

基于边缘特征引导学习的SAR目标检测

倪康^{1,2}, 孙笠焜¹, 邹旻瑞¹

(1. 南京邮电大学计算机学院, 南京 210023; 2. 南京航空航天大学雷达成像与微波光子技术教育部重点实验室, 南京 211106)

摘要: 合成孔径雷达 (Synthetic aperture radar, SAR) 图像目标通常具有不明显的边缘特征, 且不同尺度下的目标边缘特征并不完全相同。边缘特征可以提供目标物体的形状和轮廓信息, 增强模型对目标物体的定位能力。现有的 SAR 目标检测方法对于边缘特征的学习仍然不足, 导致模型对于边缘特征的感知能力较弱。基于此, 提出一种基于边缘特征引导学习 (Edge feature guided learning, EFGL) 的 SAR 目标检测方法, 该方法基于 FCOS (Fully convolutional one-stage) 目标检测框架, 主要利用目标边缘特征引导特征金字塔网络 (Feature pyramid networks, FPN) 的特征学习, 通过在 FPN 中引入边缘算子模块, 显式增强网络对不同尺度目标边缘特征的学习能力; 另外, 在多尺度特征融合过程中, 构建边缘特征引导融合模块, 利用融合边缘特征的空间注意力模块实现边缘特征引导的相邻层级特征融合。在 MSAR 数据集和 SAR-Aircraft-1.0 数据集上, 所提方法在 AP^{'07} 标准下分别达到了 68.68% 和 67.44% 的检测精度, 比基础网络分别提升了 1.34% 和 4.81%。与其他相关算法相比, 所提方法能够更好地进行目标定位, 且 SAR 目标检测性能更优。

关键词: 合成孔径雷达; 目标检测; 深度学习; 注意力机制; 特征融合

中图分类号: TP753 **文献标志码:** A

SAR Target Detection Based on Edge Feature Guided Learning

NI Kang^{1,2}, SUN Likun¹, ZOU Minrui¹

(1. School of Computer Science, Nanjing University of Posts and Telecommunications, Nanjing 210023, China; 2. Ministry of Education Key Laboratory of Radar Imaging and Microwave Photonics Technology, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China)

Abstract: Synthetic aperture radar (SAR) image targets typically exhibit subtle edge features, which can vary across different scales. Edge features provide crucial information about the shape and contour of target objects, improving the model's localization capabilities. However, existing SAR object detection methods often underperform in learning edge features, limiting their ability to accurately perceive target edges. To address this, we propose a SAR target detection method based on edge feature guided learning (EFGL). This approach builds upon the fully convolutional one-stage (FCOS) object detection framework and leverages edge features to guide the learning process in feature pyramid networks (FPN). By integrating an

基金项目: 国家自然科学基金 (62101280); 江苏省自然科学基金 (BK20210588); 中国博士后科学基金 (2023M731781); 江苏省航空对地探测与智能感知工程中心开放基金 (JSECF2023-05); 雷达成像与微波光子技术教育部重点实验室 (南京航空航天大学) 基金 (NJ20230005)。

收稿日期: 2024-04-26; **修订日期:** 2024-07-18

edge operator module into FPN, the network's capacity to learn multi-scale edge features is explicitly enhanced. Additionally, during multi-scale feature fusion, we introduce an edge feature-guided fusion module that incorporates a spatial attention mechanism to enable edge-guided fusion across adjacent feature levels. On the MSAR and SAR-Aircraft-1.0 datasets, the proposed method achieves detection accuracies of 68.68% and 67.44% under the AP'07 standard, showing improvements of 1.34% and 4.81% over the baseline network, respectively compared to other related algorithms, this method demonstrates superior target localization and overall performance in SAR target detection.

Key words: synthetic aperture radar (SAR); target detection; deep learning; attention mechanism; feature fusion

引 言

合成孔径雷达(Synthetic aperture radar, SAR)是主动式高分辨率微波成像雷达,具有全天候、全天时工作的特性。该雷达利用合成孔径原理、脉冲压缩技术和先进的信号处理方法,通过真实的小孔径天线实现双向高分辨率遥感成像。随着数字信号处理等相关算法的不断发展,SAR系统可在不同频段、不同极化下获取目标的高分辨率雷达图像,可提供重要的空地海目标信息。因此,SAR在军事、经济和科技等领域得到了广泛应用。

SAR目标检测是SAR图像智能解译的关键部分,其在军事目标打击、海洋资源管理等领域发挥着重要作用。目前,基于恒虚警率(Constant false alarm rate, CFAR)的SAR目标检测方法应用广泛。该方法基于以下假设:在背景杂波情况下,目标反射的信号强度会明显高于背景杂波。CFAR算法的目标是在不知道目标信号强度的情况下,通过统计分析来确定适当的门限值,以便在保持虚警率恒定的同时检测到目标。文献[1]提出了一种基于双边修剪统计的鲁棒恒虚警率(Robust CFAR, RCFAR)检测器,针对复杂海洋场景中的杂波问题,通过双边修剪统计方法自动修剪异常值,采用自适应修剪深度和最大似然估计器闭式解,成功提高了检测性能,降低了虚警率,在多目标环境、近海和油污海域等场景中取得了显著效果。文献[2]提出了一种基于超像素级别的CFAR检测器,建立了一种针对纯杂波超像素和混合超像素进行标记的过程,采用了无监督聚类方法进行区分,并通过一种非局部拓扑策略,以自适应地确定足够数量的纯杂波超像素,用于检测阈值的估计。但是,在SAR图像中常常存在各种复杂的地物和杂波,这些地物的反射特性可能与目标相似,导致CFAR算法难以准确区分目标和背景之间的差异。另外,目标之间的干扰、目标和背景的反射特性随时间产生的变化、不同尺寸和形状各异的目标都增加了检测的难度,可能使CFAR算法因门限设置不当或计算复杂度较高而产生虚警或漏检的现象。

