

引用格式: ZHOU Zilong, ZHOU Jie, LUO Hong, *et al.* Remote sensing image target detection based on Local Adaptive Feature Weighting Algorithm[J]. Remote Sensing Technology and Application, 2024, 39(3): 633-641. [周子龙, 周杰, 罗宏, 等. 基于局部自适应特征加权算法的遥感图像目标检测[J]. 遥感技术与应用, 2024, 39(3): 633-641.]

DOI: 10.11873/j.issn.1004-0323.2024.3.0633

基于局部自适应特征加权算法的遥感图像目标检测

周子龙¹, 周杰¹, 罗宏¹, 徐蕾¹, 邵根富²

(1. 南京信息工程大学 电子与信息工程学院, 江苏 南京 210044;

2. 杭州电子科技大学 通信学院, 浙江 杭州 310000)

摘要: 遥感图像目标检测已成为目标检测领域的重要组成部分。为了解决复杂背景下高分辨率遥感图像中小尺度目标的漏检与误检问题, 结合YOLOv5s算法并在检测端提出一种局部自适应特征加权算法, 通过学习已有的标注框信息将含有目标特征的前景与背景分离, 获得前景中具有关键信息的目标局部特征, 并自适应计算各层局部特征的空间尺度和权重。同时在主干部分引入全局注意力机制用于加强通道与空间之间的跨维度特征信息交互能力, 增强特征间的关联性, 以弥补检测端局部特征的全局信息丢失, 从而降低目标的漏检率和误检率。实验表明: 该改进算法的准确率和召回率均得到一定的提升, 其均值平均检测精度mAP达到72.33%, 相较于传统YOLOv5s算法提升了2.66%。

关键词: 深度学习; 遥感图像; 目标检测; 局部自适应特征加权

中图分类号: TP753 **文献标志码:** A **文章编号:** 1004-0323(2024)03-0633-09

1 引言

遥感图像目标检测^[1-2]是目标检测任务中的重要组成部分。随着卫星技术的快速发展, 遥感卫星的重要性日渐突出, 在自然资源监察, 城市发展, 土地利用, 卫星导航等众多领域应用广泛。遥感卫星不同于其它平台, 易受自然天气的影响, 其主要通过在轨遥感卫星进行对地面目标区域的拍摄, 具有全天候, 全疆域等优势, 但遥感卫星拍摄的图片覆盖面积广, 分辨率高, 目标种类繁多且尺度不一, 导致小目标在图像中占取的像素很少, 在检测时易漏检。研究表明^[3], 遥感图像中像素低于40×40的目标在进行特征提取时, 很可能会混灭在复杂的背景噪声之中。因此提高遥感图像中目标的检测精度具有很高的研究意义和价值。

目标检测方法主要划分为两种类型: 一种是基

于传统的目标检测算法, 另一种是基于深度学习的目标检测算法。其中基于传统的目标检测^[4]通过区域选择、特征提取、分类判别器以及非极大值抑制进行目标定位与分类, 而基于深度学习的算法则是以卷积神经网络为载体进行目标定位与分类。基于深度学习的目标检测算法主要包括单阶段(One-stage)算法和两阶段算法(Two-stage)算法。单阶段算法如SSD^[5]、YOLO系列^[6-7]等具有高速率, 高鲁棒性等优点, 在实时目标检测中应用广泛, 但检测精度不高。两阶段算法如Faster R-CNN^[8]、Cascade R-CNN^[9]等具有高精度, 但检测速率较慢。为了进一步推动单阶段检测算法的发展, 提高模型的检测精度, 黄开启等^[10]提出一种基于YOLOv3的改进算法, 利用K-means聚类算法对锚框尺寸进行调整, 以提升预测框与目标框之间的匹配。提高目标

收稿日期: 2022-11-09; 修订日期: 2024-03-20

基金项目: 国家自然科学基金面上项目(61971167), 国家自然科学基金面上项目(62101275), 国家自然科学基金面上项目(62101274)。

作者简介: 周子龙, 硕士研究生, 主要从事深度学习方面的研究。E-mail: zhouzilong2021@163.com

通信作者: 徐蕾, 博士, 硕士研究生导师, 主要从事通信与信号处理方面的研究。E-mail: xulei@nuist.edu.cn

检测准确率;Wang等^[11]提出了一种改进的YOLOv5s算法,引入通道特征融合与反卷积(CFFI)策略,优化了特征融合过程,减少了特征信息损失,再引入自适应空间特征融合(ASFF)模块,优化了多尺度特征融合能力,提高了对不同尺寸目标的识别精度;Song等^[12]提出了一种基于多尺度特征融合算法,通过有效融合不同尺度的特征,提高模型的检测精度;Lin等^[13]提出了一种新的单阶段端到端算法Retinanet,通过在残差网络中引入新的Focal Loss函数,有效抑制检测端正负样本比例的失衡,降低模型的目标误检率;Liu等^[14]提出了一种深度可分离残差网络,通过将特征送入多层卷积层,获得输出进行局部响应后归一化操作,再将跨层特征进行融合,提高了目标的表征能力。

综上所述,众多学者的研究方向着重于解决普通生活环境下自然图像的目标检测问题,而对遥感图像目标检测的研究相对较少,直接移植自然图像目标检测的检测算法并不能满足遥感图像目标检测的性能要求。实验旨在研究基于YOLOv5s算法的遥感图像目标检测在拥有检测速率快的优势的同时,还能获得较高的目标检测精度。因此,提出了一种局部自适应特征加权算法,将模型检测端的关注点集中于遥感图像中的局部特征,减少复杂背景噪声对目标的影响,同时引入全局注意力模块,提高特征之间的全局信息交互,减少特征信息的丢失,降低目标的误检率和漏检率。

