处理能力等优势,逐渐替代传统目标检测算法,成为当前目标检测的主流算法。相较于传统目标检测算法,基于深度学习的方法具有性能较高、实际应用场景广、使用方便简洁等诸多优点。

基于深度学习的目标检测算法主要有两种,一是以Fast R-CNN (convolutional neural networks)、Faster R-CNN 等为代表的Two-Stage算法<sup>[6-7]</sup>。这类算法主要分为两个步骤,第一步提取感兴趣区域;第二步结合主干特征,提取网络输出,进行目标精准定位及类别判定。另一种是以SSD (single shot multibox detector)、YOLO (you only look once) 等为代表的One-Stage算法<sup>[8-10]</sup>。这类算法直接在特征提取层做目标检测,没有生成候选区域,大大节省了检测时间,且检测精度没有明显劣势。随着对检测精度及构建更深层次网络的不断追求,国内外研究者做了大量研究,Cheng G.等<sup>[11]</sup>提出一种学习旋转不变卷积神经网络模型RICNN (rotation-invariant convolutional neural networks),该模型在现有CNN结构的基础上引入和学习新的旋转不变结构来提升检测性能。Cheng G.等<sup>[12]</sup>提出了COPD (collection

of part detectors)模型,通过一组代表性的分类器为多类地理空间目标检测提供了一种实用模型,能有效提高分类精度。Li K.等<sup>[13]</sup>从特征融合角度提出了RI-CAO (rotation-insensitive and context-augmented object detection)模型,通过局部上下文特征增强和双通道特征融合网络,提高模型检测性能。而Wang C.等<sup>[14]</sup>通过优化检测框筛选机制提出了DODN (deconvolutional object detection network)模型,该模型通过构建二级反卷积网络取代锚定框机制,再通过区域生成网络 (region proposal network, RPN)生成感兴趣区域,提高了模型检测精度。Ultralytics公司提出了YOLOv5算法,其采用跨阶段局部网络 (cross stage partial network, CSP)作为主要结构,并结合空间金字塔池化 (spatial pyramid pooling, SPP)和路径聚合网络 (path aggregation network, PANet)对网络结构进行了优化,同时数据集进行马赛克增强和更改损失函数为GIoU loss<sup>[15]</sup>,有效地提高了目标检测精度。本文拟将基于YOLOv5算法的优化模型应用于遥感图像数据集,在开源的NWPU-VHR 10数据集<sup>[16]</sup>和RSOD数据集<sup>[17]</sup>上进行训练测试,同时针对遥感数据集的特点进行数据增强,对模型损失函数进行改进,优化目标检测方法,以期提高遥感图像的检测精度。

## 2 YOLOv5 算法介绍

### 2.1 YOLO 算法思想

YOLO算法创造性地将目标检测当作回归问题来解决,直接从特征提取层获取检测框边界和类别概率。与使用RPN网络或滑动窗口的方法不同,YOLO算法将候选区域和检测阶段合二为一,将输入划分成 $s \times s$ 个窗格,每个格子会生成不同长宽比的 $B$ 个边界框 (bounding box)。当目标框中心点落在某个窗格,则此目标的检测由该窗格负责,同时预测出边界框的位置信息、分类概率及置信度 (confidence)。

