

分类号：TP751

密 级：公开

学校代码：10363

学 号：2220320180



安徽工程大学

Anhui Polytechnic University

硕士学位论文

题目 基于深度学习的光学遥感
舰船检测

论 文 作 者 朱圣博

校 内 导 师 魏利胜(教授)

校 外 导 师 刘小琿

专业学位类别 电子信息

研究方向(领域) 人工智能

论文提交日期：2025 年 5 月 10 日

分类号: TP751

密 级: 公开

单位代码: 10363

学 号: 2220320180

题 目 基于深度学习的光学遥感舰船检测

题 目 Ship Detection for Optical Remote Sensing

Based on Deep Learning

学 生 姓 名: 朱圣博

校 内 导 师: 魏利胜(教授)

校 外 导 师: 刘小琿

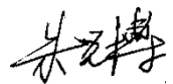
专业学位类别: 电子信息

研究方向(领域): 人工智能

论文答辩日期: 2025年5月10日

安徽工程大学学位论文原创性声明

本人郑重声明：我恪守学术道德，崇尚严谨学风。所呈交的学位论文，是本人在导师的指导下，独立进行研究工作所取得的成果。除文中已明确注明和引用的内容外，本论文不包含任何其他个人或集体已经发表或撰写过的作品及成果的内容。论文为本人亲自撰写，我对所写的内容负责，并完全意识到本声明的法律后果由本人承担。

学位论文作者签名：

日期：2025 年 5 月 10 日

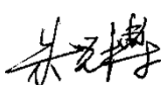
安徽工程大学学位论文版权使用授权书

学位论文作者完全了解学校有关保留、使用学位论文的规定，同意学校保留并向国家有关部门或机构送交论文的复印件和电子版，允许论文被查阅或借阅。本人授权安徽工程大学可以将本学位论文的全部或部分内容编入有关数据库进行检索，可以采用影印、缩印或扫描等复制手段保存和汇编本学位论文。

保密 ☐，在 ____ 年解密后适用本版权书。

本学位论文属于

不保密 ☒。

学位论文作者签名：

日期：2025 年 5 月 10 日

指导教师签名：

日期：2025 年 5 月 10 日

基于深度学习的光学遥感舰船检测

摘 要

舰船作为远洋航运重要载体之一，其智能检测技术在军事和民用领域具有巨大应用潜力。依托光学遥感影像丰富的空间纹理特征和契合人类视觉习惯的成像效果，基于光学遥感影像的舰船检测研究引起国内外学者的广泛关注。然而，由于光学遥感舰船目标尺度微小、细粒度类别众多、数据样本稀少等难点，光学遥感舰船检测面临技术瓶颈。鉴于此，本文针对光学遥感图像的舰船目标检测方法开展深入研究，具体研究内容如下：

(1) 针对光学遥感无向舰船检测的成像不明显、目标占比小的问题，探讨了一种基于通道特征提取的无向小型舰船检测算法。首先，在颈部层引入浅层特征图的基础上修改预测层，平衡浅层位置信息和深层语义信息的权重，增强模型对小目标的关注度；其次，采用融合 FasterBlock 和高效多尺度注意力机制的通道特征提取模块，利用通道分组和跨通道信息交互，加强对小型舰船的特征提取，并降低模型参数；最后，采用动态检测头模块，在不同预测层级上提高模型对不同空间尺度、任务目标的检测能力。实验表明，改进模型在 MASATI、DOTA-Ship、DOTA-Small Vehicle 数据集上的精度和复杂度指标取得最优，实现轻量化、高精度的小型舰船检测。

(2) 针对光学遥感有向舰船检测的细粒度特征混淆、误检漏检的问题，研究了一种基于候选框增强的有向细粒度舰船检测算法。首先，设计横向尺度均衡特征金字塔网络，为候选框的生成提供均衡的尺度特征，以生成精确候选框；其次，设计多通道融合网络，提取候选框中独立且具有鉴别力的特征表示；最后，设计定向动态检测网络，采用动态训练策略为模型训练提供指导，有效利用高质量

候选框识别定位舰船。实验表明，在 ShipRSImageNet、FGSD2021、DOSR 数据集上验证改进方法先进的细粒度舰船检测性能，在此基础上，在 MAR20 数据集上验证了改进方法出色的泛化性。

(3) 针对光学遥感有向舰船检测的旋转框标注成本高、注释少的问题，探究了一种基于动态尺度增强的无向监督有向舰船检测算法。首先，设计动态尺度特征增强模块，通过多尺度特征提取、融合及自适应增强，为自监督分支与弱监督分支提供精准的舰船尺度特征；其次，引入神经架构搜索封装头部，以搜索头部结构最优解；最后，设计外切 Scylla 交并比损失，通过方向角度感知向量引导模型准确预测舰船，并加速训练。实验表明，在 HRSC2016、DOSR 数据集上，与全监督方法相比，所提方法在弱监督下取得最佳检测性能，且在 SSDD 数据集上验证了所提方法的优异泛化性。

综上所述，本文通过研究基于深度学习的光学遥感舰船检测方法，解决无向舰船检测、有向舰船检测、无向监督有向舰船检测的现存问题，对光学遥感舰船检测研究与应用具有重要的参考价值。

关键词：舰船检测；光学遥感；小目标检测；细粒度检测；弱监督检测；深度学习

Ship Detection for Optical Remote Sensing

Based on Deep Learning

ABSTRACT

As one of the key instruments for ocean transportation, ship intelligent inspection technology has great potential for application in military and civil fields. Relying on the advantages that optical remote sensing images can provide the extensive texture information and the clear view aligns with human visual perception, the research on ship detection based on optical remote sensing images has attracted the attention of worldwide researchers. However, optical remote sensing ship detection technology faces the bottleneck stemming from inherent limitations: the small scale of ships, numerous fine-grained categories, and scarce annotation data samples. In light of these challenges, this paper conducts an in-depth investigation into ship object detection method for optical remote sensing images. The specific research contents are as follows:

(1) Aiming at the problems of obscure imaging and small object occupation in remote sensing undirected ship detection, an undirected small ship detection algorithm based on channel feature extraction is explored. Firstly, the prediction layer is modified based on the introduction of shallow feature maps in the neck to balance the weights of shallow positional information and deep semantic information, and enhance focus of the model on small objects. Secondly, the channel feature extraction module, which incorporates FasterBlock and efficient multi-scale attention mechanism, is adopted to enhance feature extraction of small ships by utilizing channel grouping and cross-channel information interactions, and

reduce model parameters. Finally, the dynamic head module is used to improve the capability for different spatial scales and mission objects at different prediction levels. The experiments show that the improve model obtains the optimal accuracy and complexity indexes on MASATI, DOTA-SHIP, DOTA-Small Vehicle datasets, and achieves the lightweight and high-precision small ship detection effectively.

(2) Aiming at the problems of feature confusion, mis-detection and missed-detection in optical remote sensing directed ship detection, an oriented fine-grained ship detection algorithm based on the proposal enhancement is designed. First, the lateral scale-equalizing feature pyramid network is devised to deliver equalized scale features for the subsequent generation of more precise proposals. Secondly, the multi-channel fusion network is designed to extract discriminative feature representation in proposal. Finally, the oriented dynamic prediction network is presented, which utilizes the dynamic training strategy to provide guidance of training, to utilize high-quality proposals validly for identification and localization. The experiments show that advanced fine-grained ship detection capability of the improved model is verified on ShipRSImageNet, FGSD2021, and DOSR datasets. On this basis, the excellent generalization of the improved model is verified on MAR20 dataset.

(3) Aiming at the problems of high cost and less annotation of rotated boxes for optical remote sensing directed ship detection, an undirected supervised directed ship detection algorithm based on dynamic scale feature fusion is designed. Firstly, the dynamic scale feature enhancement module is proposed, which provides precise scale ship features for the self-supervised branch and weakly-supervised branch via feature extraction,

fusion, and adaptive enhancement on multiple scales. Secondly, the neural architecture search is introduced to design the head module to find the optimal structure for the output layers. Finally, the circumscribed scylla intersection over union is designed, which guides the model to predict ship objects accurately and accelerates training through angle-aware vectors. The experiences show that the proposed model achieves the best detection performance under weakly-supervised conditions compared to the fully supervised model on the HRSC2016 and DOSR datasets, and the excellent generalization of the proposed method is verified on SSDD dataset.

In summary, the existing problems of undirected, directed, and undirected-supervised-directed ship detection are solved in this paper based on the deep learning of optical remote sensing ship detection method, which has important reference value for the research and application of optical remote sensing ship detection.

Zhu Shengbo (Electronic Information)

Supervised by Wei Lisheng

Key words: Ship detection; Optical remote sensing; Small object detection; Fine-grained detection; Weak-supervised detection; Deep learning

目 录

摘 要.....	I
ABSTRACT.....	III
第 1 章 绪论.....	1
1.1 研究背景及意义.....	1
1.2 国内外研究现状.....	2
1.2.1 无向舰船检测研究.....	2
1.2.2 有向舰船检测研究.....	3
1.2.3 无向监督有向舰船检测研究.....	5
1.3 主要研究内容.....	6
1.4 本章小结.....	8
第 2 章 基于深度学习的无向小型舰船检测.....	9
2.1 引言.....	9
2.2 无向小型舰船检测模型架构.....	9
2.3 基于通道特征提取的无向小型舰船检测模型.....	10
2.3.1 优化模型结构.....	11
2.3.2 通道特征提取模块.....	12
2.3.3 动态检测头模块.....	15
2.4 实验结果分析.....	16
2.4.1 实验数据集.....	17
2.4.2 实验配置与测评标准.....	17
2.4.3 对比实验.....	17
2.4.4 消融实验.....	19
2.5 本章小结.....	21
第 3 章 基于深度学习的有向细粒度舰船检测.....	22
3.1 引言.....	22
3.2 有向细粒度舰船检测模型架构.....	22
3.3 基于候选框增强的有向细粒度舰船检测模型.....	23
3.3.1 横向尺度均衡特征金字塔网络.....	24
3.3.2 多通道融合网络.....	26
3.3.3 定向动态预测网络.....	28
3.4 实验结果分析.....	30
3.4.1 实验数据集.....	30

3.4.2 实验配置与测评标准.....	31
3.4.3 对比实验.....	31
3.4.4 消融实验.....	37
3.5 本章小结.....	41
第 4 章 基于深度学习的无向监督有向舰船检测.....	42
4.1 引言.....	42
4.2 无向监督有向舰船检测模型架构.....	42
4.3 基于动态尺度增强的无向监督有向舰船检测模型.....	43
4.3.1 动态尺度特征增强模块.....	44
4.3.2 神经架构搜索头部.....	46
4.3.3 外切 Scylla 交并比损失.....	48
4.4 实验结果分析.....	49
4.4.1 实验数据集.....	49
4.4.2 实验配置与测评标准.....	50
4.4.3 对比实验.....	50
4.4.4 消融实验.....	54
4.5 本章小结.....	58
第 5 章 总结与展望.....	59
5.1 总结.....	59
5.2 展望.....	60
参考文献.....	61
攻读学位期间发表的学术论文.....	66
攻读学位期间获得的奖励.....	66
致谢.....	67

插图清单

图 1-1 研究内容框架图	7
图 2-1 基于 YOLOv8s 的无向小型舰船检测模型图	10
图 2-2 基于通道特征提取的无向小型舰船检测网络图	10
图 2-3 模型优化的结构图	11
图 2-4 EMA 的结构图	13
图 2-5 C2f 的结构图	14
图 2-6 C2f-FE 的结构图	14
图 2-7 Dyhead 的结构图	15
图 2-8 MASATI 的定量对比图	18
图 2-9 特征提取效果的对比图	21
图 3-1 基于 Oriented R-CNN 的有向细粒度舰船检测模型图	23
图 3-2 基于候选框增强的有向细粒度舰船动态检测网络图	24
图 3-3 SEPC 结构图	25
图 3-4 LSE-FPN 结构图	26
图 3-5 MCFN 结构图	27
图 3-6 ShipRSImageNet 的定量对比图	33
图 3-7 FGSD2021 的定量对比图	34
图 3-8 DOSR 的定量对比图	35
图 3-9 MAR20 的定量对比图	36
图 3-10 检测混淆矩阵的对比图	38
图 3-11 训练损失曲线的对比图	41
图 4-1 基于 H2RBox-v2 的无向监督有向舰船检测模型图	43
图 4-2 基于动态尺度增强的无向监督有向舰船检测网络图	44
图 4-3 DSFEM 的结构图	45
图 4-4 NAS-Head 的结构图	47
图 4-5 最佳头部采样池的结构图	47
图 4-6 不同回归损失的结构图	48
图 4-7 检测混淆矩阵的对比图	51
图 4-8 HRSC2016 的定量对比图	52
图 4-9 DOSR 的定量对比图	52
图 4-10 SSDD 的泛化对比图	54
图 4-11 特征提取和特征融合效果的对比图	56
图 4-12 回归损失曲线的对比图	58

插 表 清 单

表 2-1 MASATI 的定性对比实验结果	18
表 2-2 DOTA-Ship 的定性对比实验结果	19
表 2-3 DOTA-Small Vehicle 的定性对比实验结果	19
表 2-4 不同模型结构的消融实验结果	20
表 2-5 各模块有效性的消融实验结果	20
表 3-1 具体舰船类别 AP ₅₀ 的对比实验结果	32
表 3-2 ShipRSImageNet 的定性对比实验结果	32
表 3-3 FGSD2021 的定性对比实验结果	34
表 3-4 DOSR 的定性对比实验结果	35
表 3-5 MAR20 的定性对比实验结果	36
表 3-6 各模块有效性的消融实验结果	37
表 3-7 复杂度和检测速率的对比实验结果	39
表 3-8 SEPC 和横向捷径的消融实验结果	39
表 3-9 SEPC 堆叠数量的消融实验结果	39
表 3-10 监督因子的消融实验结果	40
表 3-11 超参数的消融实验结果	40
表 4-1 HRSC2016 和 DOSR 的定性对比实验结果	50
表 4-2 SSDD 的泛化对比实验结果	53
表 4-3 各模块有效性的消融实验结果	54
表 4-4 DSFEM 膨胀率的消融实验结果	55
表 4-5 NAS-Head 卷积的消融实验结果	56
表 4-6 CircumSIoU 损失的消融实验结果	57
表 4-7 CircumSIoU 超参数 θ 的消融实验结果	57

第 1 章 绪论

1.1 研究背景及意义

遥感技术是一种基于航天平台电磁波探测技术，对地表实施数字化感知的前沿科技，作为研究地球表层系统及三维建模的关键工具，其在环境监测、地理勘测等领域具有应用价值。近年来，遥感技术的发展使人类进入多层次、多方位对地观测的新时代^[1]，其图像以高分辨率呈现。通过高清晰化的图像技术，遥感数据不仅为军用领域提供数据支持，而且为民用领域提供生活保障。

在 21 世纪的今天，世界各国对领海主权极为重视，我国在拥有广袤的陆地面积的同时，拥有超 300 万平方公里的海洋面积及近 20000 公里的海岸线，海洋资源丰富，领海国土广阔。“向海而生，背海而衰”，随着“一带一路”倡议的纵深推进与海军“深蓝战略”的加快布局，海洋技术创新在战略需求牵引下呈现出多领域突破态势。其中，随着视觉感知技术的成熟，目标检测成为海洋遥感领域的关注热点，作为广袤海域中的关键运输工具，舰船位置和航行信息备受关注，而目标检测的应用可实现方位及航向的精准识别。因此，遥感舰船目标检测^[2]可在以下方面为我国领海的发展做出贡献：一是强化领海边境安防能力，舰船检测可实现对海洋舰船的高精度识别与动态追踪，提升领海主权管控；二是优化海洋资源管理效能，通过监测舰船活动以保障海洋资源的合理开发利用，为资源统筹规划与生态保护提供支撑；三是促进国际海事协作，舰船检测可构建领海权益的保障体系，助力跨境海洋治理的合作与交流。

遥感舰船检测技术以遥感图像形式可划分为三类^[3]：基于合成孔径雷达(Synthetic Aperture Radar, SAR)图像的检测、基于红外图像的检测、基于光学遥感的检测。SAR 图像因全天候性和隐蔽性等特点，相关舰船检测研究较为广泛，但其成像受雷达波束制约，成像易产生模糊效应致使分辨率下降；红外图像虽有能源利用高和抗干扰性强等特点，却存在灰度层次弱化、辨识纹理细节不足等局限。与上述图像相比，光学图像研究起步较晚，但其凭借成像面积大、宏观性强的优势，为舰船检测提供丰富的物体信息和背景空间信息。

随着硬件性能的提升，利用计算机视觉和深度学习方法相结合的技术进行舰船检测具有巨大优势：深度学习模型通过学习海量数据，自动学习舰船特征，

并对舰船进行准确检测；深度学习模型具备自主优化能力，通过动态调整网络结构与参数配置适应多样化需求；深度学习模型可构建抗干扰机制，有效克服图像畸变与噪声干扰，确保系统适应性与结果可信度；深度学习模型具备结构扩展特性，可依托迁移学习机制应对不同实际场景的挑战。此外，深度学习模型能自动完成特征提取、选择与分类等任务，降低人工干预并提高检测的效率。综上所述，利用深度学习研究方法研究光学遥感舰船检测符合检测技术的发展趋势，对实现光学遥感舰船的高效检测具有重要意义。

1.2 国内外研究现状

光学遥感舰船检测领域的研究已发展多年。在过去，学者致力设计高效手工特征^{[4][5]}进行检测。然而，手工设计视觉特征易受时间、精力及专业程度的限制，难以适应现代多样化的数据。近年来，深度学习为光学遥感舰船检测技术的研究开辟新思路，并演化为核心方法。目前，基于深度学习的舰船检测范式主要分为三类：无向舰船检测、有向舰船检测、无向监督有向舰船检测。

1.2.1 无向舰船检测研究

与自然图像的目标检测任务类似，遥感无向舰船检测采用水平框(Horizontal Box, HBox)以实现海洋舰船目标精确定位，其以单阶段方法、双阶段方法为主。本节分别从单阶段和双阶段对无向舰船检测方法的研究近况进行归纳。

(1) 基于单阶段的无向舰船检测

针对光学遥感图像的高分辨率和复杂背景的特性，单阶段检测方法采用直接执行框回归的方式进行训练，并输出目标信息和类别，满足快速实时检测的实际应用需求。因此，单阶段方法的问世引起了光学遥感舰船检测领域的关注。众多通用单阶段的无向目标检测模型现已迁移至舰船检测领域，如 RetinaNet^[6]、YOLO^[7]、RT-DETR^[8]等，在光学遥感舰船检测领域均取得优异成绩。

面对遥感无向舰船检测存在的挑战，如多尺度舰船等特定问题，单阶段的无向检测模型表现并不出众，为此研究者已探索针对性的解决方案。文献[9]设计特征融合网络保持生成增强语义信息特征，并提出混合注意力增强舰船信息；文献[10]探索相似性掩码结构，根据元素间相似性过滤特征图的有效区域，并结合自注意力和通道注意力以实现全局舰船信息捕获；文献[11]集成 YOLOv7 算法和 Jetson Nano 平台，提升其在动态海洋环境下检测舰船的潜力；文献[12]

构建渠道特征金字塔(Efficient Channel Attention Pyramid, ECAP), 利用转置卷积进行最佳上采样, 提升小尺度舰船的特征提取, 并修改预测层以降低计算消耗; 文献[13]构建残差膨胀卷积块以获取更大感受野, 优化对狭长变化目标的轮廓定位能力, 并结合全局注意力与局部注意力实现对多尺度舰船的关注; 文献[14]提出特征引导模块(Feature-Refinement Module, FRM), 引导金字塔平衡不同层次的语义特征信息, 减轻复杂场景下海岸线对舰船检测的影响。

(2) 基于双阶段的无向舰船检测

与单阶段方法不同, 双阶段方法将检测过程分为两个步骤: ①采用滑动窗口机制构建并生成不同比例的候选框; ②针对候选框执行框回归与分类。尽管双阶段的无向目标检测方法的实时性不高, 但其对目标特征的敏感度更高, 如 R-CNN 系列算法^{[15][16]}等双阶段方法在遥感舰船检测领域被广泛应用。