2 算法原理

2.1 YOLOv5s算法

YOLOv5s网络由4个模块构成。第一模块为输入端,利用Mosaic数据增强技术,对高分辨率图像进行随机拼接、缩放和裁剪处理,引入自适应锚框计算,通过聚类算法找到数据集的最优锚框尺寸,引入自适应图像缩放则是将输入的图像统一缩放到设定的尺寸大小。第二模块为Backbone,包括Focus结构,空间金字塔池化结构(Spatial Pyramid Pooling, SPP)^[15]和CSPDarknet53结构,其中Focus结构通过对输入图像分割来增加图像的维度,SPP结构通过引入池化操作,扩展了网络对不同尺度图片的适应能力,从而实现了任意尺度图片的输入处理。CSPDarknet53结构以残差块为主体网络,减少计算量并降低图片传输的损失。第三模块为Neck,主要由PANET网络^[16]构成,PANET网络是

以特征金字塔网络(Feature Pyramid Network, FPN)^[17]为基础并增加额外的特征融合路径,以提高网络的多尺度特征预测的能力。第四模块为预测端,该模块包括两个关键组成部分:边界框损失函数和加权非极大值抑制(NMS)。边界框损失函数采用CIoU(Complete Intersection over Union)损失函数,该函数在IoU(Intersection over Union)损失函数的基础上引入了宽高比惩罚项。加权NMS则用于去除冗余的目标框,以提升目标检测的准确性。

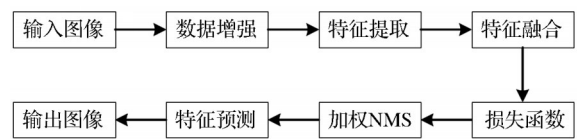


图1 YOLOv5s检测流程图

Fig.1 YOLOv5s detection flow chart

2.2 算法改进

2.2.1 图像预处理

遥感图像目标检测任务中存在图像分辨率高、目标尺寸不一以及局部小目标易漏检的问题。为了解决此问题,在图像预处理阶段使用带有重叠区域(gap)的单尺度分割和多尺度分割方法,其中,单尺度重叠分割放缩比设定为1,gap为300,多尺度重叠分割设定放缩比分别为0.5、1.0和1.5,gap为300来对数据集进行多尺度分割。单尺度重叠分割的使用既减少遥感图像在分割处目标信息的丢失,也更好地捕捉到边缘区域的目标像素,降低边缘区域目标的漏检率。多尺度重叠分割按一定的放缩比例将高分辨率遥感大图像裁剪成多个小分辨率子图像,使模型能够在输入端获得多个尺度不一的目标特征,增加模型鲁棒性的同时与模型主干网络中的FPN进行多尺度联合特征提取,获得更多的深层目标特征语义信息,从而使模型的输出端在特征融合时获得更多细致的目标特征纹理,进而提高目标检测率。如图2所示,输入一张原始分辨率为1100×1100的遥感图像,对其进行多尺度重叠分割处理,经分割后,每一层在不同的尺度分割下生成多张不同尺度的子图片,将其作为输入送入模型。

2.2.2 注意力机制算法优化

目标检测算法中使用注意力机制旨在不影响检测速率前提下,提高目标检测精度。Woo等^[18]提出了一种卷积块注意力机制模块(Convolutional Block Attention Module, CBAM),该模块引入通道

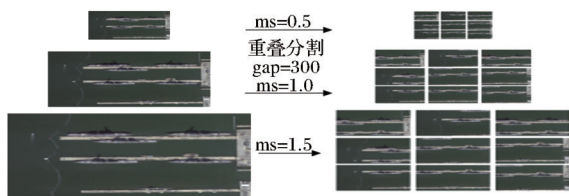


图2 图片预处理后的子图片

Fig.2 Subpicture after pretreatment

和空间注意力机制,但该机制的模块各自独立忽略了通道和空间之间的特征信息交互的重要性,所以本文中引入全局注意力模块(Global Attention Module, GAM)^[19],该模块结合空间注意力机制和通道注意力机制之间的跨维度特征信息交互并引入多层感知机的三维置换,获得更多的目标特征细节信息,加强通道与空间之间的联系。

GAM模块由通道与空间注意力共同组成, F_1 表示输入特征图, $F_1 \in R^{C \times H \times W}$,特征 F_2 、 F_3 分别表示特征 F_1 、 F_2 经通道模块的输出, G_C 表示通道注意力模块, G_S 表示空间注意力模块,公式如(1)和(2):

$$F_2 = \delta \left(\bar{\rho} \left(MLP \left(MLP \left(\rho \left(F_1 \right) \right) \right) \right) \right) \otimes F_1 \quad (1)$$

$$F_3 = \delta \left(f_{Conv}^{7 \times 7} \left(f_{Conv}^{7 \times 7} \left(F_2 \right) \right) \right) \otimes F_2 \quad (2)$$

其中: δ 表示为sigmoid函数; ρ 表示3维排列; $\bar{\rho}$ 表示3维逆排列;MLP表示多层感知机; $f_{Conv}^{7 \times 7}$ 表示由 7×7 的卷积核、批量归一化和ReLU激活函数组成的卷积网络。

2.2.3 局部自适应特征加权算法

在遥感图像目标检测任务中,检测端是影响模型检测性能的关键一步。由于YOLOv5s算法是一个单阶段端到端的检测方法,虽然其检测检测速率较快,但检测端对中小目标的检测能力较差,导致在高分辨率,背景复杂的遥感图像上的目标检测准确率难以达到任务要求,如图7中基线算法与改进算法的Layer23层的特征图对比所示,基线算法的检测端对小尺度目标特征的检测较差,因而在算法检测端引入局部自适应特征加权算法(Local Adaptive Feature Weighting, LAFW)。不同于其他基于特征金字塔^[20]的单阶段检测器依赖于逐元素法或拼接方法来融合多尺度特征的方式,LAFW算法通过将Head中含有真实框标注信息的特征前景与背景进行分离,提取不同尺度特征层中的局部目标特征并进行降维,将其降维至相同的尺度,之后将降维的各层特征尺度进行求均值操作。通过计算出每一个特征层中不同目标所占取得空间尺度大小,

