

基于频域感知与原型对齐的多源域泛化SAR舰船检测方法

艾加秋^{1,2}, 朱雅喃^{1*}, 郭 丹¹, 庄 硕¹, 陈立福³

(1. 合肥工业大学计算机与信息学院, 安徽合肥 230009; 2. 安徽省智能互联系统实验室(合肥工业大学), 安徽合肥 230009;
3. 长沙理工大学人工智能学院, 湖南长沙 410114)

摘 要: 合成孔径雷达(Synthetic Aperture Radar, SAR)是一种主动微波遥感成像技术,具备全天时、全天候的工作能力,在海洋环境感知中具有不可替代的作用。SAR舰船检测在海洋监管、海上安全保障和海上交通管理等领域发挥着重要作用,是海上目标感知和安全检测中的重要技术手段,尤其是在复杂海洋环境下。但入射角、海况和噪声水平会导致成像条件不同, SAR舰船数据存在明显的分布差异,导致模型在未知场景下不能达到稳定的舰船检测性能。域泛化技术是指仅利用源域数据来学习域不变特征表示,进而提升模型在未见目标域中的舰船检测能力。多源域泛化为学习鲁棒的舰船表征提供了必要的条件,但同时也加大了域间偏移,提升了优化难度。针对上述问题,本文提出一种基于频域感知与原型对齐的多源域泛化SAR舰船检测方法。该方法从频域建模、语义约束和源域加权优化这三个方面联合优化,以综合提高模型的跨域鲁棒性能。具体而言,由于SAR图像上的斑点噪声、海杂波与舰船结构在频域中有不同表现,本文设计了频域感知特征增强模块,利用傅里叶变换对特征进行分解,然后将低频结构信息和高频噪声成分重新加权,从而增强与舰船结构有关的域不变特征,抑制域相关扰动。其次,为缓解不同源域中的舰船目标由成像条件所导致的语义偏移,本文基于候选特征来构建各个源域的舰船语义原型,然后在特征空间里进行显式的对齐,进而达到跨域一致的语义表示,以减少舰船目标之间的语义偏差。最后,考虑到各个源域给模型带来的贡献大小是不一样的,部分源域与未知目标域的分布更加接近,但是部分源域由于噪声过大、样本质量差等原因会对模型造成影响,本文提出了多源自适应加权策略。该策略会根据训练损失来动态分配权重,进而有效地降低不可靠源域所带来的负迁移影响。实验结果表明,本文所提方法在多个SAR舰船检测数据集上显著优于现有方法,且舰船检测精度和泛化能力均取得了显著提升。

关键词: SAR舰船检测;多源域泛化;频域感知;原型对齐;多源自适应加权;负迁移抑制

基金项目: 国家自然科学基金面上项目(No. 62471171, No. 62071164);合肥市揭榜挂帅重大项目(No. 2024SZD004);中央高校基本科研业务费专项资金(No. PA2025HSL0111)

中图分类号: TN958; TP391.41 **文献标识码:** A **文章编号:** 0372-2112(XXXX)XX-0001-12
电子学报 URL: <http://www.ejournal.org.cn> **DOI:** 10.12263/DZXB.20260607

Multi-Source Domain Generalization for SAR Ship Detection Based on Frequency-Aware and Prototype Alignment

AI Jiaqiu^{1,2}, ZHU Yanan^{1*}, GUO Dan¹, ZHUANG Shuo¹, CHEN Lifu³

(1. School of Computer Science and Information Engineering, Hefei University of Technology, Hefei, Anhui 230009, China;
2. Systems Laboratory of Anhui Province, Hefei University of Technology, Hefei, Anhui 230009, China;
3. School of Artificial Intelligence, Changsha University of Science and Technology, Changsha, Hunan 410114, China)

Abstract: Synthetic Aperture Radar (SAR) is an active microwave remote sensing technique with all-day and all-weather imaging capabilities, playing an irreplaceable role in maritime environment perception. SAR ship detection is essential for target perception and maritime safety monitoring in complex marine environments. However, variations in imaging conditions, including incidence angle, sea state, and noise level, often cause substantial distribution shifts in SAR data, leading to unstable detection performance in unseen scenarios. Domain generalization addresses this issue by learning domain-invariant representations from source domains to improve detection in unseen target domains. In multi-source settings, diverse source distributions offer richer supervision for robust ship representation learning, but also introduce more complex inter-domain discrepancies and optimization challenges. To tackle these challenges, we propose a frequency-aware and prototype-aligned multi-source domain generalization framework for SAR ship detection. The framework enhances cross-do-

main robustness from three aspects. Specifically, considering that speckle noise, sea clutter, and ship structures exhibit different characteristics in the frequency domain, a frequency-aware feature enhancement module is designed. This module decomposes features using Fourier transform and reweights low-frequency structural information and high-frequency noise components, thereby enhancing domain-invariant ship-related features and suppressing domain-specific disturbances. Moreover, to alleviate semantic shifts caused by different imaging conditions across source domains, ship semantic prototypes are constructed from candidate features and explicitly aligned in the feature space, leading to more consistent cross-domain semantic representations. Finally, considering that different source domains contribute unequally to model training and that some domains may negatively affect optimization due to severe noise or poor sample quality, a multi-source adaptive weighting strategy is introduced. This strategy dynamically assigns weights according to training losses, effectively reducing negative transfer caused by unreliable source domains. Experiments on multiple SAR ship detection datasets show that the proposed method consistently outperforms existing approaches in cross-domain scenarios, achieving improved detection accuracy and generalization capability.

Keywords: SAR ship detection; multi-source domain generalization; frequency-domain awareness; prototype alignment; multi-source adaptive weighting; negative transfer suppression

Foundation Item(s): National Natural Science Foundation of China (No.62471171, No.62071164); Key Common Technology R&D Project in Hefei (No.2024SZD004); Fundamental Research Funds for the Central Universities of China (No.PA2025IISL0111)

0 引言

作为一种主动微波遥感成像技术,合成孔径雷达(Synthetic Aperture Radar, SAR)具备全天时、全天候的工作能力,在海洋环境感知中具有不可替代的作用^[1-2]。SAR舰船检测技术在海洋监管、海上安全保障和海上交通管理等领域发挥着重要作用,是海洋智能感知系统中不可或缺的重要环节^[3-6]。目前深度学习的发展使SAR舰船检测技术有了很大的进步,但是由于海况的变化、传感器参数的不同、成像角度的不同等原因,SAR图像都会出现复杂的统计特性^[7-8]。该特点会造成舰船检测模型在不同的场景之间泛化能力受到限制,从而造成检测精度下降。

为了解决上述问题,一些研究会依赖于收集更多标注数据集或对目标场景进行有针对性的模型适配^[9-11]。例如,一类方法通过扩展训练数据的规模并对特定目标域进行训练,在一定程度上提升模型性能。但是,这在SAR场景中存在很大的局限性。一方面,由于SAR图像需要专业知识的人员进行特定的标注,标签获取的成本非常高。另一方面,目标域往往是未知且无法提前获取的,因此以上基于目标域数据的适配方法无法适用。那么,在不依赖目标域数据的情况下,如何才能使模型具有更好的跨场景适应性和泛化能力?这是SAR舰船检测中急需解决的问题。

在此背景下,域泛化^[12-13]为解决该问题提供了一种有效的思路,其核心目标是在仅利用源域数据来训练的条件下,让模型能够在未见过的目标域上仍然能保持良好的性能。相比于依赖目标域数据的域自适应方法,域泛化更加符合实际的应用需求。多源域泛