针对光学遥感无向舰船检测的现存挑战, 双阶段的无向检测方法同样衍生出诸多变体。文献[17]采用区域生成网络(Region Proposal Network, RPN)生成舰船候选特征, 并设计选择性滤波机制, 实现多尺度特征向统一表征空间的动态映射, 为多尺度舰船生成通用特征; 文献[18]受深度可分离卷积(Depthwise Separable Convolution, DSConv)的启发, 设计单次聚合与深度可分离网络, 在特征提取的同时最大程度轻量化模型; 文献[19]通过分析遥感与自然图像的差异, 结合图像上采样与特征金字塔网络(Feature Pyramid Network, FPN)策略, 弥补检测小尺度舰船的不足; 文献[20]使用深层卷积神经网络作为特征提取器以获取目标特征, 并采用 RPN 划分舰船, 根据给定的固有形状对目标执行边界框回归; 文献[21]设计基于 Mask R-CNN 的近岸舰船检测方法, 完全分离舰船目标与图像背景, 并将平滑非极大值抑制策略引入检测框架以提高近岸舰船检测的鲁棒性; 文献[22]鉴于深度森林方法可实现训练资源代价的平衡, 通过结合 RPN 网络和深度森林方法, 降低了复杂场景下不同舰船形状的检测难度。

然而, 在无向舰船检测领域, 由于舰船在图像中目标占比小、特征不明显, 面向遥感图像前景和背景难以区分的海洋场景, 现阶段研究在检测分辨率较小的舰船时易出现漏检和误检, 难以兼顾轻量化和准确率的平衡。

1.2.2 有向舰船检测研究

在光学遥感影像中, 舰船多呈现狭长形态且方向分布任意, HBox 不能满足舰船检测的需求, 无向检测方法在预测密集排列的舰船时性能较差, 同时轨迹

预测与舰船编队均需包含角度的边界框，有向舰船检测采用旋转框(Rotated Box, RBox)紧密贴合目标边界以克服上述问题。与无向舰船检测相同，其以单阶段、双阶段为主。本节从单阶段和双阶段的角度归纳有向舰船检测方法的研究。

(1) 基于单阶段的有向舰船检测

单阶段的有向目标检测方法在通用目标检测方法基础上，在预测层中加入角度预测分支，使其可预测相对图像坐标系的边界框角度，在处理飞机、舰船等有方向性目标尤为重要。因此，单阶段有向检测器，如 $S^2ANet^{[23]}$ 、 $R^3Det^{[24]}$ 、Oriented RepPoints^[25]等，在光学遥感舰船数据集上表现优异。

面对遥感有向舰船检测所存在的特定问题，学者基于单阶段有向检测方法进行针对性的研究。文献[26]设计固定几何损失，在训练阶段学习舰船的几何信息，并采用傅里叶轮廓建模策略正确预测舰船中心点，使其生成精准的 RBox 以实现高精度检测；文献[27]采用形状描述符聚类短边缘中心、长边缘中心和目标中心三种分散特征关键点，并探索软训练分配策略，助力形状描述符训练及隐式学习舰船形状信息；文献[28]设计坐标注意力模块(Coordinate Attention Module, CAM)捕获方向感知和位置敏感特征，增强对舰船的关注度，并引入方向不变模块(Orientation Invariant Module, OIM)生成深度旋转不变特征，提高任意方向舰船的适应性；文献[29]提出有向样本选择策略(Shape Adaptive Selection Measurement, SASM)，通过融合几何知识和特征分布动态选择样本，并为选定正样本增加信息质量以提高检测精度；文献[30]设计相移编码器(Phase-Shifting Coder, PSC)，通过将不同周期的旋转周期性映射至不同频率的相位，解决边界不连续和方形问题；文献[31]将 RBox 转化为二维高斯分布，并利用高斯沃瑟斯坦距离度量分布间的距离损失，使其能根据目标特性动态调整参数梯度。

(2) 基于双阶段的有向舰船检测

双阶段有向目标检测方法在单阶段有向目标检测方法的基础上，引入旋转区域生成网络(Rotation Region Proposal Network, RRPN)，通过旋转感兴趣区域(Rotation Region of Interest, RRoI)池化以提取旋转区域特征。因此，双阶段有向目标检测方法所采用的旋转锚框思想，如 Oriented R-CNN^[32]、ReDet^[33]、RoI-Transformer^[34]等模型，在检测舰船时取得较好的检测效果。

为追求更高的舰船检测准确率，双阶段的有向目标检测方法成为研究热点。文献[35]采用顶点偏移法计算 RBox 顶点和外接矩形顶点偏移量，通过优化偏移

量提升 RBox 生成质量，并设计特征再融合模块，根据细粒度舰船的不同贡献系数映射特征，提供丰富的语义和空间信息；文献[36]研究多层次特征融合识别网络(Multi-Level Feature Fusion Recognition Network, MLFFRN)，按照整体、局部子区域两个层次从候选框提取特征，并计算各子区域的判别性显著度，挖掘关键子区域；文献[37]设计密集连接感受野模块(Densely Connected Receptive Field, DCRF)，以覆盖密集感受野特征，并探索基于注意力机制的特征融合结构(Attention-Guided Feature Fusion, AFF)，通过尺度调整、加权融合突出目标位置，解决旋转舰船检测效率低的问题；文献[38]设计自适应旋转区域生成网络，关注舰船中心位置生成 RBox，以获取优质候选框；文献[39]设计位置与全局特征对齐模块(Local-Global Features Alignment Module, LAM)，通过不同步幅卷积对齐舰船位置特征和全局特征，实现精细舰船检测；文献[40]为捕获和丰富多尺度舰船特征信息，采用旋转等变双向特征融合特征金字塔(Rotation-Equivariant Bidirectional Feature Fusion Pyramid Network, ReBiFPN)，有效提高检测性能。

然而，在有向舰船检测领域，受限于实际场景下遥感影像的舰船分布密集与特征相似的特性，现阶段研究在区分类间相似的细粒度舰船时易出现特征混淆的情况，难以在高分辨率图像中实现准确方向定位和类别判定。

1.2.3 无向监督有向舰船检测研究

上述内容介绍了通用无向与有向的全监督检测方法的研究现状，其性能在极大程度上受限于训练数据集。虽采用 RBox 进行目标检测更为实用，但数据集的缺乏限制有向舰船检测的进一步发展。同时，与 HBox 标注成本相比，标签软件的 RBox 标注更为繁杂，成本约为 HBox 标注的三倍。

为解决上述问题，文献[41]提出的 H2RBox 通过构建双塔结构，结合弱监督分支与自监督分支以共享主干和 FPN。其中，自监督分支仅以特征图作为输入，施加水平外接矩形约束、尺度约束、角度约束，以指导模型生成准确结果；弱监督分支输入 HBox 标注，通过添加角度参数，实现 RBox 输出。由于尺度变化与标注敏感，模型在检测舰船时效果不佳；文献[42]提出的 H2RBox-v2 通过应用对称感知学习的思想，在 H2RBox 的自监督分支上引入对称性范式，将对称轴方向作为目标朝向。与 H2RBox 相比，模型检测舰船性能得到明显提升，但与通用有向舰船检测方法仍存在差距；文献[43]在此基础上引入 LSKNet 网络提取图像中显著背景特征，并添加跨空间多尺度注意力模块捕捉跨空间特征区域，

为无向监督有向舰船检测提供有效解决方案；文献[44]提出一种弱监督学习 (Weakly-Supervised Learning, WSL) 范式，采用无向监督有向的方式，设计方向初始化解码器解码粗略的 RBox，利用精细的预测头部对 RBox 进行提炼，并联合高质量伪标签和 HBox 标注进行训练，使 WSL 范式同主流的全监督有向检测方法保持同一水准。此外，在光学遥感舰船检测领域，针对无向 HBox 监督有向 RBox 的研究较少，是非常具有潜力的研究方向。

然而，在无向监督有向舰船检测领域，由于无向监督有向的检测算法采用弱监督分支预测和自监督分支学习进行舰船预测，使现阶段研究并不成熟，在检测长宽比悬殊、尺度变化大的舰船目标时准确率较低。

1.3 主要研究内容

本文主要研究基于深度学习的光学遥感舰船检测算法，研究内容主要分为三个方面：①针对光学遥感无向舰船检测成像不明显、目标占比小的问题，研究基于通道特征提取的无向小型舰船检测算法；②针对光学遥感有向舰船检测的细粒度特征混淆、误检漏检的问题，探讨基于候选框增强的有向细粒度舰船检测算法；③针对光学遥感有向舰船检测中旋转框注释不足、成本较高的问题，探究基于动态尺度增强的无向监督有向舰船检测算法。整体框架如图 1-1 所示。

第 1 章：论述了论文的选题背景及意义，并通过调研综述了无向舰船检测、有向舰船检测、无向监督有向舰船检测的国内外研究现状，为研究指明方向。

第 2 章：探讨了一种基于通道特征提取的无向检测方法，实现无向小尺度舰船检测。首先，优化模型结构以提高对小目标的尺度感知能力；其次，引入通道特征提取模块实现通道重塑和维度分组；最后，引入动态检测头以提升区分前景和背景能力。实验表明，在 MASATI、DOTA-Ship、DOTA-Small Vehicle 数据集上对改进模型进行消融和对比实验，验证了改进模型的优越性。

第 3 章：研究了一种动态候选框增强的有向检测方法，实现有向细粒度舰船检测。首先，设计横向尺度特征金字塔网络，缓解特征淡化以增强候选框生成；其次，设计多通道融合网络，跨级融合增强特征表示，避免子任务混淆以增强候选框表示；最后，设计定向动态预测网络，自适应拟合候选框分布变化以增强候选框利用。实验表明，在 ShipRSImageNet、FGSD2021、DOSR 和 MAR20 数据集上进行消融和对比实验，证明了该方法的有效性。

第4章：探究了一种动态尺度增强的无向监督有向检测方法，实现 HBox 标注的有向舰船检测。首先，设计动态尺度特征增强模块，为弱监督和自监督分支提供精确特征信息；其次，引入神经架构搜索策略封装头部，并分别集成至角度、回归、分类子网络，自适应搜索最佳结构；最后，研究外切 Scylla 交并比损失，引导弱监督分支准确预测角度值与框回归。实验表明，在 HRSC2016、DOSR、SSDD 数据集上进行消融和对比实验，验证了所提方法的优势。

第5章：总结了本文的主要研究内容，并在此基础上对未来工作的方向与展望展开分析，以期为从事光学遥感舰船检测的研究人员提供参考。

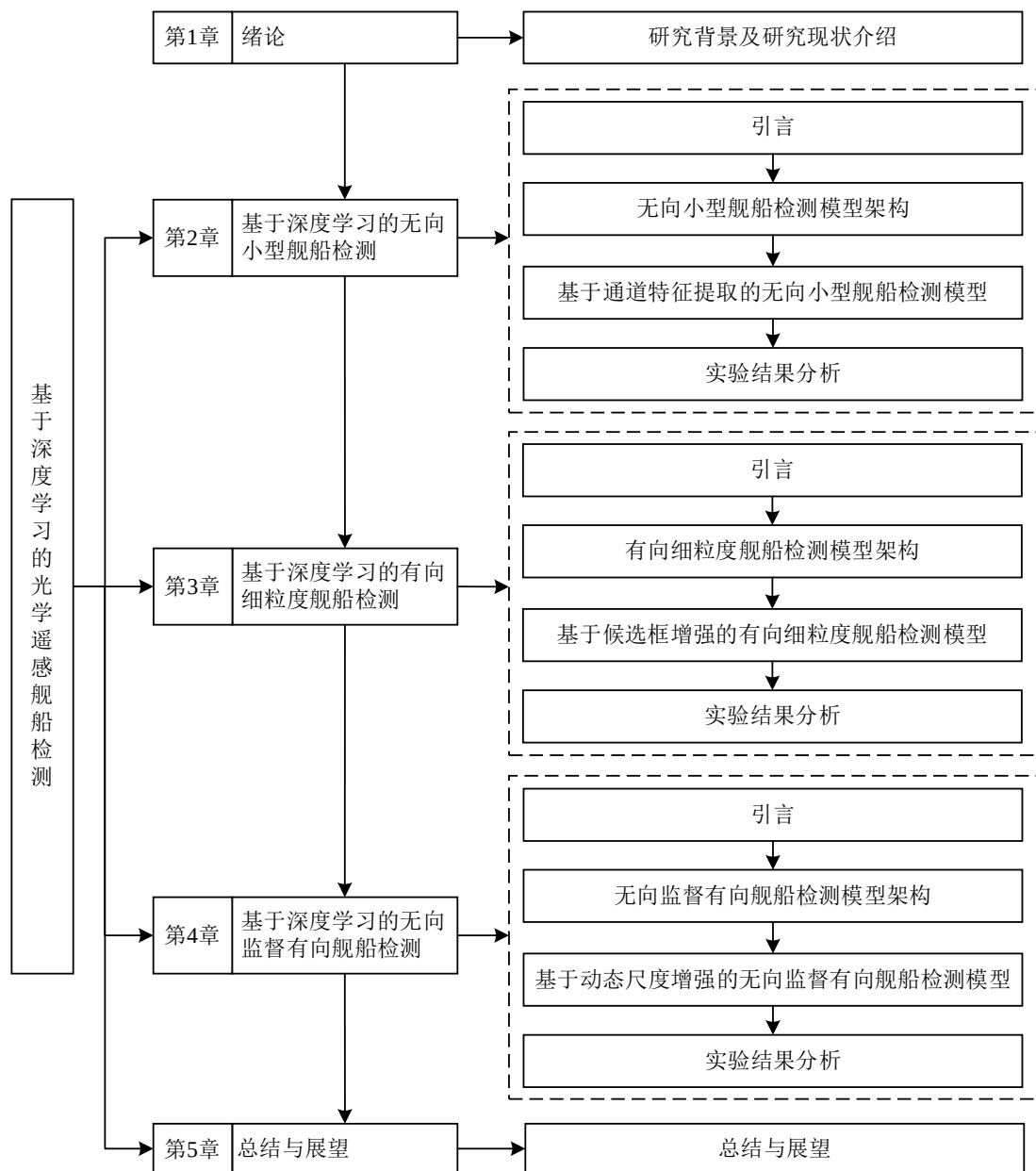


图 1-1 研究内容框架图

1.4 本章小结

首先，针对光学遥感舰船检测课题的研究背景以及现实意义进行详细介绍；其次，从无向舰船检测、有向舰船检测、无向监督有向舰船检测三个角度出发，对国内外光学遥感舰船检测的研究现状和技术方法展开论述；最后，概括介绍总体研究内容框架，简要阐述各章节的主要内容，为后续详细研究作铺垫。

第2章 基于深度学习的无向小型舰船检测

2.1 引言

随着世界经济体系的依存度加深及跨国商贸活动的发展，遥感舰船监测体系因在航道安全预警与海洋权益维护的作用，正成为海洋领域的热门研究方向。光学遥感影像以其广覆盖、全局感知的特点，为舰船识别提供丰富的目标信息和背景空间信息。然而，受制于气象条件与光照变化的干扰，光学遥感影像中舰船目标占比较低、纹理特征弱化，针对海洋环境中目标与背景分界模糊的特征，现有检测方法在低分辨率舰船目标识别时面临漏检率和误检率高的双重挑战，难以兼顾轻量化和准确率的平衡。因此，面向光学遥感场景下无向小型舰船目标检测的研究具有重要的理论意义与应用价值。

考虑到小型舰船目标的角度方位不易确定，在实际工程应用中多数采用无向检测方法。鉴于此，本章首先介绍了无向小型舰船的原始 YOLOv8s 检测模型；其次，改进基于 YOLOv8s 的无向小型舰船检测算法；最终，将本章方法与主流方法在公开数据集上进行实验验证，证明本章方法的有效性。

2.2 无向小型舰船检测模型架构

基于卷积神经网络的目标检测模型，根据网格划分、先验框生成、分类层将其分为三个部分。针对遥感图像高分辨率和复杂背景的特性，单阶段检测算法采用端到端的方式进行训练。与双阶段检测算法相比，其减少生成候选框等操作。因此，单阶段目标检测算法更适合无向小型舰船检测。

YOLO 系列的不断优化改进，使 YOLO 算法来到 v8 版本。与主流单阶段无向目标检测模型相比，YOLOv8s 在复杂度与准确率间实现较好平衡。因此，本章选取 YOLOv8s 作为无向小型舰船检测的基线模型，如图 2-1 所示。YOLOv8s 由主干网络、颈部网络和头部网络组成，主干网络用于提取输入图像的多尺度特征，由复合卷积(Convolution-BatchNorm-SiLU, CBS)模块、特征提取 C2f 模块、快速空间金字塔池化(Spatial Pyramid Pooling Fast, SPPF)模块组成；颈部网络负责提炼和融合由主干网络提取的多尺度特征，YOLOv8s 优化了路径聚合网络(Path Aggregation Network, PANet)，以改善不同特征级别的信息流；头部网络则

负责从精炼特征中生成最终预测，包括预测框坐标、置信度得分和类别标签。此外，YOLOv8s 采用无锚点的方式预测边界框，简化预测过程。

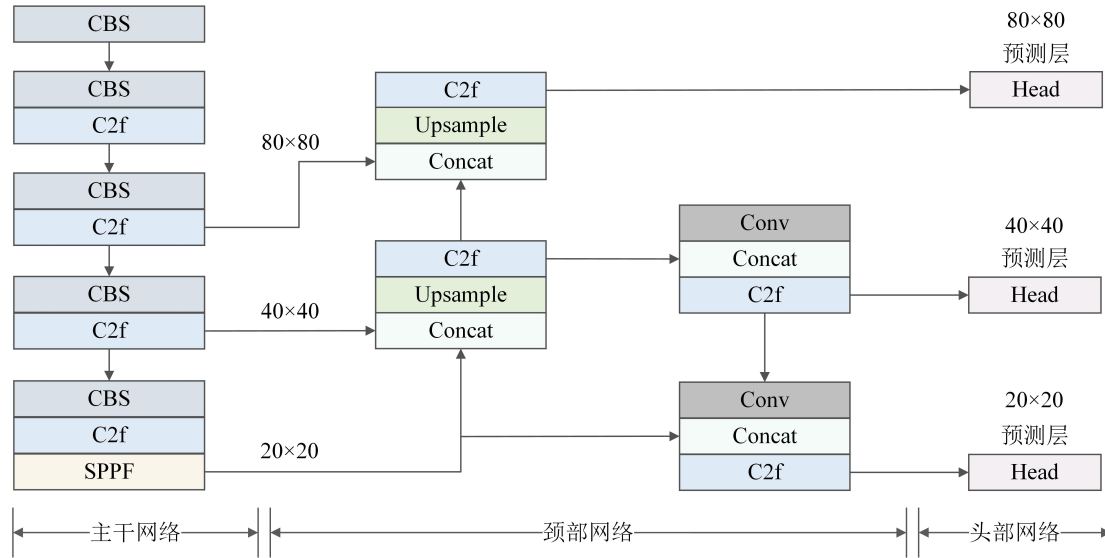


图 2-1 基于 YOLOv8s 的无向小型舰船检测模型图

2.3 基于通道特征提取的无向小型舰船检测模型

考虑到遥感图像的舰船目标尺度小、密度比例大，YOLOv8s 不能满足更高的无向小型舰船检测需求。针对这一问题，从模型结构、特征提取、目标预测三个角度出发，对原 YOLOv8s 进行优化，如图 2-2 所示。

