

一种形态学引导的解耦式SAR舰船有向检测框架

汪泽宇 王青松*

(中山大学电子与通信工程学院 深圳 518107)

摘要: 合成孔径雷达(SAR)以其全天候、全天时的观测能力,在遥感检测中得到了广泛应用。然而,受限于标注精度,目前主流的SAR目标检测方法多依赖水平框标注,难以实现精确的目标角度和尺度估计。同时,尽管弱监督学习在光学图像中的角度预测取得了进展,但其忽视了SAR特有的成像几何,难以有效泛化。为解决上述挑战,该文提出一种融合SAR成像机理的有向舰船检测新框架,核心思想在于将检测任务解耦为定位与方向估计两个独立的子模块。其中,定位模块可以直接利用任意现有的、在水平框标注上训练的检测器;而方向估计模块则在一个专门构建的形态学合成二值数据集上进行全监督训练。该框架的优势在于无需修改原有检测器结构和重新训练的前提下,即插即用赋予模型高精度的有向框预测能力。实验验证了所提方法在多个数据集上相较于现有仅依赖水平框监督的方法表现出更优的性能,部分场景中甚至超越全监督方法,体现出强大的有效性与工程应用价值。

关键词: 有向目标检测; 形态学仿真; 仿真驱动学习; 解耦检测

中图分类号: TP75; TP391.41

文献标识码: A

文章编号: 1009-5896(2025)12-4851-11

DOI: 10.11999/JEIT250979

CSTR: 32379.14.JEIT250979

1 引言

SAR图像中的舰船目标检测是图像解译的关键任务之一。传统检测方法如恒虚警率算法(Constant False Alarm Rate, CFAR)及其变体^[1,2],在背景噪声变化时力求保持恒定的虚警率,但其性能高度依赖于背景统计模型假设,在复杂场景下泛化能力受限。因此,研究界逐渐转向数据驱动的深度学习方法。

近几年来,深度学习算法由于其强大的泛化能力,在大数据的加持下能够取得显著的结果,在各类视觉任务当中得到了广泛应用。目前已有许多通用的光学目标检测器算法已经迁移至SAR舰船领域当中。例如文献^[3]提出的两阶段检测器区域卷积神经网络(Region-based Convolutional Neural Networks, RCNN),文献^[4,5]提出的1阶段检测器YOLO(You Only Look Once)算法,文献^[6]提出的基于视觉变换器(Vision Transformers, ViT)的算法。这类算法凭借着深度学习强大的特征提取能力能够取得不错的性能。然而这些方法面临着一些挑战,阻碍了它们在SAR舰船检测中的有效性。首先,大部分的检测器都没有从SAR图像本身的成像机理出发来探讨目标检测问题,不可避免地忽略了自然场景图像与SAR图像之间巨大的图像域差距。

最后,目前大部分的检测器都是采用的全监督的训练范式,这往往需要有高质量的标注数据,从而导致成本高昂、耗时费力。尤其是在SAR领域当中,实测人工标注的样本数量不足,难以覆盖实际情况。此外,由于SAR图像的几何畸变和相干斑噪声影响,很多时候SAR图像是难以确定具体边缘,导致人工标注的精度无法保证。

为了缓解这一问题,研究界开始探索减少对标注数据依赖的学习范式。其中,弱监督目标检测(Weakly Supervised Object Detection, WSOD)旨在利用部分标注数据实现对目标的精确标注,能够在降低标注成本的同时提升模型的表征能力,展现出了一定的潜力。例如,文献^[7,8]通过对图像进行已知的几何变换(如旋转、翻转)并利用几何关系约束,试图从水平标注数据中学习目标的角度和宽高信息。然而,此类方法在运用在SAR领域则会面临灰度特性和固有噪声带来显著的性能损失。还有研究还探讨了面向集群的框架^[9],这代表了处理复杂遥感任务的另一种视角。

为从新的角度应对上述挑战,本文跳出传统的弱监督学习范式,提出一种基于形态学仿真驱动SAR舰船有向检测新框架。其核心是解耦定位与方向估计任务,并利用形态学仿真进行方向学习。本方法首先采用标准水平检测器对SAR图像进行粗粒度目标定位。针对方向估计,本文设计并训练了一个方向预测子网络,本网络通过形态学仿真生成的二值图像数据集进行全监督学习。本数据集通过对几何基元进行腐蚀、膨胀等一系列形态学操作,模

收稿日期: 2025-09-24; 改回日期: 2025-12-10; 网络出版: 2025-12-25

*通信作者: 王青松 wangqs5@mail.sysu.edu.cn

基金项目: 国家自然科学基金(62273365), “小米青年学者”项目
Foundation Items: The National Natural Science Foundation of China(62273365), Xiaomi Young Talents Program

拟SAR图像中舰船目标的边缘模糊、结构变形等关键形态,使子网络能够有效捕捉与方向相关的特征,并辅助修正目标的宽高比。最终,通过融合主网络的定位结果与子网络的高精度方向预测,实现在仅有水平框标注下的精准有向目标检测。本文的主要贡献如下:

(1)提出一种新颖的解耦式SAR舰船有向检测框架。本方法的核心思想在于将目标的定位任务与方向估计任务进行解耦,从而在仅有水平框标注的条件下,通过对仿真数据的监督学习,能够有效预测目标的旋转方向。

(2)设计并实现了一种即插即用的检测器兼容框架。本框架能够借助现有的主流水平框检测器实现无缝集成,使其无需任何额外训练或结构修改,即可具备旋转目标检测能力,极大地降低了算法的应用门槛、训练开销与迁移成本。

(3)针对直接利用真实SAR图像进行学习的困难,设计并构建了一个形态学仿真二值图像数据集。本数据集通过模拟目标的结构轮廓信息,旨在有效缓解SAR图像固有挑战对模型学习的干扰,从而显著增强了检测网络对真实SAR数据的适应能力与鲁棒性。

2 相关工作

2.1 SAR领域的全监督目标检测

在全监督设定下,即拥有大规模精确标注数据的前提下,基于深度学习的目标检测已成为SAR图像解译的主流技术。研究者们围绕SAR的固有难题,对检测模型进行了大量针对性的改进。首先,SAR图像中最显著的特征是固有的相干斑噪声,它严重降低了图像的信噪比,模糊了目标边缘。在早期的做法当中通常是将图像送入深度学习网络之前,采用传统的散斑滤波算法进行预处理,例如文献[10],但这种方式可能会平滑掉对检测有用的细节。文献[11]通过散射点之间的位置语义关系设计了散射拓扑网络,通过结合目标的散射特性和构建散射点之间的拓扑关系,有效解决了成像变化性和散射点离散性的问题。在此基础上,文献[12]则从散射特性出发,构建了一个两阶段检测模块,其核心在于一个由散射点信息引导的区域提议网络(Region Proposal Network, RPN),用以更精准地预测潜在的关键散射点位置。为了进一步利用散射点间的空间关系,文献[13]提出了一种散射区域拓扑结构金字塔模型,该模型通过构建强散射点之间的拓扑几何结构,有效地增强了舰船目标与复杂近岸背景之间的可分离性。这些方法通过不同方式利用

散射点信息,均在SAR目标检测任务上取得了显著的性能提升。然而,值得注意的是,上述模型的改进思路主要集中在对检测模型本身的优化调整,而较少直接从SAR图像数据固有的成像特性、噪声统计或信息表达的根本层面出发,探索提升检测性能的途径。