化就是利用多个不同分布的源域数据,使模型可以从更多的分布中学习共享特征,进而提高它的泛化能力。

多源域设置在提供更丰富的训练信息的同时,也带来了新的挑战。第一,SAR图像的域差异主要是由频域噪声和成像机制造成的,舰船结构信息主要集中在低频区,高频噪声随着场景的变化而变化。目前的方法很难区分结构信息和噪声扰动,进而影响到模型的跨域鲁棒性。第二,不同源域内同类型舰船目标受到成像条件、背景等方面因素的影响,在特征空间里会出现语义偏差,使得模型很难形成统一的、稳定的语义表示。第三,多源数据在分布、质量上存在较大差异,各个源域对模型训练的贡献不均等,等权训练会导致负迁移,影响模型收敛。综上,本文的主要贡献有以下几点。

(1)针对SAR图像中频域噪声和成像机制造成的域差异问题,本文提出一种频域感知特征增强方法。该方法对特征进行频域分解,然后将低频结构信息和高频噪声成分重新加权,从而增强与舰船结构有关的域不变特征,抑制域相关扰动,从特征层提高SAR舰船检测模型的跨域鲁棒性。

(2)针对跨域场景下的舰船语义偏差问题,本文提出一种舰船语义原型对齐机制。该机制构建并对齐多源域中的舰船语义中心,约束各个域中的同类目标特征分布,进而减轻跨域语义偏差问题,实现更一致和更稳定的舰船语义表示。

(3)为了解决多源数据分布差异与负迁移问题,本文提出一种多源自适应加权策略。该策略会根据各个源域的训练状态,动态地改变各个源域的优化权

重,从而抑制不可靠源域带来的负迁移影响,进而提高多源域训练过程中的稳定性和整体泛化性能。

1 相关工作

1.1 SAR舰船检测

SAR舰船检测在海洋资源管理、海上安全保障以及海上交通检测等方面有着十分重要的作用和价值^[14-15],而且是遥感领域中重要的研究方向。和可见光成像相比,SAR具有全天时全天候、抗遮蔽等特点,在恶劣天气和夜间都可以稳定地获得海上目标信息,所以SAR舰船检测在海上目标感知任务中具有不可替代的作用。近几年,深度学习的发展使依靠卷积神经网络的SAR舰船检测方法取得明显进步。早期的方法大多使用Faster R-CNN^[16]、YOLO^[17]等经典的检测结构,再用适合SAR图像特点的特征提取和多尺度融合策略来提高检测效果。另外,一些研究用特征金字塔^[18]、多尺度特征融合^[19]等机制来加强对不同尺度目标的表示能力。除此之外,由于SAR图像中的舰船目标具有较强的散射特性和方向的不确定性,一些研究也开始采用旋转框检测或方向敏感特征建模的方式,提高对任意方向舰船的检测能力^[20]。

随着研究的深入,越来越多的方法开始重视SAR图像特有的成像机理和统计特性。为了解决SAR图像中经常出现的斑点噪声、背景杂波等干扰问题,一些算法会对已有的特征表示加以改良,或者利用注意力机制来抑制无用的干扰信息,从而优化目标检测的精确度^[21-22]。另外,一类算法会将上下文信息或者结构先验结合起来,对舰船的几何形状和空间分布进行建模,以提高模型对复杂场景的适应性^[23]。近年来,一些依靠不确定性建模、对比学习策略、多尺度融合机制等技术的成果也越来越多。这些成果可以减轻不同尺度、不同背景分布所引发的舰船目标检测偏差问题^[24-25]。

1.2 域泛化

域泛化是指模型只用源域数据进行训练,就能在未知的目标域上取得较好的效果,并且不需要使用目标域的数据来对模型进行适配^[26-27]。近些年,域泛化在计算机视觉范畴内得到了广泛的关注,按照任务类型可以分为基于图像分类的域泛化方法和基于目标检测的域泛化方法。目标检测任务同时包含分类和定位两个部分,因此它的跨域泛化难度要比分类问题更大^[28]。目前大部分研究都是从特征学习、训练策略、模型结构这三个方面进行的。

一些方法主要是学习域不变特征来提高泛化能力,例如,Vidit等人^[29]提出了一种面向目标检测的单域泛化方法,通过减小训练和潜在目标域之间的差

距,有效提高了模型在未见域的目标定位。Danish等人^[30]提出通过源域多样性增强和特征对齐相结合的策略,有效提升模型在未见域中的检测鲁棒性。该方法主要针对单源域目标检测的泛化问题。Al-Emadi等人^[31]从真实分布偏移出发,揭示了现有检测器在未见域中的性能退化问题,并为提升跨域检测性能提供了重要的参考。另一种方法就是用元学习的思想来模拟域间分布的变化来提高模型的泛化能力。例如,Dou等人^[32]提出了将源域划分成多个子域,在训练过程中进行跨域验证,以提高模型对未知域的适应性。

虽然域泛化在目标检测领域取得了一定的进展,但是现有方法仍多集中于自然图像场景之中^[33-35],难以直接应用于SAR图像中。SAR图像一般会存在斑点噪声大、纹理信息少、散射特性复杂等特点,因此它在不同场景下的分布差别要比自然图像更大。另外,不同的传感器类型、成像条件的变化都会使数据分布不一致的程度变大,进而使模型的泛化难度变大。

2 频域感知与原型对齐的多源域泛化SAR舰船检测方法

2.1 问题定义与总体框架

本文研究多源域泛化条件下的SAR舰船检测问题,本文所提方法的整体框架图如图1所示。具体流程为输入多源SAR图像,图像在经过共享的特征提取器之后可以获得由频域感知特征增强模块所得到的重构特征。该重构特征首先进入检测头进行目标检测,其次该重构特征会提取ROI特征用于构建语义原型,并进行对齐约束。在此基础上,模型会对各个源域的检测损失计算动态权重,并对多源损失进行加权融合,从而实现检测模型端到端的优化。

首先,定义本文所需的数据集表示,具体如下所示。假设共有 k 个源域,那么数据集表示为

$$D_s = \{D_1, D_2, \dots, D_k\} \quad (1)$$

其中,第 k 个源域定义为

$$D_k = \{\mathbf{x}_i^k, \mathbf{y}_i^k, \mathbf{b}_i^k\}_{i=1}^{N_k} \quad (2)$$

其中, $\mathbf{x}_i^k$ 表示输入的SAR图像, $\mathbf{y}_i^k$ 为类别标签, $\mathbf{b}_i^k$ 表示对应的边界框标注, N_k 是该源域样本数量。目标域数据记为 D_t ,其分布与所有源域均存在差异,且在训练过程中不可见。在训练阶段,模型被输入多个源域数据进行训练,但在测试阶段,模型仅在一个未知且训练阶段不可访问的目标域上进行测试和评估。