### 2.2 YOLOv5 算法结构

目前, YOLO 系列中, 最强大的模型为 YOLOv5

模型, 其具有轻量级的模型大小及超高的准确率等优点。具体模型结构如图 1 所示。

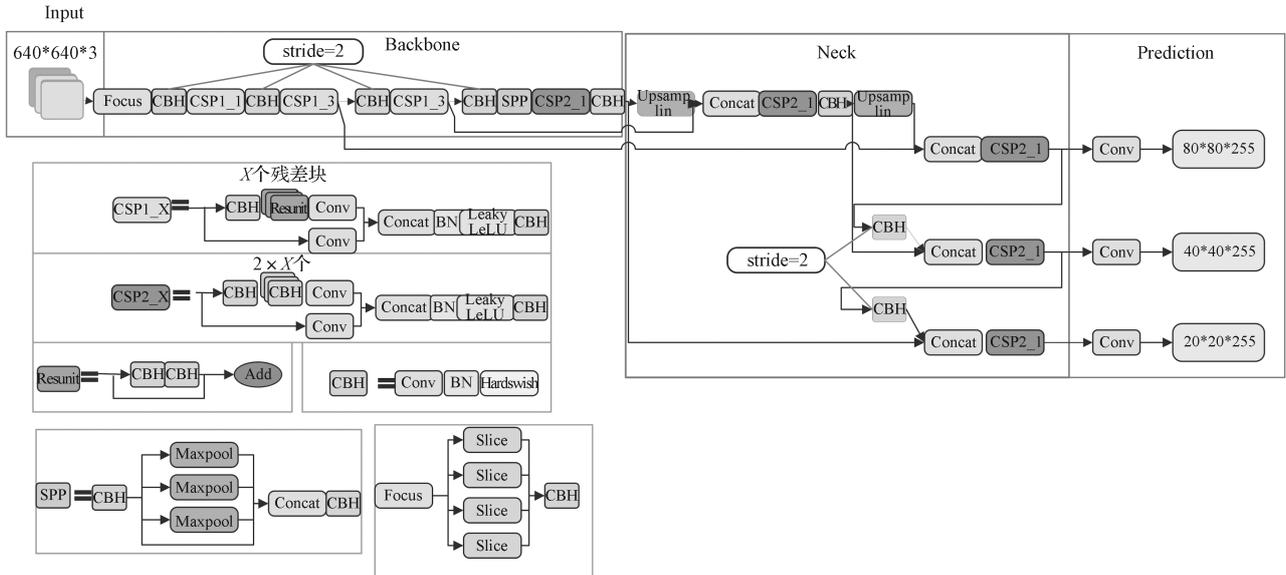


图 1 YOLOv5 模型结构图

Fig. 1 Structure diagram of YOLOv5 model

YOLOv5 共有 YOLOv5s、YOLOv5m、YOLOv5l、YOLOv5x 4 种不同大小的模型, 其中 YOLOv5s 模型最小, 也是速度最快的一个, 其他 3 种模型是在此基础上不同程度地增加模型宽度与深度。为了提高检测精度的同时也保证检测速率, 课题组选择了 YOLOv5s 模型。YOLOv5s 模型在第二次和第三次 CSP 结构中采用了 3 个残差组件, 其余的 CSP 结构只采用一个残差结构, CSP 结构的使用缓解了大量推理时间问题。同时在主干特征提取网络前还加入了 Focus 模块, 主要是对图像进行切片操作, 最大程度减少了信息损失和推理时间, 并进行了下采样操作, 较完整地保留了图像下采样信息。在主干特征提取网络中借鉴了 SPP-Net 中的 SPP 模块, 实现了多重感受野特征的融合, 最大化提高了特征利用率。针对多尺度目标检测问题, 采取了自上而下的侧向特征融合的 FPN 结构和自底向上的 PAN 结构, 以此提高检测能力, 结构如图 2 所示。

课题组选择的 YOLOv5s 模型, 其结构较浅, 在模型不够深的情况下, 通过充足的侧向链接可以保证对不同分辨率的检测目标的检测。特别是对于高空拍摄的遥感图像, 这类图像分辨率高且目标复杂, 而侧向链接有助于解决该类遥感图像的检测难题。

### 3 基于 YOLOv5 的遥感图像检测

#### 3.1 马赛克数据增强

马赛克数据增强原理如图 3 所示。

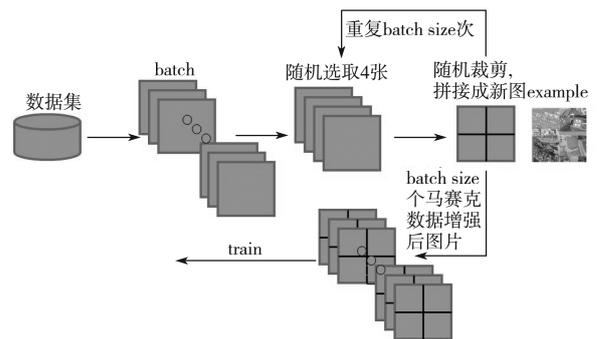


图 3 马赛克数据增强原理图

Fig. 3 Mosaic data enhancement schematic diagram

课题组在数据集 NWPU-VHR 10 和数据集 RSOD 上进行训练和测试, 数据集相对较小, 为了提升检测效果, 采用马赛克数据增强手段, 提高模型的泛化能力。马赛克数据增强是从一个 batch 中随机选取 4 张图片, 将 4 张图片通过随机缩放、裁剪、排布等操作随机拼接成一个设定边长的训练样本。这样做丰富了