近年来,随着深度学习的发展,卷积神经网络(Convolutional neural network, CNN)^[3-4]在SAR目标检测领域得到广泛应用。深度学习模型可自动学习并提取图像中的目标特征,克服了传统方法中对目标特征的手工设计和选择。同时,深度学习模型能够适应复杂的场景和目标变化,具有较强的泛化能力。将深度学习与SAR图像目标检测结合可有效提升检测效率和检测效果^[5-9]。文献[10]提出了一种基于稀疏SAR图像数据集的目标检测和分类框架,利用复杂近似消息传递(Complex approximate message passing, CAMP)算法生成稀疏SAR图像数据集,并结合卷积神经网络进行目标检测和分类,提升了稀疏SAR图像目标检测的准确性和效率。文献[11]使用改进的卷积神经网络进行SAR图像目标检测,应用迁移学习方法和快速的基于区域的卷积神经网络(Fast region-based CNN, Faster R-CNN)来增强目标检测精度,相较于原有的需要大量经验知识的SAR图像目标检测方法体现出显著的有效性。文

献[12]提出一种基于图注意力网络(Graph attention network, GAT)的半监督SAR舰船目标检测方法,该方法利用包含两层注意力层的图注意力网络,通过对有标记节点的分类交叉熵损失训练,并将无标记节点分类结果中识别为舰船目标的超像素块重新定位到原始SAR图像中,获得精细检测结果。文献[13]提出一种基于YOLOv3的多尺度检测框架,引入多尺度注意力模块融合相邻两层的空间特征,提高网络对小尺度目标的敏感性,提升目标的显著性特征,这表明通过注意力机制对于不同层级的特征融合是有意义的;但该文献并未探究特征提取与注意力机制融合的具体效果。文献[14]提出多特征融合网络,通过小波变换提取空间和纹理特征,并采用二值域特征金字塔融合特征,再通过区域提议网络(Region proposal network, RPN)从原始图像中检测船舶目标。该文献单独提取不同层级的空间与纹理特征并融合以达到更好的检测效果,证明了不同层级的特征引导对于不同尺度的目标特征学习有效果。文献[15]提出了一种Yolo v4-tiny的形态特征金字塔轻量级网络,结合形态学网络进行去噪和边缘提取,得到与原始图像相同大小和相对位置的特征图并与原始图像结合,减少斑点噪声并增强边缘信息,有效改善多尺度目标检测的性能。这表明对于目标边缘特征的提取学习能够提升目标检测效果,但该文献未探究不同层级边缘特征的融合学习对SAR目标检测的影响。

通过对上述方法的分析,深度特征增强与学习可提高SAR目标可辨别性,有效提升SAR目标定位精度。然而,尽管深度学习方法在提高SAR目标检测的可辨别性和定位精度方面取得了显著进展,但现有方法在考虑SAR目标边缘特征时存在一定的不足,特别是在目标与背景杂波相似或目标尺寸较小时,目标的边缘往往受到噪声和干扰的影响而显得不明显。此外,针对不同尺度下的SAR图像目标,其边缘特征可能存在差异,这些差异性可能会影响模型的检测精度。

针对以上问题,本文提出一种基于边缘特征引导学习(Edge feature guided learning, EFGL)的SAR目标检测方法。该方法基于FCOS(Fully convolutional one-stage)^[16]目标检测框架,引入边缘特征增强学习模块强化对不同尺度下的边缘特征显式学习,在特征金字塔中引入边缘特征引导融合模块实现高层级的边缘特征提取,并通过与低层级特征乘积实现边缘特征引导的相邻层级特征融合。所提出的边缘特征增强与引导融合模块能够有效学习目标边缘特征,并更好地通过高层级边缘特征引导低层级特征的学习,促进不同层级间边缘特征的融合,从而有效提升模型的泛化能力。

1 基于边缘特征引导学习的SAR目标检测

基于边缘特征引导学习的SAR图像目标检测方法如图1所示。网络结构包括主干网络、边缘特征

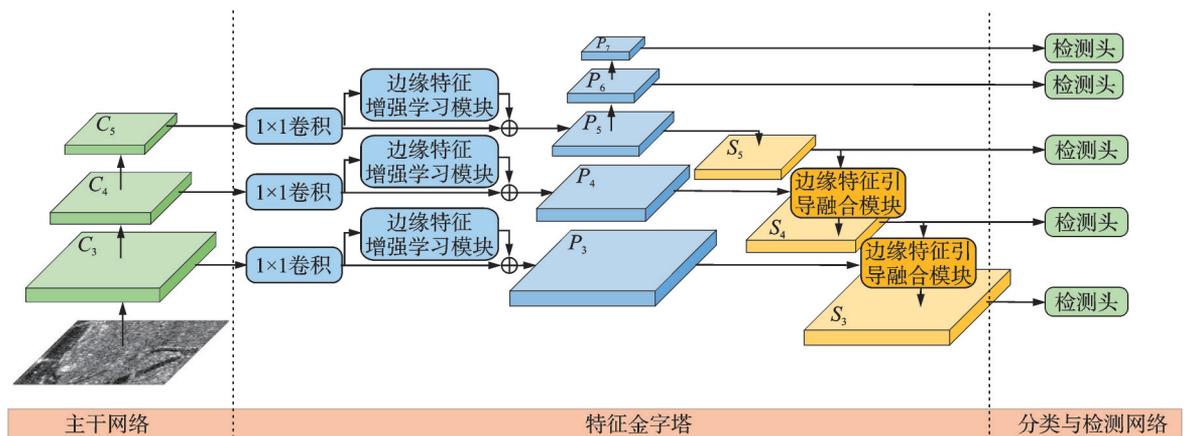


图1 基于EFGL的SAR目标检测方法

Fig.1 EFGL-based SAR target detection method

增强学习模块、边缘特征引导融合模块和检测头。主干网络对图像进行特征提取并传入边缘特征增强学习模块;边缘特征增强学习模块和边缘特征引导融合模块捕捉不同尺度上的目标特征并传入检测头;检测头对不同层级的特征进行分类与回归,得到最终的检测结果并输出。

1.1 网络结构

主干网络为深度为50层的ResNet网络。该主干网络首先通过64个卷积核大小为 7×7 ,卷积步长为2,填充为3的卷积核对输入数据进行处理,后接批归一化(Batch normalization, BN)层,修正线性单元(Rectified linear unit, ReLU)层和池化层进一步处理。输出数据最终经过4个残差块,得到通道数逐层倍增、大小逐层减半的4层特征层。