2.2 SAR领域的弱监督目标检测

尽管全监督方法在SAR目标检测中取得了显著进展,但其对大规模精确标注的依赖限制了其在实际应用中的推广。因此,能够降低标注成本的学习范式在SAR图像解译领域引起了越来越多的关注。

其中一个主要的研究方向是弱监督目标检测。在SAR弱监督目标检测研究中,由于目前公开数据集多以水平框(Horizontal Bounding Boxes, HBBs)作为标注形式,因此相关方法主要集中在基于水平框的检测。

基于水平框生成旋转框范式的研究旨在利用相对容易HBBs标注作为监督信息,来训练能够预测有向框(Oriented Bounding Boxes, OBBs)的检测器。该领域的工作最早被应用于光学遥感图像领域。文献[7]通过对原始图像进行随机角度的旋转,并进一步引入了翻转与旋转操作下的一致性约束,从而更有效地弥合了HBBs监督与OBBs预测之间的差距。然而,该工作主要针对光学遥感图像设计,其图像旋转等策略在直接应用于具有独特散射机理和相干斑噪声的SAR图像时,效果和适用性会受到严峻挑战。针对SAR图像,文献[14]提出了一种通用的弱监督检测方法。该方法的核心思想是通过挖掘图像中舰船的粗略定位信息来生成高质量的伪旋转框标签,并结合HBBs标注共同训练网络。尽管该方法在SAR数据集上取得了出色的表现,但其伪标签生成流程单一固定,可能导致在某些复杂场景或特定目标形态下生成的伪标签不够精确,从而影响最终的检测性能。此外,一些工作如文献[8]提出了统一的弱监督检测框架,旨在实现在图像级标签、不完整标注等多种弱标注条件下的有向目标检测。虽然这类方法具有一定的通用性,但它们并非严格意义上基于水平框生成旋转框的范式,其设计目标是处理更多样化的弱标签,可能未针对从HBBs到OBBs的特定转换进行深度优化。

综上所述,传统弱监督范式在SAR领域存在一些共有的局限性。首先,多数水平框生成旋转框方法主要基于光学图像设计,未能充分考虑SAR图像独特的成像机理和复杂的电磁散射特性,其有效性在SAR场景下受到限制。其次,即便文献[14]为针对SAR图像的弱监督方法,其伪标签的生成质量和泛

化能力仍有提升空间。更为关键的是，现有方法通常需要设计一个全新的网络架构，并从头开始或进行大规模的重新训练，未能充分利用当前在水平框检测领域已趋于成熟、性能饱和的检测器。它们往往无法直接赋予这些已有的高性能水平检测器预测有向框的能力，造成了计算资源和已有模型知识的浪费。

3 系统模型

本文聚焦于利用水平框标注信息，实现SAR图像中舰船目标有向检测的精确定位与方向估计。本方法将检测任务解耦，其核心在于构建一个完全基于定制化的形态学合成二值数据集，并基于此训练方向估计模块。本策略通过从形态学仿真的结构特征中学习鲁棒的方向线索，结合宽高修正机制，并利用性能已趋于饱和的水平框检测网络，在弱监督条件下实现了精确的有向目标检测。所提出方法总览图说明可见图1。其中(c)模块训练集来源于形态学仿真的数据，在推理阶段将(a)模块的二值化切片作为输入得到最终的有向检测结果。

3.1 轴对齐检测器

在仅有水平框标注的有向目标检测任务中，弱标签信息无法提供区分不同方向的细粒度信息，容易导致模型训练不稳定，并且与强监督任务相比存在着性能差距^[7,8]。因此，本模块的核心动机在于解耦复杂的有向目标检测任务。本文将任务分解为两个相对独立的子问题，而轴对齐定位模块即为该解耦框架的第1阶段，其核心功能用于提供重要的空间信息先验。

本模块采用一种常规的目标检测网络架构，任务是生成水平(轴对齐)边界框。它的核心功能在于执行粗粒度定位，识别出可能包含潜在舰船目标的感兴趣区域，而无需确定这些目标的精确方向。对于数据集中的图像 $\mathbf{X} \in R^{1 \times W \times H}$ ，其中 $F(\cdot)$ 表示端到端的水平框检测器，网络输出为 $\mathcal{B} \in \mathcal{B}_1, \mathcal{B}_2, \dots, \mathcal{B}_N$ ，其中 N 表示检测到的目标数量。对于任意的目标区域 $\mathcal{B}_i = (x_i, y_i, w_i, h_i)$ ， (x_i, y_i) 表示目标框中心坐标， w_i, h_i 分别表示目标框的宽度和高度

$$x_{\text{left}} = x_i - \frac{w_i}{2}, y_{\text{left}} = y_i - \frac{h_i}{2} \quad (1)$$

$$x_{\text{right}} = x_i + \frac{w_i}{2}, y_{\text{right}} = y_i + \frac{h_i}{2} \quad (2)$$

接下来，从原始图像 \mathbf{X} 中提取出子图像 \mathbf{X}_{sub} ，其表示为

$$\mathbf{X}_{\text{sub}} = \mathbf{X} [x_{\text{left}} : x_{\text{right}}, y_{\text{left}} : y_{\text{right}}] \quad (3)$$

3.2 形态学仿真数据集

尽管原始SAR图像信息复杂，但舰船目标在经过理想情况下的分割或二值化后，其呈现的二值轮廓和内部结构依然保留了大量可供区分方向的几何与形态线索。基于此观点，本文构建了一个基于形态学仿真合成的数据集。

基准目标生成：首先定义一组不同长宽比的二值矩形作为基础几何基元。这些基元在生成时被赋予随机的尺寸、初始方向(作为后续训练的地面真实标签)。设合成数据画布空间 $\Omega \in R^{H_c \times W_c}$ ，其中 H_c 和 W_c 分别表示画布的高度和宽度。其取值范围