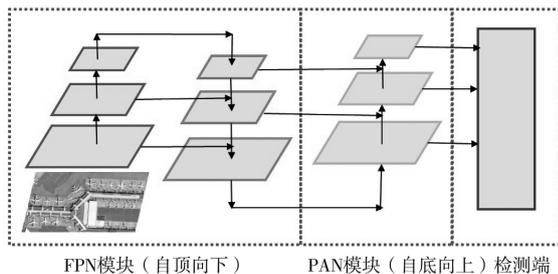


图 2 FPN+PAN 结构图

Fig. 2 Structure diagram of FPN + PAN

数据集,同时增加了许多小目标,对遥感图像检测更有利,能有效缓解遥感图像样本不足的问题,加强了网络鲁棒性。模型可以一次性对4张图像进行训练,这样使batch size不需要很大,减少了对GPU的要求,一个GPU也可以训练出较好的模型。

### 3.2 YOLOv5 算法损失函数的优化

YOLOv5的损失函数包括3个部分,分类损失、定位损失与置信度损失。通过对这3个损失值进行加权相加得到总损失值。

#### 3.2.1 分类损失函数优化

本文采用Focal loss函数作为分类损失函数,在One-Stage detector中会生成成千上万个候选框,但一张图片中的检测目标有限,这就导致大部分的候选框是负样本,带来了严重的正负样本不均匀问题。针对该问题,课题组使用Focal loss作为分类损失函数,其定义如式(1)所示,式中, $\hat{p} \in (0, 1)$ 代表预测值, $y$ 表示真实值, $\gamma > 0$ 为可调节因子。令 $p_t$ 表示预测类别为真实标签的概率,定义如式(2)所示,代入式(1)后Focal loss函数可用式(3)表达。从式(3)可看出Focal loss函数是对交叉熵损失函数的优化,相比于交叉熵损失函数,Focal loss函数多了一个调制因子 $(1-p_t)^\gamma$ ,当 $\gamma=0$ 时,Focal loss就是交叉熵损失。

$$FL = \begin{cases} -(1-\hat{p})^\gamma \log(\hat{p}), & \text{if } y=1, \\ -\hat{p}^\gamma \log(1-\hat{p}), & \text{if } y=0. \end{cases} \quad (1)$$

$$p_t = \begin{cases} \hat{p}, & \text{if } y=1, \\ 1-\hat{p}, & \text{其他。} \end{cases} \quad (2)$$

$$FL = -(1-p_t)^\gamma \log(p_t). \quad (3)$$

该损失函数的主要思想就是给正负样本加上一个权值,度量难分样本和易分样本的损失贡献,调节正负样本损失贡献的权值。图4为不同 $\gamma$ 取值的Focal loss函数曲线对比图。

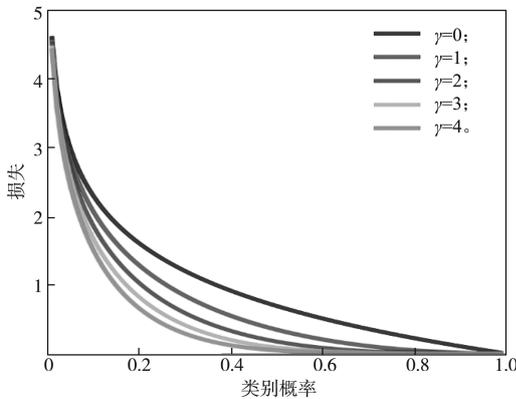


图4 Focal loss 函数曲线

Fig. 4 Focal loss function curve

当类别概率 $p_t \rightarrow 1.0$ 时,表明预测值 $\hat{p}$ 高度接近

真实值 $y$ ,此类样本属于易分样本,这些易分样本大部分是负样本和极少数与真实框高度吻合的正样本,模型能够快速准确地识别,此时 $(1-p_t) \rightarrow 0.0$ ,则以此来降低这些易分样本的权值。对于难分样本 $(1-p_t) \rightarrow 1.0$ ,此时Focal loss逼近于交叉熵函数,总体来说Focal loss函数降低了易分样本的损失贡献,使模型更关注于难分样本。如图所示,调节因子 $\gamma$ 能有效降低简单样本的损失值, $\gamma$ 越大则简单样本的损失值越小。对于置信度损失,也就是有无物体的损失,模型通过输出confidence表示图像中是否有目标,有目标则为1,预测时用sigmoid函数转化为(0, 1)的概率值,最后通过二值交叉熵损失函数(binary cross entropy loss)进行反向传播。