如图1所示,特征金字塔模块输入主干网络中的 C_3 至 C_5 层,经过卷积核大小为 1×1 的卷积后通过边缘算子模块学习边缘特征,将原始数据与其对应特征相加,以引入边缘信息,得到 P_3 至 P_5 层。 P_5 层通过执行两次卷积操作得到额外的两层特征层 P_6 和 P_7 层,形成5层的特征金字塔。然后 P_5 和 P_4 层分别通过最邻近差值方法得到尺寸加倍的特征图,并通过融合边缘特征引导的空间注意力计算出注意力图。将上述操作得到的注意力图分别与 P_4 和 P_3 层相乘,以强化边缘特征学习。继而, P_4 和 P_3 层与尺寸加倍的特征图相加进一步融合不同层次的特征。

分类与检测网络完成目标检测结果的输出,5个特征层的检测头共享权重,每个检测头通过分类分数、边界框预测和中心度来计算分类损失、边界框回归损失和中心度损失,预测出最终的目标检测结果。

1.2 边缘特征引导学习

1.2.1 边缘特征增强学习模块

通过引入边缘特征增强学习模块强化模型对于边缘特征的学习,该模块结构如图2所示。

为了强化网络对不同尺寸目标边缘特征的提取学习,将边缘特征增强学习模块^[17]引入FCOS网络中以提高模型对于边缘特征的学习。边缘特征增强学习模块可有效提高模型对于边缘特征的学习,该模块主要包含 3×3 卷积模块、 1×1 卷积- 3×3 卷积模块、Sobel算子模块和拉普拉斯滤波模块。首先分别通过 3×3 卷积模块和 1×1 卷积- 3×3 卷积模块进行深度卷积特征学习,具体公式如下

$$F_1 = K_1 * F_{in} + B_1 \quad (1)$$

$$F_2 = K_3 (K_2 * F_{in} + B_2) + B_3 \quad (2)$$

式(1)对应 3×3 卷积模块,其中 K_1 和 B_1 分别为权重和偏置。式

(2)对应 1×1 卷积- 3×3 卷积模块,其中 1×1 卷积将通道数扩大为2倍, 3×3 卷积将通道数还原,其中 K_2 、 K_3 和 B_2 、 B_3 分别为第一个和第二个卷积的权重与偏置。然后原始输入通过 1×1 卷积-Sobel_x和 1×1 卷积-Sobel_y模块分别提取水平边缘特征和竖直边缘特征,具体公式如下

$$F_x = (S_x \cdot \text{Sobel}_x) \otimes (K_x * F_{in} + B_x) + B_{D_x} \quad (3)$$

$$F_y = (S_y \cdot \text{Sobel}_y) \otimes (K_y * F_{in} + B_y) + B_{D_y} \quad (4)$$

式中:“ \otimes ”和“ $*$ ”分别代表深度可分离卷积(Depth-wise convolution, DWConv)和普通卷积操作;“ \cdot ”代表通道间的广播乘法; F_x 和 F_y 分别为水平和竖直边缘特征, S_x 和 S_y 为随机的缩放参数,这一参数是可学习的; K_x 、 K_y 和 B_x 、 B_y 分别为水平和竖直边缘特征中 1×1 卷积的权重与偏置; B_{D_x} 和 B_{D_y} 分别为水平和竖直边缘特征的偏置;Sobel_x和Sobel_y分别为水平和竖直边缘特征提取的Sobel算子,具体如下

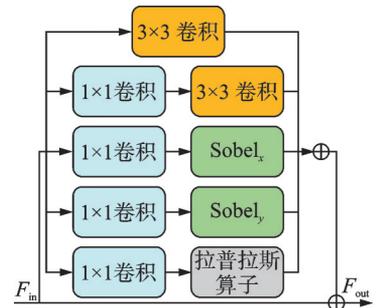


图2 边缘特征增强学习模块

Fig.2 Edge feature enhancement learning module

$$\text{Sobel}_x = \begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix}, \text{Sobel}_y = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} \quad (5)$$

最后原始输入通过 1×1 卷积-Laplacian模块强化特征提取,具体公式如下

$$F_4 = (S_4 \cdot \text{Lap}) \otimes (K_4 * F_{in} + B_4) + B_{\text{Lap}} \quad (6)$$

式中:Lap为拉普拉斯算子, B_{Lap} 为偏置,用于特征增强提取,具体如下

$$\text{Lap} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix} \quad (7)$$

如式(8)所示,将这5个模块提取的特征相加,即可得到最终边缘特征,将其与输入数据线性相加以达到保留卷积特征的同时引入边缘特征信息,即

$$F_{\text{out}} = F_1 + F_2 + F_x + F_y + F_4 \quad (8)$$

通过引入边缘算子模块,从主干网络中输出的特征被赋予了额外的边缘特征引导,模型能够更好地学习不同尺度下的SAR图像边缘特征。

1.2.2 边缘特征引导的相邻层级特征融合

边缘特征引导的相邻层级特征融合具体过程如图3所示。原始FPN中通过将高层特征图上采样后直接与低层特征图相加达到特征融合的目的,这样简单的线性相加并不能很好地发挥出高层级边缘特征对于低层级特征的引导作用,由此本文提出该模块促进相近层级边缘特征的融合与引导学习。边缘特征引导融合模块的公式具体如下

$$F_{\text{att}} = \text{SA}(\text{NU}(f_{\text{high}})) \quad (9)$$

$$F_{\text{low}} = f_{\text{low}} * F_{\text{att}} \quad (10)$$

对于从P5开始的高层特征 f_{high} ,首先对其进行最邻近差值上采样(Nearest neighbor interpolation upsampling, NU)得到 F_{high} ,使其尺寸与低层特征具有相同的特征图尺寸,继而通过边缘特征引导融合模块提取高层特征的注意力图 F_{att} 。边缘特征引导融合模块如图4所示,包含边缘特征提取和空间注意力(Spatial attention, SA)学习两部分。