再利用Softmax激活函数对其进行归一化求权重,以获得局部自适应特征的动态权重值。

在FPN层中提取的低层级相同坐标点的目标特征有可能会丢失,LAFW算法在检测端进行特征融合时自动赋予其高层级所学习到的动态权重。如图3所示,设Head模块中经FPN提取的3层特征为 $C_l, l \in \{1, 2, 3\}$,将 C_l 中非 l 层的特征设为 $x_{i,j}^n$ 。由于不同层中的特征具有不同的尺度和通道数,因此,先对特征图进行降维,再求局部特征的平均值。公式如式(3)所示:

$$F_{i,j}^n = \frac{1}{HW} \cdot \sum_{i=1}^H \sum_{j=1}^W |f_{conv}^{1 \times 1} (x_{i,j}^n)| \quad (3)$$

其中: H 表示特征长度; W 表示特征宽度; n 表示该特征层数; (i,j) 表示该特征的位置; $f_{conv}^{1 \times 1}$ 表示 1×1 卷积,降低输入特征的维度, $F_{i,j}^n$ 表示 n 层上 (i,j) 点的局部特征平均值,接着,通过平均值来求该特征的动态权重,公式如下:

$$\alpha_{i,j}^n = \frac{e^{F_{i,j}^n}}{\sum_{k=1}^l e^{F_{i,j}^k}} = \frac{e^{F_{i,j}^n}}{e^{F_{i,j}^1} + e^{F_{i,j}^2} + e^{F_{i,j}^3}} \quad (4)$$

其中: $\alpha_{i,j}^n$ 表示通过Softmax激活函数得到的第 n 层特征的权重值,且 $\alpha_{i,j}^n \in (0, 1)$,同时防止模型过拟合而导致模型性能下降,令 $\sum_{n=1}^3 \alpha_{i,j}^n = 1$,之后利用已获得的特征权重值来求模型检测端的输出 $y_{i,j}^l$,如公式(5)所示:

$$y_{i,j}^l = \sum_{n=1}^3 \alpha_{i,j}^n \cdot x_{i,j}^{n \rightarrow l} \quad (5)$$

$$= \alpha_{i,j}^1 \cdot x_{i,j}^{1 \rightarrow l} + \alpha_{i,j}^2 \cdot x_{i,j}^{2 \rightarrow l} + \alpha_{i,j}^3 \cdot x_{i,j}^{3 \rightarrow l}$$

其中: $x_{i,j}^{n \rightarrow l}$ 表示第 n 层到 l 层的特征映射; $y_{i,j}^l$ 表示模型检测输出层特征; $\alpha_{i,j}^1$ 、 $\alpha_{i,j}^2$ 、 $\alpha_{i,j}^3$ 分别表示自适应提取第1、2、3层特征的空间动态权重。

3 实验结果与分析

3.1 实验平台

本次实验平台的硬件环境是CPU处理器: Intel (R) Core i7-7700HQ CPU@2.8GHZ, GPU处理器: Nvidia A40, 内存为64GB, 显存为48GB, 操作系统为Linux操作系统, 实验所使用的软件环境为: CUDA10.2+ cudnn 7.6.5+ Python 3.9+ Pytorch 1.10.1。