#### 3.2.2 定位损失函数优化

课题组对定位损失函数也进行了优化,采用了更全面的CIoU loss函数。在目标检测中,需要对比检测框与真实框之间的检测效果,YOLOv5原作者采用的GIoU loss,定义如下:

$$IoU = \frac{A \cap B}{A \cup B}, \quad (4)$$

$$GIoU = IoU - \frac{|C - A \cup B|}{|C|}, \quad (5)$$

$$L_{GIoU} = 1 - GIoU. \quad (6)$$

式(4)~(5)中: $A$ 、 $B$ 分别为预测框与真实框; $C$ 为 $A$ 与 $B$ 的最小外接矩; $IoU$ 为真实框与预测框之间的交并比(intersection over union); $GIoU$ 为对 $IoU$ 的优化,考虑了真实框与预测框之间的重叠面积; $L_{GIoU}$ 代表 $GIoU$ 的损失。

当预测框包含于目标框时,GIoU loss退化成了IoU loss,无法区分相对位置。一个好的定位损失函数应该不止考虑到重叠面积,而CIoU还考虑到了距离、尺度以及长宽比。CIoU loss定义如下:

$$v = \frac{4}{\pi^2} \left( \arctan \frac{w^{st}}{h^{st}} - \arctan \frac{w}{h} \right)^2, \quad (7)$$

$$CIoU = IoU - \frac{\rho^2(b, b^{st})}{C^2} - \alpha v, \quad (8)$$

$$L_{CIoU} = 1 - CIoU. \quad (9)$$

式(7)~(9)中: $\alpha$ 为权值; $v$ 为衡量长宽比的相似性; $b$ 、 $b^{st}$ 分别为预测框与真实框; $w$ 、 $h$ 分别为预测框的宽和高; $w^{st}$ 、 $h^{st}$ 分别为真实框的宽与高; $c$ 为两个矩形框的最小外接矩的对角线; $\rho$ 为求两个矩形框中心点的欧氏距离; $CIoU$ 为对 $IoU$ 的一种更全面的优化; $L_{CIoU}$ 为 $CIoU$ 的损失。

定位损失函数如图5所示。

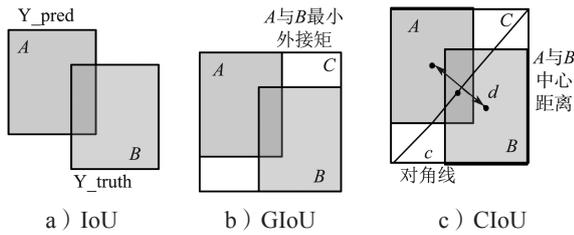


图5 定位损失函数示意图

Fig. 5 Locating loss function diagram

CIoU loss 优点如下:

- 1) 当预测框与目标框不重叠时, 边界框仍然可以学习往目标框移动;
- 2) 直接优化两个目标框的距离, 收敛速度更快;
- 3) 对于预测框包含于目标框的情况, GIoU 恒相等, 无法学习到具体相交情况, 而 CIoU 可以加速框回归;
- 4) 考虑了长宽比, 使预测框更快地回归到目标框。

### 3.3 自适应方法

#### 3.3.1 自适应锚框

由于遥感图像目标尺寸复杂, 因此训练时修改默认锚框为自适应锚框, 生成更符合数据集的锚框。一个锚框由面积 ( $s$ ) 和长宽比 ( $ratio$ ) 定义, 计算公式如下:

$$\begin{cases} w \times h = s, \\ \frac{w}{h} = ratio. \end{cases} \Rightarrow \begin{cases} w = ratio \times h, \\ ratio \times h^2 = s. \end{cases} \quad (10)$$

式中:  $w$  为边框的宽;  $h$  为边框的高。

在训练初期, 如果召回率大于阈值则不改变默认锚框, 如果召回率小于阈值, 则利用  $K$ -mean 均值和遗传算法对数据集进行分析, 自动学习新的预定锚框, 获得更适合数据集的预设锚框。