2.2 频域感知特征增强模块

如图2所示,从SAR成像特性来说,舰船目标一般是由船体轮廓、甲板结构、强散射部件共同构成,

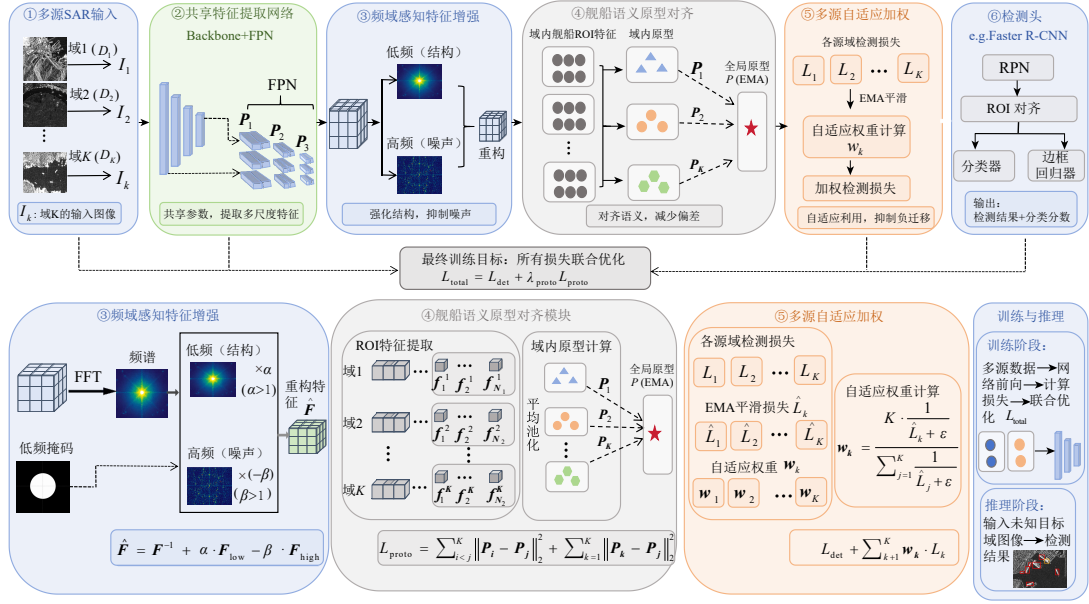


图1 基于频域感知与原型对齐的多源域泛化SAR舰船检测框架图

Figure 1 Overall framework of the proposed frequency-aware and prototype-aligned multi-source domain generalization for SAR ship detection

主体形态和空间布局比较稳定,可以反映目标整体结构信息。这类结构信息在图像频域里一般表现为低频或者中低频部分。与之相比,SAR图像上普遍存在的斑点噪声、海杂波和由于海面粗糙散射造成的随机纹理变化,具有很强的局部突变性、不确定性,在频域中更容易表现出高频扰动。

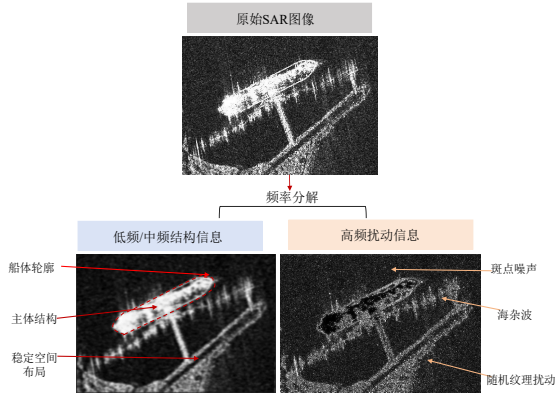


图2 SAR图像中低频结构和高频扰动物理特性示意图

Figure 2 Illustration of low-frequency structural components and high-frequency disturbance characteristics in SAR images

根据以上分析可知,本文设计出频域感知特征增强模块,加强低频结构响应,削减部分高频随机干扰,使模型更多地关注舰船目标的轮廓、尺度以及主体散射结构这类域间相对稳定的特征信息,并且降低了斑点噪声、海杂波这些域相关干扰对特征表示造成的干扰。本模块中假设 Backbone 和 FPN 输出的某一个尺度特征图为

$$\mathbf{F} \in \mathbb{R}^{C \times H \times W} \quad (3)$$

其中, C 为通道数, H 和 W 分别表示特征图的空间尺寸。对其进行二维傅里叶变换,得到频域表示:

$$\hat{\mathbf{F}} = \mathcal{F}(\mathbf{F}) \quad (4)$$

其中, $\mathcal{F}(\cdot)$ 表示快速傅里叶变换(FFT)。为了区分低频与高频信息,构造一个中心对称的低频掩码 $\mathbf{M} \in \mathbb{R}^{H \times W}$, 其中心区域取值为1, 外围区域取值为0。于是,低频和高频频谱被分别表示为

$$\hat{\mathbf{F}}_{\text{high}} = (1 - \mathbf{M}) \odot \hat{\mathbf{F}} \quad (5)$$

其中, $\odot$ 表示逐元素乘法。通过逆傅里叶变换,可以得到对应的空间域特征:

$$\mathbf{F}_{\text{low}} = \mathcal{F}^{-1}(\hat{\mathbf{F}}_{\text{low}}), \mathbf{F}_{\text{high}} = \mathcal{F}^{-1}(\hat{\mathbf{F}}_{\text{high}}) \quad (6)$$

因此,原始特征可以近似表示为 $\mathbf{F} \approx \mathbf{F}_{\text{low}} + \mathbf{F}_{\text{high}}$, 基于上述分解,本文构造频域增强后的特征为

$$\tilde{\mathbf{F}} \approx \mathbf{F} + \alpha \mathbf{F}_{\text{low}} - \beta \mathbf{F}_{\text{high}} \quad (7)$$

其中, $\alpha > 1$ 和 $\beta > 0$ 分别表示低频结构增强系数和高频噪声抑制系数。本文实验中低频增强系数 α 、高频抑制系数 β 在所有的数据集以及所有的源域组合中都是相同的,没有对特定的目标域进行单独调节。本文对 FPN 的各个尺度特征分别使用这个模块,在不同的感受野里进行频域去噪和结构增强。

2.3 舰船语义原型对齐模块

虽然前面的频域感知特征增强模块可以在一定程度上抑制高频噪声,增强结构信息,提高特征的稳定性,但是在多源域情况下,各个源域的舰船目标在特征空间里仍会存在较大的语义偏差。这些偏差主

要是由于成像条件不同所造成的,比如入射角、分辨率、背景散射等因素的不同,使得模型容易被域特有模式影响而做出判断,进而降低它在未知域中的泛化能力。因此,只依靠特征增强不能保证跨域语义的一致性,需要从语义层面对特征分布进行约束。本文采用舰船语义原型对齐机制,建立各个源域的舰船语义中心,在特征空间里进行显式的对齐,进而达到跨域一致的语义表示。

为了降低训练初期低质量候选框对语义原型的影响,本文只用训练阶段被标记为舰船类别的正样本ROI特征来构建原型,而不会使用背景ROI或未匹配的候选区域。具体而言,对于第 k 个源域,首先通过检测框对应的ROI对齐操作来提取舰船区域的特征。本文的ROI特征是检测头分类分支和回归分支之前共享特征层得到的。经过频域感知特征增强模块处理过的FPN特征,再经过RPN得到候选区域,最后用ROI Align从对应尺度的特征图中提取候选区域特征。该特征又送入检测头的共享全连接层做特征变换,本文选择最后一层共享全连接层输出的特征向量作为舰船ROI表示,用作后续语义原型的构建。由于各个源域的SAR图像散射强度、对比度以及背景分布存在差别,直接用未归一化的特征会造成原型被特征幅值的变化所影响,在计算原型之前,本文对每一个ROI特征做L2归一化处理,使原型对齐过程更注重特征方向上的语义一致性。

$$\mathcal{R}_k = \{\mathbf{r}_i^k\}_{i=1}^{N_k^r}, \mathbf{r}_i^k \in \mathbb{R}^d \quad (8)$$