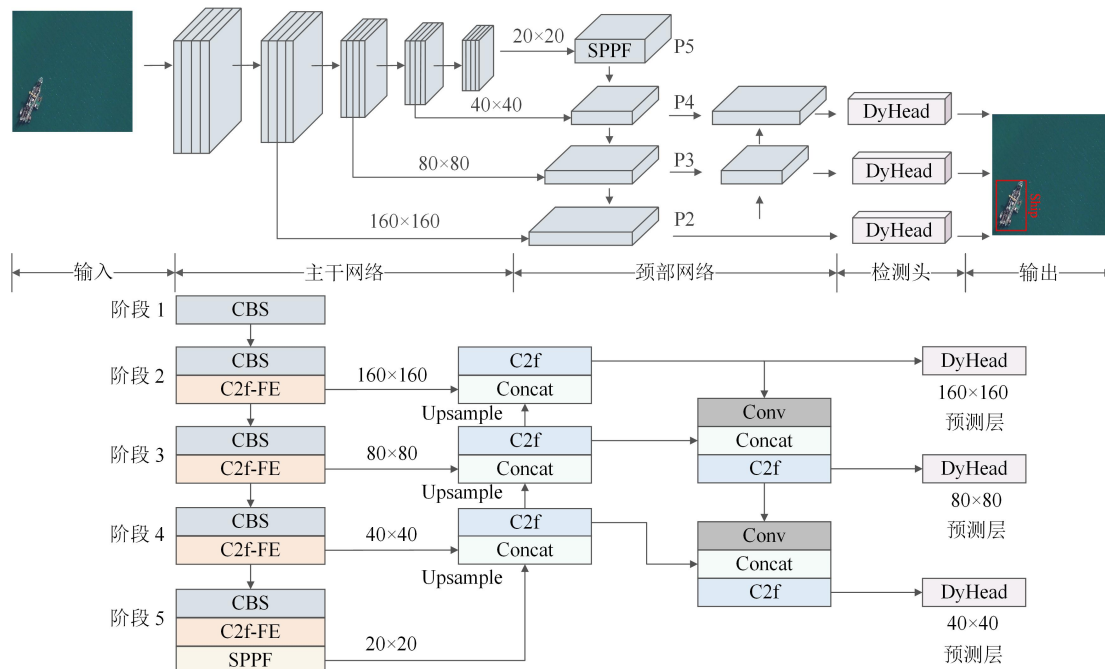


图 2-2 基于通道特征提取的无向小型舰船检测网络图

首先，在颈部网络添加小目标上采样层，引入浅层特征图，通过与主干网络的通道交互，整合浅层空间信息与深层语义特征，同步调整头部网络的预测层，剔除大目标预测层并添加小目标预测层，实现小目标尺度判别优化；其次，结合轻量级网络 FasterNet 的 FasterBlock 与 EMA 注意力机制，设计通道特征提取 C2f-FE 模块，在降低模型参数的同时加强特征提取能力；最后，在头部网络引入动态检测头模块，增强检测解耦头对小尺度特征层级的适应能力，进一步提升模型预测能力。

2.3.1 优化模型结构

在目标检测中，特征图通过颈部层融合跨层级特征，并在预测层构建判别与定位特征。根据网络深度差异，特征图可分为浅层特征图与深层特征图。浅层特征图分布于网络前端，保留高分辨率空间细节，对小型舰船定位具有优势，但受限于卷积深度不足，其所获取的语义信息稀疏。深层特征图则经过多次下采样后生成，通过扩展感受野捕获更多上下文语义信息，但随着图像信息被卷积降维，特征图分辨率逐渐衰减，削弱了小目标的定位感知能力。

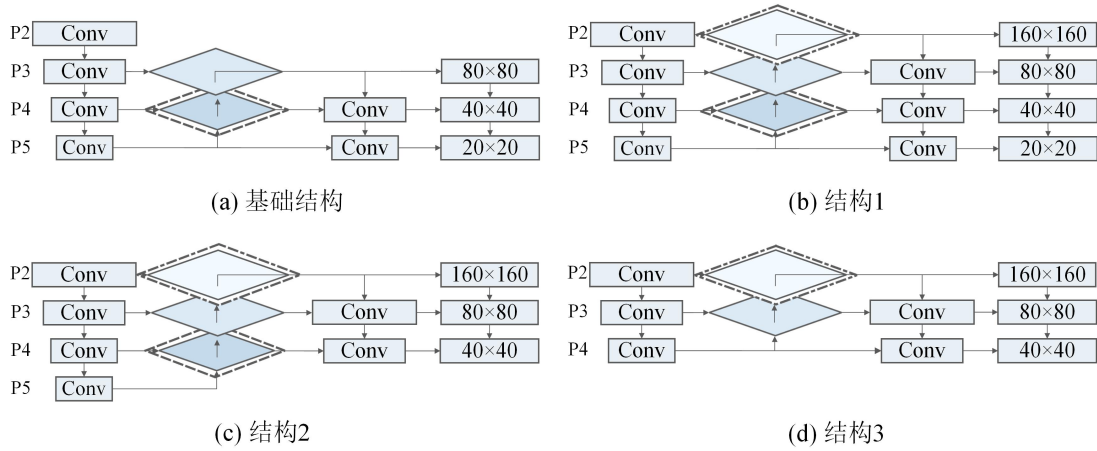


图 2-3 模型优化的结构图

在原始颈部层中，所遍历的连续上采样卷积层较少，致使上采样补偿难以匹配融合的特征尺寸，当遍历连续下采样的卷积层时，易使模型丢失位置特征。此外，在原始预测层中，利用三个深层特征图进行预测，导致特征图越向深层网络传递，保留小型舰船的信息越少，削弱了模型对小型舰船的尺度感知能力。为解决上述问题，优化模型结构，调整参与目标预测与特征融合的深层特征图与浅层特征图权重，增强模型对小目标的尺度感知。图2-3展示了基础结构与权重调整的变体结构，其中菱形框、虚线框、点划线框分别表示不同尺度特征图、

20×20 浅层特征图、160×160 深层特征图。结构 1 在基础结构上保留 20×20 深层特征图和大目标预测层，并引入 160×160 浅层特征图和小目标预测层；结构 2 在结构 1 基础上剔除 20×20 大目标预测层；结构 3 在结构 2 基础上剔除 20×20 深层特征图。通过调整特征图的连通性，从结构 1 至结构 2 至结构 3，深层特征图的权重不断降低，且浅层特征图的权重依次增加。

为确定最佳模型结构的同时确保较小的计算资源占用，经独立的实验验证，得到结构 2 为最佳模型结构。其中，模型 2 通过添加浅层特征图和小目标预测层，有效平衡目标细节信息和语义信息，避免细粒度数据的丢失。同时，深层特征图保留控制模型复杂性的能力，在一定程度上降低了引入浅层特征图的噪声累积。最终，大目标预测层的剔除，进一步轻量模型体积，避免了跨尺度舰船目标间的特征混淆与尺度误判。

2.3.2 通道特征提取模块

随着网络层数的加深与特征图通道数的扩增，致使模型重复提取相似特征信息，导致特征冗余，并降低模型在小目标特征提取过程的敏感度。因此，通过结合高效尺度注意力(Efficient Multi-Scale Attention, EMA)模块^[46]与轻量级网络 FasterNet 的 FasterBlock 结构^[45]，以实现 Faster-EMA Block 结构，进一步设计通道特征提取 C2f-FE 模块替换原始 C2f 模块，借助通道重塑和维度分组分别实现目标的特征保留和模型的轻量化。

(1) EMA 注意力机制

注意力机制通过增强模型对小目标的感知能力，促进关键特征信息的挖掘。EMA 注意力机制提出并行网络和跨空间学习思想，即保持通道维度完整性，在每个并行子网络中构建局部跨通道交互，在降低运算复杂度的同时实现轻量化。图 2-4 展示 EMA 注意力机制的结构图，其通过编码全局信息校准多分支通道权重，跨空间学习聚合并行分支特征，实现像素级成对关系的捕获。

首先，针对输入特征 $X \in \mathbb{R}^{C \times H \times W}$ 采用特征分组(Group)，按通道维度 C 划分为 G 个子特征以学习舰船的语义特征；其次，引入并行网络的思想，双路径沿 1×1 分支依次经过平均池化(Avg Pool)、通道拼接(Concat)、激活函数(Sigmoid)、矩阵相乘(Matmul)实现跨通道特征交互，单路径沿 3×3 分支经过 3×3 标准卷积捕获多尺度特征表示；随后，引入跨空间学习的思想，一方面，利用 Avg Pool 对 3×3 分支编码全局空间信息，与经 Sigmoid 函数的 1×1 分支输出聚合后统一维

度为 $R_3^{1 \times 1 \times C//G} \times R_1^{C//G \times H \times W}$ ，另一方面，利用 Avg Pool 对 1×1 分支编码全局空间信息，与经 Softmax 函数的 3×3 分支的输出聚合后统一维度为 $R_1^{1 \times 1 \times C//G} \times R_3^{C//G \times H \times W}$ 。其中，Avg Pool 的定义如下所示：

$$Z_c = \frac{1}{H \times W} \sum_j^H \sum_i^W x_c(i, j) \quad (2-1)$$

式中， C 代表输入特征图通道数， H 、 W 分别代表输入特征图高度和宽度。

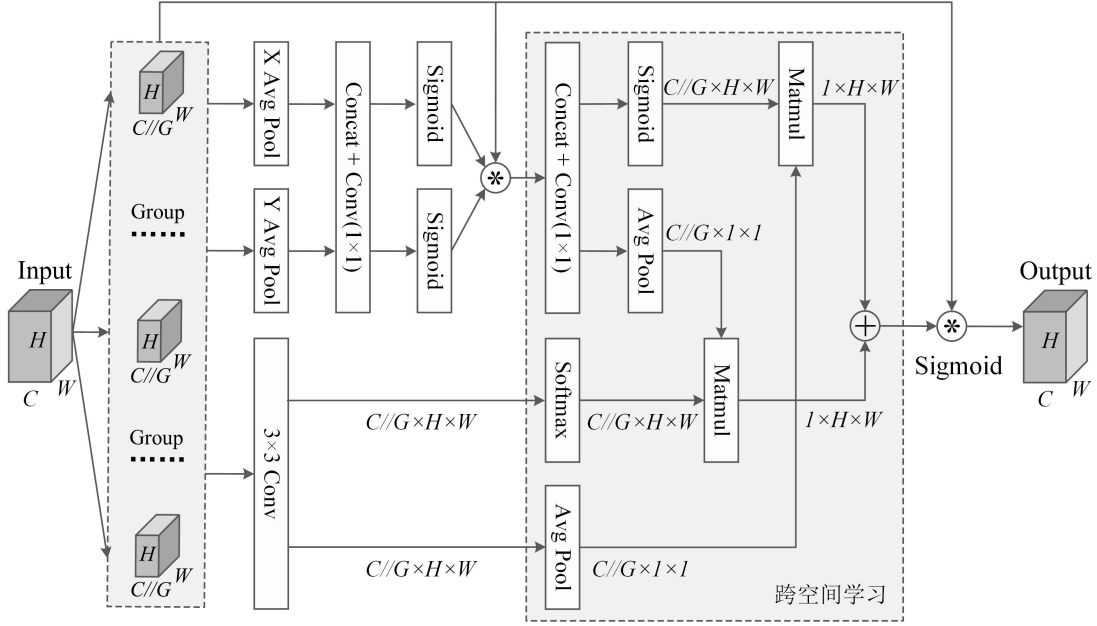


图 2-4 EMA 的结构图

最后，针对各组特征采用 Matmul 操作，计算输出特征映射，并自适应生成各分支的空间注意力权重集合，经过 Sigmoid 函数后得到最终输出。EMA 注意力根据单路径与双路径间的特征重要性校准分支权重，强化小型舰船的关注度，借助跨空间学习思想整合多分支特征信息，改善了模型在检测无向小型舰船时有限感受野导致局部信息交互不良的问题。

(2) C2f-FE 模块

在原始 C2f 模块中，如图 2-5 所示，假设首个 CBS 模块输入通道数为 $2x$ 的特征图，经 Split 层将其分为 2 份通道数为 x 的特征图，后续串联 n 个 Bottleneck 模块的输入均作为 x 个通道的特征图，且每个 Bottleneck 的输出传递为下一个 Bottleneck 的输入。最后， n 个 Bottleneck 模块沿通道特征拼接得到维度为 $(2+n)x$ 的特征图，经第二个 CBS 模块将通道压缩为 $2x$ 输出。其中，原始 Bottleneck 模块由 2 个 CBS 模块与 1 个残差结构组成。

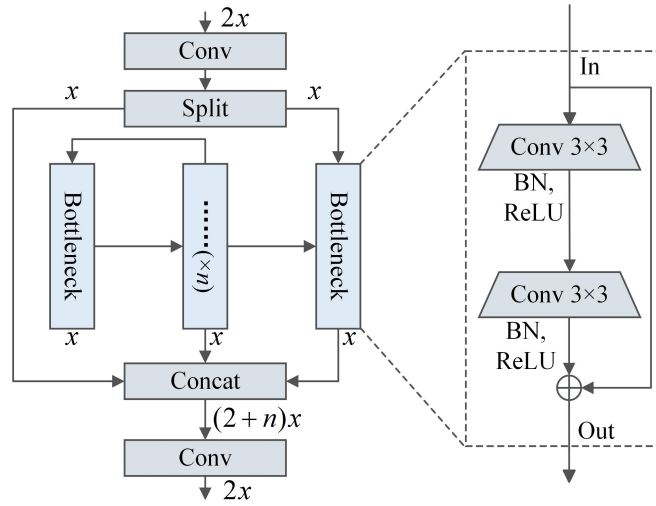


图 2-5 C2f 的结构图

受轻量级主干网络 FasterNet 与 EMA 注意力机制的启发，与原始 C2f 模块相比，通道特征提取 C2f-FE 模块剔除原始的 Bottleneck 结构，并采用融合 FasterBlock 和 EMA 注意力机制的 Faster-EMA Block 结构进行替换，在降低模型参数的同时提高模型对小型舰船的特征提取能力，如图 2-7 所示。与原始的 C2f 不同，首先，为最大化保留原始通道信息的同时降低计算消耗，C2f-FE 模块的 Faster-EMA Block 利用 3×3 部分卷积(Partial Convolution, PConv)选取 $1/4$ 通道执行卷积运算，并将其同余下 $3/4$ 通道进行拼接输出；其次，为聚焦通道信息提取的同时缓解通道分流导致特征信息的丢失，CBR 模块利用 1×1 逐点卷积双倍扩张高层输出特征图的通道数；最后，为确保经卷积运算后的特征图与捷径分支具有相同维度，利用 1×1 逐点卷积还原高层输出特征图的通道数。同时，为实现其余通道信息的充分利用和跨通道舰船特征信息的畅通交互，在 FasterBlock 的 1×1 卷积后加入 EMA 注意力机制结合生成 Faster-EMA Block。

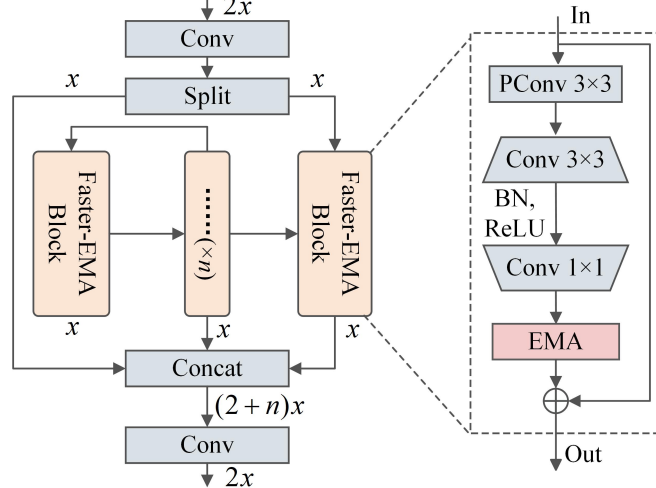


图 2-6 C2f-FE 的结构图

在原始 C2f 模块中，尽管串联 Bottleneck 可实现跨尺度特征的融合与多层级信息的整合，增强特征表达能力，但是每增加一个 Bottleneck 将带来额外参数量的增长。为解决上述问题，在改进 C2f-FE 模块中，采用 Faster-EMA Block 替换 Bottleneck，每个 Faster-EMA Block 仅对 1/4 通道执行卷积运算，而 1×1 卷积较原始 3×3 卷积参数量更小，同时 EMA 注意力的融入最大化限制参数量增长。最终，Faster-EMA Block 的参数量近似为 Bottleneck 的 1/16。

2.3.3 动态检测头模块

在遥感影像中，海洋背景与舰船目标具有相似特征，使模型难以准确区分目标与背景。此外，空间位置和尺度差异进一步加剧预测时分类与定位的难度。动态检测头(Dynamic Head, Dyhead)模块^[47]通过设计分离式注意力思想，分别在尺度感知的特征层级、空间感知的空间层级、任务感知的输出通道层级上部署不同的感知注意力，并整合至同一框架。因此，Dyhead 在保持计算资源不变的前提下，有效提升检测头对目标与背景的判别能力。

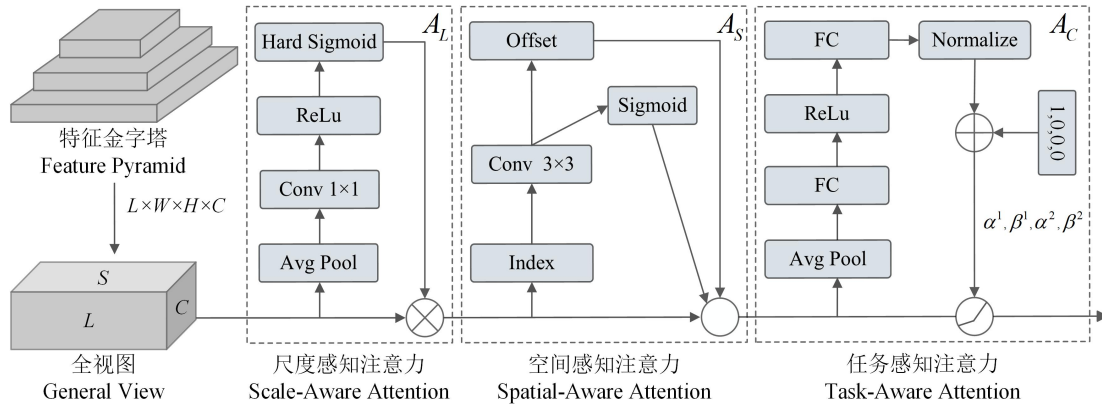


图 2-7 Dyhead 的结构图

由图 2-7 可知，特征金字塔输出的四维张量 $L \times W \times H \times C$ ，经全视图后将其调整为三维张量 $L \times S \times C$ ，即特征层级、空间层级、输出通道层级。随后，在对应层级部署尺度感知注意力(Scale-Aware Attention)、空间感知注意力(Spatial-Aware Attention)、任务感知注意力(Task-Aware Attention)，并依次经过三种分离式注意力。其中，尺度感知注意力根据语义重要性动态整合多尺度舰船的特征；空间感知注意力通过关注舰船的空间位置与特征层级间的共通分辨区域，实现舰船目标位置精准定位；任务感知注意力通过特征通道的动态闭合机制，优化舰船的不同特征，以应对各种检测任务。通过分离式注意力，Dyhead 模块实现多维度的特征感知，提升了无向小型舰船检测过程的鲁棒性。

因此，针对给定的三维特征张量 $H \in \mathbb{R}^{L \times S \times C}$ ，Dyhead 模块将全局注意力分解为三个序列化的注意力，每个注意力仅聚焦单一视角，其定义如下：

$$W(H) = A_C(A_S(A_L(H) \cdot H) \cdot H) \cdot H \quad (2-2)$$

式中， $A_L(\cdot)$ 为尺度感知注意力， $A_S(\cdot)$ 为空间感知注意力， $A_C(\cdot)$ 为任务感知注意力。尺度感知注意力 A_L 的定义如下所示：

$$A_L(H) \cdot H = \sigma\left(f\left(\frac{1}{SC} \sum_{S,C} H\right)\right) \cdot H \quad (2-3)$$

式中， $f(\cdot)$ 是一个 1×1 卷积的线性变换， $\sigma(\cdot)$ 为 Hard-Sigmoid 激活函数。空间感知注意力 A_S 的定义如下所示：

$$A_S(H) \cdot H = \frac{1}{L} \sum_{l=1}^L \sum_{k=1}^K w_{l,k} \cdot F(l; p_k + \Delta p_k; c) \cdot \Delta m_k \quad (2-4)$$

式中， K 为稀疏采样位置个数， Δp_k 和 Δm_k 分别为可变形卷积的偏移量和偏移权重，两者由中间层的输入特征学习得到。任务感知注意力 A_C 的定义如下：

$$A_C(H) \cdot H = \max(\alpha^1(H) \cdot H_c + \beta^1(H), \alpha^2(H) \cdot H_c + \beta^2(H)) \quad (2-5)$$