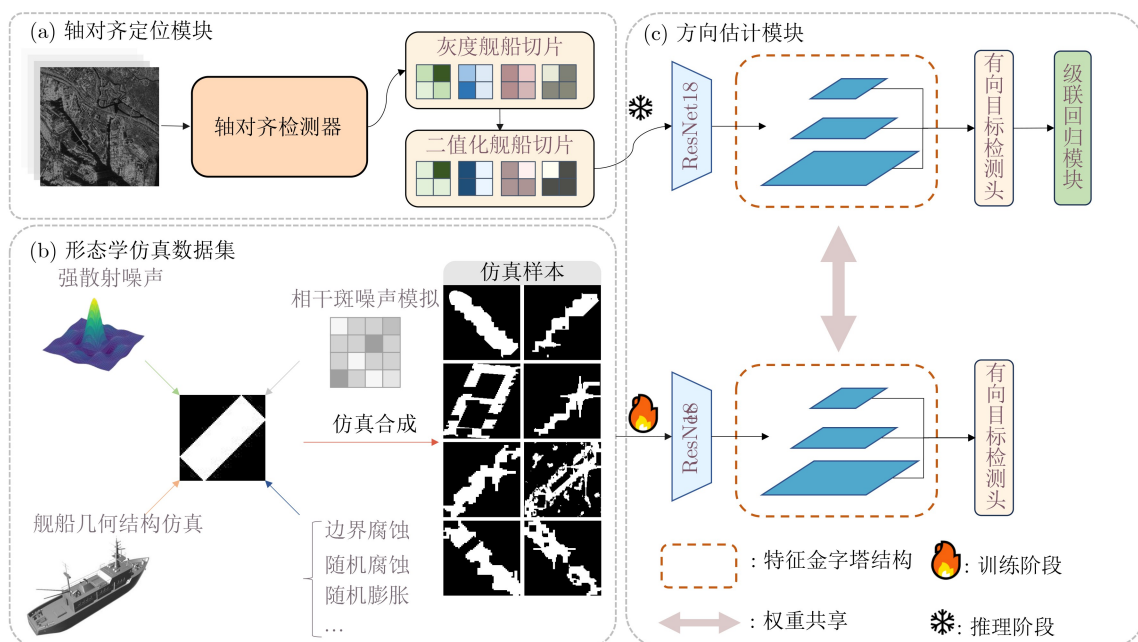


图1 模型总览图

根据训练集中目标尺寸的统计分布进行随机抽样。即： $H_c \in [H_{\min}, H_{\max}]$, $W_c \in [W_{\min}, W_{\max}]$ 。在该空间内，生成的旋转矩形需要满足内切条件，确保目标完全包含于画布空间。即

$$\begin{pmatrix} H_c \\ W_c \end{pmatrix} = \begin{pmatrix} \sin \theta & \cos \theta \\ \cos \theta & \sin \theta \end{pmatrix} \cdot \begin{pmatrix} H_s \\ W_s \end{pmatrix} \quad (4)$$

其中， H_s , W_s 分别为仿真目标的宽高。旋转矩形内部的像素值被设为255，以模拟目标区域。接着为模拟SAR目标散射特性，设计形态学微分操作：

(1) 区域腐蚀

$$\begin{aligned} & \text{Erode}(R_{xy}(\theta), S_{\text{rect}}(l)) \\ &= R_{xy}(\theta) \ominus S_{\text{rect}}(w_d, h_d), w_d, h_d \sim \mathcal{U}(0, l_{\max}) \end{aligned} \quad (5)$$

(2) 区域膨胀

$$\begin{aligned} & \text{Dilate}(R_{xy}(\theta), S_{\text{rect}}(l)) \\ &= R_{xy}(\theta) \oplus S_{\text{rect}}(w_d, h_d), w_d, h_d \sim \mathcal{U}(0, l_{\max}) \end{aligned} \quad (6)$$

其中， $R(\theta)$ 为仿真的旋转矩形， \ominus 表示形态学腐蚀操作， \oplus 表示形态学膨胀操作， S_{rect} 表示膨胀结构元为矩形结构， w_d , h_d 分别表示膨胀结构元素的宽度与高度。 $l_{\max} = \min(H_c, W_c)/10$ ， $\mathcal{U}(0, l_{\max})$ 表示服从 $[0, l_{\max}]$ 的均匀分布。

形态学调制：在基准目标的基础上，进一步引入形态学调制，以生成更加贴近真实SAR目标二值形态的训练样本，从而提升子网络的泛化能力。该过程主要包括以下3个方面：

(1) 目标随机化：

(a) 目标尺寸：通过对真实数据集中目标尺寸的统计分析，建立目标尺寸的概率分布，依据该分布，从预设范围 $[H_{\min}, H_{\max}]$ 和 $[W_{\min}, W_{\max}]$ 内随机采样，确定目标尺寸。

(b) 目标角度：考虑到舰船目标在SAR图像中的角度存在任意性、矩形框存在对称性，角度值从 $0^\circ \sim 180^\circ$ 进行均匀采样。

(2) 几何形状调制：

(a) 模拟SAR目标矩形边缘腐蚀特性的微分形态学操作：在 $R(\theta)$ 轮廓线上随机选取10%~20%的坐标点进行腐蚀操作。 C 表示 $R(\theta)$ 的边缘轮廓像素坐标， $\text{Select}(\cdot)$ 为随机选择函数， p 为选择概率。即

$$\begin{aligned} R^1(\theta) &= \{\text{Erode}(R_{x_i, y_i}(\theta), S_{\text{square}}(l)) | (x_i, y_i) \\ &\in \text{Select}(C, p)\} \end{aligned} \quad (7)$$

(b) 边缘膨胀：用于模拟SAR目标矩形边缘膨胀特性的微分形态学操作，进一步模拟图像边界的不确定性。在 $R(\theta)$ 轮廓线上随机选取10%~20%的坐标点进行膨胀操作，即

$$\begin{aligned} R^2(\theta) &= \{\text{Dilate}(R_{x_i, y_i}^1(\theta), S_{\text{rect}}(w_d, h_d)) \\ &| (x_i, y_i) \in \text{Select}'(C, p)\} \end{aligned} \quad (8)$$

(c) 弧形船首：舰船的船首通常为尖形，在SAR图像中表现为弧形，在仿真过程中，采用椭圆方程来模拟弧形结构。弧形结构的半短轴、半长轴设定如下，其中 a 为半短轴， b 为半长轴

$$a = \frac{\min(H_s, W_s)}{2}, b = \text{random}\left(a, \frac{\max(H_s, W_s)}{2}\right) \quad (9)$$

(d) 结构内随机腐蚀：在目标基元的内部区域随机选择小块进行腐蚀操作，形成孔洞或断裂，以模拟舰船甲板上不同设备、结构间的分隔，或由于遮挡、多路径效应导致的内部结构信号缺失。

(e) 内部截断：不同的舰船目标存在不同的几何结构，舰船特殊部分的结构会导致强后向散射，例如桅杆、栏杆与船舱表面形成的二面角。而这些强散射点会导致后续二值化过程中将部分弱后向散射区域出现截断现象。在仿真过程中，根据仿真目标尺寸 (H_s, W_s) ，以其短边 $\min(H_s, W_s)$ 为基准截取10%~20%的舰船结构。

(3) 噪声随机化：

(a) 强散射十字噪声：该噪声通常来源于SAR图像中目标的强反射体，这种噪声通常表现为条纹状、十字形或放射状的干扰，这些干扰在图像中具有明显的方向性，通常垂直和水平方向上都会出现。在合成数据集中，通过随机选点添加对称的十字斑噪声。十字斑的生成方式如式(10)所示

$$T = M - \left\lfloor M \cdot \frac{\ln(i+1)}{\ln(L+1)} \right\rfloor \quad (10)$$

其中， T 表示十字斑厚度， M 代表最大厚度， L 代表十字斑的最大长度， i 为小于 L 的整数， $\lfloor x \rfloor$ 表示向下取整函数。

(b) 柏林噪声：在真实SAR图像经过二值化处理后，斑点噪声通常表现出较强的空间相关性与随机分布特征，且在部分场景中呈现出特定的几何形态。为有效模拟这类噪声特性，本文引入柏林噪声作为生成机制。在二值化仿真切片生成过程中，本文对生成的柏林噪声图进行阈值化处理，具体方法为对噪声值设置截断阈值 $\tau = 0.5$ ，以强化噪声的局部随机性和空间离散性。