其中, N_k^r 表示该源域中舰船目标的ROI数量, d 为特征维度。基于该特征集合,本文定义源域 k 的舰船语义原型为

$$\mathbf{P}_k = \frac{1}{N_k^r} \sum_{i=1}^{N_k^r} \mathbf{r}_i^k \quad (9)$$

该原型可以被看作该源域中舰船类别的一阶统计中心,反映的是该域下面舰船目标的语义特征分布。源域原型在训练时用小批次均值和指数滑动平均相结合的方式更新。目前批次中第 k 个源域的有效舰船ROI特征先被求平均,得到批次层级的原型,再用它更新历史源域原型。如果当前批次中没有该源域有效的舰船ROI,就保留它的原型。但由于各个源域分布不同,分别计算得到的原型 $\mathbf{P}_k$ 往往存在偏移。因此,为了进一步获得跨域共享的语义表示,本文引入全局原型:

$$\bar{\mathbf{P}} = \frac{1}{K} \sum_{k=1}^K \mathbf{P}_k \quad (10)$$

并通过指数滑动平均(Exponential Moving Average, EMA)对模型进行更新:

$$\mathbf{P} \leftarrow r\mathbf{P} + (1-r)\bar{\mathbf{P}} \quad (11)$$

其中, r 为动量系数。该设计可以使训练过程中的全局语义信息不断积累,不会因为单次批次的波动造成原型不稳定,进而得到更加鲁棒的语义中心表示。在此基础上,为了同时约束域间的一致性和全局的一致性,本文提出了如下原型对齐损失:

$$L_{\text{proto}} = \sum_{i < j} \|\mathbf{P}_i - \mathbf{P}_j\|_2^2 + \sum_{k=1}^K \|\mathbf{P}_k - \mathbf{P}\|_2^2 \quad (12)$$

其中,第一项是使任意两个源域原型的距离最小化,即缩小不同域之间语义的差异,第二项是把各个源域的原型拉到全局原型上,在全局上约束语义的一致性。双重约束可以很好地防止局部对齐所引起的偏移。设第 k 个域的条件分布为 $\mathbf{p}_k(r|y = \text{ship})$,则原型 $\mathbf{P}_k$ 实质上对应该分布的一阶统计量,即经验均值:

$$\mathbf{P}_k \approx E_{\mathbf{p}_k(r|y = \text{ship})}[\mathbf{y}] \quad (13)$$

因此,最小化 L_{proto} 等价于使不同源域下舰船类条件分布的一阶统计量趋于一致,即缩小跨域同类样本的内部差异。

2.4 多源自适应加权策略

在多源域泛化的场景中,各个源域给模型带来的贡献大小是不一样的。部分源域与未知目标域的分布更加接近,但是部分源域由于噪声过大、样本质量差等原因会对模型造成影响。本文提出了一种多源自适应加权策略,在训练过程中根据源域学习的状态,动态调整权重因子来提高整体优化的效果。对于第 k 个源域,其检测损失定义为

$$L_k = L_{\text{rpn_cls}}^{(k)} + L_{\text{rpn_reg}}^{(k)} + L_{\text{roi_cls}}^{(k)} + L_{\text{roi_reg}}^{(k)} \quad (14)$$

其中, $L_{\text{rpn_cls}}^{(k)}$ 和 $L_{\text{rpn_reg}}^{(k)}$ 分别表示区域建议网络(RPN)的分类损失与回归损失, $L_{\text{roi_cls}}^{(k)}$ 和 $L_{\text{roi_reg}}^{(k)}$ 分别表示ROI检测头的分类损失和回归损失。为了降低瞬时波动给域权重计算带来的影响,本文用EMA对每一个源域的检测损失进行平滑:

$$\hat{L}_k \leftarrow \rho \hat{L}_k + (1-\rho)L_k \quad (15)$$

其中, $\rho \in [0, 1]$ 是平滑系数。本文用源域的训练损失来度量它的可靠性。直观上来说,如果某一个源域上的损失小,那么模型在这个域上更容易学习到稳定的、一致的判别模式,它所对应的分布就越“干净”,或越接近共享的语义空间。反之,如果损失很大,那么该源域的噪声就比较多或分布偏移较大。本文用源域损失的倒数来衡量可靠性,也就是损失越小的源域,它的可靠性就越高。因此,定义第 k 个源域的可靠性分数为

$$r_k = \frac{1}{\hat{L}_k + \varepsilon} \quad (16)$$

其中, $\hat{L}_k$ 表示第 k 个源域的经验风险(采用指数滑动平均来进行平滑,以达到减少训练震荡的目的), $\varepsilon =$

10^{-6} 是防止数值不稳定的平滑项。该定义使可靠性分数和损失呈单调递减的关系,保证损失小的源域有较高的可靠性。在此基础上对各个源域的可靠度进行归一化处理,得到动态权重如下:

$$\mathbf{w}_k = \frac{K \cdot r_k}{\sum_{j=1}^K r_j} \quad (17)$$

其中, k 表示源域数量。低损失源域并不总是代表它和未知目标域最接近。本文用源域训练损失来衡量源域在当前训练阶段学习的稳定性、监督的可靠性,而不是直接估计它和目标域的距离。为了避免模型过于偏向某个简单的源域,本文使用归一化的相对权重分配方式,保证所有的源域一起参与优化。低损失源域只在当前阶段得到相对较高的优化权重,不会完全取代其他的源域。另外,本文使用指数滑动平均对各个源域损失进行平滑,防止由于单个批次的低损失或高损失造成权重急剧变化,使源域权重可以随着较为稳定的训练趋势逐渐更新,提高多源域训练过程的稳定性。上式归一化形式保证了权重满足如下条件:

$$\sum_{k=1}^K \mathbf{w}_k = K \quad (18)$$

因此,整体损失的尺度和原始未加权的形式是一致的,有利于训练过程的稳定。需要指出的是,权重的设计存在明显的单调性,即某一个源域的损失下降时,对应该源域的权重 $\mathbf{w}_k$ 会增大,在后面优化中起到更加重要的作用。对于损失较高的源域,它的权重会自动减小,从而有效抑制它对模型更新的不利影响。最终的多源加权检测损失可以写成如下形式:

$$L_{\text{det}} = \sum_{k=1}^K \mathbf{w}_k L_k \quad (19)$$

其中, L_k 表示第 k 个源域上的检测损失。使用该种加权机制之后,在参数更新过程中就会更加重视训练稳定、风险小的源域,从而达到增强“可靠信息”、抑制“噪声信息”的目的。

2.5 总体优化目标

本文从特征建模、语义对齐、多源优化这三个方面来建立统一的多源域泛化 SAR 舰船检测框架。整体训练目标由检测损失和语义对齐损失一起构成,用公式表示为

$$L_{\text{total}} = L_{\text{det}} + \lambda_{\text{proto}} L_{\text{proto}} \quad (20)$$

其中, L_{det} 表示多源加权检测损失, L_{proto} 表示舰船语义原型对齐损失, λ_{proto} 为平衡两者的权重系数。进一步展开,上式可写为

$$L_{\text{total}} = \sum_{k=1}^K \mathbf{w}_k L_k + \lambda_{\text{proto}} \left(\sum_{i < j} \| \mathbf{P}_i - \mathbf{P}_j \|_2^2 + \sum_{k=1}^K \| \mathbf{P}_k - \mathbf{P} \|_2^2 \right) \quad (21)$$

其中, L_k 表示第 k 个源域上的检测损失, $\mathbf{w}_k$ 为对应的自适应权重, $\mathbf{P}_k$ 为第 k 个源域的舰船语义原型, $\mathbf{P}$ 为

全局语义原型。 λ_{proto} 是用来控制语义原型对齐损失在整个优化目标中所占的比重。如果 λ_{proto} 取值太小,那么语义原型对齐约束就比较弱,不同的源域舰船 ROI 特征之间的语义偏差不能得到很好的抑制,模型学到的舰船表征仍然会受到源域特有分布的影响。如果 λ_{proto} 取值过大,那么模型就会过分看重源域原型之间的一致性,从而削弱检测分类、边界框回归等主任务的优化效果,而且会压缩舰船目标特征的局部差异,影响目标细节的表达和定位性能。因此, λ_{proto} 要同时满足语义一致性约束和检测任务的优化。

