



全等级上下文压缩激励的 SAR 舰船实例分割

张天文¹, 张晓玲^{1*}, 邵子康¹, 曾天娇²

(1. 电子科技大学信息与通信工程学院 成都 611731; 2. 电子科技大学航空航天学院 成都 611731)

【摘要】 现有深度学习 SAR 舰船实例分割方法未考虑特征全等级信息和目标上下文信息, 导致了较低实例分割精度。针对上述问题, 提出了一种基于全等级上下文压缩激励感兴趣区域 ROI 提取器的 SAR 舰船实例分割方法 FL-CI-SE-ROIE。FL-CI-SE-ROIE 实现了全等级 ROI 提取, 可保留全等级信息, 增强了网络多尺度描述能力。FL-CI-SE-ROIE 实现了上下文 ROI 扩充, 可获取目标上下文信息, 增强了网络背景鉴别能力。FL-CI-SE-ROIE 引入了压缩激励 SE 模块来平衡不同范围的上下文 ROI, 可抑制背景干扰, 进一步提高了实例分割精度。在公开像素级多边形分割 SAR 舰船检测数据集 PSeg-SSDD 上的实验结果表明, 所提方法的 SAR 舰船实例分割精度高于现有其他 9 种对比模型。

关键词 深度学习; 实例分割; 感兴趣区域提取器; 合成孔径雷达
中图分类号 TN958 **文献标志码** A **doi**:10.12178/1001-0548.2022415

A Full-Level Context Squeeze-and-Excitation ROI Extractor for SAR Ship Instance Segmentation

ZHANG Tianwen¹, ZHANG Xiaoling^{1*}, SHAO Zikang¹, and ZENG Tianjiao²

(1. School of Information and Communication Engineering, University of Electronic Science and Technology of China Chengdu 611731;
2. School of Aeronautics and Astronautics, University of Electronic Science and Technology of China Chengdu 611731)

Abstract The current deep-learning based SAR ship instance segmentation models fail to consider the full level information of features and the context information of targets, which leads to low instance segmentation accuracy. In order to address this problem, a SAR ship instance segmentation method based on a full-level context information squeeze-and-excitation region of interest (ROI) extractor is proposed. This method proposes a novel ROI extractor(ROIE), called FL-CI-SE-ROIE. First of all, FL-CI-SE-ROIE can extract ROI at all levels, which can retain the full level features of the target, thus enhancing the multi-scale description capability of the network. Then, FL-CI-SE-ROIE expands the ROI context information, which can gain the context information of targets, thus enhancing the background identification capability. Finally, FL-CI-SE-ROIE introduces a squeeze and excitation(SE) module, which can balance ROI context information in different ranges, thus suppressing background interference, and further improving the accuracy of instance segmentation. The experimental results on the public polygon segmentation SAR ship detection dataset (PSeg-SSDD) show that the SAR ship instance segmentation accuracy of the proposed method is higher than that of the current 9 comparison models.

Key words deep learning; instance segmentation; ROI extractor; synthetic aperture radar

合成孔径雷达 (synthetic aperture radar, SAR) 是一种先进的主动式微波传感器, 被广泛应用于地质勘测、灾害救援、海洋监视等领域^[1]。作为一种典型海洋监视任务, 海面舰船检测在民用和军事领域具有重要应用价值。在民用方面, 舰船检测广泛应用于交通调度、舰船搜救等^[2]。在军事方面, 海面舰船检测广泛应用于战略侦查、敌舰打击等^[3]。

相比其他遥感系统, SAR 能够全天时全天候运行, 适合监视气候多变的海洋^[4]。因此, 近年来, 利用 SAR 实现舰船检测受到广泛关注。

传统 SAR 舰船检测方法的显著特点是基于手工提取特征^[5]。这些方法可大致分为基于恒虚警的方法^[6]、基于全局阈值的方法^[7]、基于极化分解的方法^[8]、基于变换域的方法^[9]、基于视觉显著性的

收稿日期: 2022-12-10; 修回日期: 2023-03-10

基金项目: 国家自然科学基金 (61571099)

作者简介: 张天文 (1994-), 男, 博士, 主要从事雷达目标检测识别跟踪、SAR 图像智能解译等方面的研究。

*通信作者: 张晓玲, E-mail: xlzhang@uestc.edu.cn

方法^[10]以及基于超像素的方法^[11]等。然而,这些传统手工提取特征具有设计繁琐、迁移能力差等缺点^[12]。近年来,随着深度学习的发展,很多学者开始关注基于深度学习的 SAR 舰船检测^[13-18]。如文献 [13] 改进了 Faster R-CNN 和特征金字塔网络 (feature pyramid network, FPN), 并融合了极化信息进行舰船检测。文献 [14] 改进了单阶段检测器 YOLOv5^[15], 实现了复杂场景下轻量化 SAR 舰船检测。文献 [16] 将水平检测框改进为旋转检测框, 减少了近岸密集排列舰船之间检测框重叠, 提高了检测精度。文献 [17] 将 SAR 传统特征与深度学习相结合, 提供了更具有可解释性的检测结果。文献 [18] 提出一种轻量级的 SAR 舰船检测网络, 能够实现舰船目标快速检测。虽然这些模型都具有良好精度, 但仅实现了框等级检测, 未能实现 SAR 舰船像素级精细化表征。

由于检测框无法实现 SAR 舰船像素级精细化表征, 一些学者将实例分割引入 SAR 舰船领域中。文献 [19] 发布了 HRSID 数据集并提供了一系列研究基准。文献 [20] 提出了一种可用于高分辨率 SAR 图像舰船实例分割的网络, 但该网络未结合舰船上下文信息, 导致了较多虚警。文献 [21] 提出了质心距离损失以增强舰船实例分割的学习效果, 但该方法对小型舰船实例分割的精度较低, 未能考虑多尺度特征。文献 [22] 将 Mask R-CNN^[23] 应用于 SAR 舰船实例分割, 但未考虑上下文信息, 影响了实例分割的精度。文献 [24] 重点研究了 SAR 图像中的舰船多尺度特性, 提高了多尺度舰船的实例分割精度。以上这些方法均未考虑特征全等级信息和目标上下文信息, 可能阻碍精度改善。