式中， H_c 为第 c 个通道特征切片， $\theta(\cdot) = [\alpha^1, \alpha^2, \beta^1, \beta^2]^T$ 为学习控制激活阈值的超函数。 θ 在维度 $L \times S$ 上进行全局平均池化，经全连接层和归一化层，结合移位的 Sigmoid 函数归一化输出为 $[-1, 1]$ 。

在较大的遥感图像感知场中，原始 YOLOv8s 的解耦头结构缺乏动态感知的能力。因此，通过引入动态检测头 Dyhead 模块，增强无向小型舰船的识别性能。此外，Dyhead 虽有叠加三种感知注意力的能力，但是鉴于计算资源的占用过高，改进方法仅堆叠使用一个 Dyhead 模块。

2.4 实验结果分析

本节针对改进的 YOLOv8s 模型，共分为四个部分进行实验验证与分析：①介绍无向小型舰船检测所采用的数据集；②介绍无向小型舰船检测模型的实验配置和测评标准；③对优化模型结构进行独立消融实验，并从模型结构、通道特征提取模块、动态检测头三方面进行整体消融实验；④将改进方法与经典无向目标检测方法开展对比分析，探讨改进方法的检测效果是否具有优越性。

2.4.1 实验数据集

MASATI 数据集^[48]由阿利坎特大学构建，通过微软必应地图收集。数据集包含 7389 张光学遥感图像，分辨率均为 512×512 像素，不仅包括各种小型舰船，而且包含不同的场景，如港口、海域、礁石。通过去除无小型舰船目标的图像，选取包含小型舰船的 2368 张图像作为数据集，并将训练集、验证集、测试集按 7:2:1 比例进行随机划分。

DOTA 数据集^[49]由中国资源卫星与谷歌地球应用中心联合构建，包含 2806 张航空影像，涵盖舰船、飞机、各类体育场馆及小型车辆等 15 类目标，共标注 188282 个实例，图像分辨率为 800~4000 像素。本节选取 DOTA 中舰船与小型车辆两个类别，分别构建 DOTA-Ship 数据集和 DOTA-Small Vehicle 数据集，并根据图像金字塔对原始数据进行尺度归一化，统一尺寸为 1024×1024 像素。其中，构建 DOTA-Small Vehicle 数据集以验证模型在检测其余小目标时的泛化性。

2.4.2 实验配置与测评标准

本章实验采用的硬件配置为 i5-12490F 处理器，NVIDIA GeForce RTX 3060 12G 显卡，操作系统为 Windows 11。实验平台是基于 PyCharm 的 PyTorch 1.12.0 框架，批次大小为 16，迭代次数为 300 个 epoch，选取权重衰减为 1e-2 的随机梯度下降(Stochastic Gradient Descent, SGD)作为优化器，初始学习率为 0.01。

本章选取的测评指标为复杂度和检测精度。其中，选取参数量(Parameters, Params)、计算量(Floating Point Operations, FLOPs)反映模型复杂度，选取准确率(Precision, Pre)、召回率(Recall, Re)、平均准确率(mean Average Precision, mAP)衡量检测精度，其定义分别如下：

$$Precision = \frac{TP}{TP + FP}, Recall = \frac{TP}{TP + FN} \quad (2-6)$$

$$mAP = \frac{1}{n} \sum_{i=1}^n \int_0^1 Precision(Recall) d(Recall) \quad (2-7)$$

式中， TP 为真正例， FN 为假反例， FP 为假正例。

2.4.3 对比实验

为验证改进方法在无向小型舰船检测中的性能优势，本节将改进方法与 YOLOv5s、YOLOv7、YOLOv8s、RT-DETR^[8]、文献[9]、文献[10]、文献[11]在 MASATI 数据集上进行定性对比实验，如表 2-1 所示。

由表可知，YOLOv7 和 RT-DETR 因主干网络深度过深，在获取过多深层语义信息情况下，导致小型舰船检测性能欠佳；文献[9]设计的特征提取结构 HR-FPN，虽增强特征捕捉能力，但存在小目标特征提取不充分与计算冗余的缺陷；文献[10]构建的相似掩模机制，虽抑制无效背景干扰，却加剧小型舰船信息丢失，整体性能不突出；文献[11]在轻量化后 Re 值达到最佳 77.0%，但模型压缩程度仍不足。相比之下，本章方法在 Pre、mAP₅₀、mAP_{50:95}、Params 指标上均达到最优，分别为 85.3%、81.4%、34.0%、6.4M，与基线 YOLOv8s 相比，分别提高 4.4%、4.2%、2.2%、3.7M，兼顾检测准确率与复杂度的协同优化。

表 2-1 MASATI 的定性对比实验结果

方法	Pre(%)	Re(%)	mAP ₅₀ (%)	mAP _{50:95} (%)	Params(M)
YOLOv5s	82.3	73.3	77.1	30.2	7.1
YOLOv7	75.1	57.7	60.8	20.3	36.5
YOLOv8s	80.9	74.7	77.2	31.8	11.1
RT-DETR	80.8	76.1	79.1	33.0	19.9
文献[9]	84.6	72.0	77.3	—	16.6
文献[10]	82.5	76.5	78.7	—	7.5
文献[11]	80.0	77.0	76.0	—	36.1
本章方法	85.3	74.9	81.4	34.0	6.4

此外，本节在 MASATI 数据集上对部分算法的检测效果执行定量对比实验，如图 2-8 所示，结果表明，与 YOLOv5s、YOLOv7、YOLOv8s、RT-DETR 四种方法相比，本章方法在小型舰船识别任务中具有更优检测精度。

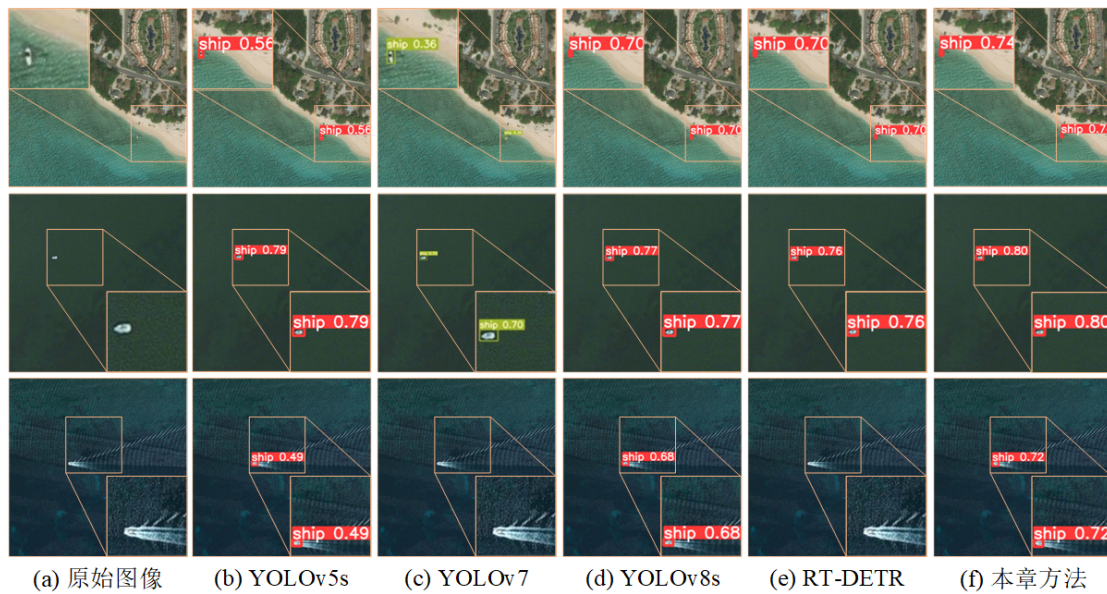


图 2-8 MASATI 的定量对比图

为验证改进方法在多个场景下检测小型舰船的优越性和检测其余小型目标泛化性,选取 YOLOv5s、RT-DETR、YOLOv7、YOLOv8s、文献[50]、文献[51]作为对比基线,分别在 DOTA-Ship 数据集和 DOTA-Small Vehicle 数据集上执行定性对比实验,如表 2-2 和表 2-3 所示。

表 2-2 DOTA-Ship 的定性对比实验结果

方法	Pre(%)	Re(%)	mAP ₅₀ (%)	mAP _{50:95} (%)
YOLOv5s	93.0	86.1	90.2	65.5
YOLOv7	92.9	85.5	89.3	63.1
YOLOv8s	92.6	87.0	91.2	67.9
RT-DETR	92.2	85.1	88.7	64.6
文献[50]	93.9	81.9	90.9	64.5
文献[51]	—	—	88.5	66.4
本章方法	93.5	87.4	91.4	69.6

表 2-3 DOTA-Small Vehicle 的定性对比实验结果

方法	Pre(%)	Re(%)	mAP ₅₀ (%)	mAP _{50:95} (%)
YOLOv5s	82.3	73.3	77.1	30.2
YOLOv7	75.1	57.7	60.8	20.3
YOLOv8s	80.9	74.7	77.2	31.8
RT-DETR	80.8	76.1	79.1	33.0
文献[50]	84.6	72.0	77.3	—
文献[51]	82.5	76.5	78.7	—
本章方法	85.3	74.9	81.4	34.0

由表 2-2 与表 2-3 可知,改进方法的 mAP₅₀ 和 mAP_{50:95} 指标均达到最优。在 DOTA-Ship 和 DOTA-Small Vehicle 数据集上,与基线方法相比,改进方法的 mAP₅₀ 指标达到 91.4%和 81.5%,分别获得 0.2%和 1.3%的增益,改进方法的 mAP_{50:95} 指标达到 69.6%和 51.5%,分别获得 1.7%和 1.4%的增益。因此,针对光学遥感目标检测任务,改进方法不仅在检测小型舰船目标时具有优越性,而且在检测其余小目标(如小型车辆)时具备出色的泛化性。

2.4.4 消融实验

为探究不同优化模型结构对光学遥感影像中无向小型舰船检测任务的影响,并确定最优模型结构,如表 2-4 所示,本节在 MASATI 数据集上对三种不同的优化模型结构执行独立的消融实验。

表 2-4 不同模型结构的消融实验结果

模型结构	Pre(%)	Re(%)	mAP ₅₀ (%)	mAP _{50:95} (%)
基础结构	80.9	74.3	77.2	31.8
结构 1	82.5	72.9	77.5	32.0
结构 2	83.9	74.7	79.4	33.2
结构 3	81.9	73.4	78.2	32.2

由表 2-4 可知，结构 1 引入浅层特征图和小目标预测层，虽在小型舰船检测精度上有所改善，但大目标预测层干扰模型的尺度辨识，整体性能的提升有限；结构 2 在结构 1 的基础上，保留深层特征图并剔除大目标预测层，既实现深层语义特征的融合，又优化预测层权重分配并加强小目标检测，与基础结构相比，mAP₅₀ 和 mAP_{50:95} 分别提高 2.2% 和 1.4%；结构 3 在结构 2 的基础上剔除深层特征图，因缺乏足够的深层语义信息，导致其检测精度较基础结构未有明显提升。因此，结构 2 为小型舰船检测的最佳结构。

为直观呈现改进方法相较基线方法的性能优势，本节在 MASATI 数据集上通过依次添加改进模块以执行消融实验，如表 2-5 所示。

表 2-5 各模块有效性的消融实验结果

方法	组 1	组 2	组 3	组 4	组 5	组 6
基线方法	√	√	√	√	√	√
C2f-FE		√			√	√
结构 2			√		√	√
Dyhead				√		√
mAP ₅₀ (%)	77.2	77.0	79.4	78.1	80.3	81.4
mAP _{50:95} (%)	31.8	31.9	33.2	32.7	32.9	34.0
Params(M)	11.1	9.7	7.4	10.9	6.0	6.4

由表 2-5 可知，实验组 1 与实验组 2 对比显示，通道特征提取 C2f-FE 模块通过跨维度交互强化通道间的特征交互，在保持精度不变的前提下实现通道维度压缩，使模型参数量得以降低；实验组 1 与实验组 3 对比显示，结构 2 平衡浅层特征与深层特征参与特征融合与预测权重，有效缓解模型的定位偏移和尺度误判，mAP₅₀ 和 mAP_{50:95} 分别获得 2.2% 和 1.4% 的增益，参数量降低 30%；实验组 1 与实验组 4 对比显示，Dyhead 模块通过动态注意力聚焦小型舰船目标预测，mAP₅₀ 和 mAP_{50:95} 均提升 0.9%；实验组 3 与实验组 5 对比显示，在应用结构 2 基础上引入 C2f-FE 模块，可获得精度与复杂度的协同增益；实验组 1 与实验组

6 对比显示，与基线方法相比，改进方法参数量降低 42.3%， mAP_{50} 和 $mAP_{50:95}$ 分别提升 4.2%和 2.2%，证明其在检测精度与实际落地的双重优势。

为清晰可视化改进方法对小目标特征提取的增强，如图 2-9(a₁-d₁)所示，输入背景各异、单多目标的图像。如图 2-9(a₂-d₂)和图 2-9(a₃-d₃)所示，基线方法存在目标关注度不足、特征提取能力薄弱的问题，导致沿岸植被出现误检，而本章方法提升了舰船特征提取的效果，实现细节特征的聚焦与误判干扰的抑制。

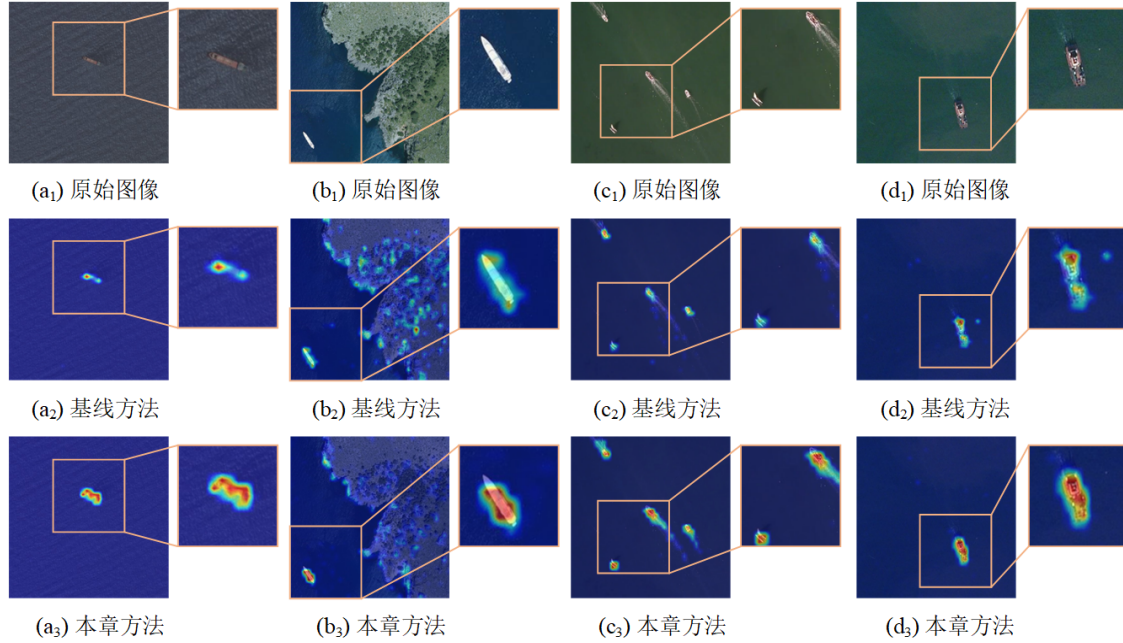


图 2-9 特征提取效果的对比图

2.5 本章小结

为解决光学遥感图像无向检测中小型舰船难以检测的问题，本章研究了一种改进 YOLOv8s 模型。其中，模型的主要贡献在三个方面：①提高小尺度判别能力（优化模型结构）；②增强通道特征提取能力（C2f-FE 模块）；③加强分类与定位能力（Dyhead 模块）。实验结果表明，在三个通用数据集上执行消融实验和对比实验，与其他主流的目标检测模型相比，改进模型在检测效果和模型大小方面均表现优异，且更易实际部署应用。

第 3 章 基于深度学习的有向细粒度舰船检测

3.1 引言

在第 2 章的无向舰船检测中，经算法设计和实验论证，实现了光学遥感影像中小型舰船目标的识别。然而，受限于实际场景下遥感影像的目标分布密集与背景错综复杂的特性，难以在高分辨率图像中进行准确方向定位和类别判定。为将舰船从背景信息分离，根据不同标注形式，存在分割标注、无向标注、有向标注三种方式进行细粒度舰船检测(Fine-Grained Ship Detection, FGSD)。

基于分割的标注形式，实现像素级的信息提取，但其存在标注成本高、检测耗时长的缺陷，不适用于端到端的 FGSD 系统；基于无向的标注形式，实现高速交并比(Intersection of Union, IoU)计算，但舰船朝向不同且密集排布，导致单目标框内重叠其余舰船部分，非极大值抑制(Non-Maximum Suppression, NMS)过程产生漏检现象；基于有向的标注形式，兼顾考虑舰船朝向不同、比例狭长的问题，引入角度属性，剔除冗余背景信息，避免产生漏检和误检。

鉴于此，本章采用有向细粒度舰船检测框架，从候选框生成、候选框表示及候选框利用的角度出发，分别设计三个候选框增强模块。经定性和定量实验，有效验证所研究的方法能实现高效的细粒度舰船识别与定位。

3.2 有向细粒度舰船检测模型架构

在 FGSD 任务中，由于舰船特征相似，使模型在预测时容易出现特征混淆的情况。与双阶段检测器相比，直接执行边界框回归和分类使其对细粒度目标的特征敏感度不高。双阶段检测方法分为两个步骤：①生成候选框并粗略定位；②细化候选框并进行细粒度识别。尽管双阶段的框架更为复杂，但其凭借高精度的优势，使双阶段检测器在 FGSD 任务更受欢迎。

为了检测遥感图像中任意方向的舰船，众多基于 CNN 的双阶段有向检测模型已迁移至 FGSD 领域。其中，Oriented R-CNN^[32]在传统 R-CNN 基础上增加对目标方向性的识别能力，以低成本的方式生成高质量有向候选框，突破有向目标检测的计算瓶颈。因此，本章选取 Oriented R-CNN 作为有向细粒度舰船检测框架，如图 3-1 所示。首先，利用主干网络 ResNet50 和颈部网络 FPN 分别对输

入的图像进行特征提取，生成并融合不同尺度的特征图；其次，在第一阶段中，定向区域生成网络(Oriented Region Proposal Network, ORPN)根据 FPN 的特征图，生成有向候选框，并通过定向感兴趣区域(Region of Interest, RoI)对齐提取固定尺寸的特征；最后，在第二阶段中，Oriented R-CNN 检测头将 FPN 提取的特征和 ORPN 生成的候选框作为输入，执行细粒度分类和精细框回归。通过上述方式，Oriented R-CNN 在保证高检测精度的同时兼顾检测速率。