#### 3.3.2 自适应图片缩放

对于输入图像, 传统的做法是将图像缩放到固定大小再进行后续操作。当图像的长宽比过大或过小时, 则会产生大量的填充区域, 造成信息冗余, 影响模型推理速度。基于遥感图像的复杂性, 采用了自适应图片缩放, 先选择图像的较长边进行缩放到固定尺寸, 本文设定为 640, 按照该比例再对较短边进行缩放, 这样得到一个小于 640 的短边长, 具体公式如下:

$$f = \begin{cases} w, & w > h, \\ h, & \text{others.} \end{cases} \quad (11)$$

$$scale = \frac{640}{f}, \quad (12)$$

$$\begin{cases} \hat{w} = w \times scale, \\ \hat{h} = h \times scale. \end{cases} \quad (13)$$

式中:  $\hat{w}$ 、 $\hat{h}$  均为图像缩放后的尺寸, 其中必定有一个值为 640, 另一个小于 640。

YOLOv5 网络模型中经过了 5 次下采样, 最后的特征层感受野为 32, 为了便于计算机计算, 对小于 640 的短边进行填充, 填充到大于短边同时为 32 的倍数的最小值。这样保留了图像原有的比例信息, 同时也最小化了填充黑边, 降低了原始填充方法带来的信息冗余问题, 在测试时采用自适应图片缩放, 缩减填充, 提高模型推理速度。

## 4 试验结果与分析

### 4.1 试验环境配置及数据准备

本项目是基于 Pytorch 框架, 使用 GPU 进行训练, 试验环境具体配置如表 1 所示。

表 1 试验环境配置

Table 1 Experimental environment configuration

项目	环境
操作系统	Windows 10(x64)
CPU	i5-9300HF CPU@2.40 GHz
GPU(显存大小)	NVIDIA GeForce GTX 1660 Ti (8 GB)
内存	16 GB
Python 版本	3.3.7
CUDA	10.1

本文基于现有的配置, batch size 设为 4, 初始学习率为 0.01, epochs 设置为 1 000。NWPU-VHR10 数据集包括 650 张含目标的图像和 150 张无目标图像, 共 800 张, 10 个类别, 分别为飞机、舰船、油罐、棒球场、网球场、篮球场、田径场、港口、桥梁和汽车。RSOD 数据集共有 976 张图像, 包含飞机、操场、立交桥和油桶 4 个类别。

### 4.2 评价标准

本试验评价体系包括召回率  $R$  (recall), 精准率  $P$  (precision), 平均精准度  $AP$  (average precision) 和均值平均精度  $mAP$  (mean average precision), 各标准的具体公式如下:

$$R = \frac{TP}{TP + FN}, \quad (14)$$

$$P = \frac{TP}{TP + FP}, \quad (15)$$

$$AP = \int_0^1 P_{\text{smooth}}(r) dr, \quad (16)$$

$$mAP = \frac{\sum_{i=1}^K AP_i}{K}. \quad (17)$$

式 (14) ~ (17) 中:  $TP$ 、 $FP$ 、 $TN$ 、 $FN$  分别代表预测为正样本且预测正确、预测为正样本但预测错误、预测为负样本且预测正确和预测为负样本但预测错

误;  $P_{smooth}(r)$  为平滑的  $P-R$  曲线。

### 4.3 试验结果比较

为了更好地评价本次试验效果, 课题组利用其他目标检测算法在数据集 NWPU-VHR10 上进行测试对比, 比较结果如表 2 所示。表中加 \* 数据为该类别  $AP$  最高值。从表 2 中可以看出, 优化后的 YOLOv5s

模型在 NWPU-VHR 10 数据集上取得了较好的结果。优化算法的所有类别精度都接近 1, 其  $mAP$  较优化前的 YOLOv5s 模型高出 2.31%, 同时比其他模型中排第二的 DODN 模型高出 8.19%。同时为了验证模型的泛化能力, 也在数据集 RSOD 上进行了训练测试。