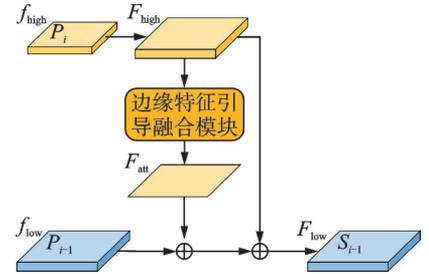


图3 边缘特征引导的相邻层级特征融合过程

Fig.3 Edge feature-guided cross-level feature fusion process

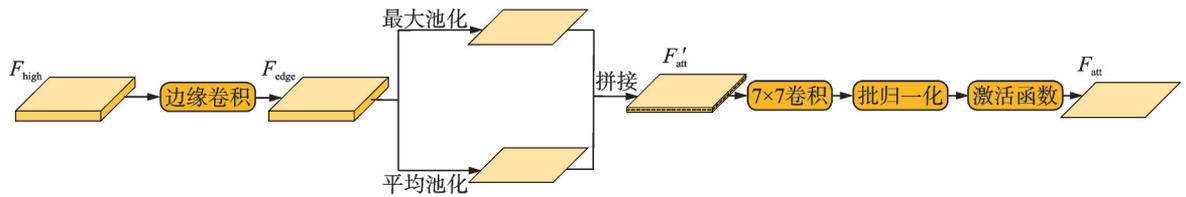


图4 边缘特征引导融合模块

Fig.4 Edge feature-guided fusion module

$$F_{\text{edge}} = \text{Conv}_{\text{edge}}(F_{\text{high}}) \quad (11)$$

$$F'_{\text{att}} = \text{Cat}[\text{MP}(F_{\text{edge}}), \text{AP}(F_{\text{edge}})] \quad (12)$$

$$F_{\text{att}} = \text{Sigmoid}(\text{BN}(\text{Conv}_S(F'_{\text{att}}))) \quad (13)$$

边缘特征提取通过一个尺寸为 3×3 的边缘特征学习卷积核 $\text{Conv}_{\text{edge}}$ 实现,该卷积核具体如下

$$\text{Conv}_{\text{edge}} = \begin{bmatrix} -1/16 & 5/16 & -1/16 \\ 5/16 & -1 & 5/16 \\ -1/16 & 5/16 & -1/16 \end{bmatrix} \quad (14)$$

该卷积核在水平和垂直方向上有较大的正权重,使得在这些方向上有明显变化的像素值会被放大,从而突出图像的边缘;中心的负权重用于减小中心像素的值,以此进一步突出其与周围像素的对比,增强边缘;周围的负权重一定程度上提供了平滑的作用,同时保持了对比,使边缘更加明显。对于所得到的突出边缘特征 F_{edge} ,可沿着通道维度对其进行最大池化(Max pooling, MP)和平均池化(Average pooling, AP),并将两种池化结果在通道维度上进行拼接得到 F'_{att} 。然后将拼接结果通过尺寸为 7×7 ,卷积步长为1,填充为3,输出通道数为1的卷积核 Conv_s 进行卷积操作,并在其后引入BN层和Sigmoid激活函数得到对应的注意力图 F_{att} 。将该可突出边缘特征的注意力图与低层特征进行广播乘法操作,得到由高层级边缘特征引导的低层特征。最后,将特征融合后的低层特征与高层特征进行线性相加,完成相邻层的特征融合。

上述特征融合过程既保留了将高层级特征融合至低层级特征的策略,又通过引入融合边缘特征的空间注意力模块强化了高层级特征对于低层级特征边缘特征学习的引导作用。同时,由于该特征融合过程自上而下进行,上层特征可叠加引导下层特征对于边缘特征的学习,这更加有利于不同层级间的特征融合。

2 实验结果与分析

2.1 数据集介绍

本文采用的数据集包括MSAR和SAR-Aircraft-1.0。MSAR数据集采用海丝一号卫星和高分三号卫星数据,数据集场景包括机场、港口、近岸、岛屿、远海和城区等;目标包括1851架桥梁、39858条船只、12319个油罐和6368架飞机。SAR-Aircraft-1.0数据集图像来自高分三号卫星,包含4368张图像和16463架飞机目标实例,包括A220、A320/321、A330、ARJ21、Boeing737、Boeing787和其他7个类别。MSAR和SAR-Aircraft-1.0数据集的图像示例分别如图5、6所示。

2.2 实验环境和数据集处理

本文实验环境是Ubuntu 22.04, CPU: I9-13900K, GPU: NVIDIA RTX 4090, RAM: 32 GB。对于MSAR数据集,将其中的图像大小调整为 $256 \text{ 像素} \times 256 \text{ 像素}$ 大小并且以50%的概率随机水平翻转。该数据集采用DAdaptAdam作为优化器,基础学习率设置为1.0并采用0.05的权重衰减,批次大小设置为64,在24和33轮学习率依次衰减为上一阶段的1/10,最大训练轮数为36轮。对于SAR-Aircraft-1.0数据集,将其中的图像大小

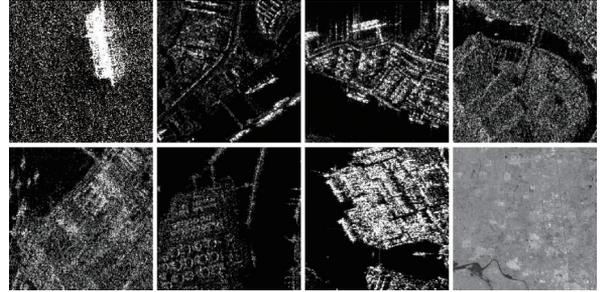


图5 MSAR数据集示例
Fig.5 MSAR dataset samples

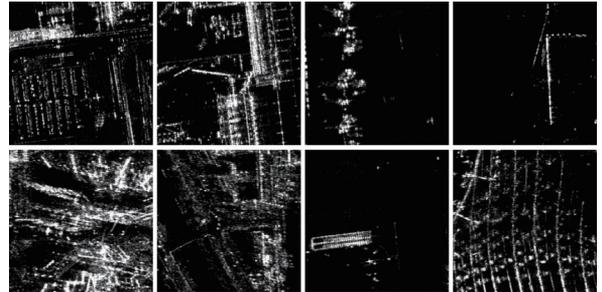


图6 SAR-Aircraft-1.0数据集示例
Fig.6 SAR-Aircraft-1.0 dataset samples

调整为 512×512 像素大小并且同样以50%的概率随机水平翻转。该数据集与MSAR数据集采用相同的优化器、基础学习率与权重衰减,另外设置批次大小为32,最大训练轮数为12轮,在8和11轮学习率依次衰减为1/10。两个数据集均采用VOC07和VOC12作为评价标准。