本文提出了一种基于全等级上下文压缩激励感兴趣区域提取器 (region of interest extractor, ROIE) 的 SAR 舰船实例分割方法 (full-level context-information squeeze-excitation region of interest extractor, FL-CI-SE-ROIE), 主要创新点如下。

1) 实现全等级 ROI 提取: ROIE 在特征金字塔的所有尺度上提取, 从而保留了全等级信息, 提高了网络多尺度描述能力。

2) 实现上下文 ROI 扩充: 在 ROIE 中生成多个尺度的感兴趣区域, 从而获取了目标上下文信息, 增强了网络背景鉴别能力。

3) 引入压缩激励模块: 在 ROIE 中引入压缩激励 (squeeze and excitation, SE) 模块来平衡不同范围

上下文 ROI 的贡献, 抑制背景干扰, 增强重要特征。

1 方法原理

FL-CI-SE-ROIE 网络结构类似于经典双阶段实例分割网络 HTC^[25], FL-CI-SE-ROIE 由骨干网络、FPN、ROIE 和检测头构成。输入图像首先由骨干网络和 FPN 生成多尺度特征图, 然后将多尺度特征图输入 ROIE 提取特征子集, 最后将特征子集输入检测头获得实例分割结果。以下详细介绍 FL-CI-SE-ROIE 各项改进。

1.1 全等级 ROI 提取

经典双阶段实例分割网络, 如 HTC, 常基于以下准则, 确定当前特征子集在 FPN 中所处等级:

$$i = \left\lfloor i_0 + \log_2 \left(\frac{\sqrt{wh}}{224} \right) \right\rfloor \quad (1)$$

式中, i 表示特征子集在 FPN 中所处等级; i_0 表示基准等级 5; w 和 h 表示检测框宽度和高度。

基于上述原则, 经典 ROIE 仅选择单个特定 FPN 等级 P_i 以提取特征子集。图 1 展示了单等级 ROIE 提取过程。如图 1 所示, 假设选中 FPN 的 P_2 等级 (实线箭头), 则 FPN 的其他等级将被抛弃^[26] (虚线箭头)。然而, 这种做法会丢失目标在不同等级下不同尺度的差异性特征, 如图 1 中的 P_3 , P_4 , P_5 等级均被丢弃。丢弃 FPN 的其他等级, 会削弱网络的多尺度特征描述能力, 从而导致实例分割精度下降^[27]。

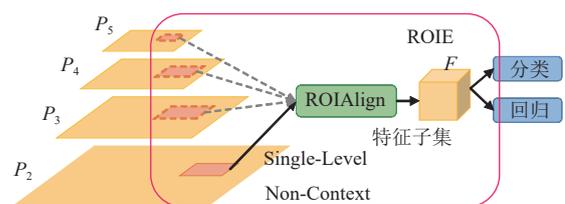


图 1 单等级 ROI 提取

因此, FL-CI-SE-ROIE 未采用式 (1) 准则, 而是在 FPN 所有等级提取特征子集, 从而保留了目标的全等级信息, 增强网络对目标的多尺度描述能力。此外, 对于不同等级提取到的特征子集, FPN 的输出特征图分别为 $[P_2, P_3, P_4, P_5]$, 分别对各等级的特征图提取特征子集, 即:

$$F_i = \text{ROIAlign}(P_i) \quad (2)$$

式中, P_i 为特征金字塔的各个等级特征图; F_i 为提取的特征子集; $\text{ROIAlign}(\cdot)$ 是将 RPN 生成的任意

尺寸的提案映射到固定尺寸特征图上的过程^[28]。
图2展示了ROIAlign的过程。

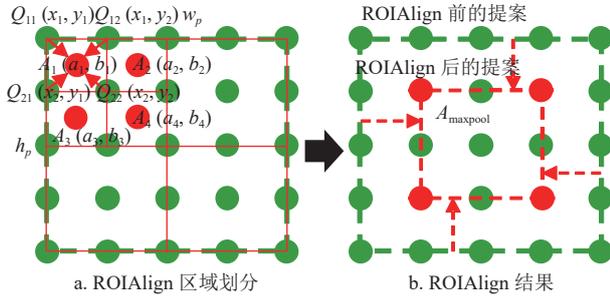


图2 ROIAlign过程示意图

如图2所示, 对于一个输入大小为 (w_p, h_p) 的提案, 对其内部划分成 $H \times W$ 个区域 (图2中划分为 2×2 个区域)。在每个区域采样4个位置 $A_1(a_1, b_1)$, $A_2(a_2, b_2)$, $A_3(a_3, b_3)$, $A_4(a_4, b_4)$, 其中 A_i 的值由双线性插值计算:

$$f(A_i) = \frac{(x_2 - x)(y_2 - y)}{(x_2 - x_1)(y_2 - y_1)} f(Q_{11}) + \frac{(x - x_1)(y_2 - y)}{(x_2 - x_1)(y_2 - y_1)} f(Q_{21}) + \frac{(x_2 - x)(y - y_1)}{(x_2 - x_1)(y_2 - y_1)} f(Q_{12}) + \frac{(x - x_1)(y - y_1)}{(x_2 - x_1)(y_2 - y_1)} f(Q_{22}) \quad (3)$$

式中, $Q_{11}=(x_1, y_1)$; $Q_{12}=(x_1, y_2)$; $Q_{21}=(x_2, y_1)$; $Q_{22}=(x_2, y_2)$, 为距离 A_i 的最近4个像素的坐标; $f(\cdot)$ 为该坐标的值。之后对4个位置 A_1, A_2, A_3, A_4 的值进行最大池化操作得到 $f(A_{\max\text{pool}})$, 即:

$$f(A_{\max\text{pool}}) = \max\text{pool}(f(A_1), f(A_2), f(A_3), f(A_4)) \quad (4)$$