3.2 实验数据集与设置

实验中使用的DOTA数据集,由Google Earth卫星拍摄,数据集由2 806张高分辨率图像组成,包涵188 282个实例,共有15种分类,分别为飞机

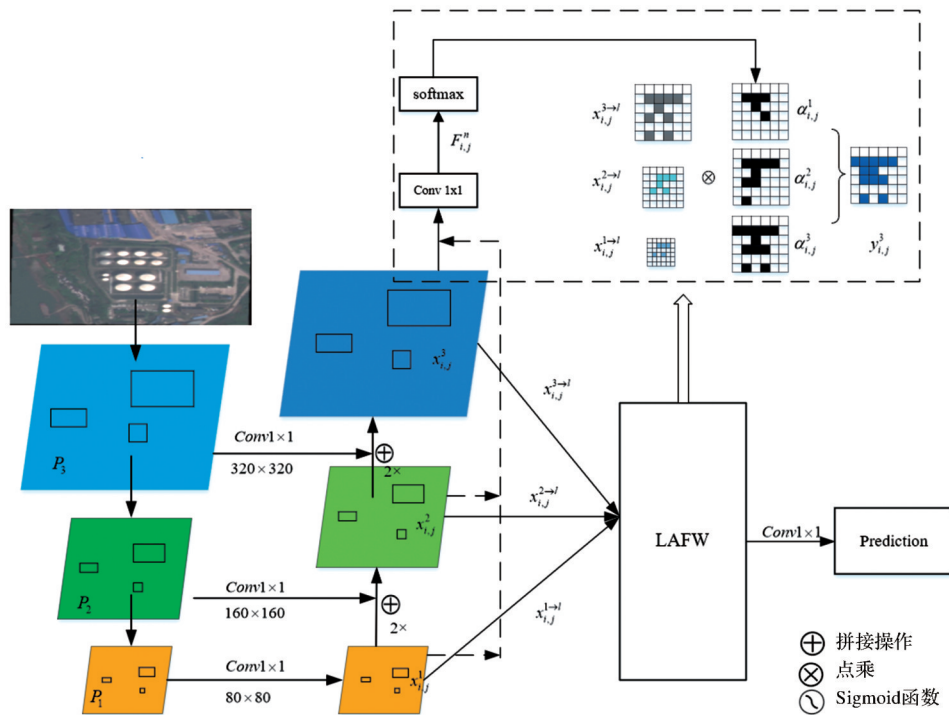


图3 局部自适应特征加权算法示意图

Fig.3 Diagram of local adaptive feature weighting algorithm

(PL)、棒球场(BD)、桥梁(BR)、操场跑道(GTF)、小型车辆(SV)、大型车辆(LV)、轮船(SH)、网球场(TC)、篮球场(BC)、储油罐(ST)、足球场(SBF)、环路(RA)、港口(HA)、游泳池(SP)和直升机(HC)。为了进一步加强目标检测模型的检测准确率和模型的鲁棒性,本文对数据集进行一系列的图像处理,生成两个实验数据集,第一个数据集中设置为单尺度重叠分割,分割尺度比 $r=1.0$,重叠尺度 $gap=300$,获得单尺度重叠分割图像训练集有15749张图片,验证集有5297张图片,测试集有923张图片;另一个数据集中设置多尺度重叠分割,分割尺度比分别为 $r=0.5, 1.0$ 和 1.5 ,重叠尺度 $gap=500$,获得多尺度分割图像训练集有103421张图片,验证集有35462张图片,测试集有5023张图片,其中训练集和验证集是严格独立的。

实验中设置输入图片的大小均为 640×640 ,设置Batch-size为16,训练迭代次数为200,初始学习率 lr 为0.01,学习率的衰减权重为0.0005,采用SGD函数优化模型梯度,采用余弦退火算法来更新模型的学习率。

3.3 模型评价指标

在遥感图像目标检测的结果分析中使用平均精度(Average Precision, AP)和均值平均精度(Mean Average Precision, mAP)作为评估检测模型

性能优异的重要指标。其中平均精度(AP)表示模型目标检测的平均性能,均值平均精度(mAP)表示模型对目标检测精度(Precision, P)和召回率(Recall, R)的综合能力衡量,其中精度P表示模型正确预测的实例占全部预测为真的实例的概率,召回率R表示模型检出为真的实例占全部正实例的概率。

平均精度(AP)公式如式(6)所示:

$$AP = \int_0^1 P(R) d(R) \quad (6)$$

均值平均精度(mAP)公式如式(7)所示:

$$mAP = \frac{1}{n} \sum_0^n AP \quad (7)$$

精度P和召回率R的计算公式如公式(8)和(9)所示:

$$P = \frac{TP}{TP + FP} \quad (8)$$

$$R = \frac{TP}{TP + FN} \quad (9)$$

其中:TP(True Positive)表示模型预测为真且实例也为真的样本个数;FP(False Positive)表示模型预测为真但实例为假的样本个数;FN(False Negative)表示模型预测为假但实例为真的样本个数。

3.4 消融实验

为了验证改进算法的有效性,在基线算法的基础上分别引入GAM模块和LAFW模块以及两个

模块均添加。如表1所示,在单独添加GAM模块的精度P、召回率R、mAP@0.5和mAP@0.5:0.95较基线算法分别提升了-0.39%、+0.28%、-0.11%和+0.53%,由于加入GAM模块,提高了通道和空间之间的信息交互,虽然模型的精度有所下降,但召回率上升,由mAP@0.5:0.95的值得出部分小目标的检出率是提高了。单独添加LAFW模块的精度P、召回率R、mAP@0.5和mAP@0.5:0.95较基线算法分别提升了+1.86%、+0.61%、+0.53%和+1.14%,表明模型在检测端提高对局部目标的捕捉。两模块均添加的精度P、召回率R、mAP@0.5和mAP@0.5:0.95较基线算法分别提升了+4.37%、+0.93%、+2.66%和+1.83%,表明模型能够更好地捕捉局部小目标的信息并在检测端将其检测筛选出来。实验结果表明,该改进算法是行之有效的。

表1 改进算法的消融实验

Table 1 Ablation Experiment of improved algorithms

基线 算法	GAM	LAFW	P/%	R/%	mAP@ 0.5/%	mAP@0.5: 0.95/%
√	×	×	74.86	66.95	69.67	43.37
√	√	×	74.47	67.23	69.56	44.57
√	×	√	76.72	67.56	70.20	44.51
√	√	√	79.23	67.88	72.33	45.20

3.5 实验结果

为了验证改进算法在目标检测中的优势。首先,通过运行结果获得各模型的精度和召回率,为了直观地展示改进算法的优势,如图4~图5中所示的各模型准确率的比较,以200个epochs为横轴,精度P和召回率R分别为纵轴。如图4所示,从第50 epoch开始,LAFW算法的精度均比基线算法以及模型1和模型2高2%左右,由于LAFW算法能够自适应的赋予不同目标特征以合适的空间权重,降低小目标的漏检,从而获得更高的准确率。

如图5所示,改进算法的召回率高于其他算法约0.8%,由于LAFW算法通过真实框将局部特征与背景噪声分离,降低冗余框对检测的影响,从而获得较高的召回率。结合准确率和召回率表明了改进算法在保持较高的检测准确率的同时,也能够保持较高的召回率。为了进一步表示LAFW算法的优势,将对比实验分为单尺度训练模型对比和多尺度训练模型对比,其中单尺度训练对比结果如表1上半表所示,模型1为YOLOv5s+CBAM模