2.3 性能指标

采用均值平均精度(Mean average percision, mAP)作为SAR目标检测性能指标,其具体计算过程如下。首先根据模型的预测结果和真实标签,计算不同置信度阈值下的精度(Precision)和召回率(Recall)。

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (15)$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (16)$$

式中:TP(True positive)表示被正确预测为正样本的数量;FP(False positive)表示被错误预测为正样本的数量;FN(False negative)表示未被正确预测为正样本的数量。对于每个召回点,选择在该点及其之后所有预测值的最大值作为插值的预测值。将插值后的预测值进行平均,即为平均精度(Average precision, AP)。假设有 N 个类别,对每个类别 i ,该类别AP为 $\text{AP}(i)$,则mAP的计算公式为

$$\text{mAP} = \frac{\sum_{i=1}^N \text{AP}(i)}{N} \times 100\% \quad (17)$$

2.4 实验结果

为了验证本文方法的有效性,在多个不同算法上测试,并在AP'07和AP'12标准下分别进行了实验结果对比。在MSAR数据集上,对边缘算子模块和融合边缘特征的空间注意力模块进行了测试,具体结果如表1所示。在AP'07的指标下本文网络的mAP能够达到68.68%,比原始的FCOS网络提升了1.34%,在AP'12的指标下本文网络的mAP能够达到69.95%,比原始的FCOS网络提升了0.59%。对比不同的主流算法,所提方法也取得了最好的检测效果,这表明其在MSAR数据集检测效果上比原先的网络有一定程度的提升。

表1 MSAR数据集实验结果
Table 1 Experimental results on the MSAR dataset

算法	Average precision(AP'07)					Average precision(AP'12)				
	mAP	Ship	Airplane	Bridge	Oil Tank	mAP	Ship	Airplane	Bridge	Oil Tank
FCOS ^[16]	67.34	89.13	41.64	75.17	63.42	69.36	90.16	41.67	77.29	68.32
CenterNet ^[18]	67.18	89.87	41.05	74.91	62.92	69.25	91.10	41.24	77.05	67.61
PVT-T ^[19]	29.72	58.88	10.14	11.53	38.34	26.25	59.64	1.52	6.77	37.05
TOOD ^[20]	66.58	89.04	42.78	71.00	63.49	68.32	89.88	42.24	73.91	67.24
VFNet ^[21]	65.31	87.94	38.80	71.03	63.49	66.56	89.21	36.74	73.12	67.16
YOLOF ^[22]	46.39	69.39	9.09	58.81	48.25	44.59	71.72	0.33	60.07	46.25
Grid R-CNN ^[23]	55.82	80.85	9.09	69.89	63.47	57.30	86.65	3.74	71.99	66.83
DETR ^[24]	8.89	13.39	1.26	6.70	14.21	3.85	5.38	0.82	2.96	6.24
DAB-DETR ^[25]	43.88	80.25	3.47	52.20	39.61	44.17	84.20	1.58	52.36	38.56
Conditional DETR ^[26]	42.26	78.63	9.80	54.13	26.47	40.01	82.86	1.21	54.34	21.64
EFGL	68.68	89.18	44.08	78.08	63.37	69.95	90.09	42.63	78.80	68.27

在 SAR-Aircraft-1.0 数据集上,同样对边缘算子模块和融合边缘特征的空间注意力模块进行了实验测试,具体结果如表 2 所示。在 AP'07 的指标下本文网络的 mAP 能够达到 67.44%,比原始的 FCOS 网络提升了 4.81%,在 AP'12 的指标下本文网络的 mAP 能够达到 68.47%,比原始的 FCOS 网络提升了 4.71%。对比其他主流算法的 mAP,本文网络取得了最好的效果,这表明本文网络在 SAR-Aircraft-1.0 数据集上同样有明显的检测效果提升。

表 2 SAR-Aircraft-1.0 数据集实验结果

Table 2 Experimental results on the SAR-Aircraft-1.0 dataset

%

算法	Average precision(AP'07)							Average precision(AP'12)								
	mAP	A220	A320	A330	ARJ21	B737	B787	Other	mAP	A220	A320	A330	ARJ21	B737	B787	Other
FCOS ^[16]	62.63	62.27	93.19	63.65	63.10	40.06	52.37	63.79	63.76	64.63	65.16	64.61	63.02	39.36	53.05	66.50
Center-Net ^[18]	64.11	58.03	93.03	97.02	61.90	36.41	41.74	60.61	64.90	58.15	96.58	97.36	62.13	34.99	42.28	62.81
PVT-T ^[19]	61.64	52.70	82.01	85.94	61.07	32.63	54.94	62.15	62.43	53.32	84.46	87.85	62.12	30.57	55.56	63.15
TOOD ^[20]	62.66	53.94	82.20	91.86	60.51	35.36	53.81	60.95	62.98	53.81	83.91	93.99	60.51	33.91	53.50	61.21
VFNet ^[21]	66.17	60.94	90.91	95.25	61.29	39.41	56.46	58.93	66.90	62.00	95.57	96.14	60.76	36.63	56.76	60.43
YOLOF ^[22]	66.25	60.24	90.75	88.81	61.00	36.30	59.64	66.99	67.71	60.97	94.85	91.31	63.37	34.35	60.42	68.71
Grid R-CNN ^[23]	64.15	54.05	98.08	97.16	60.47	28.94	47.73	62.61	64.04	54.48	98.42	97.36	59.54	26.80	49.14	62.51
DETR ^[24]	10.61	19.55	0.78	0.47	4.15	9.11	20.95	19.25	8.27	15.80	0.56	0.45	2.45	6.14	18.97	13.55
DAB-DETR ^[25]	53.62	59.82	88.39	18.84	62.26	34.52	54.34	57.15	54.61	60.80	91.10	18.97	63.61	33.20	54.04	57.42
Conditional DETR ^[26]	62.25	53.74	89.38	78.83	61.75	37.44	61.45	53.19	63.02	54.27	93.30	80.47	62.07	35.92	61.99	53.14
EFGL	67.44	65.14	96.02	95.69	65.82	36.24	48.32	64.87	68.47	69.08	97.15	96.45	66.32	36.50	48.13	65.69