对每个区域都进行上述操作即可获得提取到的特征子集 F_i 。在 FPN 全等级上使用 ROIAlign 提取特征子集后, 再根据以下特征求和方式实现高低不同等级的特征融合。计算如下:

$$F = \sum_{i=2}^5 F_i \quad (5)$$

式中, F 表示特征融合输出, 该输出能平衡高级语义信息和低等级空间位置信息。此外, 全等级 ROI 提取操作不会引入额外的参数量, 因为该操作是在 FPN 的每一个输出级别上均进行 ROI 提取, 不需要引入额外的参数, 因此不会改变模型的复杂度。但是, 由于引进了额外的 ROI 提取操作, 模型的检测速度会受到一定影响。

1.2 上下文 ROI 扩充

上下文 ROI 扩充指在 RPN 生成初始候选框的周边生成多个尺度的包含背景上下文的候选框, 从而包含更多背景上下文信息。图3展示了不同范围的舰船背景上下文信息。

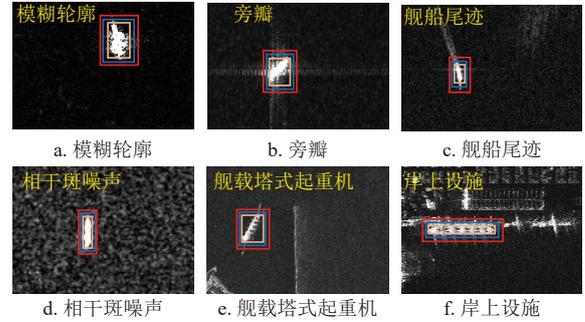


图3 不同范围舰船背景上下文信息

初始候选框记为 ROI: (w, h) , 如图3中橙色矩形框所示, 其中 w 和 h 为候选框宽度和高度。另外两个候选框分别记为 $\text{ROI}_{C1}: (w_1, h_1)$ 和 $\text{ROI}_{C2}: (w_2, h_2)$, 分别由红色矩形框和蓝色矩形框标记, 即:

$$\begin{aligned} (w_1, h_1) &= k_1(w, h) \\ (w_2, h_2) &= k_2(w, h) \end{aligned} \quad (6)$$

式中, k_1 和 k_2 表示两个大于1的扩展系数。外围上下文候选框 $\text{ROI}_{C2}: (w_2, h_2)$ 比 $\text{ROI}_{C1}: (w_1, h_1)$ 包含更多背景上下文信息。由图3可知, 舰船目标的周围具有多种复杂的背景上下文信息, 如舰船的模糊轮廓、旁瓣、舰船尾迹、相干斑噪声、舰载塔式起重机、岸上设施等, 这些背景上下文信息都会影响最终的舰船像素预测性能^[29]。使用包含更多背景上下文信息的候选框的目的在于增强网络对于这些复杂的背景上下文信息的学习能力。

经典的双阶段实例分割模型, 如 Mask R-CNN^[23], 使用 ROIAlign 来提取图3中最内部的候选框 ROI 的特征子集; 接着基于特征子集, 在候选框中进行舰船像素预测。然而, 这会导致以下问题。首先, ROI 的提取结果直接影响了像素掩模的结果。如果框检测的结果精度较低, 那么像素掩模的结果也将变差。其次, 基于 ROI 的特征子集进行舰船像素预测限制了网络的视野域。由于网络结构设计, ROI 以后的网络可操作空间均限制在 ROI 所提取的 ROI 中, 这导致了后端网络对于全局信息的感受能力有限, 对舰船目标的细节预测能力降低, 最终导致了检测精度的下降。图像背景内

容在多种场景下均会对舰船检测的结果产生影响。首先,对于近岸场景的舰船,它们往往受到复杂海岸背景的干扰,导致生成检测框的位置不准确。如图 4a 所示,紧凑的矩形框所包含的背景信息较少,网络对于复杂背景的学习能力较弱。如图 4b 所示,当扩大矩形框,像素预测过程中能够考虑更多的背景内容,结合全局信息,提高检测精度。其次,对于离岸场景的舰船,它们常常受到相干斑噪声、旁瓣和尾迹的干扰。如图 5a 所示,尽管紧凑的矩形框可以消除距离舰船较远的旁瓣,但也限制了网络对于旁瓣的学习能力,最终导致检测时出现旁瓣导致的虚警。如图 5b 所示,扩大矩形框,能够增强网络对于背景干扰的鉴别能力,从而减少虚警,提高检测精度。

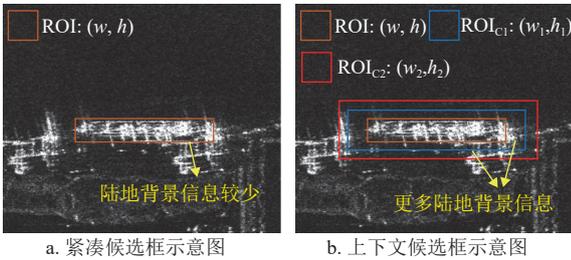


图 4 近岸场景内容扩充示意图

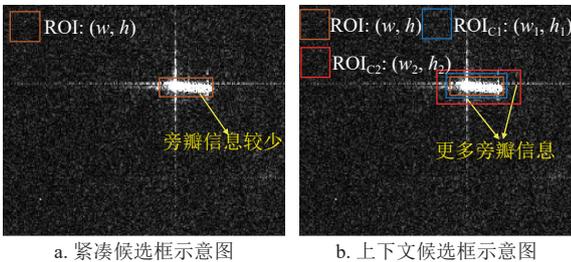


图 5 离岸场景内容扩充示意图

经典双阶段网络 ROI 生成结果如图 4 和图 5 所示,棕色矩形框表示由经典双阶段网络生成的 ROI。由图 4a 和图 5a 可知,经典双阶段网络生成的 ROI 感受野有限,这导致了网络对舰船四周的背景信息不敏感。而 FL-CI-SE-ROIE 在 ROIE 中扩充了上下文 ROI,即在初始中心候选框 ROI: (w, h) 周围设置了两种包含不同范围背景上下文 ROI,即 $ROI_{C1}: (w_1, h_1)$ 和 $ROI_{C2}: (w_2, h_2)$,如图 4b 和图 5b 所示。特别地,为了不增加过多计算成本和不影响检测速度,未设置多种不同范围的上下文 ROI。此外, k_1 被设为经验值 2, k_2 被设为经验值 3。过高的 k_1 和 k_2 会导致网络引入过多的背景干扰,导致网络因正负样本不平衡而训练难以收