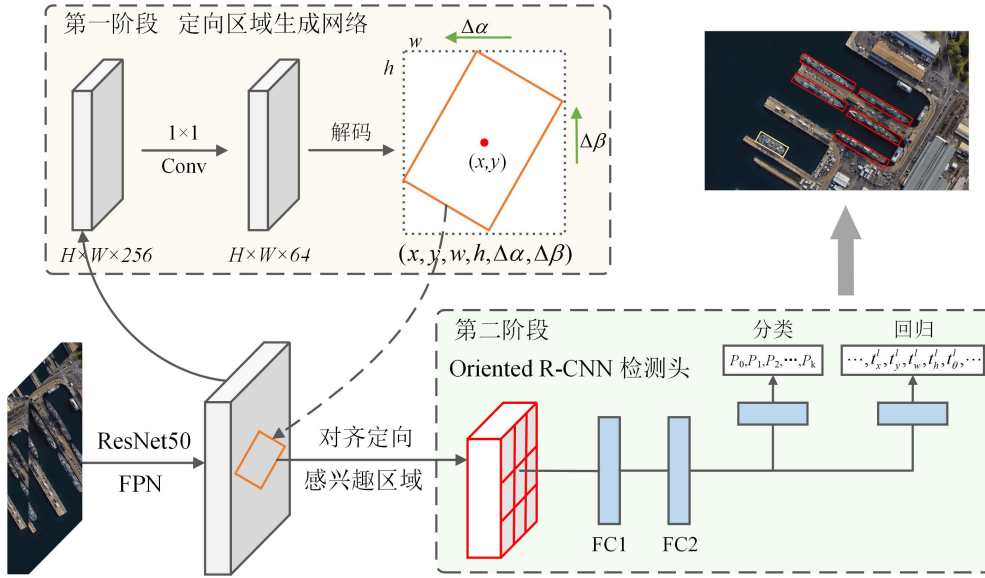


图 3-1 基于 Oriented R-CNN 的有向细粒度舰船检测模型图

3.3 基于候选框增强的有向细粒度舰船检测模型

Oriented R-CNN 模型在舰船检测仍存在问题。首先，第一阶段生成的 RPN 生成候选框作为关键任务，传统方法依赖 FPN 直接融合尺度特征，为候选框生成提供尺度和位置信息，但在尺度变换中各层级间忽略尺度归一化，导致特征淡化或丢失，从而生成粗糙的候选框；其次，第一阶段负责前背景区分和候选框粗略定位，第二阶段负责候选框细化和细粒度识别，但两个阶段的特征均从 FPN 提取，导致子任务间混淆；最后，预测网络采用分类器和回归器实现目标分类与定位，在分类器中，不同舰船类别的宽高比悬殊，边界框的细微角度差异会显著降低 IoU 分数，固定 IoU 阈值的标签分配将损害细粒度识别学习。在回归器中，损失函数会随 IoU 阈值的升高而大幅降低高质量样本的贡献，使高质量正样本得不到充分利用，固定回归损失限制了回归器的定位能力。

为解决上述问题，受候选增强策略^[52]启发，本章研究了一种基于候选框增

强的有向细粒度舰船动态检测网络(Proposal Enhancement Dynamic Network, PEDNet), 从候选框生成、候选框表示、候选框利用三个角度出发, 分别设计横向尺度均衡特征金字塔网络(Lateral Scale-Equalizing Feature Pyramid Network, LSE-FPN)、多通道融合网络(Multi-Channel Fusion Network, MCFN)、定向动态预测网络(Oriented Dynamic Prediction Network, ODPN), 如图 3-2 所示。

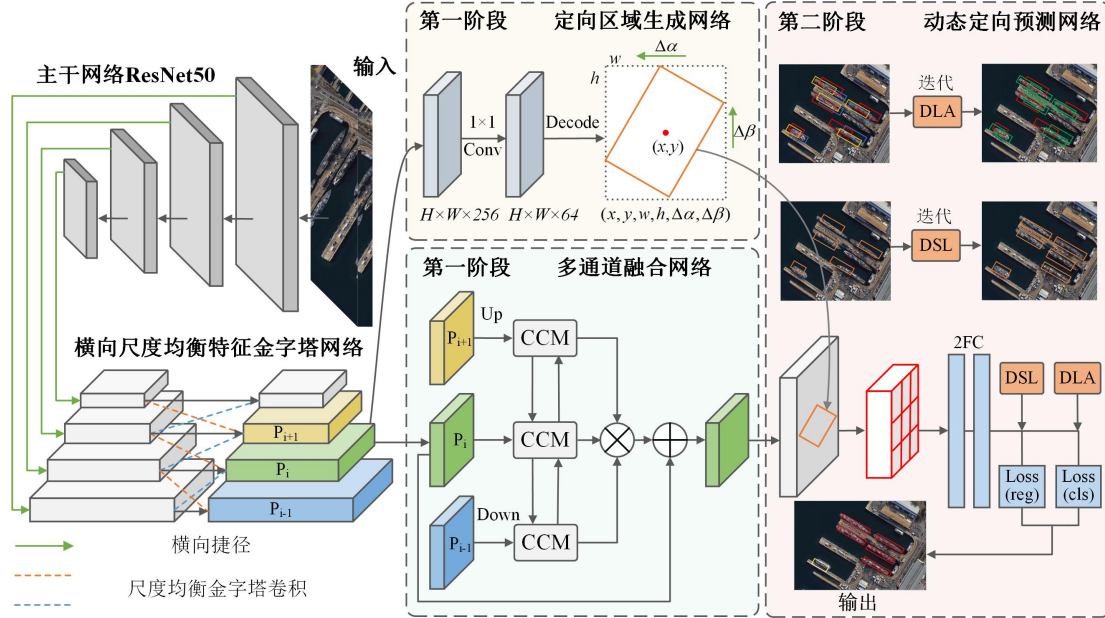


图 3-2 基于候选框增强的有向细粒度舰船动态检测网络图

首先, ResNet50 负责提取目标的语义信息, 并生成不同尺度特征图, LSE-FPN 融合不同尺度的特征图, 以获取全面的特征信息, 经多次信息提取和融合后, 输出包含各种语义信息的特征映射至 ORPN; 其次, 在第一阶段中, ORPN 根据 LSE-FPN 所输出的特征信息, 采用具有预定义角度和长宽比的有向预测框, 为目标舰船生成候选框, 同时 MCFN 通过多通道信息跨级融合, 输出独立的判别特征至下一阶段的候选框, 以增强候选区域的代表性; 最后, 在第二阶段中, ODDN 采用动态分类器与动态回归器, 实现候选框类别分布变化的动态适应和高质量候选框权重的提升, 使模型专注细粒度舰船识别的学习。

3.3.1 横向尺度均衡特征金字塔网络

在有向细粒度舰船检测中, 目标尺度作为细粒度的判别标准之一尤为重要。优秀的特征金字塔网络可从海量数据中融合有效的尺度特征和位置特征, 为高质量候选框的生成提供保障, 使模型准确完成各种视觉任务。本节对 FPN 进行分析, 并设计一种名为 LSE-FPN 的高效 FPN。

(1) 尺度均衡金字塔卷积

针对多尺度舰船检测，FPN 选择直接融合多尺度特征，以增强特征表达能力。然而，FPN 忽略各层级的尺度归一化，导致尺度不匹配，造成大量负样本，不利于高质量候选框的生成。因此，引入尺度均衡金字塔卷积(Scale-Equalizing Pyramid Convolution, SEPC)^[53]为 LSE-FPN 的一环，进一步提取与融合尺度不变特征，解决特征淡化问题，利于 ORPN 生成候选框。

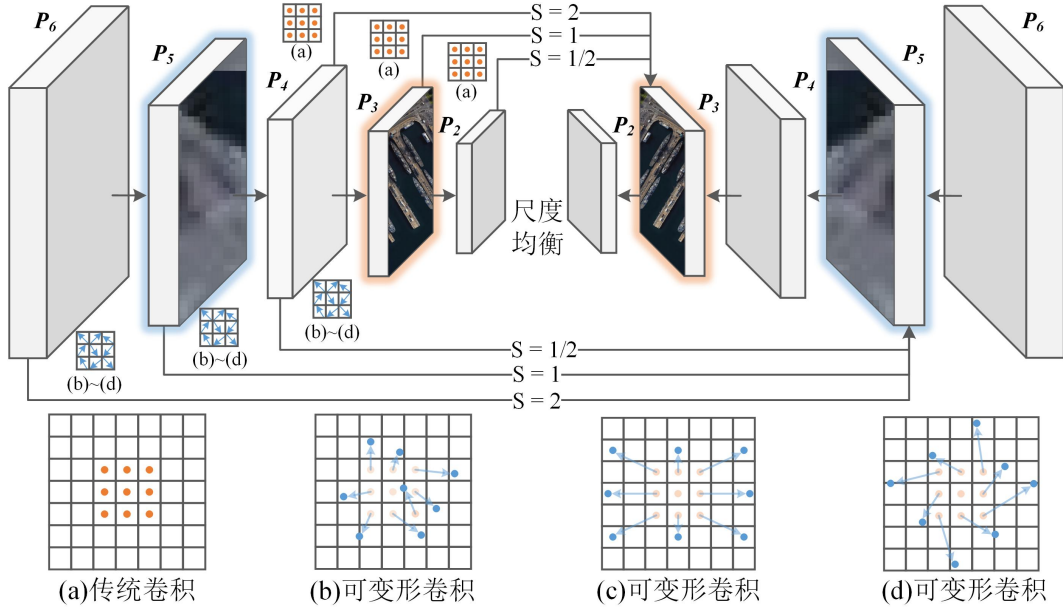


图 3-3 SEPC 结构图

SEPC 由金字塔卷积(Pyramid Convolution, PyConv)和可变形卷积构成，如图 3-3 所示，左侧为经 FPN 的特征图，右侧为经 SEPC 的特征图。一方面，为了解决各级特征图上目标尺度不匹配的问题，PyConv 使用不同步长的卷积进行校准。以 P_3 层级为例，PyConv 在 P_4 、 P_3 、 P_2 三个层级上通过步长为 1/2、1、2 的卷积进行尺度均衡化，并将结果拼接至 P_3 层级，输出结果如下：

$$Y_l = W_1 * S_{1/2} X_{l+1} + W_0 * S_1 X_l + W_{-1} * S_2 X_{l-1} \quad (3-1)$$

式中， l 表示金字塔层级， X 表示输入特征图， W_1 、 W_0 、 W_{-1} 表示三个独立的 2D 卷积核， $*S_2$ 表示步幅为 2 的卷积。由于步幅为 1/2 的卷积难以实现，故通过双线性上采样，采用步幅为 1 的卷积替换，输出结果如下：

$$Y_l = \text{Upsample}(W_1 * S_1 X_{l+1}) + W_0 * S_1 X_l + W_{-1} * S_2 X_{l-1} \quad (3-2)$$

另一方面，经过非线性和多层卷积操作，高层特征的模糊效应更为严重。

因此, SEPC 对 PyConv 的提取特征过程进行均衡化, 在低层特征图上固定 3×3 标准卷积, 在高层特征图上固定可变形卷积进行偏移预测, 以缓解 FPN 的高层模糊效应, 缩减特征金字塔与高斯金字塔的差异。

(2) 横向尺度均衡特征金字塔

作为经典的特征金字塔结构, FPN 通过选择不同层次特征图进行上采样等操作, 以融合深层特征和浅层特征。然而, 由于舰船的特征区域多变, FPN 将图像划分为多个尺度的同时忽略尺度归一化, 致使目标尺度不匹配。此外, 特征融合降低特征图的清晰度, 多次融合淡化了尺度特征, 而稀疏特征不能提供足量信息。因此, 通过优化 FPN 的特征融合机制, 缓解信息丢失的问题, 为后续 ORPN 生成高质量候选框提供精准的尺度信息。

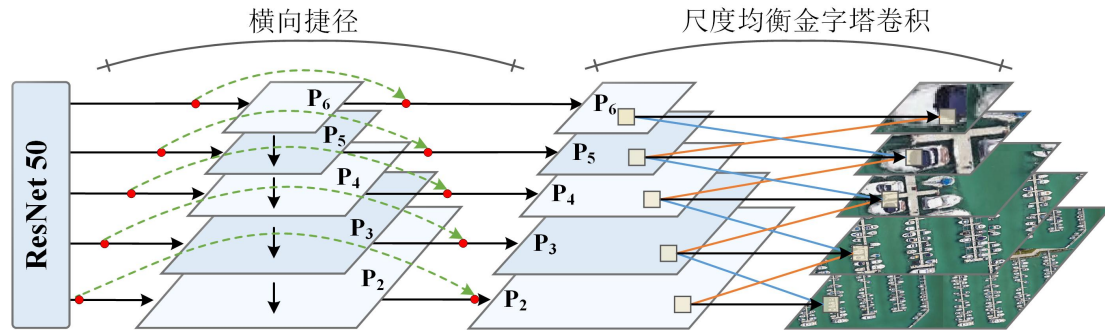


图 3-4 LSE-FPN 结构图

图 3-4 展示了 LSE-FPN 结构。一方面, LSE-FPN 添加横向捷径连接 (绿色虚线), 在高分辨率和低分辨率特征图之间建立连接, 使不同层级的特征信息顺畅流通, 增强特征表示; 另一方面, LSE-FPN 引入 SEPC 模块, 在捕获不同层级上下文信息的同时, 提取用于细粒度判别的尺度不变特征, 进一步处理输出特征, 为第一阶段的 ORPN 提供足量精确的特征信息。此外, SEPC 模块能同 FPN 结构进行多次堆叠, 考虑到计算开销因素, 经过独立模块的消融验证, 在仅堆叠一个 SEPC 模块时, 模型性能达到最优。

3.3.2 多通道融合网络

本节设计了一种多通道融合网络 MCFN, 通过切片分别结合低级、中级及高级特征图, 为候选框提取独立的细粒度舰船判别特征。在 FGSD 中, FPN 既为候选框的生成提供粗粒度特征, 又为候选框的细化提供细粒度特征, 导致子任务混淆。同时, 从 FPN 提取的特征未进行特征解耦, 不利于舰船细粒度检测。

此外，高级特征包含更多语义信息，低级特征对局部纹理特征更为敏感，预测阶段不仅需要语义信息进行精确框回归，而且需要高分辨率空间信息进行细粒度识别。尽管 FPN 已提供特征信息，但候选框的特征表示仍不够丰富。通过多通道跨层融合的操作，MCFN 可以有效解决上述问题。

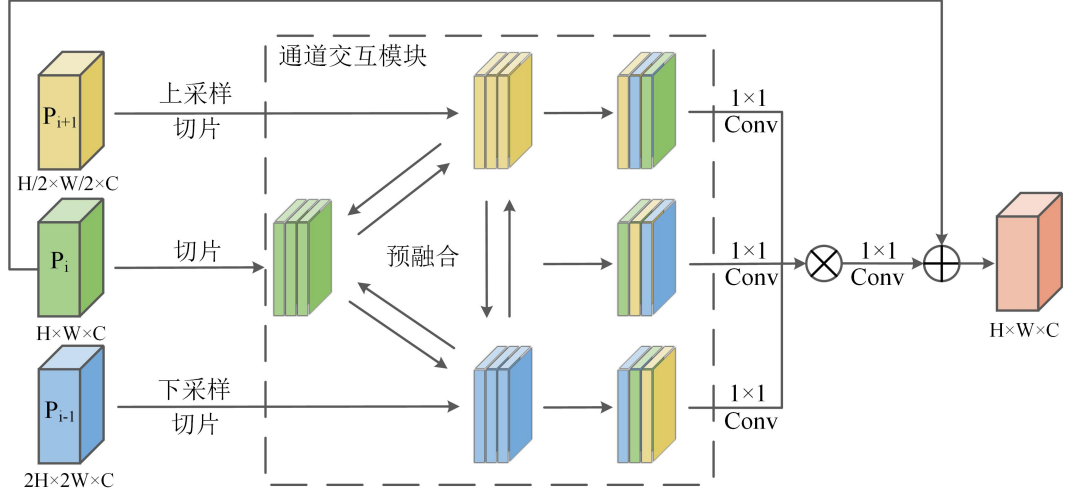


图 3-5 MCFN 结构图

如图 3-5 所示，首先，MCFN 将三组 FPN 相邻层级的特征图作为输入，并进行通道切片，以生成独立的增强跨层级特征，如式(3-3)所示。其中，假设有 L 个层级，当 $i=1$ 或 $i=L$ 时，式(3-3)的第一行和最后一行将被忽略：

$$\begin{aligned} [X_{11}, X_{12}, X_{13}] &= \text{Slicing}(\downarrow P_{i-1}) \\ [X_{21}, X_{22}, X_{23}] &= \text{Slicing}(P_i) \\ [X_{31}, X_{32}, X_{33}] &= \text{Slicing}(\uparrow P_{i+1}) \end{aligned} \quad (3-3)$$

式中， $P_{i-1} \in \mathbb{R}^{H/2 \times W/2 \times C}$ 、 $P_i \in \mathbb{R}^{H \times W \times C}$ 、 $P_{i+1} \in \mathbb{R}^{2H \times 2W \times C}$ 分别为低级、中级、高级特征图， H 、 W 、 C 分别为特征图的宽度、高度、通道数， $\text{Slicing}(\cdot)$ 代表在同一层级的通道切片操作， $\downarrow$ 、 $\uparrow$ 分别为邻近插值下采样、上采样操作。

随后，通道交叉互换模块(Channel Crossover Module, CCM)进行通道预融合，在不增加训练参数的情况下执行通道交换，其表达式如下所示：

$$\begin{aligned} X_L &= \text{Concat}([X_{11}, X_{23}, X_{32}]) \\ X_M &= \text{Concat}([X_{21}, X_{12}, X_{33}]) \\ X_H &= \text{Concat}([X_{31}, X_{22}, X_{13}]) \end{aligned} \quad (3-4)$$

式中， L 、 M 、 H 分别为低层、中层、高层特征， $\text{Concat}(\cdot)$ 为通道拼接操作。

其次，受双线性通道融合网络(Bilinear Channel Fusion Network, BCFN)^[52]的启发，基于 CCM 的输出，执行多通道融合(Multi-Channel Fusion, MCF)实现跨层空间和语义信息的结合，增强候选框的特征表示。然而，在 FGSD 任务中，二阶特征的计算需要占用大量计算资源。为解决此问题，通过结合 BCFN 与多模态低秩双线性池化(Multimodal Low-Rank Bilinear, MLB)网络^[54]，采用三个线性映射与两个 Hadamard 乘积，实现 MCF 的线性运算，如公式(3-5)所示：

$$MCF(X_L, X_M, X_H, l, m, h, f) = f(l(X_L) \odot m(X_M) \odot h(X_H)) \quad (3-5)$$

式中， $l(\cdot)$ 、 $m(\cdot)$ 、 $h(\cdot)$ 、 $f(\cdot)$ 分别是由 1×1 卷积实现的通道线性映射， $\odot$ 代表两个矩阵的 Hadamard 乘积。MCF 作为门控线性单元(Gated Linear Unit, GLU)^[55]的变体，使 MCFN 能在模型聚焦 RoI 之前，即在 RoI Pooling 前应用，为候选框预测生成独立的增强型空间特征，实现细粒度特征的精准判别。

最后，由于中层特征图可实现语义和空间信息的平衡，MCFN 将原始的中层特征图进行捷径拼接，得到线性融合特征映射 $O_i \in \mathbb{R}^{H \times W \times C}$ 。更具体地，将 LSE-FPN 的 $\{P_2, P_3, P_4, P_5, P_6, P_7\}$ 引入 MCFN，生成步长为 $\{4, 8, 16, 32, 64\}$ 的 $\{O_2, O_3, O_4, O_5, O_6\}$ ，以支持下一阶段的特征解耦和特征丰富。

3.3.3 定向动态预测网络

本节设计了一种定向动态预测网络 ODPN，在分类器与回归器中引入优化动态标签分配(Dynamic Label Assignment, DLA)和动态 SmoothL1 损失(Dynamic SmoothL1 Loss, DSL)，自适应拟合候选框分布变化，充分利用高质量候选框。

(1) 动态标签分配

在细粒度预测网络中，分类器通过固定标签分配策略预测正负样本，根据动态训练特性，候选框分布将随时间而不断变化。由于不同舰船类别的宽高比悬殊，固定阈值将大幅影响 IoU 分数，损害了细粒度识别的学习。因此，每类目标的最佳 IoU 阈值应不同。为解决上述问题，受到 Dynamic R-CNN^[56]的启发，引入 DLA 策略并在此基础上优化，以适应候选框分布变化。

为更好提取各舰船类别的最佳 IoU 阈值，自适应候选框的分布变化，根据固定标签分配策略，优化的 DLA 策略定义如下：

$$DLA = \begin{cases} 1, & \text{if } \max \text{IoU}(A, G) \geq T_{now} \\ 0, & \text{if } \max \text{IoU}(A, G) < T_{now} \end{cases} \quad (3-6)$$

式中， A 和 G 分别代表锚框(Anchor)和真实框(Ground-Truth, GT)，1 和 0 分别代表正样本和负样本， T_{now} 代表当前 IoU 阈值。