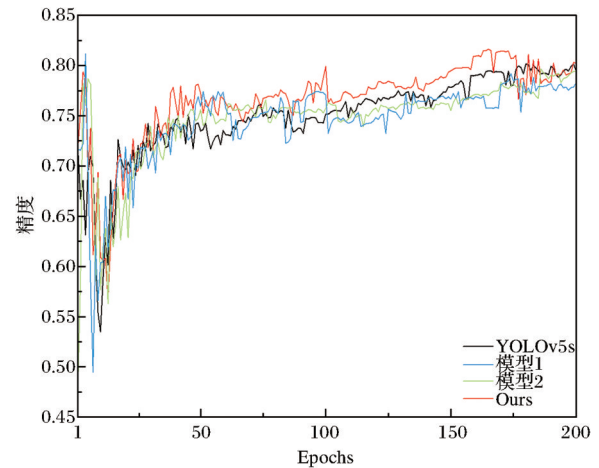


图4 各模型的精度比较

Fig.4 Comparison of precision of each model

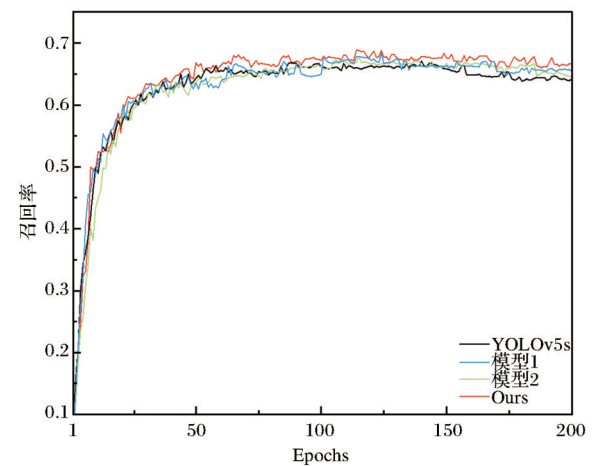


图5 各模型的召回率比较

Fig.5 Comparison of recall of each model

块;模型2为在YOLOv5s+GAM模块;改进算法在引入GAM模块的基础上在检测端融入局部自适应特征加权算法。结果表明,改进算法的mAP较TO-SO^[21]、PIoU^[22]、模型1、模型2、DAL^[23]分别提高了12.33%、3.68%、2.6%、2.77%和0.55%。多尺度训练结果对比如表1的下半部分所示,改进后的算法经多尺度训练后的平均检测精度mAP较ROI-Trans^[24]、P-RSDet^[25]、SCRDet^[26]、R³Det^[27]、S²A-Net^[28]分别提升了4.98%、4.78%、1.98%、0.88%和0.46%。

利用Grad-CAM算法作为可视化工具对不同模型的检测进行分析。如图6所示,a组红色方框内标注的是局部小目标区域,b、c、d组是不同模型的热力图结果,直观地表示不同模型在该区域内的检测性能。通过c组热力图可视化结果分析,引入GAM模块后,相较于a组的传统算法能够更好的定位中大尺度目标信息并且能够获取部分区域内小

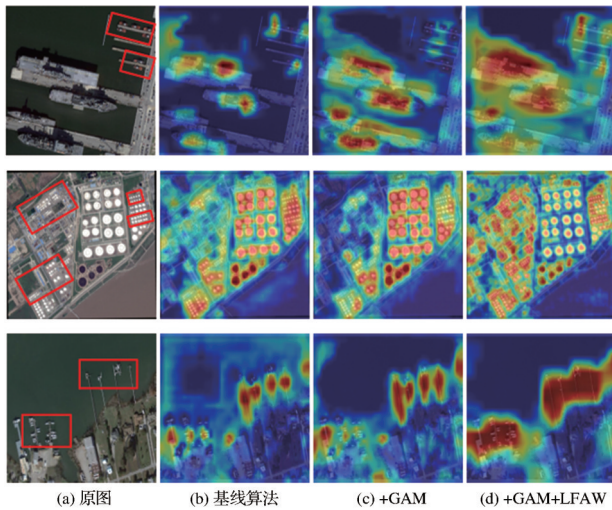


图6 不同算法下的热力图可视化对比图

Fig.6 Comparison of heatmap visualization under different algorithms

目标的定位信息,但对小目标关键像素点的定位并不显著。在此基础上加入LAFW算法,如d组所示,相较于前两种算法,模型的Neck端利用特征金字塔层对不同层不同尺度的特征进行自适应加权,一方面,LAFW算法进一步加强了对中大尺度目标的信息定位能力,另一方面,通过对局部特征加权,使模型获得更多局部特征的关键像素点,并对该像素点进行加权,进而获得更加精准的局部小目标的定位信息,从而提高目标检测的准确率。

3.6 检测结果对比

如图7检测结果所示,通过对ABCD 4组的遥

感图像进行检测,图像类别分别包括密集油罐区,停车场,海面船只与港口以及飞机等,经统计得4组中小尺度目标占比分别为22.5%、10%、16.32%和20.68%。每一组中第一行为基线算法的特征图,第二行则是改进算法的特征图。分析Layer3、Layer20和Layer23层可视化特征图,在特征图7(b)中,LAFW算法和基线算法在对目标特征区域的特征较为一致;在特征图7(c)中,LAFW算法对目标特征的关键区域像素点的提取比基线算法更好;在特征图7(d)中,LAFW算法对局部目标特征自适应赋予权重,使目标特征区域更加突出,特征边缘更加清晰,小尺度目标特征得到很好的保留,相邻区域的噪声点也得到很好的抑制。A组中改进算法较基线算法检测准确率提升16%,小尺度目标检测准确率提高55.55%,在A组中目标密集,大小目标夹杂在一起,且目标形状类似,使得基线算法漏检较高,但改进算法通过加入GAM模块加强通道和空间之间的特征信息交互;B组的改进算法较基线算法的检测准确率分别提升3.2%和12.50%,由于检测器加强对图像边缘处目标的检测;C组的改进算法较基线算法的检测准确率分别提升了26%和26.66%,在特征融合时,赋予局部特征一个动态权重值,使得小目标可以保留特征信息;D组的改进算法较基线算法检测准确率分别提升9.7%和50%,通过对FPN的各层进行自适应特征加权,突出图像中局部小目标的特征,使边缘处目标信息更加