敛,最终导致较低检测精度。因此,引入上下文内容并不是无限制的引入,引入过多的上下文反而会影响检测精度。

图 6 展示了 FL-CI-SE-ROIE 的上下文扩充信息提取和融合过程。

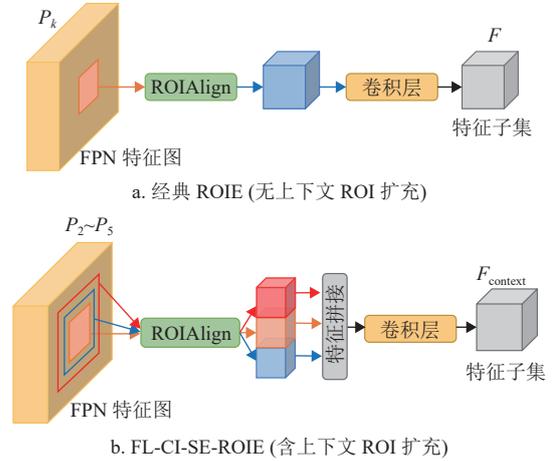


图 6 上下文扩充信息提取和融合过程

图 6a 展示了无上下文 ROI 扩充的经典 ROIE 特征子集提取过程,可描述为:

$$F = \text{conv}_{3 \times 3} \{ \text{ROIAlign}(\text{ROI}) \} \quad (7)$$

式中, F 表示经典 ROIE 提取的特征子集; ROIAlign 采样尺寸设置为 Mask R-CNN 默认值 14×14 , 输入通道宽度为 FPN 每个等级特征图的默认通道宽度 256, 因此最终 F 维度为 $14 \times 14 \times 256$ 。

在图 6b 中,当加入上下文背景信息扩充时,FL-CI-SEE-ROIE 的特征子集提取过程可描述为:

$$F_{\text{context}} = \text{conv}_{3 \times 3} \{ \text{ROIAlign}(\text{ROI}) \odot \text{ROIAlign}(\text{ROI}_{C1}) \odot \text{ROIAlign}(\text{ROI}_{C2}) \} \quad (8)$$

式中, \odot 表示特征拼接操作; F_{context} 表示提取的特征子集。此处特征子集的特征图被直接拼接,输入至一个 3×3 卷积层来实现降维,从而得到与式 (7) 中 F 相同的特征维度。通过上下文 ROI 的扩充、提取和融合,增加了感受野,增强了网络对背景信息的鉴别能力。

1.3 压缩激励模块

在 1.2 节中提出的上下文 ROI 扩充能够有效地扩大网络的感受野,获取目标上下文信息,但同时也增加了一定的背景干扰,限制了检测精度进一步提高。为平衡不同范围上下文 ROI 的贡献,抑制无用背景干扰,受文献 [30] 启发,引入 SE 模块。SE 指对输入特征图通道信息进行压缩并对各个通道信息的重要性建模,并根据建立模型对原始特征

图通道进行加权输出。SE 模块能够抑制部分无用的通道信息, 增强激励部分重要的通道信息。

在 FL-CI-SE-ROIE 中, SE 模块用于平衡 ROI: (w, h) 、ROI_{C1}: (w_1, h_1) 和 ROI_{C2}: (w_2, h_2) 的贡献, 关注重要的上下文信息, 避免 FL-CI-SE-ROIE 可能因引入过多上下文信息导致训练震荡的问题。该过程可描述为:

$$F'_{\text{context}} = \text{SE}(F_{\text{context}}) \quad (9)$$

图 7 为 SE 模块的具体结构。

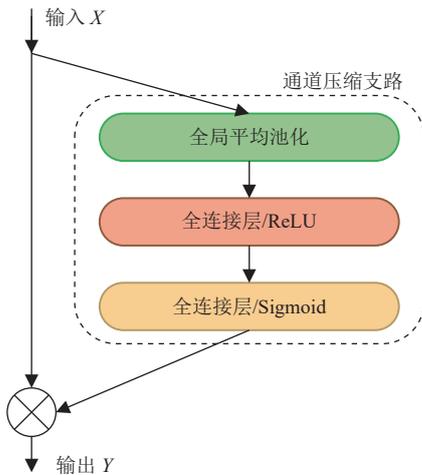


图 7 压缩激励模块结构

由图 7 可知, SE 模块实现方式类似于文献 [30]。输入 X 的维度为 $W \times H \times C$, 首先使用全局平均池化层 (global average pooling, GAP) 对输入进行全局平均池化处理, 即:

$$\text{GAP}(X) = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W X(i, j) \quad (10)$$

将 GAP 的输出 $\text{GAP}(X)$ 输入全连接层 (FC_1) 和线性整流激活函数 (linear rectification function, ReLU) 中, 在 FC_1 层中将输入信息压缩 r 倍。 r 设置为 3, 表示平衡输入 ROI、ROI_{C1} 和 ROI_{C2} 3 种不同范围的特征子集贡献。全连接层 $\text{FC}_1 \in \mathbb{R}^{(G/r) \times C}$, 得到维度为 $1 \times 1 \times (C/r)$ 的输出; 再使用另一个全连接层 (FC_2) 和 sigmoid 激活函数得到通道重要性权重 S , 该激励权重表示 ROI、ROI_{C1} 和 ROI_{C2} 的特征子集的通道重要性, 同时在 FC_2 层中将通道数还原为 C 。此处全连接层 $\text{FC}_2 \in \mathbb{R}^{C \times (C/r)}$, 得到维度为 $1 \times 1 \times C$ 的输出。最终, 在通道维度上, 将原始输入特征图 X 和通道重要性权重 S 使用乘法运算得到输出 $Y \in \mathbb{R}^{W \times H \times C}$:

$$Y = XS = X \text{sigmoid} \{ \text{FC}_2 [\text{ReLU} [\text{FC}_1 (\text{GAP}(X))]] \} \quad (11)$$