在训练中，DLA 根据候选框的统计数据自动更新 T_{now} 。考虑到各类候选框分布会随时间而变化，基于 DLA 优化 T_{now} 的更新方式，优化的 T_{now} 定义如下：

$$T_{now} = \begin{cases} \max(\text{mean}(S_I), T_{last}), & \text{mean}(S_I) > T_{ori} \\ T_{ori}, & \text{mean}(S_I) < T_{ori} \end{cases} \quad (3-7)$$

式中， T_{ori} 为初始 IoU 阈值， S_I 代表候选框与 GT 框的 IoU 矩阵 I 中第 K_I 个最大 IoU 值的集合， $\text{mean}(\cdot)$ 为均值， T_{last} 代表监督因子，为上一轮的 T_{now} 。

在最开始时， $T_{now} = T_{ori}$ 。随着训练推进， T_{now} 每经过 C 轮迭代更新，计算 S_I 的均值，以更新不同 IoU 阈值。与初始 DLA 相比，引入监督因子 T_{last} 进行阈值监督，确保每轮迭代的 IoU 均处于上升趋势，以增强训练鲁棒性。

(2) 动态 SmoothL1 损失

在细粒度预测网络中，回归器根据损失函数定位候选框，在训练期间，随着候选框的质量不断提高，固定归一化参数将导致模型对高质量候选框的关注度不足，不利于细粒度识别的训练。同时，随着 IoU 阈值的升高，正样本质量进一步提高，损失函数则进一步降低高质量候选框贡献。为解决上述问题，引入 DSL 策略并在此基础上优化，以充分利用高质量候选框。

为更好补偿高质量候选框的贡献，充分利用不同质量的候选框，动态调整回归 SmoothL1 损失，优化的 DSL 损失定义如下：

$$\text{DLS}(x, \beta_{now}) = \begin{cases} 0.5|x|^2 / \beta_{now}, & \text{if } |x| < \beta_{now} \\ |x| - 0.5\beta_{now}, & \text{otherwise} \end{cases} \quad (3-8)$$

式中， x 代表回归标签， β 代表控制 L1 损失函数软硬程度的超参数，固定的 β 值将使其贡献的梯度逐渐降低。 β_{now} 表示当前超参数值。

与 DLA 类似，DLS 根据反映定位精度的回归标签的统计信息以改变 β_{now} 值。考虑到候选框贡献值的动态变化，基于 DLS 优化 β_{now} 的更新方式：

$$\beta_{now} = \begin{cases} \min(\text{median}(S_E), \beta_{last}), & \text{median}(S_E) < \beta_{ori} \\ \beta_{ori}, & \text{median}(S_E) > \beta_{ori} \end{cases} \quad (3-9)$$

式中， β_{ori} 为初始超参数值， S_E 代表候选框与 GT 框的回归标签矩阵 E 中第 K_β 个最小 β 值的集合， $\text{median}(\cdot)$ 为中位值， β_{last} 代表监督因子，为上一轮 β_{now} 。

在最开始时, $\beta_{now} = \beta_{ori}$ 。类似地, 后续 β_{now} 每经过 C 轮迭代更新, 采用每个批次 S_E 的中位数促使梯度更新。与初始 DSL 相比, 引入监督因子 β_{last} 监督 β_{now} , 确保梯度保持下降趋势。针对 FGSD 任务, ODPN 能自适应候选框分布变化, 更关注高质量少样本的类别。算法 1 描述了设计的 ODPN, 其中引入三个超参数: IoU 阈值的 top-k K_I 、 β 的 top-k K_β 、更新迭代批次 C 。

算法 1 定向动态预测网络 ODPN

输入:

- 1: 候选框设置为 P , 真实框设置为 G
- 2: K_I 为 IoU 阈值的 top-k, K_β 为 β 的 top-k, C 为更新迭代批次

输出:

- 1: 初始化 IoU 阈值 T_{ori} 和 SmoothL1 的 β_{ori} , 分别设为 T_{now} 、 β_{now}
 - 2: 建立两个空集合 S_I 、 S_E , 以记录 IoU 值和回归标签值
 - 3: **循环** $i = 0$ 至最大迭代批次
 - 4: 根据 P 、 G , 获取匹配的 IoU 矩阵 I 和回归标签矩阵 E
 - 5: 根据 K_I 、 K_β , 选择阈值 I_k 、 E_k
 - 6: 记录对应值, 添加 I_k 至 S_I , 添加 E_k 至 S_E
 - 7: **如果** $i \% C == 0$
 - 8: 更新 IoU 阈值: $T_{now} = \max(\text{mean}(S_I), T_{last})$
 - 9: 更新 SmoothL1 β : $\beta_{now} = \min(\text{median}(S_E), \beta_{last})$
 - 10: 设置 $S_I = \emptyset$ 、 $S_E = \emptyset$
 - 11: 采用参数 T_{now} 、 β_{now} 训练网络
 - 12: 记录 T_{now} 、 β_{now} 作为下一轮迭代的 T_{last} 、 β_{last}
 - 13: **返回** 定向动态预测网络 ODPN
-

3.4 实验结果分析

3.4.1 实验数据集

ShipRSImageNet 数据集^[57]是一个大规模细粒度舰船检测数据集, 包含 3435 张图像和 17573 个舰船实例, 分辨率为 0.12m 至 6m, 图像大小从 930×930 至 1024×1024 像素, 舰船被划分为 50 个细粒度类别, 简称为 $C_1 \sim C_{50}$ 。训练集、验证集、测试集分别包含 2198 张、550 张、687 张图像。

FGSD2021 数据集^[58]是一个细粒度舰船检测数据集, 包含 636 张图像和 5274 个舰船实例, 图像宽度为 157 至 7789 像素, 高度为 224 至 6506 像素, 舰船被划分为 20 个类别, 训练集、测试集分别包含 424 张、212 张图像。

DOSR 数据集^[59]是一个用于细粒度舰船识别的数据集，包含 1066 张图像和 6127 个舰船实例，图像大小为 600 至 1300 像素，舰船被划分为 20 个细粒度类别，训练集、验证集、测试集分别包含 523 张、223 张、320 张图像。

MAR20 数据集^[60]是一个遥感军用细粒度飞机识别数据集，包含 3842 张图像和 22341 个飞机实例，图像大小约为 800×800 像素，飞机被划分为 20 个细粒度类别，训练集、测试集分别包含 1311 张、2511 张图像。

ShipRSImageNet 数据集、FGSD2021 数据集、DOSR 数据集被用于评估模型的有向 FGSD 性能，MAR20 数据集被用于评估模型的泛化性能。其中，主要的消融实验在规模最大的 FGSD 数据集 ShipRSImageNet 上实现。

3.4.2 实验配置与测评标准

本章实验采用的硬件配置为 NVIDIA GeForce RTX 3060 12G 显卡，实验平台是基于 PyTorch 1.11.0 的基准工具箱 MMRotate 0.3.4^[61]，批次大小为 2，迭代次数为 36 个 epoch，在第 24 和 33 个 epoch 的学习率降低 0.1 倍，选取动量为 0.9、权重衰减为 0.0001 的 SGD 作为优化器，初始学习率为 0.005。其中，不同数据集的参数设置有所不同，ShipRSImageNet 和 DOSR 的输入大小为 1024×1024 ，MAR20 的输入大小为 800×800 ，FGSD2021 图像像素过大，将其均裁剪成 1024×1024 ，重叠 200，输入大小为 1024×1024 。

本章选取 Pascal VOC2007 度量计算的平均准确率 AP 为主要评价指标，每个类别的 AP 是基于准确率和召回率计算的。其中，采用不同 IoU 阈值的 AP（AP₅₀、AP₇₅、AP_{50:95}）用于评估 FGSD 性能，Params、FLOPs 用于反映模型的复杂度，帧数(Frames Per Second, FPS)用于反映检测速率。

3.4.3 对比实验

(1) ShipRSImageNet 的定量实验和定性实验

为探索 PEDNet 具体舰船类别的检测性能，表 3-1 列出基线 Oriented R-CNN 和 PEDNet 在 50 个舰船类别(C₁~C₅₀)的精度指标 AP₅₀。与基线相比，PEDNet 更擅长于学习能力和检测能力的培养，在大部分舰船类别均表现较为突出。其中，PEDNet 在 C₁、C₂、C₃ 等 41 个舰船类别上优于基线，在 C₄、C₂₂、C₂₄ 等 4 个类别上同基线一致，在 C₆、C₁₃、C₂₆ 等 5 个类别上低于基线。此外，由于部分类别的样本数量过少，其检测精度较低，如 C₃₇、C₄₁、C₄₅ 等舰船类别。

表 3-1 具体舰船类别 AP₅₀ 的对比实验结果

方法	C ₁	C ₂	C ₃	C ₄	C ₅	C ₆	C ₇	C ₈	C ₉	C ₁₀	C ₁₁
基线方法	37.1	55.9	79.3	100	83.3	90.4	83.3	89.3	61.1	76.8	89.7
本章方法	37.9	56.6	88.2	100	86.5	89.9	90.9	90.2	65.1	83.7	90.3
	C ₁₂	C ₁₃	C ₁₄	C ₁₅	C ₁₆	C ₁₇	C ₁₈	C ₁₉	C ₂₀	C ₂₁	C ₂₂
基线方法	87.3	99.5	73.9	61.7	89.7	73.2	36.1	74.6	90.1	78.4	100
本章方法	87.9	97.3	68.4	66.5	90.5	80.7	42.4	87.0	90.9	89.4	100
	C ₂₃	C ₂₄	C ₂₅	C ₂₆	C ₂₇	C ₂₈	C ₂₉	C ₃₀	C ₃₁	C ₃₂	C ₃₃
基线方法	83.9	100	97.4	87.0	90.9	96.8	41.7	77.3	50.1	97.0	88.7
本章方法	86.5	100	100	86.6	90.9	98.7	46.4	82.1	51.5	98.5	90.9
	C ₃₄	C ₃₅	C ₃₆	C ₃₇	C ₃₈	C ₃₉	C ₄₀	C ₄₁	C ₄₂	C ₄₃	C ₄₄
基线方法	87.0	84.8	86.6	11.4	59.6	86.6	73.8	17.7	55.1	48.4	79.6
本章方法	86.6	85.9	89.5	14.5	67.2	90.7	69.6	20.7	58.7	48.8	79.9
	C ₄₅	C ₄₆	C ₄₇	C ₄₈	C ₄₉	C ₅₀	AP ₅₀	-	-	-	-
基线方法	17.7	39.2	63.8	62.0	21.0	59.9	71.50	-	-	-	-
本章方法	17.8	43.3	73.4	65.6	23.8	67.2	74.58	-	-	-	-

为探索 PEDNet 与其余 FGSD 检测器的性能差异, 表 3-2 采用 11 种主流 FGSD 检测器, 同 PEDNet 进行包括所有舰船类别的定性对比分析, 其中, 除 ConvNext、ReDet、LSKNet 外, 其余方法采用 ResNet50 作为主干网络。

表 3-2 ShipRSImageNet 的定性对比实验结果

方法	主干	AP _{50:95} (%)	AP ₅₀ (%)	AP ₇₅ (%)
单阶段检测器				
RetinaNet [6]	R-50-FPN	17.36	27.58	19.04
R ³ Det [24]	R-50-FPN	20.61	35.15	22.57
ConvNext [62]	Conv-T-FPN	24.10	36.89	28.34
S ² ANet [23]	R-50-FPN	32.47	55.14	35.04
FCOS [63]	R-50-FPN	34.45	50.70	40.50
双阶段检测器				
Faster R-CNN [16]	R-50-FPN	29.34	56.89	26.94
Gliding Vertex [64]	R-50-FPN	30.06	59.89	27.59
RoI Transformer [34]	R-50-FPN	32.67	61.64	29.40
ReDet [33]	ReR-50-FPN	38.10	67.02	37.11
LSKNet [65]	LSK-S-FPN	48.07	69.33	57.35
Oriented R-CNN [32]	R-50-FPN	51.28	71.50	62.25
PEDNet (本章方法)	R-50-L-FPN	56.25	74.58	68.34

与单阶段检测器相比, 多数双阶段检测器的结果更为优异, 而 PEDNet 的 AP₅₀、AP₇₅、AP_{50:95} 以 74.48%、68.34%、56.25% 的优势超越主流的有向目标检测器。在相同实验设置下, 与基线 Oriented R-CNN 相比, PEDNet 的 AP₅₀、AP₇₅、AP_{50:95} 分别提高 3.08%、6.09%、4.97%。与 ReDet 相比, PEDNet 的 AP₅₀、AP₇₅、AP_{50:95} 分别提高 7.56%、31.23%、18.15%, 即采用旋转不变网络的 ReDet 不适

用 FGSD。此外，PEDNet 优于 SOTA 模型 LSKNet， AP_{50} 、 AP_{75} 、 $AP_{50:95}$ 分别高出 5.25%、10.99%、8.18%，而 LSKNet 采用体量更大的 LSK-S 调整感受野。

为更直观展示 FGSD 性能，图 3-6 展示 ShipRSImageNet 的定量结果。由于缺乏充足语义信息，基线方法常出现误检和漏检，而本章方法既减少了假负例目标，又提升了高质量样本的权重，漏检误检得以改善，利于细粒度舰船识别。此外，PEDNet 加强了定位舰船能力，进一步证明其 FGSD 性能。

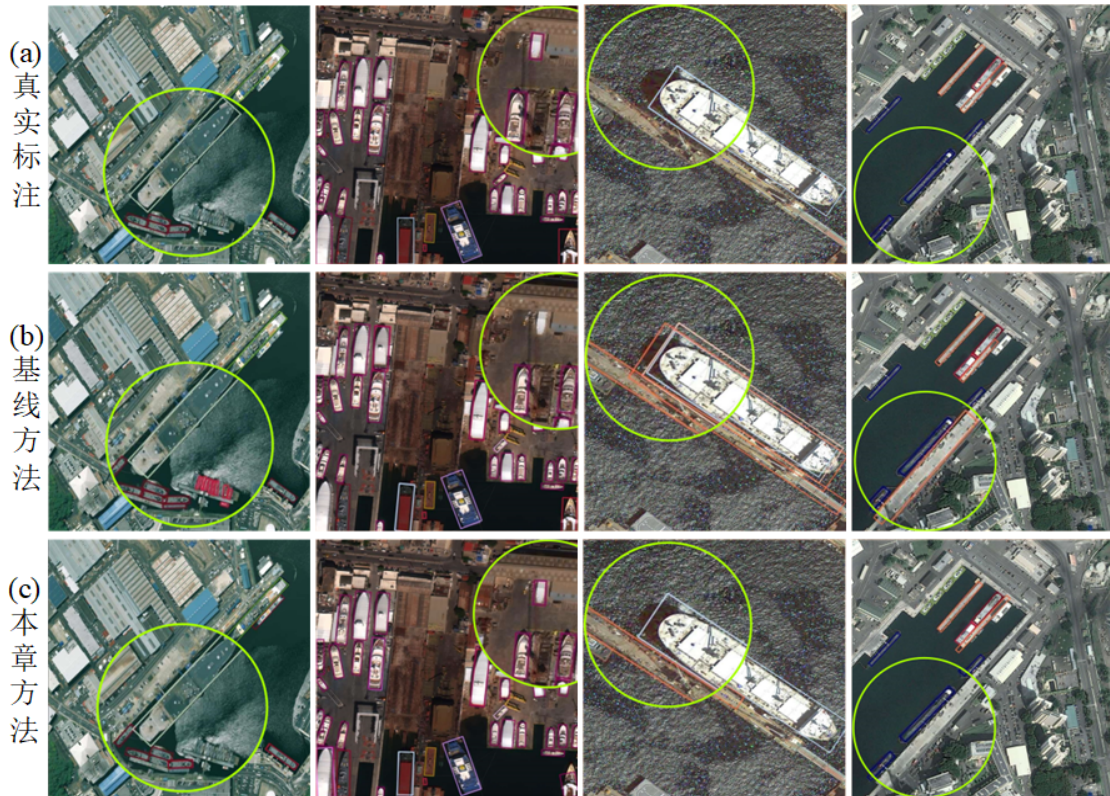


图 3-6 ShipRSImageNet 的定量对比图

(2) FGSD2021 的定量实验和定性实验

为进一步验证 PEDNet 在 FGSD 上的有效性，对于 FGSD2021，采用 11 种主流 FGSD 检测器，同 PEDNet 进行所有类别的定性比较分析，如表 3-3 所示。

与 ShipRSImageNet 相比，FGSD2021 所包含的 20 个舰船类别使其细粒度识别难度相对较低，而 PEDNet 同样在 FGSD2021 达到最先进的水平。尽管单阶段的 S^2ANet 和 FCOS 表现优异，但由于缺乏候选框信息，双阶段的 LSKNet 和 Oriented R-CNN 的性能更突出。然而，PEDNet 的 AP_{50} 、 AP_{75} 、 $AP_{50:95}$ 以 92.50%、86.26%、68.78% 的优势超过所有模型。与基线相比， AP_{50} 、 AP_{75} 、 $AP_{50:95}$ 提高 2.44%、4.35%、2.90%，体现本章方法在 FGSD 任务上优越性。

表 3-3 FGSD2021 的定性对比实验结果

方法	主干	AP _{50:95} (%)	AP ₅₀ (%)	AP ₇₅ (%)
单阶段检测器				
RetinaNet [6]	R-50-FPN	36.37	56.23	41.64
R ³ Det [24]	R-50-FPN	34.97	56.73	39.29
ConvNext [62]	Conv-T-FPN	45.81	67.16	54.15
S ² ANet [23]	R-50-FPN	54.12	84.23	63.93
FCOS [63]	R-50-FPN	55.39	79.85	67.61
双阶段检测器				
Faster R-CNN [16]	R-50-FPN	44.43	80.34	42.90
Gliding Vertex [64]	R-50-FPN	45.20	80.69	44.47
RoI Transformer [34]	R-50-FPN	53.29	86.56	58.70
ReDet [33]	ReR-50-FPN	51.58	87.52	54.46
LSKNet [65]	LSK-S-FPN	65.09	91.19	81.49
Oriented R-CNN [32]	R-50-FPN	65.88	90.06	81.91
PEDNet (本章方法)	R-50-L-FPN	68.78	92.50	86.26

图 3-7 展示了 FGSD2021 的定量结果。针对漏检，基线漏检率较高，而 LSE-FPN 使模型提取充分语义信息，漏检得到改善；针对误检，MCFN 为候选框提供独立判定特征，使模型对舰船类别的判定更准确。此外，ODCN 使模型预测比基线更接近真实框的细粒度类别，有力地证明本章方法的有效性。

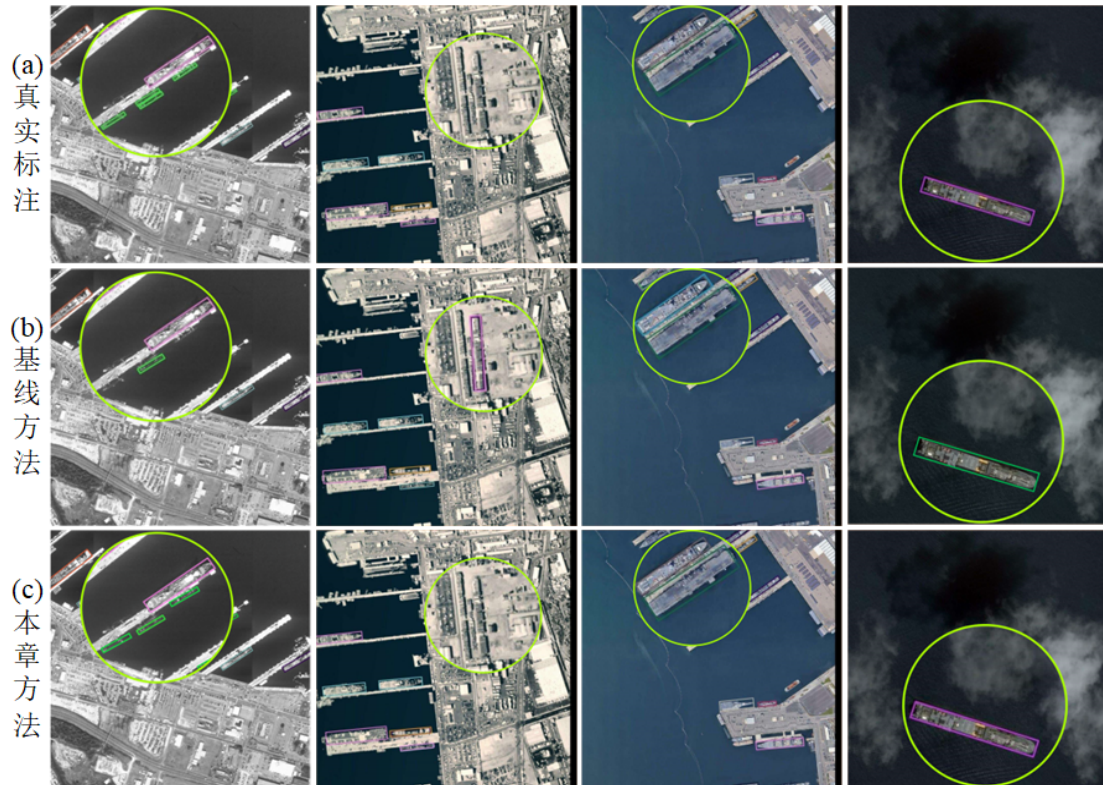


图 3-7 FGSD2021 的定量对比图

(3) DOSR 的定量实验和定性实验

为验证 PEDNet 在复杂 FGSD 任务上的性能，在具有类间不平衡和类间相似性的 DOSR 上，同 9 种 FGSD 检测器进行相似舰船定性分析，如表 3-4 所示。