通过上述的仿真流程可以提供高度可控和可扩展的数据来源，允许高效生成远超人力采集标注规模的训练样本，从而提升模型的泛化能力。精心设计的形态学调制操作亦致力于复现SAR目标在二值化后可能呈现的独特结构表征，以期在一定程度上减小仿真数据与真实数据二值化后在关键结构特征上的域差异。各条件单独作用于标准矩形框的仿真效果可见图2。通过上述流程，本文最终生成了一个包含5 000张仿真二值图像的训练数据集。

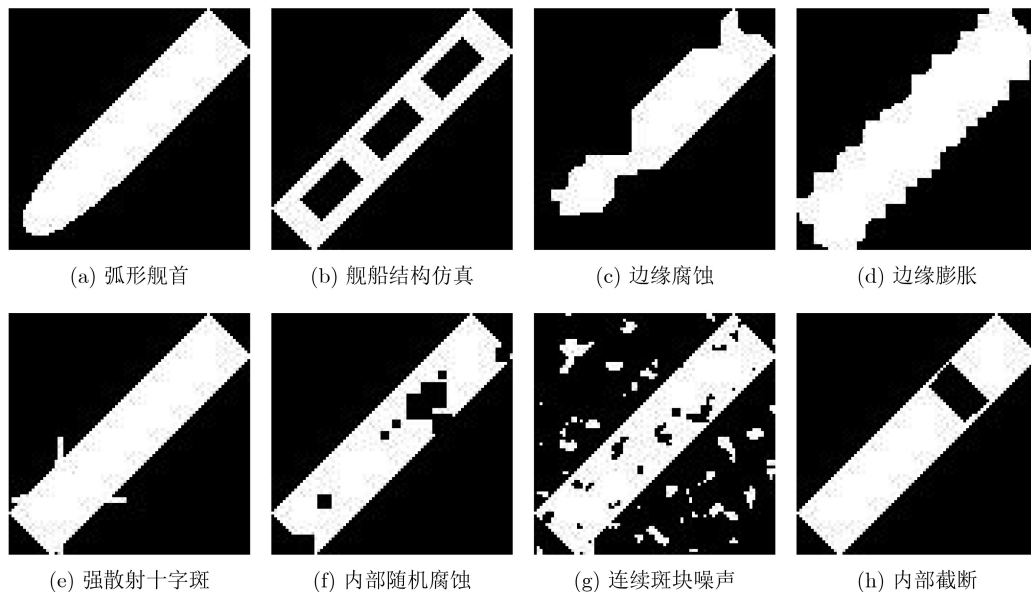


图2 不同仿真条件单独作用时的效果图

3.3 方向估计模块

本模块完全在二值的形态学合成数据集上进行训练。这种设计使得网络能够专注于学习目标的几何轮廓与结构特征来进行方向判断。为了确保网络在训练和推理阶段的输入数据具有分布上的一致性，本文对从真实SAR图像中提取的候选区域图像块执行了一个关键的二值化预处理步骤。具体而言，在将这些图像块送入方向估计模块的主干网络之前，采用文献[14]所提出的方法对其进行二值化处理。该方法旨在有效地从包含斑点噪声和复杂灰度分布的SAR图像块中提取出目标的二值结构掩模。通过此操作，推理阶段的输入样本在形态结构上与其训练数据保持了高度的一致性。

模型架构：该模块主要基于一个轻量级的卷积神经网络并结合了多尺度特征处理能力。值得注意的是本文选用了ResNet-18作为特征提取的主干网络。选择ResNet-18的原因在于其相对较浅的层数和较少的参数量，能够在保证足够特征提取能力的同时，降低计算开销。在训练过程中所使用的数据均来自于上一节得到的合成数据集。使用高度抽象和简化的二值图像进行训练，能够让网络更加关注SAR图像中的结构信息，简化了网络的学习任务。

损失函数：完整的损失函数定义为

$$L_{\text{ws}} = \frac{\lambda_1}{\sum_{\text{cn}_{\text{pos}}} \sum_{(x,y)} \mathbb{I}_{\{c(x,y)>0\}} \text{cn}(x,y)} \cdot L_{\text{reg}} \left(\text{rbox}_{(x,y)}^p, \text{syn}_{(x,y)}^{\text{gt}} \right) + \frac{\lambda_2}{N_{\text{pos}}} \sum_{(x,y)} L_{\text{cn}} \left(\text{cn}'_{(x,y)}, \text{cn}(x,y) \right) \quad (11)$$

其中 L_{reg} 为旋转IoU损失， L_{cn} 为交叉熵损失， c 表

示标签的分类值， cn 与 cn' 分别表达预测值与合成数据真值的Center-ness， N_{pos} 代表正样本数目， $\text{rbox}_{(x,y)}^p, \text{syn}_{(x,y)}^{\text{gt}}$ 分别为预测有向矩形框与合成数据真值矩形框。超参数设置为 $\lambda_1 = 1.5, \lambda_2 = 1$ 。

级联回归模块：在分别通过轴对齐定位模块获取目标的初始水平框参数以及通过方向估计模块预测出目标的旋转角度和调整后的尺寸后，需要将这两个阶段的结果进行有效融合，以生成最终的精确有向边界框。有向目标边界框在原始SAR图像坐标系下的参数 $(x_{\text{obb}}, y_{\text{obb}}, w_{\text{obb}}, h_{\text{obb}}, \theta_{\text{obb}})$ 通过式(12)的方式组合得到

$$\begin{aligned} & (x_{\text{obb}}, y_{\text{obb}}, w_{\text{obb}}, h_{\text{obb}}, \theta_{\text{obb}}) \\ & = (x_{\text{left}} + x', y_{\text{left}} + y', w', h', \theta') \end{aligned} \quad (12)$$

其中 $(x', y', w', h', \theta')$ 为方向估计模块的输出坐标。

4 实验与分析

4.1 数据集介绍

为了验证所提形态学仿真驱动SAR舰船有向检测方法的有效性，本研究在包含有向框标注的数据集上进行了实验，具体包括HRSID和SSDD数据集。在所有实验中，训练阶段仅使用水平框标注信息和形态学仿真数据，而数据集中的有向框标注信息仅在测试阶段使用。

HRSID数据集[15]是一个用于高分辨率SAR图像中船舶检测、语义分割和实例分割任务的数据集。该数据集涵盖有5 604张高分辨率SAR图像，涵盖0.5 m, 1 m, 3 m的分辨率、HH, HV, VV和VH 4种极化方式，包括有16 951个船舶实例，每个实例都提供了水平检测框、有向检测框、实例分割掩膜。

SSDD数据集[16]是一个专门针对SAR图像船

船舶检测任务而设计的公开数据集。该数据集涵盖有1 160张不同分辨率的SAR图像,数据主要有RadarSat-2,TerraSAR-X和Sentinel-1传感器,HH, HV, VV和VH 4种极化方式,共包括有2 456个船舶实例,每个实例都提供检测、有向检测框、实例分割掩膜。

4.2 实现细节

本文中的所有实验使用Pytorch框架,在一个2.1 GHz Intel Xeon Silver 8352V CPU 和一个24 GB Nvidia GeForce RTX 4090 GPU。对于每个数据集,本文使用交并比阈值为0.5时的平均准确率进行评估。对于HRSID数据集前馈网络的重采样尺寸为 800×800 像素、SSDD数据集的重采样尺寸为 256×256 像素。