表2 各模型在DOTA数据集上的比较

Table 2 Comparison of models in DOTA dataset

方法	BB	PL	BD	BR	GTF	SV	LV	SH	TC	BC	ST	SBF	RA	HA	SP	HC	mAP
单尺度训练																	
TOSO	R-101	80.17	65.59	39.82	39.82	49.71	65.01	53.58	81.45	44.66	78.51	48.85	56.73	64.4	64.25	36.75	60.50
PIoU	DLA34	80.90	69.70	24.10	60.20	38.30	64.40	64.80	90.90	77.20	70.40	46.50	37.10	57.10	61.90	64.00	68.65
模型1	C3	91.40	72.22	48.20	62.59	67.30	85.00	86.90	92.91	58.10	68.13	53.40	61.41	83.00	60.52	54.90	69.73
模型2	C3	91.50	73.51	47.43	62.63	64.92	85.73	86.55	92.64	61.76	68.81	50.71	57.24	84.23	61.32	54.43	69.56
DAL	R-101	88.61	79.69	46.27	70.37	65.89	76.10	78.53	90.84	79.98	78.41	58.71	62.02	69.23	71.32	60.65	71.78
Ours	C3	92.09	76.53	51.66	67.58	69.79	88.98	88.25	93.08	65.79	72.55	54.63	58.97	84.95	62.79	57.37	72.33
多尺度训练																	
ROI-Tran	R-101	88.6	78.5	43.4	75.9	68.8	73.7	83.6	90.7	77.3	81.5	58.4	53.5	62.8	58.9	47.7	69.6
P-RSDet	R-101	89.0	73.7	47.3	72.0	70.6	73.7	72.8	90.8	80.1	81.3	59.5	57.9	60.8	65.2	62.2	69.8
SCRDet	R-101	90.0	80.7	52.1	68.4	68.4	60.3	72.4	90.9	87.9	86.9	65.0	66.7	66.3	68.2	65.2	72.6
R ³ Det	R-101	89.5	81.2	50.5	66.1	70.9	78.7	78.2	90.8	85.3	84.2	61.8	63.8	68.2	69.8	67.2	73.7
S ² A-Net	R-50	89.11	82.84	48.37	71.11	78.11	78.39	87.25	90.83	84.90	85.64	60.36	62.60	65.26	69.13	57.94	74.12
Ours*	C3	93.71	80.40	52.42	66.70	72.23	84.62	88.10	94.43	72.20	76.26	63.40	65.10	85.02	60.34	63.81	74.58

注:BB表示主干网络,R-101表示ResNet-101,R-50表示ResNet-50,C3表示CSPDarkNet53,DLA34表示DLA框架中的一种,*表示进行多尺度训练与测试

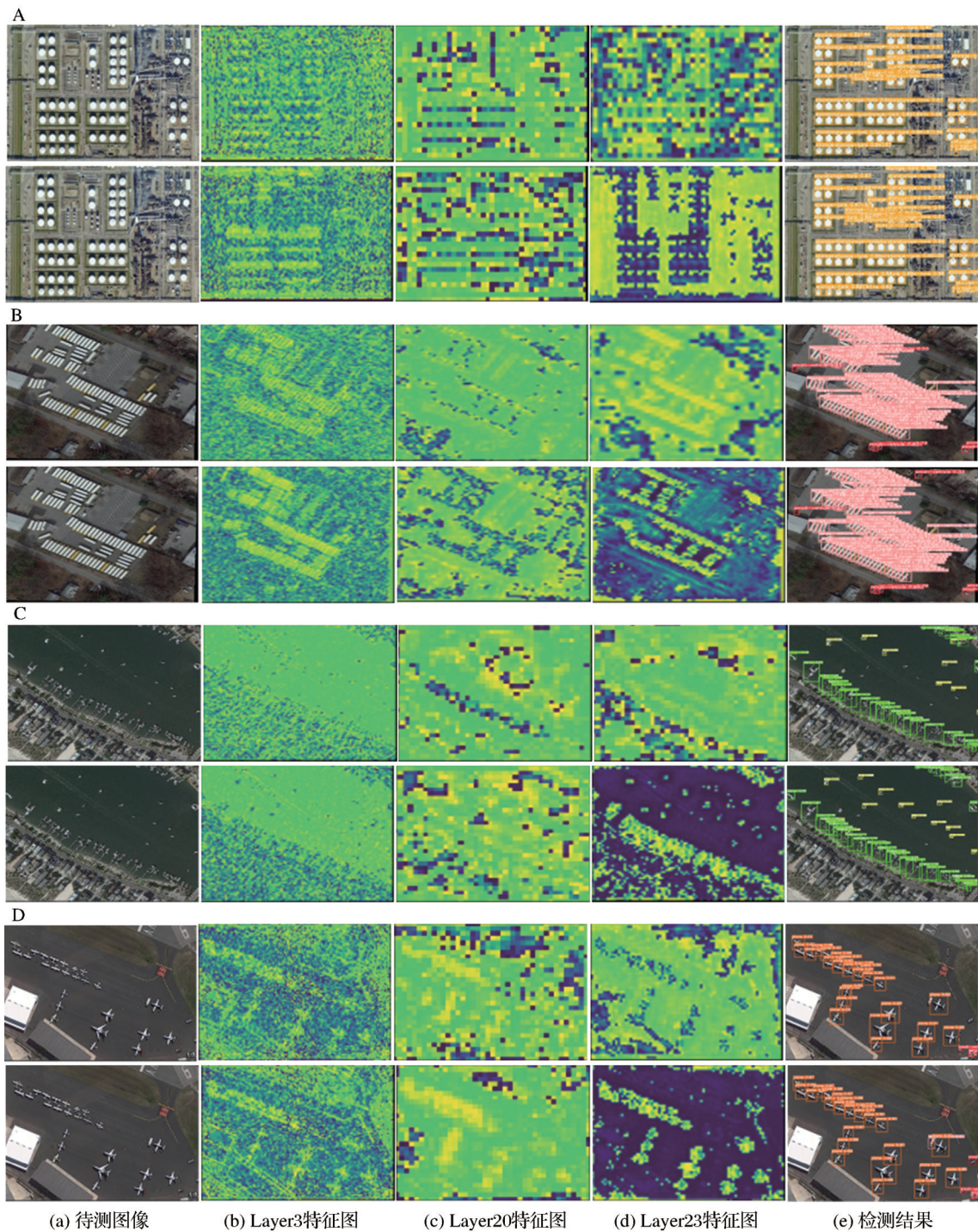


图7 改进实验结果对比

Fig.7 Comparison of improved experimental results

清晰。

4 结 论

本研究中,针对遥感图像复杂的背景、较高的分辨率以及目标尺度不一的问题。首先,在图像预处理中通过单尺度和多尺度重叠分割,利用不同的分割比例以及不同的重叠度提高不同尺度的目标