表 3-4 DOSR 的定性对比实验结果

方法	主干	AP _{50:95} (%)	AP ₅₀ (%)	AP ₇₅ (%)
单阶段检测器				
ConvNext [62]	Conv-T-FPN	12.12	21.65	11.55
S ² ANet [23]	R-50-FPN	16.53	30.74	15.81
FCOS [63]	R-50-FPN	18.39	29.45	20.62
双阶段检测器				
Faster R-CNN [16]	R-50-FPN	20.78	44.20	15.81
Gliding Vertex [64]	R-50-FPN	20.20	44.32	15.65
RoI Transformer [34]	R-50-FPN	27.38	56.97	21.74
LSKNet [65]	LSK-S-FPN	31.06	51.95	35.20
Oriented R-CNN [32]	R-50-FPN	37.08	58.52	42.59
PEDNet (本章方法)	R-50-L-FPN	40.17	61.28	46.19

单阶段的 ConvNext、FCOS 等直接生成候选框的方式，缺少足量判别特征，导致检测结果较差，而 PEDNet 在遇到类间不平衡与相似问题时，仍表现出较好性能。与基线相比，PEDNet 的 AP₅₀、AP₇₅、AP_{50:95} 提高 2.76%、3.60%、3.09%。此外，双阶段方法由于缺乏增强策略，类间相似性导致检测效果不佳。

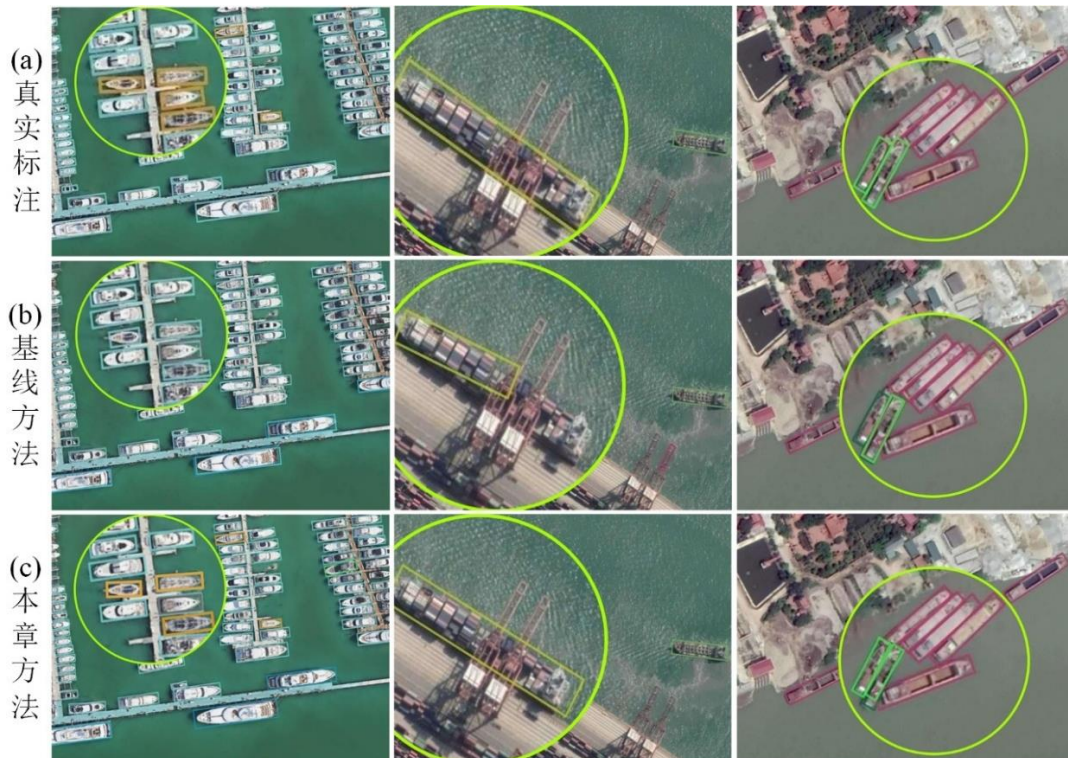


图 3-8 DOSR 的定量对比图

图 3-8 展示了 DOSR 的定量结果。针对密集和小型舰船，本章方法既降低了在密集场景下的舰船目标漏检率，又提高了小型舰船目标的检测效果。根据上述三组定量实验结果，验证了候选框增强策略的有效性。

(4) MAR20 的定量实验和定性实验

为评估 PEDNet 在其余细粒度目标类别的检测性能，采用包含 20 种飞机类别的 MAR20 数据集进行泛化性测试，如表 3-5 所示。

表 3-5 MAR20 的定性对比实验结果

方法	主干	AP _{50:95} (%)	AP ₅₀ (%)	AP ₇₅ (%)
单阶段检测器				
RetinaNet [6]	R-50-FPN	42.37	69.63	45.54
ConvNext [62]	Conv-T-FPN	49.13	77.09	53.26
S ² ANet [23]	R-50-FPN	46.48	76.43	49.80
FCOS [63]	R-50-FPN	38.75	62.14	45.32
双阶段检测器				
Faster R-CNN [16]	R-50-FPN	47.87	75.81	53.14
Gliding Vertex [64]	R-50-FPN	37.05	73.61	31.84
RoI Transformer [34]	R-50-FPN	50.93	75.39	61.04
LSKNet [65]	LSK-S-FPN	53.68	79.88	67.11
Oriented R-CNN [32]	R-50-FPN	52.69	77.18	65.55
PEDNet (本章方法)	R-50-L-FPN	55.54	80.84	69.18

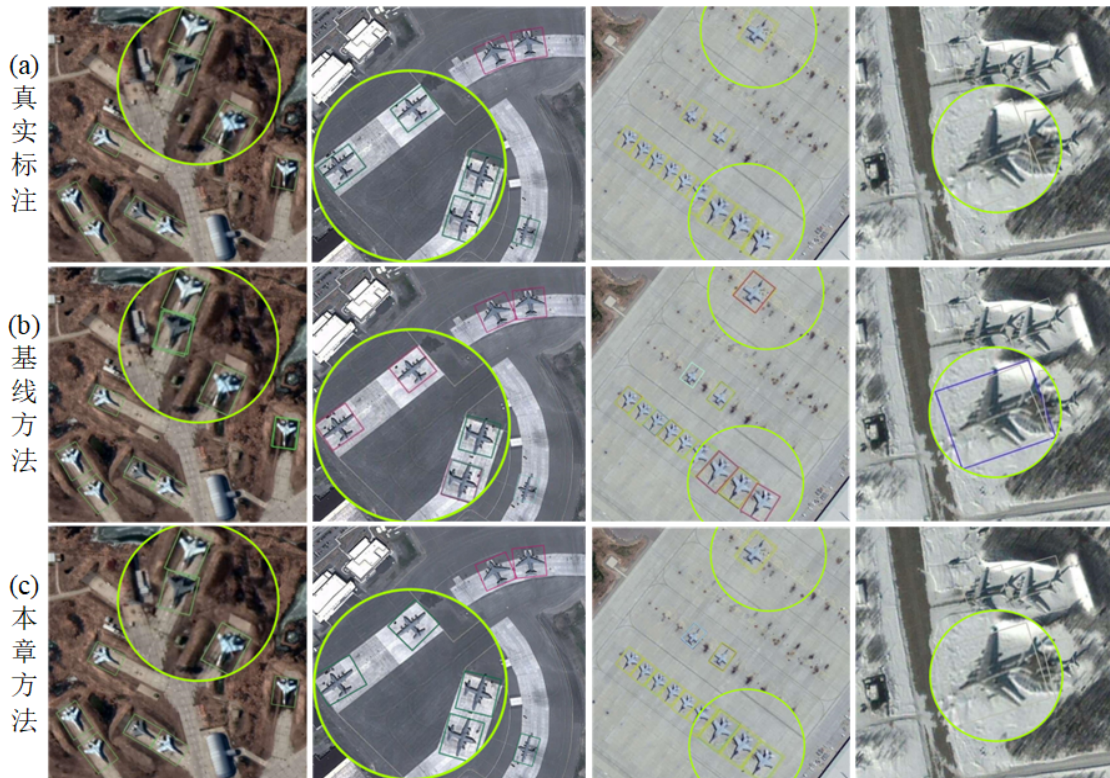


图 3-9 MAR20 的定性对比图

ConvNext 采用更深的主干，在单阶段方法中表现出色，而 PEDNet 比 ConvNext 更优异。与 LSKNet 相比， AP_{50} 、 AP_{75} 、 $AP_{50:95}$ 提高了 0.96%、2.07%、1.86%。与基线方法相比，PEDNet 的 AP_{50} 、 AP_{75} 、 $AP_{50:95}$ 提高了 3.66%、3.63%、2.85%，达到 80.84%、69.18%、55.54%。因此，本章方法具有良好泛化性。

图 3-9 展示了 MAR20 的定性结果。针对多种飞机类别，尽管 PEDNet 对个别飞机出现误检情况，但其在判断类别、定位目标方面均优于基线。

3.4.4 消融实验

(1) 各模块有效性的消融实验

为彻底分析三个模块的有效性，执行模块消融实验，如表 3-6 所示，模块 LSE-FPN、MCFN、ODPN 对各数据集的性能均是有益的。

表 3-6 各模块有效性的消融实验结果

数据集	LSE-FPN	MCFN	ODPN	$AP_{50:95}(\%)$	$AP_{50}(\%)$	$AP_{75}(\%)$
ShipRS ImageNet	√	√	√	51.28	71.50	62.25
				52.82	71.64	65.18
				54.41	73.59	66.17
				56.25	74.58	68.34
FGSD2021	√	√	√	65.88	90.06	81.91
				67.17	90.43	84.58
				67.86	91.55	84.65
				68.78	92.50	86.26
DOSR	√	√	√	37.08	58.52	42.59
				37.79	59.98	44.63
				38.91	60.66	45.25
				40.17	61.28	46.19
MAR20	√	√	√	52.69	77.18	65.55
				53.64	78.19	66.80
				54.11	79.24	66.88
				55.54	80.84	69.18

在三个舰船数据集上，LSE-FPN 为候选框生成提取多尺度特征，使 AP_{50} 提高 0.14%、0.37%、1.46%， AP_{75} 提高 2.93%、2.67%、2.04%， $AP_{50:95}$ 提高 1.54%、1.29%、0.71%；MCFN 为候选框表示提供判别特征，使 AP_{50} 提高 1.95%、1.12%、0.68%， $AP_{50:95}$ 提高 1.59%、0.69%、1.12%；ODPN 为候选框利用拟合分布变化，使 AP_{50} 提高 0.99%、0.95%、0.62%， AP_{75} 提高 2.17%、1.61%、0.94%， $AP_{50:95}$ 提高 1.84%、0.92%、1.26%。在 MAR20 上，各模块使 AP_{50} 提高 1.01%、1.05%、1.60%，使 AP_{75} 提高 1.25%、0.08%、2.30%， $AP_{50:95}$ 提高 0.95%、0.47%、1.43%。

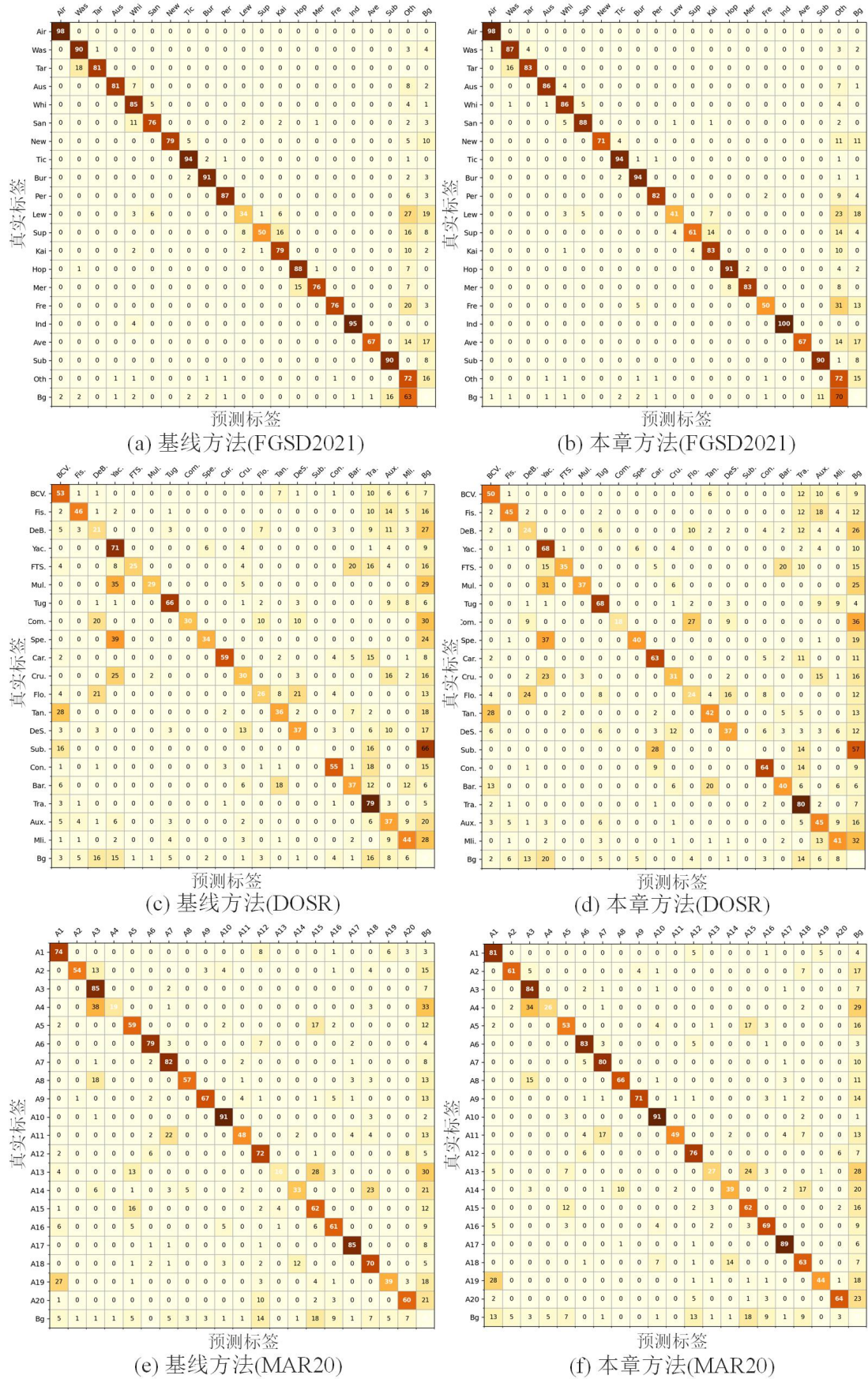


图 3-10 检测混淆矩阵的对比图

图 3-10 展示了 FGDS2021、DOSR、MAR20 测试的混淆矩阵。与基线相比，PEDNet 在多数细粒度类别的性能优于基线，但在个别类别中表现较差，这是由于 AP 综合考虑 Pre 和 Re，而混淆矩阵只关注 Pre，忽略了置信度。因此，本章方法倾向在置信度较低的情况下，做出更多的假负例预测，导致真正例的比例降低。在实际应用中，通过动态阈值调整可消除低置信度的预测。

表 3-7 复杂度和检测速率的对比实验结果

方法	FLOPs(G)	Params(M)	FPS(S)
基线方法	198.6	43.2	9.4
本章方法	203.5	44.3	8.9

表 3-7 展示了基线方法和 PEDNet 的复杂度和检测速率。由于引入了 LSE-FPN 模块和 MCFN 模块，本章方法在增加少量 Params、FLOPs 的复杂度成本和牺牲少量 FPS 的检测速率成本下，实现了更可观的精度增益。

(2) LSE-FPN 模块的消融实验

在 LSE-FPN 中，分别引入 SEPC 模块以增强多尺度特征融合能力、设计横向捷径以促进不同层级特征信息流通。为探索上述两种方式在 LES-FPN 中的有效性，表 3-8 展示独立模块有效性的消融实验，其中，不采用 MCFN 和 ODPN 模块。由表可知，单独应用 SEPC 模块和横向捷径方式均可以提高模型的 FGSD 能力，而结合应用 SEPC 模块和横向捷径方式时，增益达到最大化，有效证明 LSE-FPN 的两个模块为候选框生成提供良好的特征基础。

表 3-8 SEPC 和横向捷径的消融实验结果

SEPC (N=1)	横向捷径	AP _{50:95} (%)	AP ₅₀ (%)	AP ₇₅ (%)
		51.28	71.50	62.25
√		52.26	71.43	64.32
	√	51.79	71.28	63.53
√	√	52.82	71.64	65.18

表 3-9 SEPC 堆叠数量的消融实验结果

数量	AP _{50:95} (%)	AP ₅₀ (%)	AP ₇₅ (%)	FLOPs(G)	Params(M)
N=1	52.82	71.64	65.18	200.63	44.33
N=2	51.97	71.36	64.34	202.68	46.23
N=3	51.75	70.59	62.90	204.73	48.12
N=4	50.34	69.90	60.97	206.79	50.02

为探索 SEPC 的重要性，考虑到堆叠的 SEPC 数量同时关联检测精度与计算资源，在应用横向捷径的基础上，对 SEPC 的堆叠数量进行研究，如表 3-9 所示。随着 SEPC 的堆叠，计算量与参数量逐渐增加，但是特征融合层次的过度加深，导致尺度特征融合能力下降， AP_{50} 、 AP_{75} 、 $AP_{50:95}$ 依次降低。因此，堆叠数量过多不利于 FGSD 任务，在仅堆叠一个 SEPC 模块时，精度收益达到最优。

(3) ODPN 模块的消融实验

在 ODPN 中，在分类器和回归器上分别设计了带有监督因子 T_{last} 和 β_{last} 的 DLA 策略与 DSL 策略。为评估监督因子 T_{last} 和 β_{last} 的有效性，表 3-10 列出监督因子的消融实验，其中，均采用 LSE-FPN 和 MCFN 模块。与没有监督因子相比，DLA 应用 T_{last} 增强了对舰船识别能力，DSL 应用 β_{last} 提高了对舰船定位能力，在 DLA 和 DSL 中同时应用 T_{last} 和 β_{last} ，模型的鲁棒性达到最高。

表 3-10 监督因子的消融实验结果

T_{last}	β_{last}	$AP_{50:95}(\%)$	$AP_{50}(\%)$	$AP_{75}(\%)$
		54.89	72.85	66.07
√		55.57	73.87	67.38
	√	55.49	73.92	67.27
√	√	56.25	74.58	68.34

表 3-11 超参数的消融实验结果

K_I	K_β	$AP_{50:95}(\%)$	$AP_{50}(\%)$	$AP_{75}(\%)$
70	12	55.74	73.84	66.65
70	10	55.89	73.98	67.23
70	8	56.25	74.58	68.34
75	12	55.10	73.08	66.89
75	10	55.29	73.25	67.27
75	8	55.66	73.28	67.61
80	12	54.54	72.81	66.63
80	10	54.78	73.27	66.19
80	8	55.31	73.83	66.53