所有方法在开源的MMRotate框架^[17]下实现,本文采用以ResNet-50为Backbone的YOLOX^[4]作为前馈网络用来获取水平框切片并进行后续的二值化数据处理。前馈网络训练500epoch。所有实验优化器均采用AdamW优化器,参数设置 $\beta_1 = 0.9$, $\beta_2 = 0.999$, $lr = 5e-5$,权重衰减为0.05。最后,本文采用ResNet-18为Backbone的方向预测子网络来获取对二值化切片的角度值。

4.3 实验结果综合分析

4.3.1 对比实验

为了验证所提方法在目标检测任务中的有效性,本文在SSDD和HRSID上进行了大量的实验,

并与现有的完全监督方法和其他仅依赖水平框监督的方法进行了对比。实验的评估指标为平均精度(Average Precision, AP50)和召回率(Recall, R)以全面评估各方法的性能。选取全监督检测算法有:R-RetinaNet^[18],GWD (Gaussian Wasserstein Distance)^[19],S2ANet (Single-Shot Alignment Network)^[20],R-Faster-RCNN^[3],RoI Transformer (Region of Interests Transformer)^[21]。选取的仅依赖水平框监督的算法有:H2Rbox-v2 (Horizontal Bounding Boxes to Rotate Bounding Boxes)^[7],Wholly-WOOD (Wholly Weakly supervised Oriented Object Detector)^[8,14]。通过与第1类方法的比较,可以衡量本文方法在更低标注成本下所能达到的性能水平;通过与第2类方法的比较,可以验证本文所提仿真驱动解耦框架相较于其他同类监督范式的优越性;实验对比结果可见图3直观地展示了本文方法与其他仅依赖水平框监督的方法在HRSID数据集上的可视化结果对比。其中,第1列中的绿色矩形代表数据集的矩形框标注,第2~3列中的绿色矩形代表网络正确预测的目标,红色矩形代表网络的虚警。

在HRSID数据集上,本文方法在AP50这一性能指标上优于R-RetinaNet, S2ANet, RoI Transformer等所有对比的全监督方法。同时,本文方法在召回率方面也始终高于所有对比方法。与现有的基于水平框标注的算法相比,本文方法在AP50上提高了2.8%,在召回率上提升了6.9%。详见表1所示。

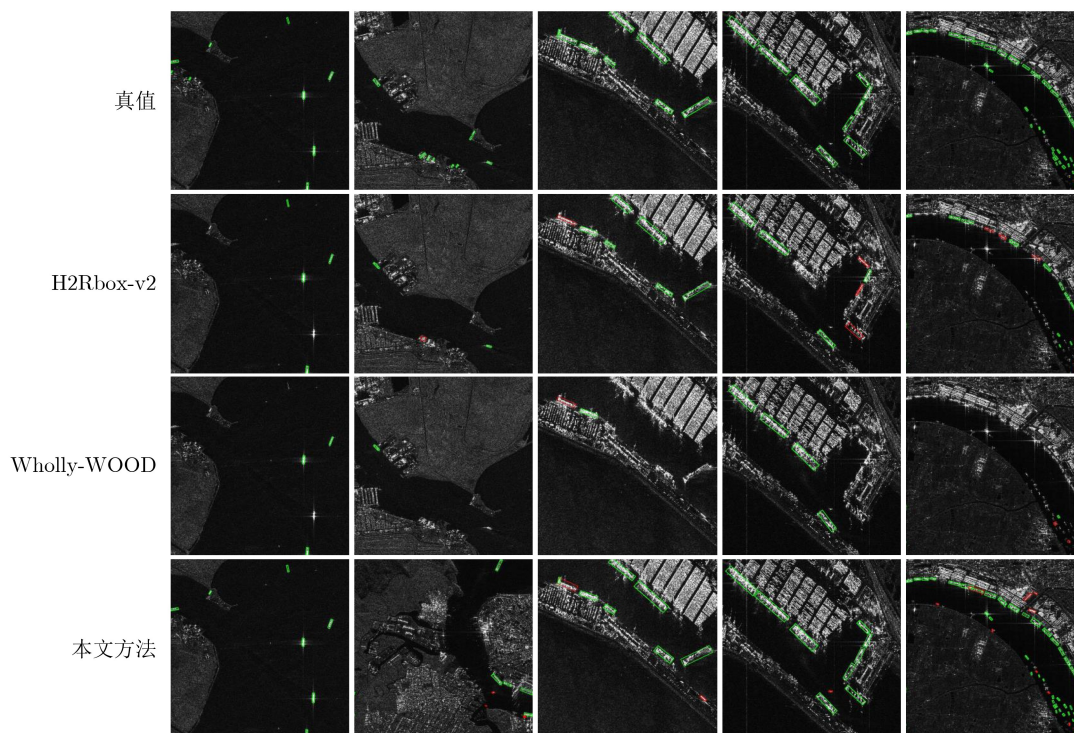


图3 不同方法的检测可视化结果对比图

在SSDD数据集上,本文的方法同样表现出较强的竞争力。与其他基于水平框标注的算法相比,AP50提升了2.1%(89.4% vs. 87.3%),召回率提升了5.9%(95.6% vs. 89.7%)。详见表2所示。

从可视化结果可以观察到,本文方法在漏检情况下较对比方法明显更少。这主要归因于两点:(1)第1阶段的水平框网络训练采用强监督方式,使模型能够有效识别小目标;(2)第2阶段的方向信息源于形态学仿真数据集,能够准确捕捉小目标的方向信息,进一步降低漏检率。因此,本文方法在召回率上的显著提升增强了检测器在复杂环境下的鲁棒性和适应性。

4.3.2 不同轴对齐检测器对实验精度的影响

在本实验中,探讨了轴对齐检测器的精度对后续子网络最终检测精度的影响。具体而言,轴对齐检测器的输出作为子网络的输入,对于整个检测系统的级联精度具有重要作用。本研究选取了4种广泛使用的水平框检测器进行比较:全卷积1阶段检测器(Fully Convolutional One-Stage object detection, FCOS^[22]),YOLOX^[4],CenterNet^[23]和Faster-RCNN^[3]。各检测器的性能可见表3。

可以观察到,在实际应用中,由于后续子网络的方向预测存在一定的错误概率,导致无论选择哪种轴对齐检测器,其AP50值和召回率在级联后的

输出精度上都有一定的下降。然而,尽管存在这种下降,由于轴对齐检测器的性能逐渐趋于饱和,通常仍能够保持较高的召回率。因此,尽管存在误差,级联后的网络仍能在一定程度上保证较高的召回率,进而保证对目标的全面检测。

具体来说,召回率较高的优势在于,尽管可能会引入一定的误检测(假阳性),但能够确保大部分目标被准确检测到。在实际应用中,这一特点尤为重要,尤其在需要高召回率而对误检容忍度较高的任务中。

4.3.3 仿真数据集样本多样性对实验精度的影响

形态学仿真二值样本与真实世界的二值样本的数据分布存在差异,为探究合成数据集样本多样性对实验结果的影响,设置不同的形态学仿真调制条件,一共分为了8个实验组分别与全条件仿真进行对照。8个实验组设置:A:不使用弧形结构仿真;B:不使用舰船结构仿真;C:不进行形态学边缘腐蚀;D:不进行形态学边缘膨胀;E:不进行随机腐蚀;F:不进行十字斑噪声仿真;G:不进行柏林噪声仿真;H:不进行结构截断。