3 实验结果与分析

为了检验所提出的方法在多源域泛化 SAR 舰船检测任务中是否有效,本文在多个公开的 SAR 舰船检测数据集上进行了实验。选择 SSDD、AIRSAR、HRSID、SARShip 四个有代表性的数据集来建立多源域的实验环境。该数据集在成像传感器类型、分辨率、场景复杂程度等方面都存在较大的差别,可以很好地模拟出实际使用过程中由于成像条件发生变化所造成的数据分布偏移问题,为模型跨域泛化能力的评价提供良好的基础。其中, SSDD 数据集来源于 Sentinel-1 卫星影像,分辨率适中,场景简单,可以作为基础源域。AIRSAR 数据集由机载 SAR 获取,分辨率高,散射特性复杂,有较强的斑点噪声,域特性差异大。HRSID 是高分辨率数据集,包含港口、近岸、远海等多种复杂场景,目标尺度变化大,分布密集,对模型鲁棒性要求高。SARShip 数据集场景多样、背景复杂,与其他数据集存在较大差别,适合用来评价模型在未知域中的泛化能力。

实验设置上使用多源域泛化评价协议,用不同的源域组合来检验模型在未知的目标域下对舰船的检测性能。第一组以 SSDD 和 AIRSAR 作为源域进行训练,并分别在 HRSID 和 SARShip 上测试。第二组以 HRSID 和 SARShip 作为源域进行训练,并分别在 SSDD 和 AIRSAR 上进行测试。该设置可以从不同的分布组合出发来对模型的跨域泛化能力进行评价。本文以 Faster R-CNN 为基础搭建了检测模型,并使用 ResNet 作为骨干网络。模型训练 30 个 epoch, batch size 设置为 8,并且采用随机梯度下降 (SGD) 优化算法,学习率、动量和权重衰减分别设为 0.02、0.9 和 0.000 1。评价时,本文采用 IoU=0.5 作为正确检测的判定阈值,并在该阈值下计算 Precision、Recall、 F_1 和 mAP0.5 指标。所有实验均在 Intel (R) Core (TM) i9-13900KF CPU、64 GB 内存及 NVIDIA GeForce RTX 3090 GPU 环境下完成。

3.1 实验对比

我们将 P 、 R 、 F_1 和 $mAP_{0.5}$ 作为衡量不同检测器性能的度量指标,定义如下:

$P = \frac{TP}{TP + FP}$	(22)
$R = \frac{TP}{TP + FN}$	(23)
$AP = \int_0^1 P(R) dR$	(24)
$F_1 = 2 \times \frac{P \times R}{P + R}$	(25)

其中,真阳性(TP)指的是正确预测并满足交并比(IoU)阈值的船舶检测框,否则它们被视为假阳性(FP)。假阴性(FN)表示尽管存在真实的船舶,但是模型没有检测到。 P 和 R 分别代表准确率和召回率。式(24)表示平均精度(AP)是针对船舶类别的精度-召回率曲线(PRC)的积分。 $mAP_{0.5}$ 是在IoU=0.50时计算的船舶类别,本文中统一简写为mAP。式(25)中, F_1 是评价分类模型性能的一个指标,它把准确率 P 和召回率 R 结合起来考虑。 F_1 值的取值范围为0~1,数值越大代表模型性能越高。为了全面评价提出的方法对SAR舰船检测任务的跨域泛化能力,本文分别在单源域泛化和多源域泛化两种情况下做对比实验,实验结果见表1和表2。需要注意的是,虽然DMSN、PMT属于多源域适应方法,但是为了保证本文多源域泛化任务设置的一致性,在本文的实验中所有的对比方法都采用源域训练设置,即训练阶段只使用源域数据,不访问任何目标域图像或者目标域标签信息。对于原方法中包含目标域数据参与训练的部分,在复现实验中都予以去除或关闭,从而保证各个方法在相同的实验条件下进行公平的比较。

表1、表2中的“/”前后的结果分别表示用单个数

据集进行训练,在没有见到目标域的情况下进行泛化测试得到的检测结果。根据结果可知,源域和目标域分布差异越大,舰船检测性能就越差。例如,在表1中由SSDD迁移至AIRSAR时,DGFRNN-RS的mAP仅为19.6%,相比在SSDD上的56.8%下降超过37个百分点。类似地,在表2中由AIRSAR迁移至HRSID时,G-NAS的mAP从52.6%下降至16.2%,性能退化明显。这就说明单源训练不能覆盖复杂的SAR成像条件下分布的变化,容易造成域偏移,造成跨域舰船检测性能不稳定。

表1和表2的下半部分给出了多源域泛化的结果。在表1中,以HRSID和SARShip作为源域训练,并在未见域SSDD和AIRSAR上进行测试。与PMT相比,本文方法在平均mAP上从62.8%提高到66.3%,提高了3.5个百分点;在平均 F_1 上从63.6%提高到67.3%,提高了3.7个百分点。相比于DMSN, F_1 指标提高了4.6个百分点。这说明所提的方法可以不用目标域的数据来学习跨域鲁棒的特征。类似地,在表2中,用SSDD、AIRSAR作为源域训练时,本文方法在未见域HRSID、SARShip上也取得了最好的性能。具体而言,相较于PMT方法,本文方法在平均mAP上由65.0%提升至65.8%,在平均 F_1 上由71.9%提升至72.6%。相较于DMSN方法, F_1 指标提高2.7个百分点。虽然不同的源域组合所对应的分布是不一样的,但是本文的方法仍然可以保持稳定的性能,说明它对复杂的跨域场景有较好的泛化能力。不同的目标域上性能提升幅度不一样,主要是由目标域的场景复杂度、目标尺度分布、噪声水平等因素造成的。当目标域里有复杂的港口背景、强海杂波、高斑点噪声或密集的小目标时,模型的检测难度就会增大,因此性能增益也会受到一定的限制。

表1 不同方法在未见域SSDD、AIRSAR上的SAR舰船检测性能对比

单位: %

Table 1 Performance comparison of different methods on unseen domains SSDD and AIRSAR for SAR ship detection

unit: %

Method		SSDD				AIRSAR				Avg	
		mAP	P	R	F_1	mAP	P	R	F_1	mAP	F_1
HRSID/ SARShip	DGFRNN-RS ^[36]	56.8/ 71.8	66.8/ 77.6	56.1/ 75.1	61.0/ 76.3	19.6/ 23.5	29.4/ 29.5	30.2/ 32.3	29.8/ 30.8	38.2/ 47.7	45.4/ 53.6
	OA-DG ^[37]	66.9/ 65.7	69.4/ 66.2	62.0/ 64.1	65.5/ 65.1	31.1/ 27.8	43.7/ 38.6	36.6/ 32.1	39.9/ 35.0	49.0/ 46.8	52.7/ 50.1
	G-NAS ^[38]	69.6/ 75.5	73.4/ 74.9	65.2/ 73.5	69.1/ 74.2	28.5/ 42.6	30.5/ 56.5	36.6/ 41.0	33.3/ 47.5	49.1/ 59.1	51.2/ 60.9
	DMCL ^[39]	72.9/ 77.4	75.7/ 80.5	78.2/ 72.5	76.9/ 76.3	31.9/ 37.6	48.8/ 44.2	34.5/ 41.4	40.4/ 42.7	52.4/ 57.5	58.7/ 59.5
HRSID+ SARShip	DMSN ^[40]	79.6	79.4	75.3	77.3	40.6	52.7	44.1	48.0	60.1	62.7
	PMT ^[41]	80.1	78.2	79.3	78.7	45.5	50.2	46.7	48.4	62.8	63.6
	Ours	85.4	80.8	82.0	81.4	47.2	53.6	52.8	53.2	66.3	67.3