2 实验结果与分析

2.1 实验数据

本文在像素级多边形分割 SAR 舰船检测数据集 (polygon segmentation SAR ship detection dataset, PSeg-SSDD)^[31] 上进行实验以验证所提出方法的有效性。PSeg-SSDD 数据集包含 1 160 张 SAR 图像样本。SAR 图像样本由 Radarsat-2, TerraSAR-X 和 Sentinel-1 等卫星采集。SAR 图像样本的平均尺寸为 500 像素 \times 500 像素, 样本采集位置主要位于中国山东烟台港和印度维萨卡帕特南港, 数据集包含复杂靠岸场景的舰船以及简单离岸场景的舰船。PSeg-SSDD 数据集内舰船总数量为 2 587, 其中最大尺寸舰船所占面积为 78 797 像素, 最小尺寸舰船所占面积为 66 像素。

2.2 实验设置

FL-CI-SE-ROIE 使用 ResNet-101-FPN 作为骨干网络。为了加速网络收敛, 骨干网络在 ImageNet 数据集上进行了预训练。使用双线性插值算法将 SAR 图像样本调整尺寸为 512 像素 \times 512 像素。训练优化器选为随机梯度下降优化器 (stochastic gradient descent, SGD), 学习率设为 0.012 5, 动量设为 0.9, 权重衰减设为 0.000 1, 总训练迭代次数为 12 次, 学习率在第 8 次迭代和第 11 次迭代各降低 10 倍, 训练批次大小设为 2。损失函数由分类损失和回归损失组成, 分类损失 L_{CLS} 选择经典交叉熵损失^[32], 回归损失 L_{REG} 选择为经典平滑 L_1 ^[33], 即:

$$L_{\text{CLS}} = -\frac{1}{N} \sum_{i=1}^N p_i \log(p_i^*) + (1 - p_i) \log(1 - p_i^*) \quad (12)$$

$$L_{\text{REG}} = \frac{1}{N} \sum_{i=1}^N p_i^* \text{smooth}_{L_1}(t_i - t_i^*) \quad (13)$$

式中, N 为预测框的数目; $\text{smooth}_{L_1}(\cdot)$ 为平滑 L_1 损失, 定义为:

$$\text{smooth}_{L_1}(x) = \begin{cases} 0.5x^2 & \text{if } |x| < 1 \\ |x| - 0.5 & \text{otherwise} \end{cases} \quad (14)$$

定义 p_i 为预测类别概率; p_f^* 为真实类别标签; t_i 为预测矩形框坐标; t_i^* 为真实矩形框坐标; t_i 包括 $t_{i,x}$ 、 $t_{i,y}$ 、 $t_{i,w}$ 和 $t_{i,k}$; t_i^* 包括 $t_{i,x}^*$ 、 $t_{i,y}^*$ 、 $t_{i,w}^*$ 和 $t_{i,N}^*$:

$$t_{i,x} = (x_i - x_a) / w_a \quad t_{i,y} = (y_i - y_a) / h_a \quad (15)$$

$$t_{i,w} = \log(w_i/w_a) \quad t_{i,h} = \log(h_i/h_a) \quad (16)$$

$$t_{i,x}^* = (x_i^* - x_a)/w_a \quad t_{i,y}^* = (y_i^* - y_a)/h_a \quad (17)$$

$$t_{i,w}^* = \log(w_i^*/w_a) \quad t_{i,h}^* = \log(h_i^*/h_a) \quad (18)$$

式中, 下标 a 表示锚框参数。此外, 不同于框等级检测, SAR 舰船实例分割的分类损失 L_{CLS} 还包含一个像素预测二分类损失, 该二分类损失也使用经典交叉熵损失。总分类损失 L_{CLS} 为:

$$L_{CLS} = \lambda_1 L_{class} + \lambda_2 L_{pixel} \quad (19)$$

式中, L_{class} 为类别预测损失; L_{pixel} 为像素预测损失; λ_1 和 λ_2 为权重, 分别设置为经验值 1 和 1.5。在测试过程中, 非极大值抑制被用来移除多余重复的检测框, 模型的 IOU 阈值设为经验值 0.5^[34]。实验运行在配备 RTX 3090 GPU 和 Intel i9-9900K CPU 的计算机硬件平台上, 该计算机的内存为 32 G, 实验程序语言为 Python, 实验软件环境平台为 Pytorch。

2.3 评估指标

本文采用基于 Microsoft COCO 数据集的目标分割指标来衡量 SAR 舰船实例分割性能。记舰船真值掩模为 $Mask_G$, 舰船预测掩模为 $Mask_P$, 掩模交并比 Mask-IOU 的定义为:

$$Mask-IOU = \frac{Mask_G \cap Mask_P}{Mask_G \cup Mask_P} \quad (20)$$

记实例分割平均精度为 average precision(AP), AP 定义为:

$$AP = \int_0^1 p(r) dr \quad (21)$$

式中, r 表示召回率; p 表示准确率; $p(r)$ 表示精度-召回率曲线。本文 SAR 舰船实例分割精度的评估指标为 AP_{50} 、 AP_{75} 、 AP_S 、 AP_M 、 AP_L 、AP。其中, AP_{50} 表示当 Mask-IOU=0.50 时, 所有尺寸目标的平均精度。 AP_{75} 表示当 Mask-IOU=0.75 时, 所有尺寸目标的平均精度。 AP_S 表示当 Mask-IOU=0.50 时, 像素数小于 32^2 的小尺寸目标的平均精度。 AP_M 表示当 Mask-IOU=0.50 时, 像素数大于 32^2 且小于 96^2 的中等尺寸目标的平均精度。 AP_L 表示当 Mask-IOU=0.50 时, 像素数大于 96^2 的大尺寸目标的平均精度。AP 表示当 Mask-IOU 从 0.50 到 0.95, 间隔 0.05 时, 十组平均精度的均值。AP 可以反映在不同 Mask-IOU 阈值条件下的综合检测性能, 因此它被选择作为衡量 SAR 舰船

实例分割精度的唯一核心指标, 通过以下公式计算得到:

$$AP = \frac{1}{10} \sum_{Mask-IOU=0.50}^{0.95} AP_{Mask-IOU} \quad (22)$$

式中, Mask-IOU $\in [0.50:0.05:0.95]$, 共有 10 组取值。

此外, 本文还在实验中对比了模型复杂度和处理时间。模型复杂度使用模型参数量 (parameter, #Para) 来衡量, #Para 越高, 模型越复杂。处理时间指网络检测一张图片所需要的平均时间。