表 2 不同算法在数据集 NWPU-VHR10 上的试验结果

Table 2 Experimental results of different algorithms on dataset NWPU-VHR10

类别	RICNN	COPD	RI-CAO	DODN	baseline	优化后 YOLOv5s
飞机	0.883 5	0.622 5	0.997 0	0.939 2	1.000 0*	0.999 5
轮船	0.773 4	0.693 7	0.908 0	0.929 7	0.944 6	0.998 2*
油罐	0.852 7	0.645 2	0.906 1	0.992 5	0.974 0	1.000 0*
棒球场	0.881 2	0.821 3	0.929 1	0.963 3	0.945 6	1.000 0*
网球场	0.408 3	0.341 3	0.902 9	0.961 2	1.000 0*	1.000 0*
篮球场	0.584 5	0.352 5	0.801 3	0.709 7	0.954 5	1.000 0*
田径场	0.867 3	0.842 1	0.908 1	1.000 0	0.990 5*	0.9857
港口	0.686 0	0.563 1	0.802 9	0.968 8	1.000 0*	1.000 0*
桥梁	0.615 1	0.164 3	0.685 3	0.811 5	1.000 0*	1.000 0*
汽车	0.711 0	0.442 8	0.871 4	0.800 9	0.855 5	0.912 1*
$mAP$	0.726 3	0.548 9	0.871 2	0.907 7	0.966 5	0.989 6

每类的  $P-R$  曲线如图 6 所示。从图中可以得出, 优化后的 YOLOv5s 模型在数据集 RSOD 的  $mAP$  约为 93.07%。其中优化后 YOLOv5s 模型在飞机、操场和油桶这 3 个类别上都有很好的检测精度, 但在立交桥的目标检测精度只有 79.27%, 主要原因是 RSOD 数据集中类别样本不均衡, 导致训练及测试样本不平衡, 虽然进行了数据增强操作, 增加了数据样本, 但立交桥类的样本仅占测试样本集的 1.94%。

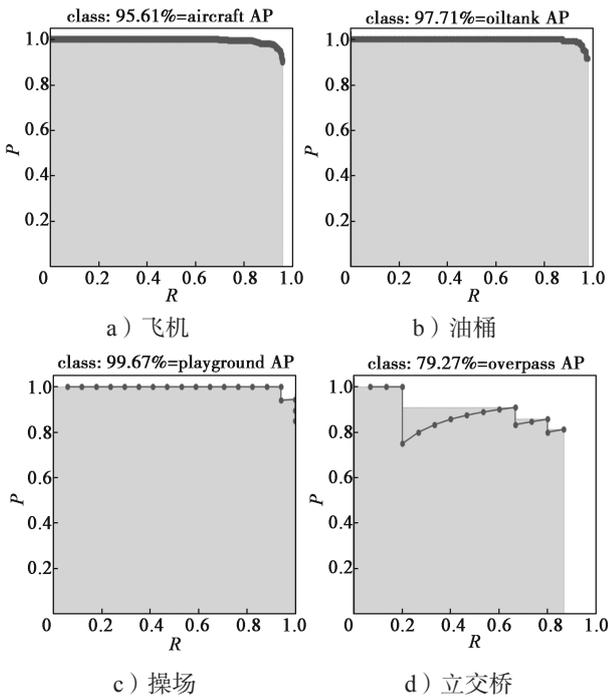


图 6 不同类别的  $P-R$  曲线

Fig. 6  $P-R$  curves of different classes

本文算法对遥感图像小目标的检测效果如图 7 所示。从图中可以看出, 本文基于 YOLOv5s 的优化模型对遥感图像小目标检测具有一定的优越性。

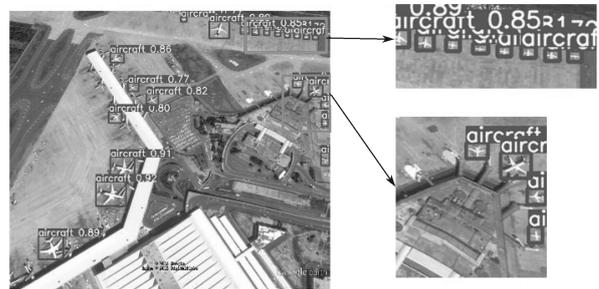


图 7 本文算法对小尺度目标的检测效果图

Fig. 7 Small target detection results with the proposed algorithm adopted

比较其他模型与优化 YOLOv5s 模型之间的均值平均精度, 发现优化 YOLOv5s 模型表现出了超高的性能, 究其根本, 在以下几个方面: 1) 在输入端对训练集进行了马赛克数据增强, 增加了样本多样性, 降低了过拟合, 提高模型泛化能力; 2) 主干特征提取层结合了 CSP 结构与 SPP 结构, 加强了特征提取效率; 3) 在进入检测端前对 3 种不同尺寸的特征图采用了自顶而下的 FPN 结构和自底而上的 PAN 结构, 充分利用了特征提取信息, 使低层次特征具有强语义信息高分辨率等特征, 有利于小目标检测; 4) 采用 Focal loss 的分类损失函数, 降低了正负样本不平衡问题, 使模型更关注于难分样本的损失; 5) 采用 CIoU loss 作为定位损失函数, 同时考虑了目标框之