3.2 可视化

为了更加直观地评价不同的方法对复杂环境下的跨域检测能力,本文选择港口场景中典型的SAR图像做可视化对比,其中红色框、绿色框、橙色框分

别代表正确的检测目标、漏检的目标和误检的目标。从图3可以看出,在复杂的港口环境中,各个对比方法都会存在一定的性能下降。DGFRNN-RS在海岸线附近容易出现误检、漏检弱散射小目标的情况。

表2 不同方法在未见域HRSID、SARShip上的SAR舰船检测性能对比

单位: %

Table 2 Performance comparison of different methods on unseen domains HRSID and SARShip for SAR ship detection

unit: %

Method		HRSID				SARShip				Avg	
		mAP	P	R	F_1	mAP	P	R	F_1	mAP	F_1
SSDD/ AIRSAR	DGFRNN-RS ^[36]	43.0/ 55.0	54.3/ 63.3	51.6/ 56.3	52.9/ 59.6	63.9/ 55.3	68.9/ 61.0	66.5/59.4	67.6/ 60.2	53.5/ 55.2	60.3/ 59.9
	OA-DG ^[37]	43.1/ 28.1	58.4/ 33.5	57.6/ 39.6	58.0/ 36.3	48.7/ 63.2	54.5/ 67.5	57.3/62.1	55.9/ 64.7	45.9/ 45.7	57.0/ 50.5
	G-NAS ^[38]	52.6/ 16.2	78.5/ 24.6	53.4/ 25.4	63.6/ 25.0	62.3/ 59.8	69.1/ 64.7	63.5/60.3	66.2/ 62.4	57.5/ 38.0	64.9/ 43.7
	DMCL ^[39]	59.9/ 39.7	88.1/ 58.3	57.9/ 42.8	69.9/ 49.4	68.8/ 47.2	70.2/ 63.1	72.4/42.6	71.3/ 50.9	64.4/ 43.5	70.6/ 50.2
SSDD+ AIRSAR	DMSN ^[40]	58.9	73.8	67.2	70.3	67.9	68.6	70.2	69.4	63.4	69.9
	PMT ^[41]	59.4	74.8	69.9	72.3	70.5	69.4	73.8	71.5	65.0	71.9
	Ours	60.3	74.8	70.4	72.5	71.3	71.2	74.2	72.7	65.8	72.6

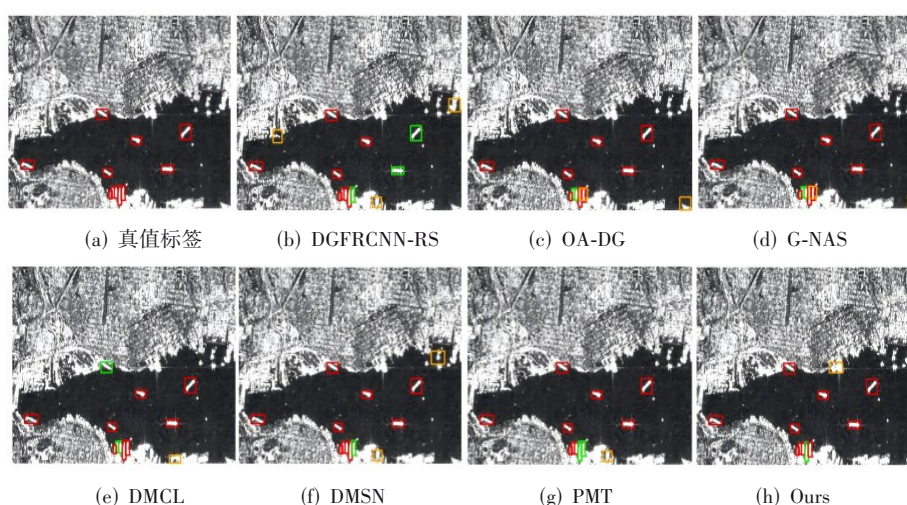


图3 港口场景下不同对比方法的域泛化检测结果视觉比较

Figure 3 Visual comparison of domain-generalized detection results of different methods in harbor scenes

OA-DG 和 G-NAS 虽然可以降低部分误检,但是对密集或者低对比的目标检测能力仍然比较弱。DMCL 方法的检测性能有所提高,但是对复杂的背景区域仍然会出错。多源方法 DMSN、PMT 整体性能较好,但是对岸线干扰和密集目标的情况仍然会出现漏检、误检的情况。本文的方法在这类场景下比其他方法检测效果更好,可以有效减少误检的情况,提高弱小、密集目标的检测能力。由于频域特征增强抑制了噪声,增强了结构信息,语义原型对齐、自适应加权策略改善了跨域特征分布,模型在复杂的港口环境中更加鲁棒、泛化。

3.3 消融实验

为了检验各个模块在多源域泛化 SAR 舰船检测任务中是否有效,本文进行了消融实验,逐层分析了频域感知特征增强模块、舰船语义原型对齐机制以及多源自适应加权策略的影响。实验以 Faster R-CNN 为基础框架,在 SSDD 和 AIRSAR 作为源域的多源设置下训练,并在未见目标域 HRSID 与 SARShip 上测

试,所得结果如表 3 所示。从 Baseline 的结果可以看出,在没有使用域泛化策略时,模型跨域检测性能很低, HRSID 上 mAP 为 47.2%, SARShip 上 mAP 为 48.8%,说明传统的检测模型不能适应 SAR 图像成像条件变化所造成的分布偏移。

引入频域感知特征增强模块(+Freq)后,模型性能显著提升, HRSID 和 SARShip 上的 mAP 分别提高至 53.8% 和 57.8%,较 Baseline 提升 6.6 个百分点和 9.0 个百分点, F_1 也提升至 68.2% 和 62.9%。结果表明,用频域分解来加强低频结构信息并抑制高频噪声,可以有效减少域相关扰动,提高特征的稳定性和跨域鲁棒性。在频域感知特征增强模块的基础上引入舰船语义原型对齐机制(+Freq+Proto)之后,模型性能又有所提高。 HRSID、SARShip 上 mAP 分别为 57.6%、67.4%,相比只用频域增强分别提高了 3.8 个百分点、9.6 个百分点, F_1 也分别提高到了 70.1%、69.3%。

这说明对齐多源域舰船语义原型可以降低舰船目标的跨域语义偏差,提高模型对舰船特征的判别能

表3 不同模块对多源域泛化SAR舰船检测性能的影响 单位: %

Table 3 Ablation study of different components on multi-source domain generalized SAR ship detection unit: %

Method	Freq	Proto	Weight	HRSID		SARShip	
				mAP	F_1	mAP	F_1
Baseline	×	×	×	47.2	51.1	48.8	54.9
+Proto	×	✓	×	51.6	59.4	55.2	60.7
+Weight	×	×	✓	49.8	56.3	52.6	58.1
+Freq	✓	×	×	53.8	68.2	57.8	62.9
+Freq+Proto	✓	✓	×	57.6	70.1	67.4	69.3
+Freq+Weight	✓	×	✓	56.1	69.5	65.3	68.4
Ours	✓	✓	✓	60.3	72.5	71.3	72.7

力。而在频域感知特征增强模块的基础上引入多源适应加权策略(+Freq+Weight)之后,模型的性能相比单独使用频域增强模块同样有所提升。HRSID和SARShip上mAP分别为56.1%、62.3%,分别比+Freq提高2.3个百分点、4.5个百分点, F_1 分别为69.5%、66.4%。由此可知,多源域自适应权重策略可以依据各个源域训练状态来动态调节优化贡献,降低不可靠源域或噪声较强的源域给模型更新造成的不良影响,进而提高模型对未知目标域的泛化能力。