仿真数据集样本多样性的结果展示可见图4。虚线代表使用了全部的仿真条件的性能指标,各实验结果根据对照组进行归一化显示处理。从图4中结果可以看出缺少任意仿真条件都会使子网络的方向预测性能下降。对于HRSID数据集,条件C、条件E和条件H对结果的影响较大,对于SSDD数据集,条件A、条件B和条件C对结果的影响较大。

4.3.4 仿真数据集样本有效性分析与仿真数目对实验精度的影响

为了直观地了解合成数据集与真实数据在二值化后的差异,本文将二维图像展平,并使用统一流形近似与投影(见图5)(Uniform Manifold Approximation and Projection, UMAP)^[24]展示了不同数据源下的分布情况,仿真样本数据集与HRSID数

表1 本文方法与典型检测方法在HRSID数据集上的性能比较(%)

方法	监督方式	AP50	R
R-RetinaNet ^[18]	有向框监督	72.7	79.2
GWD ^[19]	有向框监督	73.3	79.1
S2ANet ^[20]	有向框监督	80.8	83.9
R-Faster-RCNN ^[3]	有向框监督	77.3	82.1
RoI Transformer ^[21]	有向框监督	83.8	86.4
H2Rbox-v2 ^[7]	水平框监督	56.2	68.6
Wholly-WOOD ^[8]	水平框监督	61.5	64.5
文献[14]	水平框监督	81.5	85.0
本文方法	水平框监督	84.3	91.9

表2 本文方法与典型检测方法在SSDD数据集上的性能比较(%)

方法	监督方式	AP50	R
R-RetinaNet ^[18]	有向框监督	79.9	84.0
GWD ^[19]	有向框监督	82.5	85.9
S2ANet ^[20]	有向框监督	88.4	90.0
R-Faster-RCNN ^[3]	有向框监督	87.7	89.8
RoI Transformer ^[21]	有向框监督	89.5	91.0
文献[14]	水平框监督	87.3	89.7
本文方法	水平框监督	89.4	95.6

表3 不同轴对齐检测器对最终精度的影响(%)

数据集	轴对齐检测器	HBBs AP50	OBBs AP50	HBBs R	OBBs R
SSDD	FCOS ^[22]	81.9	78.1	91.2	88.8
	CenterNet ^[23]	90.2	88.5	96.7	94.7
	Faster-RCNN ^[3]	90.0	88.1	93.0	90.1
	YOLOX ^[4]	90.3	89.4	99.1	95.6
HRSID	FOCS ^[22]	78.4	71.2	87.6	80.3
	CenterNet ^[23]	86.6	77.5	92.3	85.1
	Faster-RCNN ^[3]	79.7	70.3	83.5	78.7
	YOLOX ^[4]	89.0	84.3	96.1	90.8

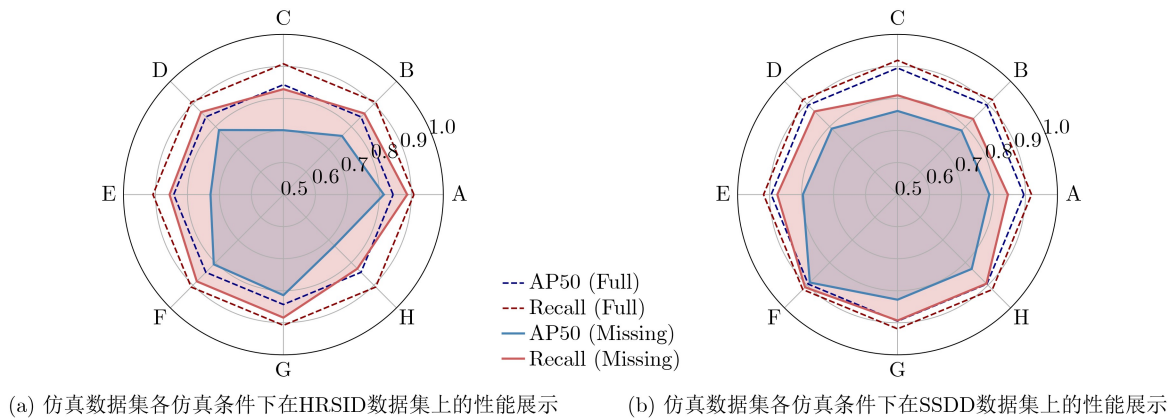


图4 仿真数据集样本多样性结果展示

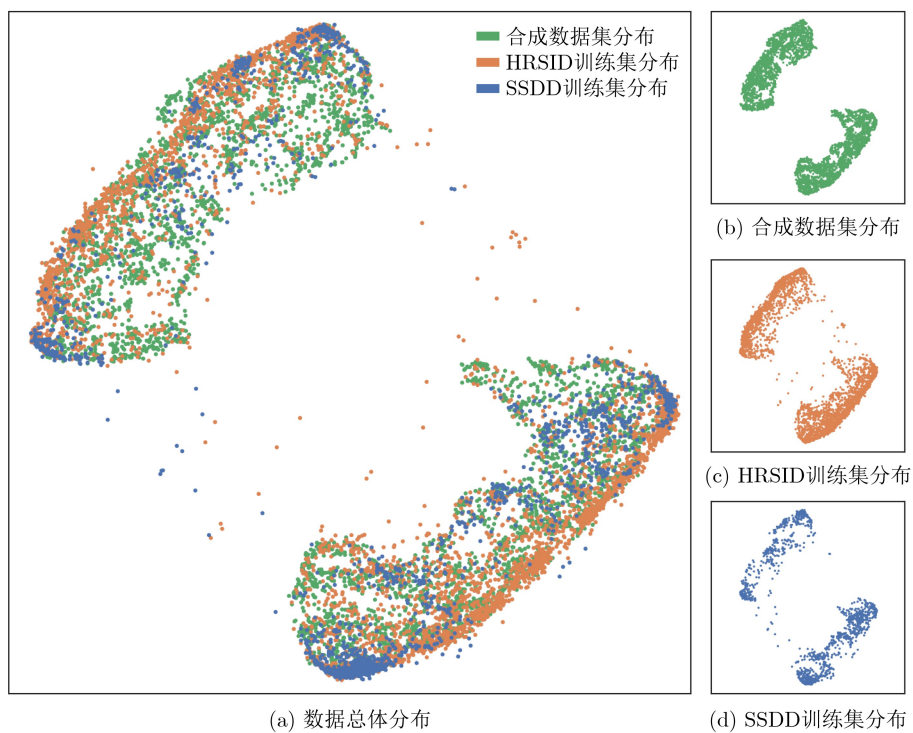


图5 不同源的数据分布降维可视化

数据集都进行随机抽样选取5 000个样本, 由于SSDD训练集样本数少于5 000, 使用了所有训练集进行降维可视化。从图中可以清晰地观察到, 尽管数据来源不同, 但仿真数据集(绿色点)与真实数据集(橙色点与蓝色点)的特征在低维空间中的分布展现出了高度的相似性与重叠性。因此, 模型在仿真数据集上学习到的特征表示能够有效地泛化应用于真实场景, 从而确保了整个框架在面对真实SAR图像时依然具备鲁棒且精准的检测性能。

同时为了探究仿真数据集大小对预测性能的影响, 本文使用1阶段表现最好的水平框检测器YOLOX作为前馈网络, 方向子网络选用以ResNet18为主干的FCOS网络。方向子网络的训练过程中除了训练数目的不同, 其他所有设置保持一致。实验