2.4 实验结果

由表 1 可知, 对比其他 9 种模型, 本文所提出的 FL-CI-SE-ROIE 提供了最高的 SAR 舰船实例分割精度, AP 达到了 61.0%, 相较于次优模型 PANet 高 1.4%, 证实了 FL-CI-SE-ROIE 具有更优越的 SAR 舰船实例分割性能。尽管 FL-CI-SE-ROIE 的 AP_L 指标低于 HQ-ISNet-w40, AP_{50} 低于 HTC, 但 FL-CI-SE-ROIE 在这两个指标上仍然取得了次优的结果。此外, 所提出的 FL-CI-SE-ROIE 的其他所有评估指标均高于对比方法, 因此总体来说 FL-CI-SE-ROIE 可以实现更准确的舰船目标实例分割, 并且能适应不同的舰船尺度, 即在 AP_{75} 、 AP_S 和 AP_M 指标上的结果均为最优。虽然, FL-CI-SE-ROIE 每张图片的处理时间略长, 且参数量略大于其他对比模型, 不过这换取了更高的检测精度。对比 YOLACT 模型, 虽然它的检测速度较快且模型参数较少, 但是它的检测精度远低于本文所提出的 FL-CI-SE-ROIE。总之, 根据以上定量对比结果, FL-CI-SE-ROIE 可以提供更优越的 SAR 舰船实例分割精度。

图 8 和图 9 展示了在像素级 PSeg-SSDD 数据集上, FL-CI-SE-ROIE 和 HTC 的 SAR 舰船实例分割定性结果对比。FL-CI-SE-ROIE 是在 HTC 模型基础上进行改进的, 因此展示了其与 HTC 的定性检测结果对比。图 8 展示了 HTC 模型的舰船实例分割结果。图 9 展示了 FL-CI-SE-ROIE 舰船实例分割结果。由图 8 和图 9 可知, 对比 HTC, FL-CI-SE-ROIE 漏检更少, 如对于图 8a, HTC 未检测到该靠岸舰船, 但是如图 9a 所示, FL-CI-SE-ROIE 成功检测到该靠岸舰船。这是因为 FL-CI-SE-ROIE 实现了上下文信息扩充, 增大了网络的感受野, 提高了对复杂海岸背景的鉴别能力, 从而成功地区分了海岸背景和舰船。而 HTC 模型由于对复杂海岸背景内容的学习能力不足, 导致其对海岸

背景和目标的鉴别能力较差, 最终导致漏检。此外, FL-CI-SE-ROIE 具有更强的抑制虚警能力。如图 8b 所示, HTC 产生了一个虚假检测 (图 8b 中棕色框标记), 但是如图 9b 所示, FL-CI-SE-ROIE 能够抑制该虚警。这是因为 FL-CI-SE-ROIE 的上下文信息能够增强网络对于噪声的鉴别能力。HTC 模型未采用上下文信息扩充, 导致图 9b 误将左上角的噪声当做舰船。而 FL-CI-SE-ROIE 在学习的过程中扩充了上下文信息, 因此网络对于噪声的鉴别能力更强, 因此正确地将其区分为噪声背景, 未

产生虚警。对于多尺度舰船和陆地背景混杂的场景, FL-CI-SE-ROIE 具有更好的实例分割性能, 如对于图 8c, HTC 产生了 5 个虚假检测 (图 8c 中棕色框标记), 包括海面虚假检测和陆地虚假检测, 而在图 9c 中, FL-CI-SE-ROIE 仅产生了 3 个虚假检测 (图 9c 中黄色框标记), 这也证明了 FL-CI-SE-ROIE 对多尺度目标具有更强的检测能力, 对于复杂背景具有更强的鉴别能力。根据以上结果, FL-CI-SE-ROIE 可以提供更优越的 SAR 舰船实例分割精度。

表 1 FL-CI-SE-ROIE 和其他方法实例分割对比

方法	AP ₅₀ /%	AP ₇₅ /%	AP _S /%	AP _M /%	AP _L /%	AP/%	网络推理耗时/s	#Para×10 ⁶
Mask R-CNN ^[23]	88.5	72.1	57.2	60.8	27.4	57.8	0.091	62.74
Mask Scoring R-CNN ^[35]	89.4	73.2	58.0	61.4	22.6	58.6	0.078	79.00
Cascade Mask R-CNN ^[36]	87.5	70.5	56.3	58.8	22.6	56.6	0.094	95.79
HTC ^[25]	91.7	73.1	58.7	<u>61.6</u>	34.8	59.3	0.086	95.92
PANet ^[27]	91.1	<u>74.0</u>	59.3	61.0	52.1	<u>59.6</u>	0.073	66.28
YOLACT ^[37]	88.0	52.1	47.3	53.5	40.2	48.4	0.065	53.72
GRoIE ^[38]	89.8	72.7	58.6	58.7	21.8	58.3	0.103	66.53
HQ-ISNet-w18 ^[20]	89.3	73.6	58.2	60.4	37.2	58.6	0.116	62.75
HQ-ISNet-w32 ^[20]	90.4	75.5	58.9	61.1	37.3	59.3	0.125	82.55
HQ-ISNet-w40 ^[20]	86.0	72.6	56.7	61.3	50.2	57.6	0.129	98.79
SA R-CNN ^[39]	90.4	73.3	<u>59.6</u>	60.3	20.2	59.4	0.073	53.75
FL-CI-SE-ROIE	<u>91.3</u>	77.6	60.9	62.6	<u>40.3</u>	61.0	0.112	107.55