表征能力,同时降低分割边缘处目标的丢失,提高目标的检出率。接着,在主干部分引入全局注意力机制,通过加强通道和空间之间的跨维度信息交互,获得更加详细的目标特征信息,从而弥补模型检测端局部特征信息的丢失。然后,提出了一种局部自适应特征加权算法,通过自适应学习不同特征层的局部特征的空间权重,并且在特征融合时给予

不同层不同尺度的目标特征更高的空间动态权重,使网络获得更加精准的目标关键像素,提高模型的检测精度。最后,通过对几种不同模型的精度P、召回率R和均值平均精度mAP的对比分析并结合特征可视化算法,表明本文所提出的改进算法能够有效降低小尺度目标的漏检率与误检率,提高遥感图像目标检测的准确率。通过相关检测结果分析表明针对遥感图像目标检测的改进算法是有效的,具体结果表述如下:

(1)通过消融实验分析,改进算法的精度和召回率较基线算法分别提高4.37%和0.93%,表明改进算法对中小目标的检测和召回性能的优异。

(2)通过各模型实验结果的分析,在单尺度模型训练下,改进算法的mAP相较于TOSO、PIoU算法分别提升了12.33%、3.68%。在多尺度模型训练下,改进算法的mAP相较于ROI-Trans、P-RSDet、SCRDet、R³Det、S²A-Net分别提升了4.98%、4.78%、1.98%、0.88%和0.46%,表明了改进算法能够有效提取复杂背景下的目标细粒度特征,并通过检测端检测出来。

(3)通过图像检测结果以及小尺度目标检测结果分析,表明了改进算法在复杂背景下能够更好的捕获图像边缘处的特征,抑制背景噪声,在检测端能够更好地抓取中小尺度目标的特征,提高了模型的检测精度。

为了进一步提高遥感图像目标检测的检测精度,在后续的研究提出以下两点展望:

(1)由于遥感图像具有较高的分辨率,受显存等因素的影响,现有硬件设备难以直接处理全尺寸的遥感图像,针对这一问题,未来研究可以考虑设计新的图像采样策略,以缓解类别间不平衡的现象。

(2)随着大模型AI时代的快速发展,PC端、服务器端的算法部署已不能满足当前社会的要求,移动设备的使用是今后的发展潮流,因此需考虑在算法剪枝与轻量化方向进行进一步的研究探索。

参考文献(References):

- [1] JIANG D, NAN X, XIA G S, *et al.* Object detection in aerial images: A large-scale benchmark and challenges[C]//2022 IEEE Transactions on Pattern Analysis and Machine Intelligence, 44:7778-7796. DOI:10.1109/TPAMI.2021.3117983
- [2] FU Han, FAN Xiangtao, YAN Zhenzhen, *et al.* Research progress of remote sensing image target detection based on deep learning [J]. Remote Sensing Technology and Application, 2022, 37(2): 290-305.[付涵,范湘涛,严珍珍,等.基于深度学习的遥感图像目标检测技术研究进展[J].遥感技术与应用, 2022, 37(2): 290-305.]
- [3] ZOU Z, CHEN K, SHI Z, *et al.* Object detection in 20 Years: A survey [J]. Proceedings of the IEEE, 2023, 111(3): 257-276. DOI: 10.1109/JPROC.2023.3238524
- [4] CHANDRA M A, BEDI S S. Survey on SVM and their application in image classification[J]. International Journal of Information Technology, 2021, 13: 1-11.
- [5] LIU W, ANGUELOV D, ERHAN D, *et al.* Ssd: Single shot multibox detector[C]//2016 14th European Conference Computer Vision, 2016: 21-37. DOI: 10.1007/97833194644802
- [6] REDMON J, FARHADI A. YOLOv3: An incremental improvement[S]. arXiv preprint arXiv:1804.02767, 2018. DOI: 10.48550/arXiv.1804.02767
- [7] ZHENG G, LIU S T, WANG F, *et al.* YOLOX: Exceeding YOLO series in 2021[S]. arXiv preprint arXiv: 2107.08430, 2021. DOI: 10.48550/arXiv.2107.08430
- [8] REN S, HE K, GIRSHICK R, *et al.* Faster R-CNN: Towards real-time object detection with regional proposal networks [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6): 1137-1149. DOI: 10.1109/TPAMI.2016.2577031
- [9] HUANG H, LI L, MA H. An improved Cascade R-CNN based target detection algorithm for UAV aerial images[C]//2022 7th International Conference on Image, Vision and Computing, 2022: 232-237. DOI: 10.1109/ICIVC55077.2022.9886321
- [10] HUANG Kaiqi, LIU Xiaorong, HUANG Maoyun. Research on small target detection method based on improved YOLOv3 [J]. Transducer and Microsystem Technologies, 2022, 41(4): 52-55.[黄开启,刘小荣,黄茂云.基于改进YOLOv3的小目标检测方法研究[J].传感器与微系统, 2022, 41(4): 52-55.]
- [11] WANG A, PENG T, CAO H, *et al.* TIA-YOLOv5: An improved YOLOv5 network for real-time detection of crop and weed in the field [J]. Frontiers in Plant Science, 2022, 13: 1091655. DOI:10.3389/fpls.2022.1091655
- [12] SONG Z Y, ZHANG Y, LIU Y, *et al.* MSFYOLO: Feature fusion based detection for small objects [J]. IEEE Latin America Transactions, 2022, 20(5): 823-830. DOI: 10.1109/TLA.2022.9693567
- [13] LIN T, GOYAL P, GIRSHICK R, *et al.* Focal loss for dense object detection [C]//IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 42(2): 318-327. DOI: 10.1109/TPAMI.2018.2858826
- [14] LIU R, YU Z, MO D, *et al.* An improved Faster-RCNN algorithm for object detection in remote sensing images [C]//2020 39th Chinese Control Conference, 2020: 7188-7192. DOI: 10.23919/CCC50068.2020.9189024
- [15] HE K M, ZHANG X Y, REN S Q, *et al.* Spatial pyramid pooling in deep convolutional networks for visual recognition [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37(9): 1904-1916. DOI: 10.1109/TPAMI.2015.2389824
- [16] LIU S, QI L, QIN H F, *et al.* Path aggregation network for instance segmentation [C]//IEEE Conference on Computer