对比所提算法和 Grid R-CNN 的实验结果可明显看到,由于本文算法采用的是单阶段检测器,对不同尺度的特征图上每一个像素位置都进行预测而不受固定大小的滑动窗口或预定义的锚框的限制,因而能够更好地适应小目标的变化尺寸和位置,在识别 MSAR 数据集中的 Airplane 类别的小目标时有十分明显的优势。另外,由于 DETR 等端到端算法采用了全局注意力机制,导致其在处理弱小目标时不够敏感,对于 MSAR 数据集中的小目标类别飞机和暗目标类别桥梁的检测效果弱于本文算法的检测效果。对比本文算法和同为单阶段检测算法的基础算法 FCOS 以及 CenterNet 等算法,本文算法在保持弱小目标检测效果的基础上引入边缘特征的学习与融合,进一步提升了小目标及暗目标的检测结果。

对于 SAR-Aircraft-1.0 数据集中存在的较大尺度目标,主流的算法都能够获得较为不错的检测效果,不过由于存在部分与背景融合度较高的暗目标以及与陆地建筑物融合度较高的亮目标,主流算法的检测效果仍然存在不足。本文算法对于不同尺度目标的边缘特征进行单独提取,并进行融合与引导学习,这使得其在这两类目标的检测效果有所提升,能够显著提高对于停放在较亮的建筑物旁和较暗的跑道上飞机的检测效果,对于 A320 和 ARJ21 等目标有最好的检测效果。

2.5 消融实验

在 MSAR 和 SAR-Aircraft-1.0 数据集上,分别对边缘算子模块和融合边缘特征的空间注意力模块进行了消融实验,实验结果分别如表 3、4 所示,消融实验折线图分别如图 7、8 所示。

表3 MSAR数据集消融实验

Table 3 Ablation study on the MSAR dataset

算法	Average precision (AP'07)	Average precision (AP'12)
FCOS ^[16]	67.34	69.36
边缘特征增强学习模块	68.52	69.82
边缘特征引导融合模块	67.94	69.61
EFGL	68.68	69.95

表4 SAR-Aircraft-1.0数据集消融实验

Table 4 Ablation study on the SAR-Aircraft-1.0 dataset

算法	Average precision (AP'07)	Average precision (AP'12)
FCOS ^[16]	62.63	63.76
边缘特征增强学习模块	66.16	67.03
边缘特征引导融合模块	65.31	65.84
EFGL	67.44	68.47

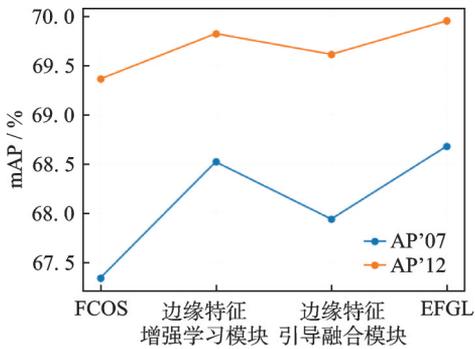


图7 MSAR数据集消融实验折线图

Fig.7 Line chart of ablation study results on the MSAR dataset

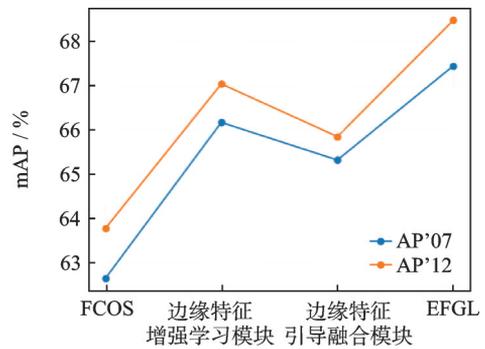


图8 SAR-Aircraft-1.0数据集消融实验折线图

Fig.8 Line chart of ablation study results on the SAR-Aircraft-1.0 dataset

在MSAR数据集上,以FCOS网络作为基准网络。FCOS网络在AP'07指标下的mAP为67.34%。在此基础上本文单独加入边缘算子模块和融合边缘特征的空间注意力模块,此时mAP分别为68.52%和67.94%,对比baseline分别提升了1.18%和0.60%。继而将两个模块同时加入FCOS网络,此时可获得68.68%的mAP,SAR目标检测效果提升了1.34%。由此可见,单独加入边缘算子模块和融合边缘特征模块都能在一定程度上提升网络对于边缘特征的学习,提升检测效果;同时加入两个模块后,边缘特征引导的相邻层级特征融合时能够充分有效地利用边缘特征增强学习模块提取的不同尺度下的边缘特征,能够更加明显地提升网络对于边缘特征的解析能力。

在SAR-Aircraft-1.0数据集上,同样分别测试了不同模块嵌入网络的目标检测效果。在AP'07指标下基准网络的mAP为62.63%,单独加入边缘算子和融合边缘特征的空间注意力模块以及二者同时加入的mAP分别为66.16%、65.31%和67.44%,分别提升了3.53%、2.68%和4.81%。在AP'12指标下基准网络的mAP为63.76%,3种算法的mAP为67.03%、65.84%和68.47%,分别提升了3.27%、2.08%和4.71%。由此可见,本文所提出的边缘特征学习模块对于SAR-Aircraft-1.0数据集中较大尺度的目标同样能够取得更好的检测效果。

2.6 可视化实验

将基础算法FCOS与本文方法的检测结果进行了可视化对比。

图9为FCOS和本文方法在MSAR数据集上的检测结果。其中第1~3行分别为真实框(Ground truth,GT)、FCOS检测效果和基于EFGL方法的检测效果。可以看出,在处理某些与背景融合度较高