结果见图6, 在HRSID数据集与SSDD数据集上分别达到了84.3%, 89.4%的AP50。可以发现当仿真数据集数目在1k数量级时, 训练结果不稳定且性能明显比其他数量级的性能要差。随着仿真数据样本数目的增加性能逐渐出现饱和, 进一步提升数据集的量级并不能带来显著性的性能优势, 反而会增加训练的成本与资源消耗。

5 结束语

综上所述, 本文提出了一种基于水平框标注的SAR舰船有向检测方法, 本方法利用形态学仿真合成数据集实现了对目标角度的精准预测, 并能够与现有的所有SAR图像水平框检测器兼容, 无需重新训练即可为网络添加角度预测功能, 实现对舰船目

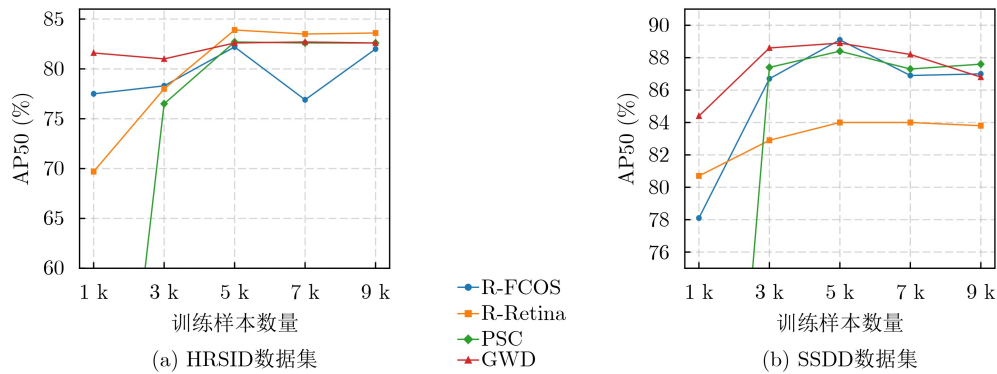


图6 不同方法使用不同训练样本量对AP50精度的影响

标的有向框检测。实验证明了本方法在SAR目标检测任务中的有效性与可推广性。若将其应用于其他类别的目标(如SAR图像中的飞机),核心框架无需改动。仅需根据新目标的几何先验知识,对形态学仿真库进行针对性调整,例如,将模拟船首的操作替换为模拟机翼、机身和尾翼的几何生成规则。这种灵活性使得本方法具备向其他遥感目标检测任务迁移的潜力。未来的研究可以进一步分析合成数据集的属性与检测精度之间的关系,并优化级联结构,尤其是针对第一段网络的输出质量。此外,提升有限监督信息下的检测性能仍是一个值得深入探索的方向。

参考文献

- [1] DI BISCEGLIE M and GALDI C. CFAR detection of extended objects in high-resolution SAR images[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2005, 43(4): 833–843. doi: [10.1109/TGRS.2004.843190](https://doi.org/10.1109/TGRS.2004.843190).
- [2] LENG Xiangguang, JI Kefeng, YANG Kai, et al. A bilateral CFAR algorithm for ship detection in SAR images[J]. *IEEE Geoscience and Remote Sensing Letters*, 2015, 12(7): 1536–1540. doi: [10.1109/LGRS.2015.2412174](https://doi.org/10.1109/LGRS.2015.2412174).
- [3] 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. doi: [10.1109/tpami.2016.2577031](https://doi.org/10.1109/tpami.2016.2577031).
- [4] GE Zheng, LIU Songtao, WANG Feng, et al. YOLOX: Exceeding YOLO series in 2021[EB/OL]. <https://arxiv.org/abs/2107.08430>, 2021.
- [5] CHEN Yuming, YUAN Xinbin, WANG Jiabao, et al. YOLO-MS: Rethinking multi-scale representation learning for real-time object detection[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2025, 47(6): 4240–4252. doi: [10.1109/tpami.2025.3538473](https://doi.org/10.1109/tpami.2025.3538473).
- [6] DOSOVITSKIY A, BEYER L, KOLESNIKOV A, et al. An image is worth 16x16 words: Transformers for image recognition at scale[C]. 9th International Conference on Learning Representations, Austria, 2021.
- [7] YU Yi, YANG Xue, LI Qingyun, et al. H2RBox-v2: Incorporating symmetry for boosting horizontal box supervised oriented object detection[C]. The 37th International Conference on Neural Information Processing Systems, New Orleans, USA, 2023: 2581.
- [8] YU Yi, YANG Xue, LI Yansheng, et al. Wholly-WOOD: Wholly leveraging diversified-quality labels for weakly-supervised oriented object detection[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2025, 47(6): 4438–4454. doi: [10.1109/TPAMI.2025.3542542](https://doi.org/10.1109/TPAMI.2025.3542542).
- [9] HU Fengming, XU Feng, WANG R, et al. Conceptual study and performance analysis of tandem multi-antenna spaceborne SAR interferometry[J]. *Journal of Remote Sensing*, 2024, 4: 0137. doi: [10.34133/remotesensing.0137](https://doi.org/10.34133/remotesensing.0137).
- [10] YOMMY A S, LIU Rongke, and WU Shuang. SAR image despeckling using refined lee filter[C]. 2015 7th International Conference on Intelligent Human-Machine Systems and Cybernetics, Hangzhou, China, 2015: 260–265. doi: [10.1109/IHMSC.2015.236](https://doi.org/10.1109/IHMSC.2015.236).
- [11] KANG Yuzhuo, WANG Zhirui, ZUO Haoyu, et al. ST-Net: Scattering topology network for aircraft classification in high-resolution SAR images[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2023, 61: 5202117. doi: [10.1109/tgrs.2023.3236987](https://doi.org/10.1109/tgrs.2023.3236987).
- [12] ZHANG Yipeng, LU Dongdong, QIU Xiaolan, et al. Scattering-point-guided RPN for oriented ship detection in SAR images[J]. *Remote Sensing*, 2023, 15(5): 1411. doi: [10.3390/rs15051411](https://doi.org/10.3390/rs15051411).
- [13] PAN Dece, GAO Xin, DAI Wei, et al. SRT-Net: Scattering region topology network for oriented ship detection in large-scale SAR images[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2024, 62: 5202318. doi: [10.1109/tgrs.2024.3351366](https://doi.org/10.1109/tgrs.2024.3351366).
- [14] YUE Tingxuan, ZHANG Yanmei, WANG Jin, et al. A weak supervision learning paradigm for oriented ship