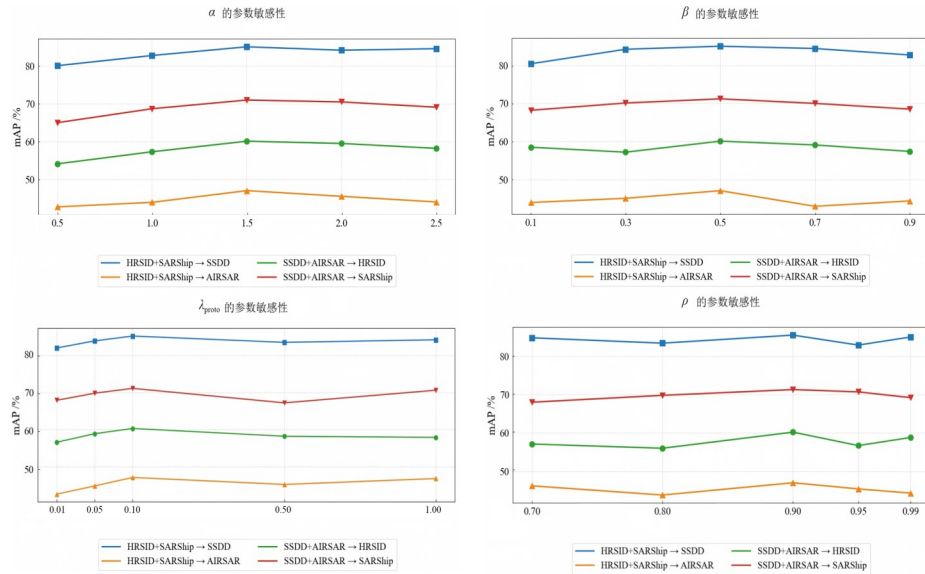
将频域感知特征增强模块、舰船语义原型对齐机制以及多源适应加权策略一起加入之后(Ours),模型就达到了最佳的效果。HRSID和SARShip上的mAP分别提升至60.3%和71.3%,较未加权模型提高了2.7个百

分点和3.9个百分点, F_1 分别达到72.5%和72.7%。从结果可知,动态调节源域权重可以很好地解决不可靠源域引起的负迁移问题,使模型更加重视稳定的、有判别力的源域信息,从而提高模型的总体泛化能力。

3.4 参数敏感性分析

为了更好地了解各关键超参数对模型性能的影响,本文又对频域感知特征增强模块的低频增强系数 α 、高频抑制系数 β 、语义原型对齐损失权重 λ_{proto} 和多源自适应加权中的平滑系数 ρ 等做了参数敏感性分析实验。实验均在相同设置下进行,即以SSDD与AIRSAR为源域训练,在未见域HRSID与SARShip上测试,同时以HRSID与SARShip为源域,在SSDD与AIRSAR上进行测试。

如图4所示, α 在[0.5,2.5]范围内变化, β 取值于[0.1,0.9], λ_{proto} 设置[0.01,1.0],而 ρ 在[0.7,0.99]区间内调整。随着 α 的增大,模型性能先提升后趋于稳定,在 $\alpha=1.5$ 时取得最佳效果,表明适度增强低频结构信息有助于突出舰船主体特征。本文最后使用了在各个源域组合中都表现稳定的参数设置,在所有的实验中保持不变,以保证方法评价的公平性和可重复性。对于 β ,当其取值为0.5时性能最优,说明在抑制高频噪声与保留有效细节之间达到了良好平衡。在 λ_{proto} 的变化过程中,模型在 $\lambda_{\text{proto}}=0.1$ 时表现最佳,表明适当的语义对齐约束是能够有效减少跨域语义偏差的,而过强的约束会影响特征判别。

图4 $\alpha, \beta, \lambda_{\text{proto}}, \rho$ 参数对模型检测准确性的影响Figure 4 Effects of parameters $\alpha, \beta, \lambda_{\text{proto}}, \rho$ on detection accuracy

除此之外,当 $\rho=0.9$ 时,模型性能最好,说明在该设置下,权重更新的稳定性和灵活性之间能够取得较

好的平衡。从整体上看,各参数的性能影响比较平缓,说明所提方法具有较好的鲁棒性,各模块之间存

在特征建模、语义约束和优化机制的协同作用,从而提高了多源域泛化SAR舰船检测的性能。

4 结论

本文针对复杂海洋环境下SAR舰船检测中由于成像条件不同造成数据分布偏移的问题,提出了一种多源域泛化SAR舰船检测方法。首先,设计频域感知特征增强模块,用傅里叶分解后重新加权低频结构信息和高频噪声成分,以增强舰船结构有关的域不变特征,抑制域相关扰动。其次,采用舰船语义原型对齐机制,建立并对齐多源域舰船语义中心,减少同类目标跨域语义偏差。最后,提出一种多源域适应加权方法,按照各源域训练损失的多少动态分配权重,以减少不可靠源域引起的负迁移。实验结果表明,在各种SAR舰船检测数据集的跨域场景下,本文提出的算法可以有效提高舰船检测的性能和泛化性。需要指出的是,本文的方法还存在不足之处。由于频域重构参数在不同的数据集、源域组合中是固定的,当实际的海况、海杂波等成像条件发生变化时,模型的性能也会受到影响。语义原型是否稳定与源域ROI特征的好坏、分布是否均匀有关。未来将对更自适应的频域建模、源域关系建模进行研究,使方法更适用于复杂的实际情况。

参考文献

- [1] 胡艳芬, 王宝发. 多极化SAR图像的降斑研究[J]. 电子学报, 2003, 31(2): 207-210.
Hu Yanfen, Wang Baofa. Speckle reduction in multipolarization SAR imagery[J]. Acta Electronica Sinica, 2003, 31(2): 207-210. (in Chinese)
- [2] 鲍厚兵, 黄顺吉, 刘钊. SAR并行成像处理的研究[J]. 电子学报, 2002, 30(12): 1745-1747.
Bao Houbing, Huang Shunji, Liu Zhao. The study of parallel processing of SAR imaging[J]. Acta Electronica Sinica, 2002, 30(12): 1745-1747. (in Chinese)
- [3] Li Dong, Liang Quanhuan, Liu Hongqing, et al. A novel multidimensional domain deep learning network for SAR ship detection[J]. IEEE Transactions on Geoscience and Remote Sensing, 2022, 60: 5203213.
- [4] Chen Yulin, Shen Yanyun, Duan Chi, et al. Robust and efficient SAR ship detection: An integrated despeckling and detection framework[J]. Remote Sensing, 2025, 17(4): 580.
- [5] Guan Tianyue, Chang Sheng, Wang Chunle, et al. SAR small ship detection based on enhanced YOLO network[J]. Remote Sensing, 2025, 17(5): 839.
- [6] Xue Weibao, Ai Jiaqiu, Zhu Yanan, et al. LMCNet: Lightweight modality compensation network via knowledge distillation for salient ship detection under missing-modality conditions[J]. IEEE Transactions on Aerospace and Electronic Systems, 2026, 62: 6547-6560.
- [7] Ren Haohao, Liang Shuyi, Miao Lei, et al. Multi-granularity feature fusion network integrating ASC for SAR target recognition[J]. IEEE Geoscience and Remote Sensing Letters, 2026, 23: 4003805.
- [8] Zhou Yun, Zhu Yinglin, Ren Haohao, et al. Refined multi-modal feature learning framework for marine target detection using radar sensor[J]. Digital Signal Processing, 2026, 170: 105816.
- [9] Ai Jiaqiu, Tian Ruitian, Luo Qiwu, et al. Multi-scale rotation-invariant Haar-like feature integrated CNN-based ship detection algorithm of multiple-target environment in SAR imagery[J]. IEEE Transactions on Geoscience and Remote Sensing, 2019, 57(12): 10070-10087.
- [10] Shi Yu, Du Lan, Guo Yuchen. Unsupervised domain adaptation for SAR target detection[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2021, 14: 6372-6385.
- [11] Huang Jinyang, Feng Yuanhao, Cui Fengqi, et al. Identifying who you are no matter what you write through abstracting handwriting style[J]. IEEE Transactions on Dependable and Secure Computing, 2026, 23(3): 6890-6905.
- [12] Li Da, Yang Yongxin, Song Yizhe, et al. Deeper, broader and artier domain generalization[C]//Proceedings of the IEEE International Conference on Computer Vision. Piscataway: IEEE, 2017: 5542-5550.
- [13] Zhong Yingyi, Zhou Wen'an, Wang Zhixian. A survey of data augmentation in domain generalization[J]. Neural Processing Letters, 2025, 57(2): 34.
- [14] Zhang Zenghui, Zhang Limeng, Wu Juanping, et al. Optical and synthetic aperture radar image fusion for ship detection and recognition: Current state, challenges, and future prospects[J]. IEEE Geoscience and Remote Sensing Magazine, 2024, 12(4): 132-168.
- [15] Han Jianhong, Yang Wenjie, Wang Yupei, et al. Remote sensing teacher: Cross-domain detection transformer with learnable frequency-enhanced feature alignment in remote sensing imagery[J]. IEEE Transactions on Geoscience and Remote Sensing, 2024, 62: 5619814.
- [16] Ren Shaoqing, He Kaiming, 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.