间的距离、重叠面积、尺度以及长宽比。

## 5 结语

本文针对遥感图像小目标占比较高且目标复杂、目标尺度不一等问题,把优化后的YOLOv5s模型应用于遥感图像的检测任务中,同时应用马赛克数据增强,有效地扩充数据样本。结合自适应的方法降低目标尺寸复杂的影响,选取了CIoU loss作为定位损失函数,提高了模型检测精度。以Focal loss作为分类损失函数,可降低训练时正负样本不均衡问题。试验结果表明,优化后的YOLOv5s模型能有效地检测遥感图像,平均精准度和均值平均精度显著提升。

### 参考文献:

- [1] 王彦情, 马雷, 田原. 光学遥感图像舰船目标检测与识别综述[J]. 自动化学报, 2011, 37(9): 1029-1039.  
WANG Yanqing, MA Lei, TIAN Yuan. State-of-the-Art of Ship Detection and Recognition in Optical Remotely Sensed Imagery[J]. Acta Automatica Sinica, 2011, 37(9): 1029-1039.
- [2] 周晓彦, 王珂, 李凌燕. 基于深度学习的目标检测算法综述[J]. 电子测量技术, 2017, 40(11): 89-93.  
ZHOU Xiaoyan, WANG Ke, LI Lingyan. Review of Object Detection Based on Deep Learning[J]. Electronic Measurement Technology, 2017, 40(11): 89-93.
- [3] TONG X Y, XIA G S, LU Q K, et al. Land-Cover Classification with High-Resolution Remote Sensing Images Using Transferable Deep Models[J]. Remote Sensing of Environment, 2020, 237: 111322.
- [4] 楚玉建, 黎武迪, 高锦辉, 等. 深度学习技术在遥感图像识别中的应用研究[J]. 电脑编程技巧与维护, 2021(4): 134-136.  
CHU Yujian, LI Wudi, GAO Jinhui, et al. Application Research of Deep Learning Technology in Remote Sensing Image Recognition[J]. Computer Programming Skills & Maintenance, 2021(4): 134-136.
- [5] CHA Y J, CHOI W, BÜYÜKÖZTÜRK O. Deep Learning-Based Crack Damage Detection Using Convolutional Neural Networks[J]. Computer-Aided Civil and Infrastructure Engineering, 2017, 32(5): 361-378.
- [6] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation[C]//2014 IEEE Conference on Computer Vision and Pattern Recognition. Columbus: IEEE, 2014: 580-587.
- [7] 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.
- [8] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single Shot MultiBox Detector[J]. Computer Vision, 2016, 9905: 21-37.
- [9] 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. Las Vegas: IEEE, 2016: 779-788.
- [10] REDMON J, FARHADI A. YOLO9000: Better, Faster, Stronger[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu: IEEE, 2017: 6517-6525.
- [11] CHENG G, ZHOU P C, HAN J W. Learning Rotation-Invariant Convolutional Neural Networks for Object Detection in VHR Optical Remote Sensing Images[J]. IEEE Transactions on Geoscience and Remote Sensing, 2016, 54(12): 7405-7415.
- [12] CHENG G, HAN J W, ZHOU P C, 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.
- [13] LI K, CHENG G, BU S H, et al. Rotation-Insensitive and Context-Augmented Object Detection in Remote Sensing Images[J]. IEEE Transactions on Geoscience and Remote Sensing, 2018, 56(4): 2337-2348.
- [14] WANG C, SHI J, YANG X Q, et al. Geospatial Object Detection via Deconvolutional Region Proposal Network[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2019, 12(8): 3014-3027.
- [15] ZHENG Z H, WANG P, LIU W, et al. Distance-IoU Loss: Faster and Better Learning for Bounding Box Regression[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2020, 34(7): 12993-13000.
- [16] CHENG G, HAN J W. A Survey on Object Detection in Optical Remote Sensing Images[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2016, 117: 11-28.
- [17] XIAO Z F, LIU Q, TANG G F, et al. Elliptic Fourier Transformation-Based Histograms of Oriented Gradients for Rotationally Invariant Object Detection in Remote-Sensing Images[J]. International Journal of Remote Sensing, 2015, 36(2): 618-644.

(责任编辑: 申剑)