- Vision and Pattern Recognition, 2018: 8759–8768. DOI: 10.1109/CVPR.2018.00913
- [17] LIN T, DOLLAR P, GIRSHICK R, *et al.* Feature pyramid networks for object detection[C]//IEEE Conference on Computer Vision and Pattern Recognition, 2017: 936–944. DOI: 10.1109/CVPR.2017.106
- [18] WOO S Y, PARK J C, LEE J Y, *et al.* CBAM: convolutional block attention module[C]//Proceedings of the European Conference on Computer Vision, 2018: 3–19. DOI: 10.1007/978-3-030-01234-2_1
- [19] LIU Y C, SHAO Z R, NICO H. Global attention mechanism: Retain information to enhance channel spatial interactions[S]. arXiv preprint arXiv:2112.05561, 2021. DOI: https://arxiv.org/abs/2112.05561
- [20] ZHANG Shen, LI Shanshan, WEI Guofang, *et al.* Object detection in remote sensing images with fine-scale multi-scale features [J]. Journal of Remote Sensing, 2023, 26 (12) : 2616–2628. [张省, 李山山, 魏国芳, 等. 面向精细化多尺度特征的遥感图像目标检测[J]. 遥感学报, 2023, 26(12): 2616–2628.]
- [21] FENG P M, LIN Y T, GUAN J, *et al.* TOSO: Student's-t distribution aided one stage orientation target detection in remote sensing images [C]//2020 IEEE International Conference on Acoustics, Speech and Signal Processing, 2020: 4057–4061. DOI: 10.1109/ICASSP40776.20.20.9053562
- [22] ZHOU L, WEI H, HAO L, *et al.* Objects detection for remote sensing images based on polar coordinates[J]. arXiv preprint arXiv:2001.02988, 2021. DOI:10.48550/arXiv.2001.02988
- [23] MING Q, ZHOU Z, MIAO L, *et al.* Dynamic anchor learning for arbitrary oriented object detection[C]//Proceedings of the AAAI Conference on Artificial Intelligence, 2021, 35 (3): 2355–2363. DOI:10.1609/aaai.v35i3.16336
- [24] DING J, XUE N, LONG Y, *et al.* Learning roi transformer for oriented object detection in aerial images[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2019: 2849–2858. DOI: 10.1109/CVPR.2019.00296
- [25] ZHOU L, WEI H, LI H, *et al.* Arbitrary oriented object detection in remote sensing images based on polar coordinates [J]. IEEE Access, 2020, 8: 223373–223384. DOI: 10.1109/ACCESS.2020.3041025
- [26] YANG X, YANG J, YAN J, *et al.* Srdet: Towards more robust detection for small, cluttered and rotated objects[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. 2019: 8232–8241. DOI: 10.1109/ICCV.2019.00832
- [27] YANG X, YAN J, FENG Z, *et al.* R3det: Refined single stage detector with feature refinement for rotating object[C]//Proceedings of the AAAI Conference on Artificial Intelligence, 2021, 35(4): 3163–3171. DOI:10.1609/aaai.v35i4.16426
- [28] HAN J, DING J, LI J, *et al.* Align deep features for oriented object detection[J]. IEEE Transactions on Geoscience and Remote Sensing, 2022, 60: 1–11. DOI:10.1109/TGRS.2021.3062048

Remote Sensing Image Target Detection based on Local Adaptive Feature Weighting Algorithm

ZHOU Zilong¹, ZHOU Jie¹, LUO Hong¹, XU Lei¹, SHAO Genfu²

(1. School of Electronics and Information Engineering, Nanjing University of Information Engineering, Nanjing 210044, China;

2. College of Communication Engineering, Hangzhou 310000, China)

Abstract: Object detection in remote sensing images has become a vital aspect of the overall object detection domain. To address the problems of missed detection and false detection of small-scale objects in high-resolution remote sensing images with complex backgrounds, a local adaptive feature weighting algorithm is proposed at the detection stage, combined with the YOLOv5s algorithm. By learning the existing label box information, the foreground containing target features is separated from the background, and the local features of the target with key information in the foreground are obtained. The spatial scale and weight of local features of each layer are calculated adaptively. Meanwhile, a global attention mechanism is proposed to enhance the interaction capability of cross-dimensional feature information between channels and spatial dimensions in the backbone, so as to strengthen the correlation between features and compensate for the loss of global information within local features at the detection stage, thereby reducing the rates of missed detection and false detection of targets. Experimental results show that the improved algorithm achieves certain improvements in precision and recall, with a mean average precision reaching 72.33%, representing an increase of 2.66% compared to the traditional YOLOv5s algorithm.

Key words: Deep learning; Remote sensing image; Target detection; Local adaptive feature weighting