- [17] Redmon J, Divvala S, Girshick R, et al. You only look once: Unified, real-time object detection[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2016: 779-788.
- [18] Lin T Y, Dollár P, Girshick R, et al. Feature pyramid networks for object detection[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2017: 936-944.
- [19] Cao Rui, Sui Jianghua. A dynamic multi-scale feature fusion network for enhanced SAR ship detection[J]. *Sensors*, 2025, 25(16): 5194.
- [20] Ding Jian, Xue Nan, Long Yang, et al. Learning RoI transformer for oriented object detection in aerial images[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2019: 2844-2853.
- [21] 王浩添, 冀振元, 化青龙, 等. 基于多分支多信息多深度复值特征融合网络的SAR舰船目标识别方法[J]. *电子学报*, 2025, 53(10): 3759-3772.
Wang Haotian, Ji Zhenyuan, Hua Qinglong, et al. Recognition method of ship targets for SAR based on M3Net[J]. *Acta Electronica Sinica*, 2025, 53(10): 3759-3772. (in Chinese)
- [22] Shan Huilin, Fu Xiangwei, Lv Zongkui, et al. SAR ship detection algorithm based on deep dense sim attention mechanism network[J]. *IEEE Sensors Journal*, 2023, 23(14): 16032-16041.
- [23] Zhu Yanan, Ai Jiaqiu, Xue Weibao, et al. Cross-modal ship detection from optical to SAR images based on pixel- and feature-level progressive transfer[J]. *IEEE Sensors Journal*, 2025, 25(8): 13344-13356.
- [24] Kim J U, Park S, Ro Y M. Uncertainty-guided cross-modal learning for robust multispectral pedestrian detection[J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2022, 32(3): 1510-1523.
- [25] Chen Xi, Wang Zhirui, Wang Wenhao, et al. CroMoDa: Unsupervised oriented SAR ship detection via cross-modality distribution alignment[J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2024, 17: 11899-11914.
- [26] Lin Mingkai, Li Wenzhong, Li Ding, et al. Multi-domain generalized graph meta learning[C]//Proceedings of the AAAI Conference on Artificial Intelligence. Palo Alto: AAAI, 2023: 4479-4487.
- [27] Liu Hong, Song Pinhao, Ding Runwei. Towards domain generalization in underwater object detection[C]//Proceedings of 2020 IEEE International Conference on Image Processing (ICIP). Piscataway: IEEE, 2020: 1971-1975.
- [28] Addepalli S, Asokan A R, Sharma L, et al. Leveraging vision-language models for improving domain generalization in image classification[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2024: 23922-23932.
- [29] Vidit V, Engilberge M, Salzmann M. CLIP the gap: A single domain generalization approach for object detection[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2023: 3219-3229.
- [30] Danish M S, Khan M H, Munir M A, et al. Improving single domain-generalized object detection: A focus on diversification and alignment[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2024: 17732-17742.
- [31] Al-Emadi S A, Yang Y, Ofli F. Benchmarking object detectors under real-world distribution shifts in satellite imagery[C]//Proceedings of the Computer Vision and Pattern Recognition Conference. Piscataway: IEEE, 2025: 8299-8309.
- [32] Dou Qi, Castro D C, Kamnitsas K, et al. Domain generalization via model-agnostic learning of semantic features[C]//Proceedings of the 33rd International Conference on Neural Information Processing Systems. New York: Curran Associates Inc., 2019: 579.
- [33] He Boyong, Ji Yuxiang, Ye Qianwen, et al. Generalized diffusion detector: Mining robust features from diffusion models for domain-generalized detection[C]//Proceedings of the Computer Vision and Pattern Recognition Conference. Piscataway: IEEE, 2025: 9921-9932.
- [34] Wu Aming, Deng Cheng. Single-domain generalized object detection in urban scene via cyclic-disentangled self-distillation[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2022: 837-846.
- [35] Lee S, Seong H, Lee S, et al. WildNet: Learning domain generalized semantic segmentation from the wild[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2022: 9936-9946.
- [36] Duraklı E, Aptoula E. Domain generalized object detection for remote sensing images[C]//Proceedings of 2023 31st Signal Processing and Communications Applications Conference (SIU). Piscataway: IEEE, 2023: 1-4.

- [37] Lee W, Hong D, Lim H, et al. Object-aware domain generalization for object detection[C]//Proceedings of the AAAI Conference on Artificial Intelligence. Palo Alto: AAAI, 2024: 2947-2955.
- [38] Wu Fan, Gao Jinling, Hong Lanqing, et al. G-NAS: Generalizable neural architecture search for single domain generalization object detection[C]//Proceedings of the AAAI Conference on Artificial Intelligence. Palo Alto: AAAI, 2024: 5958-5966.
- [39] Chen Yang, Song Pinhao, Liu Hong, et al. Achieving domain generalization for underwater object detection by domain mixup and contrastive learning[J]. Neurocomputing, 2023, 528: 20-34.
- [40] Yao Xingxu, Zhao Sicheng, Xu Pengfei, et al. Multi-source domain adaptation for object detection[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. Piscataway: IEEE, 2021: 3273-3282.
- [41] Belal A, Meethal A, Romero F P, et al. Multi-source domain adaptation for object detection with prototype-based mean teacher[C]//Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision. Piscataway: IEEE, 2024: 1266-1275.

作者简介



艾加秋 男, 1985年8月出生于江西省吉安市。现为合肥工业大学计算机与信息学院教授、博士生导师。主要研究方向为多模态遥感图像智能处理与应用。

E-mail: aijiaqiu1985@hfut.edu.cn



朱雅楠 女, 1997年12月出生于山东省青岛市。现为合肥工业大学计算机科学与技术流动站博士后。主要研究方向为域适应、跨模态遥感图像智能解译。

E-mail: 2026820020@hfut.edu.cn