- detection in SAR image[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2024, 62: 5207812. doi: [10.1109/TGRS.2024.3375069](https://doi.org/10.1109/TGRS.2024.3375069).
- [15] WEI Shunjun, ZENG Xiangfeng, QU Qizhe, *et al.* HRSID: A high-resolution SAR images dataset for ship detection and instance segmentation[J]. *IEEE Access*, 2020, 8: 120234–120254. doi: [10.1109/access.2020.3005861](https://doi.org/10.1109/access.2020.3005861).
- [16] ZHANG Tianwen, ZHANG Xiaoling, LI Jianwei, *et al.* SAR ship detection dataset (SSDD): Official release and comprehensive data analysis[J]. *Remote Sensing*, 2021, 13(18): 3690. doi: [10.3390/rs13183690](https://doi.org/10.3390/rs13183690).
- [17] ZHOU Yue, YANG Xue, ZHANG Gefan, *et al.* MMRotate: A rotated object detection benchmark using PyTorch[C]. The 30th ACM International Conference on Multimedia, Lisboa, Portugal, 2022: 7331–7334. doi: [10.1145/3503161.3548541](https://doi.org/10.1145/3503161.3548541).
- [18] LIN T Y, GOYAL P, GIRSHICK R, *et al.* Focal loss for dense object detection[C]. 2017 IEEE International Conference on Computer Vision, Venice, Italy, 2017: 2999–3007. doi: [10.1109/iccv.2017.324](https://doi.org/10.1109/iccv.2017.324).
- [19] YANG Xue, ZHANG Gefan, YANG Xiaojiang, *et al.* Detecting rotated objects as gaussian distributions and its 3-D generalization[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2023, 45(4): 4335–4354. doi: [10.1109/tpami.2022.3197152](https://doi.org/10.1109/tpami.2022.3197152).
- [20] LI Jianfeng, CHEN Mingxu, HOU Siyuan, *et al.* An improved S²A-net algorithm for ship object detection in optical remote sensing images[J]. *Remote Sensing*, 2023, 15(18): 4559. doi: [10.3390/rs15184559](https://doi.org/10.3390/rs15184559).
- [21] DING Jian, XUE Nan, LONG Yang, *et al.* Learning RoI transformer for oriented object detection in aerial images[C]. 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, USA, 2019: 2844–2853. doi: [10.1109/cvpr.2019.00296](https://doi.org/10.1109/cvpr.2019.00296).
- [22] TIAN Zhi, SHEN Chunhua, CHEN Hao, *et al.* FCOS: A simple and strong anchor-free object detector[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2022, 44(4): 1922–1933. doi: [10.1109/tpami.2020.3032166](https://doi.org/10.1109/tpami.2020.3032166).
- [23] DUAN Kaiwen, BAI Song, XIE Lingxi, *et al.* CenterNet++ for object detection[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2024, 46(5): 3509–3521. doi: [10.1109/tpami.2023.3342120](https://doi.org/10.1109/tpami.2023.3342120).
- [24] HEALY J and MCINNES L. Uniform manifold approximation and projection[J]. *Nature Reviews Methods Primers*, 2024, 4(1): 82. doi: [10.1038/s43586-024-00363-x](https://doi.org/10.1038/s43586-024-00363-x).
- 汪泽宇: 男, 硕士生, 研究方向为深度学习、目标检测、表征学习。
王青松: 男, 教授, 研究方向为遥感图像精细化处理、协同探测感知与信息融合。
- 责任编辑: 余蓉

A Morphology-guided Decoupled Framework for Oriented SAR Ship Detection

WANG Zeyu WANG Qingsong

(School of Electronics and Communication Engineering, Sun Yat-sen University, Shenzhen 518107, China)

Abstract:

Objective Synthetic Aperture Radar (SAR) is an indispensable remote sensing technology; however, accurate ship detection in SAR imagery remains challenging. Most deep learning-based detection approaches rely on Horizontal Bounding Box (HBB) annotations, which do not provide sufficient geometric information to estimate ship orientation and scale. Although Oriented Bounding Box (OBB) annotation contains such information, reliable OBB labeling for SAR imagery is costly and frequently inaccurate because of speckle noise and geometric distortions intrinsic to SAR imaging. Weakly supervised object detection provides a potential alternative, yet approaches designed for optical imagery exhibit limited generalization capability in the SAR domain. To address these limitations, a simulation-driven decoupled framework is proposed. The objective is to enable standard HBB-based detectors to produce accurate OBB predictions without structural modification by training a dedicated orientation estimation module using a fully supervised synthetic dataset that captures essential SAR ship morphology.

Methods The proposed framework decomposes oriented ship detection into two sequential sub-tasks: coarse localization and fine-grained orientation estimation (Fig. 1). First, an axis-aligned localization module based on a standard HBB detector, such as YOLOX, is trained using available HBB annotations to identify candidate

regions of interest. This stage exploits the high-recall capability of mature detection networks and outputs image patches that potentially contain ship targets. Second, to learn orientation information without real OBB annotations, a large-scale morphological simulation dataset composed of binary images is constructed. The dataset generation begins with simple binary rectangles of randomized aspect ratios and known ground-truth orientations. To approximate the appearance of binarized SAR ship targets, morphological operations, including edge-level and region-level erosion and dilation, are applied to introduce boundary ambiguity. Structured strong scattering cross noise is further injected to simulate SAR-specific artifacts. This process yields a synthetic dataset with precise orientation labels. Third, an orientation estimation module based on a lightweight ResNet-18 architecture is trained exclusively on the synthetic dataset. This module predicts object orientation and refines aspect ratio using only shape and contour information. During inference, candidate patches produced by the localization module are binarized and processed by the orientation estimation module. Final OBBs are generated by fusing the spatial coordinates derived from the initial HBBs with the predicted orientation and refined dimensions.

Results and Discussions The proposed method is evaluated on two public SAR ship detection benchmarks, HRSID and SSDD. Training is conducted using only HBB annotations, whereas performance is assessed against ground-truth OBBs using Average Precision at 0.5 intersection over union (AP50) and Recall (R). The method demonstrates superior performance relative to existing weakly supervised approaches and remains competitive with fully supervised methods (Table 1 and Table 2). On the HRSID dataset, an AP50 of 84.3% and a recall of 91.9% are achieved. These results exceed those of weakly supervised methods such as H2Rbox-v2 (56.2% AP50) and the approach reported by Yue et al.^[14] (81.5% AP50), and also outperform several fully supervised detectors, such as R-RetinaNet (72.7% AP50) and S2ANet (80.8% AP50). A similar advantage is observed on the SSDD dataset, where an AP50 of 89.4% is obtained, representing a significant improvement over the best reported weakly supervised result of 87.3%. Qualitative inspection of detection outputs supports these quantitative results (Fig. 3). The proposed method shows a lower missed-detection rate, particularly for small and densely clustered ships, relative to other weakly supervised approaches. This robustness is attributed to the high-recall property of the first-stage localization network combined with reliable orientation cues learned from the morphological dataset. To examine key methodological aspects, additional experiments are conducted. Analysis of the domain gap between synthetic and real data using UMAP-based visualization of high-dimensional features (Fig. 5) reveals substantial overlap and similar manifold structures across domains, indicating strong morphological consistency. An ablation study of the morphological components (Fig. 4) further shows that each simulation element contributes incrementally to performance improvement, supporting the design of the high-fidelity simulation process.

Conclusions A morphology-guided decoupled framework for oriented ship detection in SAR imagery is presented. By separating localization and orientation estimation, standard HBB-based detectors are enabled to perform accurate oriented detection without retraining. The central contribution is a fully supervised morphological simulation dataset that allows a dedicated module to learn robust orientation features from structural contours, thereby mitigating the annotation challenges associated with real SAR data. Experimental results demonstrate that the proposed approach substantially outperforms existing HBB-supervised methods and remains competitive with fully supervised alternatives. The plug-and-play design highlights its practical applicability.

Key words: Oriented objective detection; Morphological simulation; Simulation-driven learning; Decoupled detection