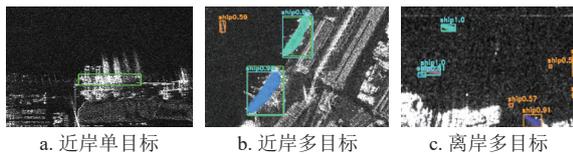


图 8 HTC 实例分割结果

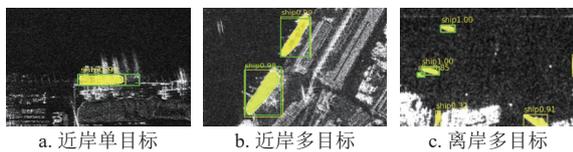


图 9 FL-CI-SE-ROIE 实例分割结果

2.5 消融实验

表 2 展示了在 FL-CI-SE-ROIE 中逐渐加入各种改进手段对 SAR 舰船实例分割精度的改善情况, 包括 3 项改进: 全等级 ROI 提取、上下文 ROI 扩充和压缩激励模块。由表 2 可知, 随着改进手段的逐渐加入, 实例分割精度被逐渐改善, 精度

从初始 59.3% AP 渐进式地上升到了最终 61.0% AP。如表 2 所示, 初始实例分割精度 AP 为 59.3%, 多尺度舰船的检测精度可以提现网络的多尺度描述能力。在加入全等级 ROI 提取后, 小尺度舰船的检测精度提高了 0.6%; 中尺度舰船的检测精度提高了 1.1%; 大尺度舰船的检测精度提高了 5.5%; 综合检测精度 AP 提升至 59.8%。由此可知, 全等级 ROI 提取增强了网络的多尺度描述能力。在加入上下文 ROI 扩充后, 扩大了感受野, 增强了对背景的鉴别能力, AP 上升至 60.7%。在加入压缩激励模块后, 平衡了不同范围上下文信息的贡献, 进一步提高了分割精度, AP 上升至 61.0%。逐渐加入改进手段时, 模型虽然复杂度更高且处理时间增加, 但复杂度并未提升过多, 且处理时间也未增加过多, 这对于精度的提升而言是可以接受的。综上, FL-CI-SE-ROIE 的各项改进均能提高 SAR 舰船实例分割的精度。

表 2 FL-CI-SE-ROIE 逐渐加入改进手段对精度影响

全等级	上下文	压缩激励	AP ₅₀ %	AP ₇₅ %	AP _S %	AP _M %	AP _L %	AP%	#Para×10 ⁶	推理耗时/s
—	—	—	91.7	73.1	58.7	61.6	34.8	59.3	95.92	0.086
√	—	—	90.4	74.6	59.3	62.7	40.3	59.8	95.92	0.091
√	√	—	91.4	75.5	60.0	63.0	30.3	60.7	106.76	0.099
√	√	√	91.3	77.6	60.9	62.6	40.3	61.0	107.55	0.112

3 结束语

针对目前深度学习 SAR 舰船实例分割方法未能考虑特征全等级信息和目标上下文信息的问题, 本文提出了一种基于全等级上下文压缩激励感兴趣区域提取器的 SAR 舰船实例分割方法 FL-CI-SE-ROIE。首先, FL-CI-SE-ROIE 实现了全等级 ROI 提取, 保留了特征的全等级信息, 增强了网络的多尺度描述能力。其次, FL-CI-SE-ROIE 实现了上下文 ROI 扩充, 获取了目标上下文信息, 增强了网络的背景鉴别能力。最后, FL-CI-SE-ROIE 引入压缩激励模块来平衡不同范围上下文 ROI, 可抑制背景干扰, 进一步提高了分割精度。在公开数据集 PSeg-SSDD 上的实验结果显示, FL-CI-SE-ROIE 的检测精度高于其他 9 种对比模型, 高于次优模型 1.4%。此外, FL-CI-SE-ROIE 在具有高检测精度的同时, 没有较大的检测速度牺牲。因此, FL-CI-SE-ROIE 虽然速度略慢于部分模型, 但获得了最优的检测精度。消融实验证明了 FL-CI-SE-ROIE 每项改进的有效性。

参 考 文 献

- [1] 李健伟, 曲长文, 彭书娟. 基于级联 CNN 的 SAR 图像舰船目标检测算法[J]. *控制与决策*, 2019, 34(10): 2191-2197.
LI J W, QU C W, PENG S J. A ship detection method based on cascade CNN in SAR images[J]. *Control and Decision*, 2019, 34(10): 2191-2197.
- [2] ZHANG T W, ZHANG X L, KE X, et al. LS-SSDD-v1.0: A deep learning dataset dedicated to small ship detection from large-scale sentinel-1 SAR images[J]. *Remote Sensing*, 2020, 12(18): 2997.
- [3] ZHANG T W, ZHANG X L. HTC+ for SAR ship instance segmentation[J]. *Remote Sensing*, 2022, 14(10): 2395.
- [4] ZHANG T W, ZHANG X L, KE X, et al. Quad-FPN: A novel quad feature pyramid network for SAR ship detection[J]. *Remote Sensing*, 2021, 13(14): 2771.
- [5] ZHANG T W, ZHANG X L, SHI J, et al. Balance scene learning mechanism for offshore and inshore ship detection in SAR images[J]. *IEEE Geoscience and Remote Sensing Letters*, 2022, 19: 4004905.
- [6] LENG X G, JI K F, YANG K, et al. A bilateral CFAR algorithm for ship detection in SAR images[J]. *IEEE Geoscience and Remote Sensing Letters*, 2015, 12(7): 1536-1540.
- [7] ELDHUSET K. An automatic ship and ship wake detection system for spaceborne SAR images in coastal regions[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 1996, 34(4): 1010.
- [8] TOUZI R, CHARBONNEAU F, HAWKINS R K, et al. Ship-Sea contrast optimization when using polarimetric SARs[C]//IGARSS 2001 Scanning the Present and Resolving the Future. Sydney: IEEE, 2001: 426-428.
- [9] SCHWEGMANN C P, KLEYNHANS W, SLAMON B P. Synthetic aperture radar ship detection using haar-like features[J]. *IEEE Geoscience and Remote Sensing Letters*, 2017, 14(2): 154-158.
- [10] XIE T, ZHANG W K, YANG L N, et al. Inshore ship detection based on level set method and visual saliency for SAR images[J]. *Sensors*, 2018, 18(11): 3877.
- [11] LIN H P, CHEN H, JIN K, et al. Ship detection with superpixel-level fisher vector in high-resolution SAR images[J]. *IEEE Geoscience and Remote Sensing Letters*, 2020, 17(2): 247-251.
- [12] ZHANG T W, ZHANG X L, LIU C, et al. Balance learning for ship detection from synthetic aperture radar remote sensing imagery[J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2021, 182: 190-207.
- [13] ZHANG T W, ZHANG X L. Squeeze-and-Excitation Laplacian pyramid network with dual-polarization feature fusion for ship classification in SAR images[J]. *IEEE Geoscience and Remote Sensing Letters*, 2022, 19: 4019905.
- [14] 李男, 叶晓东, 王昊, 等. 基于改进 YOLOv5 的复杂场景下 SAR 图像船舶检测方法[J]. *信号处理*, 2022, 19(5): 4019905.
LI N, YE X D, WANG H, et al. A ship detection method for SAR images in complex scene based on improved YOLO v5[J]. *Journal of Signal Processing*, 2022, 19(5): 4019905.
- [15] WU W T, LIU H, LI L L, et al. Application of local fully convolutional neural network combined with YOLO v5 algorithm in small target detection of remote sensing image [EB/OL]. [2021-10-29]. <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0259283>.
- [16] SHAO Z K, ZHANG X L, ZHANG T W, et al. RBFA-Net: A rotated balanced feature-aligned network for rotated SAR ship detection and classification[J]. *Remote Sensing*, 2022, 14(14): 3345.