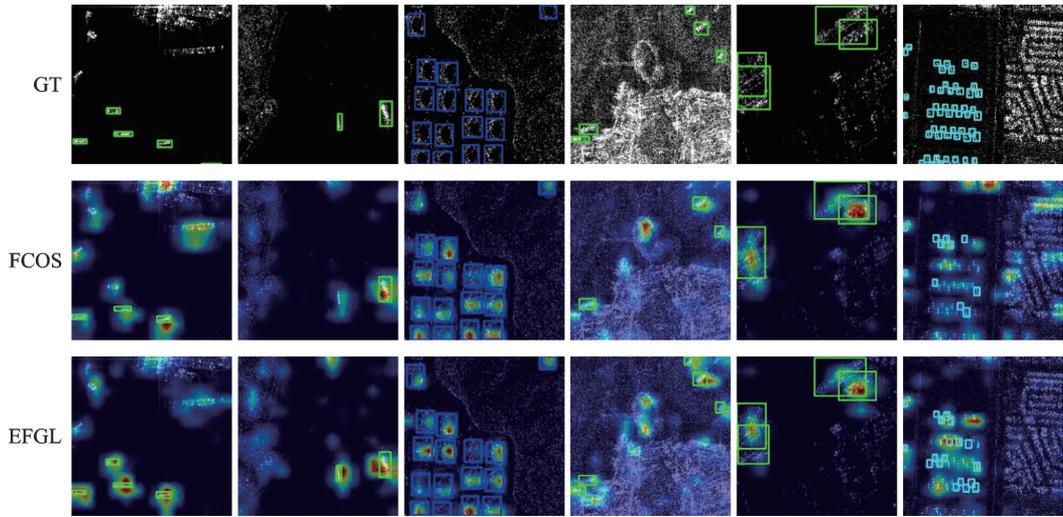


图9 MSAR数据集检测结果

Fig.9 Detection results on the MSAR dataset

的飞机时,FCOS难以将其识别导致漏检。同时,对于某些具有较高重合度的目标,FCOS简单地将其检测为一个完整的大目标,而本文算法由于对于边缘特征进行了学习,能够正确识别出两个重叠目标。MSAR数据集中存在一些集群小目标,它们集中在某些区域,会干扰网络对这些目标的检测。FCOS算法只能够识别出其中的一小部分目标,而本文方法能够识别出更多的目标,且能够更加准确地对识别出的目标进行定位。另外,对比注意力图可发现,FCOS给予弱小目标的注意力不足,与背景较为相似;得益于边缘特征的学习,本文方法能够给予这些区域更高的注意力,使得检测效果提升。这表明通过对于边缘特征的引导学习能够有效减少漏检率,提升对重叠目标的检测,提高网络对于一些弱小目标的检测效果。

图10为FCOS和本文方法在SAR-Aircraft-1.0数据集上的检测结果。可以发现,由于一些目标与建筑物或者跑道靠近,容易影响网络的检测,FCOS出现了一定程度的漏检,对比其注意力图可发现,这

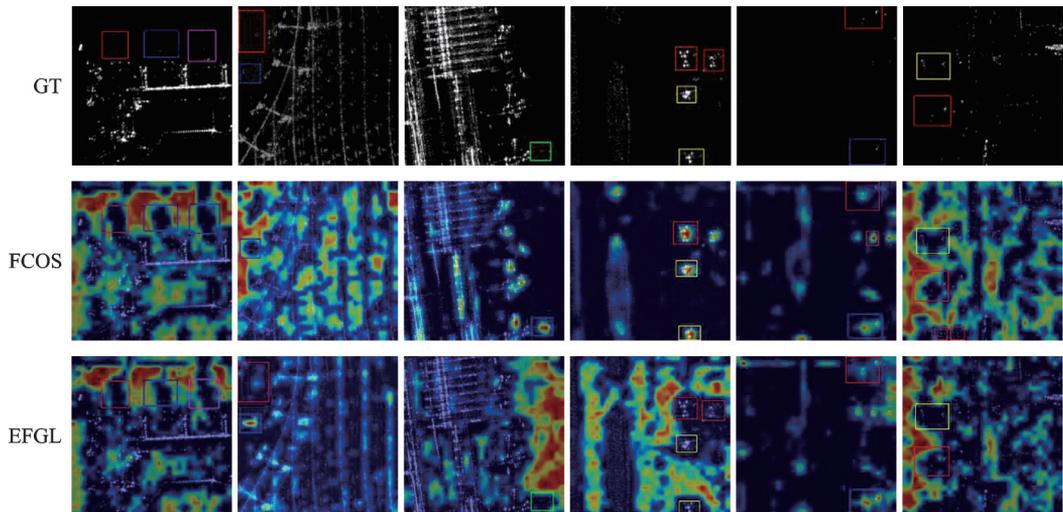


图10 SAR-Aircraft-1.0数据集检测结果

Fig.10 Detection results on the SAR-Aircraft-1.0 dataset

些目标的注意力与周围较亮的建筑物融合导致网络难以区分,而本文方法有效改善了这一情况,能够识别出更多的目标。另外,由于不同类别目标的特征较为相似以及与背景融合度较高的问题,FCOS网络中会发生的分类错误及错误识别背景区域的情况,本文方法由于对边缘特征进行了额外的学习,对于这种相似性能更好地区分,能够更准确地识别目标及其类别,这表明边缘特征学习能够使网络更精准地识别目标类别,提升检测效果。

3 结束语

本文提出了一种基于边缘特征引导学习的SAR目标检测方法,该方法通过引入边缘特征增强学习模块和边缘特征引导的相邻层级特征融合方法强化了模型对于边缘特征的学习。通过边缘特征增强学习模块提取数据边缘特征,通过边缘特征引导融合模块获取高层级边缘特征并与低层级特征融合,有效地提升了模型的边缘特征学习能力,促进了不同层级特征的更好融合。实验结果表明,本文方法可有效提升SAR图像目标检测效果,提高模型的泛化学习能力。不同尺度目标的边缘特征的差异性特征学习,将是未来工作之一;另外,在未来工作中,将着重考虑如何将边缘特征信息嵌入至损失函数中,以达到更加精确地定位目标的目的。

参考文献:

- [1] AI Jiaqiu, MAO Yuxiang, LUO Qiwu, et al. Robust CFAR ship detector based on bilateral-trimmed-statistics of complex ocean scenes in SAR imagery: A closed-form solution[J]. IEEE Transactions on Aerospace and Electronic Systems, 2021, 57(3): 1872-1890.
- [2] LI Mingdian, CUI Xingchao, CHEN Siwei. Adaptive superpixel-level CFAR detector for SAR inshore dense ship detection[J]. IEEE Geoscience and Remote Sensing Letters, 2021, 19: 1-5.
- [3] LECUN Y, BOTTOU L, BENGIO Y, et al. Gradient-based learning applied to document recognition[J]. Proceedings of the IEEE, 1998, 86(11): 2278-2324.
- [4] LECUN Y, BOSER B, DENKER J S, et al. Backpropagation applied to handwritten zip code recognition[J]. Neural Computation, 1989, 1(4): 541-551.
- [5] GONG Yicheng, ZHANG Zhuo, WEN Jiabao, et al. Small ship detection of SAR images based on optimized feature pyramid and sample augmentation[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2023, 16: 7385-7392.
- [6] ZHAO Ming, ZHANG Xin, KAUP A. Multitask learning for SAR ship detection with gaussian-mask joint segmentation[J]. IEEE Transactions on Geoscience and Remote Sensing, 2023, 61: 1-16.
- [7] YANG Chule, ZHANG Chao, FAN Zunlin, et al. A multi-channel aggregation framework for object detection in large-scale SAR image[C]//Proceeding of ICASSP 2023—2023 IEEE International Conference on Acoustics, Speech and Signal Processing. [S.l.]: IEEE, 2023: 1-5.
- [8] ZHANG Chongqi, ZHANG Ziwen, CHONG Mingzhe, et al. A joint unsupervised super-resolution and SAR object detection network[C]//Proceeding of 2023 8th International Conference on Computer and Communication Systems. [S.l.]: IEEE, 2023: 1027-1030.
- [9] 李生辉, 李晓飞, 宋璋晗, 等. 基于改进YOLOv5的船舶多尺度SAR图像检测算法[J]. 数据采集与处理, 2024, 39(1): 120-131.
LI Shenghui, LI Xiaofei, SONG Zhanghan, et al. Multi-scale SAR image detection algorithm for ships based on improved YOLOv5[J]. Journal of Data Acquisition and Processing, 2024, 39(1):120-131.
- [10] BI Hui, DENG Jiarui, YANG Tianwen, et al. CNN-based target detection and classification when sparse SAR image dataset is available[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2021, 14: 6815-6826.
- [11] SUN Long, WU Tao, SUN Guangcai, et al. Object detection research of SAR image using improved faster region-based convolutional neural network[J]. Journal of Geodesy and Geoinformation Science, 2020, 3(3): 1-18.

- [12] 吕进东, 王彤, 唐晓斌. 基于图注意力网络的半监督SAR舰船目标检测[J]. 电子与信息学报, 2023, 45(5): 1541-1549.
LÜ JINDONG, WANG Tong, TANG Xiaobin. Semi-supervised SAR ship target detection with graph attention network[J]. *Journal of Electronics & Information Technology*, 2023, 45(5): 1541-1549.
- [13] 张冬冬, 王春平, 付强. 基于特征增强网络的 SAR 图像舰船目标检测[J]. 系统工程与电子技术, 2023, 45(4): 1032-1039.
ZHANG Dongdong, WANG Chunping, FU Qiang. Ship target detection in SAR image based on feature-enhanced network[J]. *Systems Engineering and Electronics*, 2023, 45(4): 1032-1039.
- [14] WANG Shiyu, CAI Zhanchuan, YUAN Jieyu. Automatic SAR ship detection based on multi-feature fusion network in spatial and frequency domain[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2023, 61: 1-11.
- [15] ZHAO Congxia, FU Xiongjun, DONG Jian, et al. SAR ship detection based on end-to-end morphological feature pyramid network[J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2022, 15: 4599-4611.
- [16] TIAN Zhi, SHEN Chunhua, CHEN Hao, et al. FCOS: Fully convolutional one-stage object detection[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. [S.l.]: IEEE, 2019: 9627-9636.
- [17] ZHANG Xindong, ZENG Hui, ZHANG Lei. Edge-oriented convolution block for real-time super resolution on mobile devices [C]//Proceedings of the 29th ACM International Conference on Multimedia. [S.l.]: ACM, 2021: 4034-4043.
- [18] ZHOU X Y, WANG D Q, KRÄHENBÜHL P. Objects as points[EB/OL]. (2019-04-25). <https://doi.org/10.48550/arXiv.1904.07850>.
- [19] WANG Wenhai, XIE Enze, LI Xiang, et al. Pyramid vision transformer: A versatile backbone for dense prediction without convolutions[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. [S.l.]: IEEE, 2021: 568-578.
- [20] FENG Chengjian, ZHONG Yujie, GAO Yu, et al. TOOD: Task-aligned one-stage object detection[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. [S.l.]: IEEE, 2021: 3490-3499.
- [21] ZHANG H Y, WANG Y, DAYOUB F, et al. VarifocalNet: An IoU-aware dense object detector[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. [S.l.]: IEEE, 2021: 8514-8523.
- [22] CHEN Qiang, WANG Yingming, YANG Tong, et al. You only look one-level feature[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. [S.l.]: IEEE, 2021: 13039-13048.
- [23] LU Xin, LI Buyu, YUE Yuxin, et al. Grid R-CNN[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. [S.l.]: IEEE, 2019: 7363-7372.
- [24] CARION N, MASSA F, SYNNAEVE G, et al. End-to-end object detection with transformers[C]//Proceedings of the European Conference on Computer Vision. Cham: Springer, 2020: 213-229.
- [25] LIU Shilong, LI Feng, ZHANG Hao, et al. DAB-DETR: Dynamic anchor boxes are better queries for DETR[EB/OL]. (2022-03-30). <https://doi.org/10.48550/arXiv.2201.12329>.
- [26] MENG Depu, CHEN Xiaokang, FAN Zejia, et al. Conditional DETR for fast training convergence[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. [S.l.]: IEEE, 2021: 3651-3660.

作者简介:



倪康(1990-), 通信作者, 男, 博士, 副教授, 硕士生导师, 研究方向: SAR 图像目标检测与分类, E-mail: tznikang@163.com。



孙笠焜(2003-), 男, 本科, 研究方向: SAR 图像目标检测, E-mail: 2429801165@qq.com。



邹旻瑞(1999-), 男, 硕士研究生, 研究方向: SAR 图像目标检测, E-mail: traveler_wood@163.com。

(编辑:夏道家)