- [17] ZHANG T W, ZHANG X L. Injection of traditional hand-crafted features into modern CNN-based models for SAR ship classification: What, why, where, and how[J]. *Remote Sensing*, 2021, 13(11): 2091.
- [18] ZHANG T W, ZHANG X L, SHI J, et al. HyperLi-Net: A hyper-light deep learning network for high-accurate and high-speed ship detection from synthetic aperture radar imagery[J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2020, 167: 123.
- [19] WEI S J, ZENG X F, QU Q X, et al. HRSID: A high-resolution SAR images dataset for ship detection and instance segmentation[J]. *IEEE Access*, 2020, 8: 1.
- [20] SU H, WEI S J, LIU S, et al. HQ-ISNet: High-Quality instance segmentation for remote sensing imagery[J]. *Remote Sensing*, 2020, 12(6): 989.
- [21] GAO F, HUO Y Y, WANG J, et al. Anchor-Free SAR ship instance segmentation with centroid-distance based loss[J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2021, 14: 11352-11371.
- [22] ZHANG T W, ZHANG X L, LI J W, et al. Contextual squeeze-and-excitation mask R-CNN for SAR ship instance segmentation[C]//2022 IEEE Radar Conference (RadarConf22). New York: IEEE, 2022: 1-6.
- [23] HE K M, GKIOXARI G, DOLLAR P, et al. Mask R-CNN[C]//Proceedings of the IEEE International Conference on Computer Vision. Italy: IEEE, 2017: 2961-2969.
- [24] ZHANG T W, ZHANG X L. A mask attention interaction and scale enhancement network for SAR ship instance segmentation[J]. *IEEE Geoscience and Remote Sensing Letters*, 2022, 19: 4511005.
- [25] CHEN K, PANG J M, WANG J Q, et al. Hybrid task cascade for instance segmentation[C]//IEEE Conference on Computer Vision and Pattern Recognition. Long Beach: IEEE, 2019: 4969-4978.
- [26] ZHANG T W, ZHANG X L, SHI J, et al. Depthwise separable convolution neural network for high-speed SAR ship detection[J]. *Remote Sensing*, 2019, 11(21): 2483.
- [27] LIU S, QI L, QIN H F, et al. Path aggregation network for instance segmentation[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake: IEEE, 2018: 8759-8768.
- [28] ZHANG T W, ZHANG X L. ShipDeNet-20: An only 20 convolution layers and <1-MB lightweight SAR ship detector[J]. *IEEE Geoscience and Remote Sensing Letters*, 2021, 18(7): 1234-1238.
- [29] ZHANG T W, ZHANG X L. A full-level context squeeze-and-excitation ROI extractor for SAR ship instance segmentation[J]. *IEEE Geoscience and Remote Sensing Letters*, 2022, 19: 4506705.
- [30] HU J, SHEN L, SUN G. Squeeze-and-Excitation networks[C]//Proceedings of the IEEE conference on Computer Vision and Pattern Recognition. Salt Lake: IEEE, 2018: 7132-7141.
- [31] ZHANG T W, ZHANG X L, LI J W, et al. SAR ship detection dataset (SSDD): Official release and comprehensive data analysis[J]. *Remote Sensing*, 2021, 13(18): 3690.
- [32] ZHANG T W, ZHANG X L. A polarization fusion network with geometric feature embedding for SAR ship classification[J]. *Pattern Recognition*, 2021, 123: 108365.
- [33] ZHANG T W, ZHANG X L, KE X, et al. HOG-ShipCLSNet: A novel deep learning network with hog feature fusion for SAR ship classification[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2021, 60: 5210322.
- [34] ZHANG T W, ZHANG X L. High-Speed ship detection in SAR images based on a grid convolutional neural network[J]. *Remote Sensing*, 2019, 11(10): 1206.
- [35] HUANG Z J, HUANG L C, GONG Y C, et al. Mask scoring R-CNN[C]// IEEE Conference on Computer Vision and Pattern Recognition. Long Beach: IEEE, 2019: 6409-6418.
- [36] CAI Z W, VASCONCELOS N. Cascade R-CNN: Delving into high quality object detection[C]//IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake: IEEE, 2018: 6154.
- [37] BOLYA D, ZHOU C, XIAO F Y, et al. YOLACT: Real-Time instance segmentation[C]//IEEE International Conference on Computer Vision. Seoul: IEEE, 2019: 9156-9165.
- [38] ROSSI L, KARIMI A, PRATI A. A novel region of interest extraction layer for instance segmentation [C]//International Conference on Pattern Recognition. Italy: IEEE, 2021: 2203.
- [39] ZHAO D P, ZHU C B, QI J, et al. Synergistic attention for ship instance segmentation in SAR images[J]. *Remote Sensing*, 2021, 13(21): 4384.