DLA 策略和 DSL 策略分别包含 K_I 、 K_β 、 C 三个超参数，在 Dynamic R-CNN 中，迭代 C 的最优数值为 100， K_I 的最优数值为 75， K_β 的最优数值为 10。然而，由于引入监督因子监督训练过程，需采用更小 K_I 、 K_β 适应候选框分布。为探索超参数 K_I 和 K_β 最优取值，采用区间 (70,80) 的 K_I 和区间 (8,12) 的 K_β 进行消融实验，如表 3-11 所示。无论 K_I 和 K_β 取值如何，DLA 和 DSL 的优化均在

AP 上获得收益。以 $K_I = 70$ 为例，增大 K_β 将导致 AP_{50} 、 AP_{75} 、 $AP_{50:95}$ 逐渐减小；以 $K_\beta = 8$ 为例，增大 K_I 将导致 AP_{50} 、 AP_{75} 、 $AP_{50:95}$ 逐渐减小。因此，较小的 K_I 和 K_β 利于候选框利用，当 $K_I = 70$ 、 $K_\beta = 8$ 及 $C = 100$ 时，所获收益达最优。

图 3-11 列出基线和 PEDNet 在训练间的分类损失和回归损失曲线。在应用 DLA 和 DSL 后，与基线相比，PEDNet 的分类与回归损失均能更好地收敛。

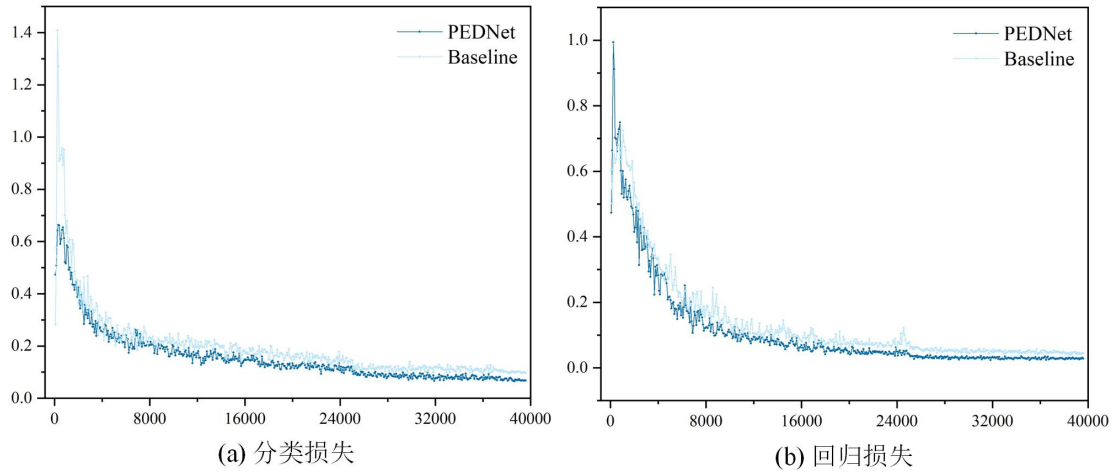


图 3-11 训练损失曲线的对比图

3.5 本章小结

为解决光学遥感图像中有向细粒度舰船检测的特征混淆、误检漏检的问题，本章通过多任务学习方式以增强候选框，研究一种改进的有向 FGSD 模型 PEDNet。其中，模型的主要贡献在三个方面：①提高候选框的生成质量（LSE-FPN 模块）；②为候选框提供多通道判别特征（MCFN 模块）；③动态拟合候选框分布，聚焦高质量候选框的利用（ODPN 模块）。实验结果表明，在四个通用数据集上执行消融实验和对比实验，与其他主流的有向检测器相比，本章方法 PEDNet 的 FGSD 性能更为有效且更具泛化性。

第 4 章 基于深度学习的无向监督有向舰船检测

4.1 引言

深度卷积神经网络的进步极大地推动了光学遥感舰船检测的发展。面向复杂海洋场景，有向舰船检测受到广泛关注，但其全监督方法的性能在很大程度上取决于训练数据集，与 HBox 相比，尽管 RBox 检测的应用范围更为广泛，但是 RBox 数据集的缺乏限制了有向舰船检测的进一步发展。

弱监督学习是一类强大的学习范式，其不需要精确标注数据即可从海量数据中学习。在目标检测中，弱监督学习在训练阶段不依赖精确的标注信息，便可准确定位目标；自监督学习通过利用数据本身的结构和特征训练模型，在训练阶段完全脱离标签注释，实现精确目标识别。因此，通过结合弱监督学习和自监督学习的思想，无向监督有向的目标检测方法应运而生，以降低 RBox 的标注成本和摆脱数据集限制。然而，现有无向监督有向的目标检测方法并不成熟，在检测长宽比悬殊、尺度变化大的舰船目标时准确率较低。

为解决上述问题，本章探究了一种无向监督有向的舰船检测方法，通过设计动态尺度特征增强模块、集成神经架构搜索的头部结构、设计外切 Scylla 交并比损失，解决无向监督有向的遥感舰船检测效果不佳的问题。

4.2 无向监督有向舰船检测模型架构

H2RBox^[41]通过弱监督(Weakly-Supervised, WS)分支利用 HBox 注释进行训练与预测，自监督(Self-Supervised, SS)分支利用输入图像的两个增强视图学习视图间 RBox 预测的一致性，通过结合 WS 分支和 SS 分支，实现了端到端 HBox 输入和 RBox 输出。H2RBox-v2^[42]则在 H2RBox 的基础上，采用对称感知学习的方式，实现更高效的无向监督有向的目标检测，如图 4-1 所示。

首先，输入图像至翻转视图和旋转视图进行数据增强后生成三个视图分支，并输入主干进行特征提取，FPN 随后通过特征融合进一步提取舰船特征；其次，在 SS 分支中，采用对称性自监督学习的思想，通过翻转和旋转一致性，计算对称一致性损失不断学习并生成视图间的角度信息，在 WS 分支中，根据 HBox 注释，分类子网络计算类别损失与中心性损失，回归子网络通过提取位置信息与尺度信息，结合 SS 分支的预测角度信息，生成粗略预测 RBox；最后，在 WS

分支中, 粗略预测 RBox 根据外切交并比(Circumscribed Intersection over Union, CircumIoU)损失, 同旋转视图的真实 RBox 标注计算圆周损耗, 经微调后生成精细预测 RBox, 并结合分类子网络的信息获取预测输出。

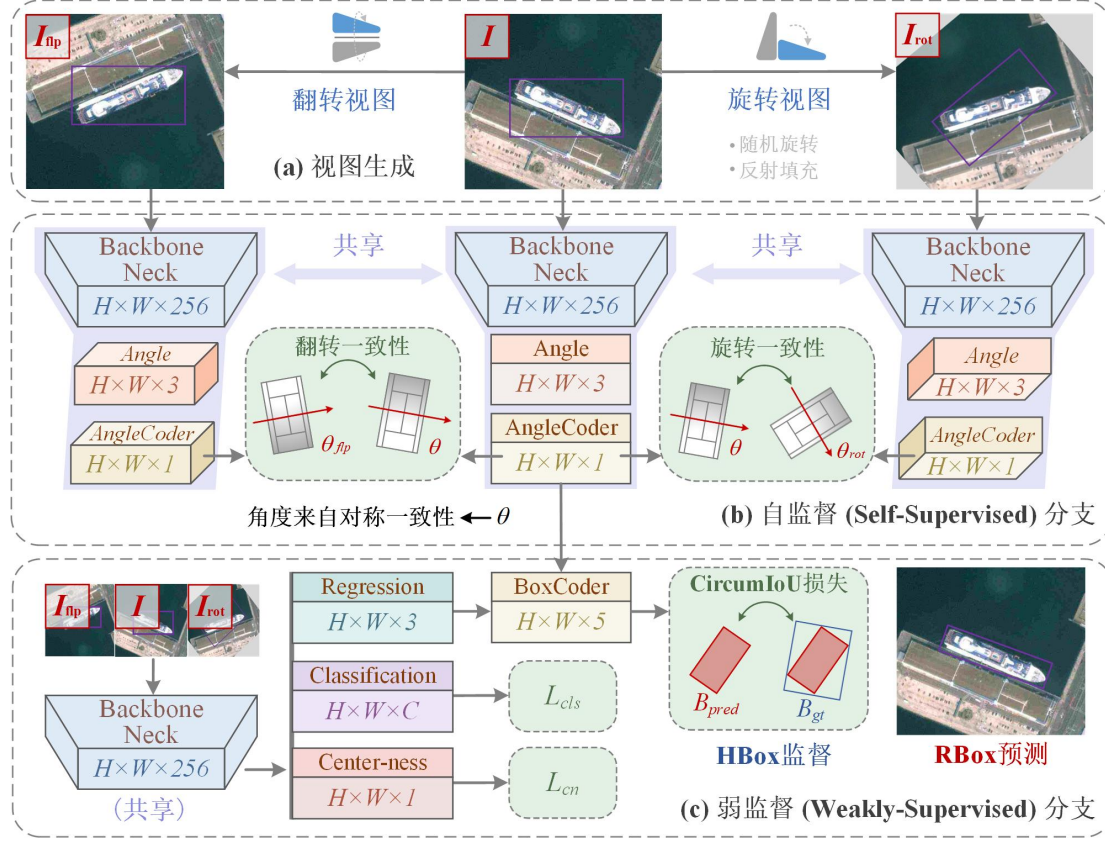


图 4-1 基于 H2RBox-v2 的无向监督有向舰船检测模型图

4.3 基于动态尺度增强的无向监督有向舰船检测模型

在 H2RBox-v2 中, 由于舰船存在不同尺度变化, 传统 FPN 通过多级特征提取以适应不同舰船尺寸的方式, 迫使单级特征跨越大范围的目标尺寸, 导致 SS 分支和 WS 分支获取不精确的角度信息和位置信息; 其次, 原始头部结构通过四个标准卷积层进行特征强化, 不能自适应调整各层级权重, 对输入的特征利用率不高, 且部署至双分支后消耗过多资源; 最后, WS 分支的 IoU 损失在计算旋转视图的真实框与预测框的 IoU 时, 忽略真实框和预测框的角度与形状差异, 导致模型无法正确判别预测框位置, 该情况在有向舰船检测时更严重。

针对上述问题, 选取 H2RBox-v2 作为基线, 探究一种基于动态尺度增强的无向监督有向舰船检测网络(Dynamic Scale Enhancement Network, DSENet), 以实现在 HBox 受限情况下的舰船检测, 如图 4-2 所示。

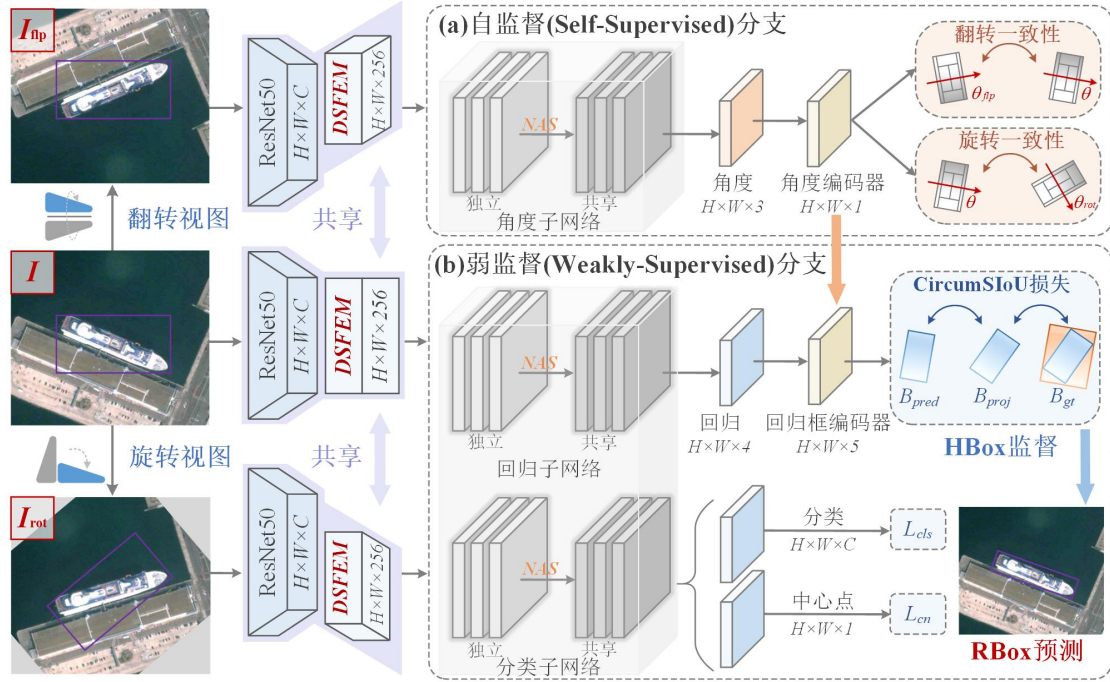


图 4-2 基于动态尺度增强的无向监督有向舰船检测网络图

首先，输入图像至翻转视图和旋转视图进行数据增强后生成三个视图分支，并输入主干 ResNet50 进行特征提取以提取尺度不变特征。随后，动态尺度特征增强模块(Dynamic Scale Feature Enhancement Module, DSFEM)提取、融合与自适应增强多尺度特征，模拟 FPN 自上而下的特征层级融合，并将输出特征映射至 SS 分支和 WS 分支；其次，在 SS 分支中，引入神经架构搜索(Neural Architecture Search, NAS)策略优化角度子网络，并通过翻转视图和旋转视图间的对称性以学习角度方向，经角度编码器编码计算对称一致性损失，不断学习和生成预测的角度信息，在 WS 分支中，分类子网络经 NAS 策略优化，根据 HBox 注释分别计算类别损失和中心性损失，回归子网络经 NAS 策略优化，提取位置和尺度信息，并结合 SS 分支的角度信息，经过回归框编码器生成粗略预测 RBox；最后，在 WS 分支中，粗略的预测 RBox 根据外切 Scylla 交并比(Circumscribed Scylla Intersection over Union, CircumSIOU)损失，同旋转视图的真实 RBox 标注计算圆周损耗，生成精细的预测 RBox，并结合分类信息得到最终输出。

4.3.1 动态尺度特征增强模块

在无向监督有向目标检测中，舰船的尺度变化给特征提取和融合带来挑战。多尺度特征提取是在 FPN 应用不同大小卷积核的同时在不同感受野中提取特征，但卷积核堆叠易造成计算冗余。多尺度特征融合则通过捕捉不同尺度间的特征

信息，提高模型对目标的理解能力。然而，不同尺度分支对预测网络的贡献是相等的，但不同舰船的尺度信息重要性并不一致，这不利于后续的学习和预测。为解决上述问题，如图 4-3 所示，设计一种动态尺度特征增强 DSFEM 模块，以替换原始 FPN，为 SS 分支的角度学习和 WS 分支的目标预测提供特征保障。

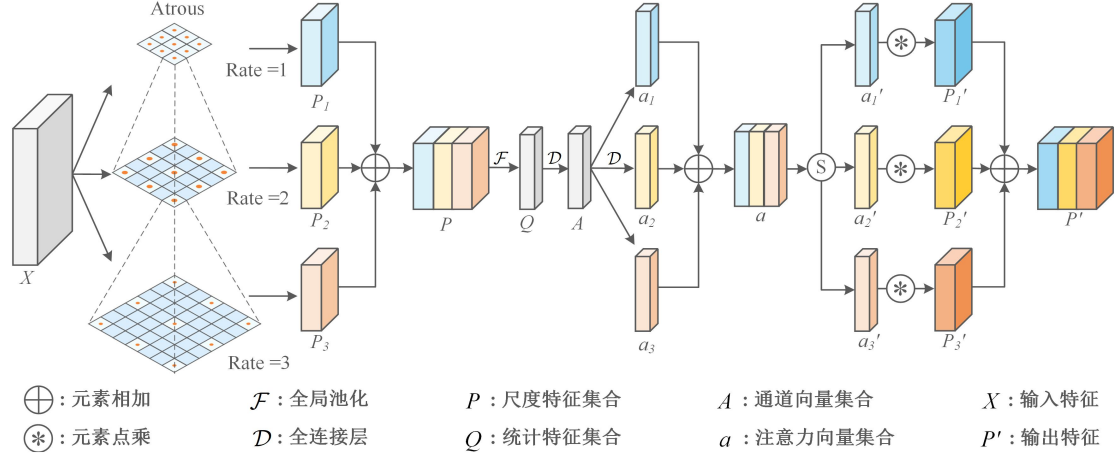


图 4-3 DSFEM 的结构图

首先，对于给定输入 $X \in \mathbb{R}^{H \times W \times C}$ ，利用具有膨胀率的空洞卷积^[66]对 X 执行变换，生成三个分支尺度信息 $P_1 \in \mathbb{R}^{H' \times W' \times C'}$ 、 $P_2 \in \mathbb{R}^{H' \times W' \times C'}$ 、 $P_3 \in \mathbb{R}^{H' \times W' \times C'}$ ，对于第 i 个分支，为获取不同尺度信息，生成分支的操作如下：

$$P_i = \mathcal{K}_i(X) \quad (4-1)$$

式中， $\mathcal{K}_i(\cdot)$ 表示空洞卷积，膨胀率分别为 1、2、3。应用空洞卷积避免计算冗余，实现轻量级多尺度特征提取。为使目标提取特征保持一致，空洞卷积核保持参数共享。随后，采用元素相加生成尺度特征集合 $P = [P_1, P_2, P_3]$ ，整合多尺度信息至同一层级，以获取不同分支信息的重要性。

其次，为避免 P 作为直接输入使 SS 分支和 WS 分支的贡献值相等，受 SKNet^[67] 的启发，添加注意力动态调整各分支的通道贡献，实现尺度特征动态加权。具体地，沿着 H' 和 W' 对 P 应用全局平均池化，获得统计特征 $Q \in \mathbb{R}^{1 \times 1 \times 3C'}$ ，实现全局信息的统计分析。对于第 K 个通道，全局平均池化的操作如下：

$$Q_k = \mathcal{F}(P_k) = \frac{1}{H' \times W'} \sum_{i=1}^{H'} \sum_{j=1}^{W'} P_k(i, j) \quad (4-2)$$

式中， H' 代表特征图的高度， W' 代表特征图的宽度， $\mathcal{F}(\cdot)$ 表示全局平均池化， P_k 代表尺度特征集合 P 中第 K 个通道的统计信息。

为建立通道关联的同时实现通道压缩，对 Q 使用全连接层转换成通道向量 $A \in \mathbb{R}^{1 \times 1 \times d}$ ，为后续提供自适应选择，全连接层转换可简述为：

$$A = \mathcal{D}(Q) = W_Q \cdot Q + B_Q \quad (4-3)$$

式中， $\mathcal{D}(\cdot)$ 为全连接层转换， W_Q 为权重矩阵， B_Q 为偏差，符号 $\cdot$ 为逐元素相乘。

由于分支间通道信息相对独立，对 A 同样使用全连接层转换成独立注意力向量 a_1 、 a_2 、 a_3 。对于第 i 个分支，执行全连接层转换如下：

$$a_i = \mathcal{D}(A) = W_i \cdot A + B_i \quad (4-4)$$

式中， W_i 表示对应分支的权重矩阵， B_i 表示对应分支的偏差。通过上述的转化，通道信息在实现分支间独立的同时又相互交织。

随后，通过元素相加将三个分支的注意力向量合并得到 $a = [a_1, a_2, a_3]$ ，并沿维度 $1 \times 1 \times 3d$ 对其执行 Softmax 操作，以导出各分支的加权分数：

$$[a'_1, a'_2, a'_3] = \text{softmax}(a) = \left[\frac{e^{a_1}}{e^{a_1} + e^{a_2} + e^{a_3}}, \frac{e^{a_2}}{e^{a_1} + e^{a_2} + e^{a_3}}, \frac{e^{a_3}}{e^{a_1} + e^{a_2} + e^{a_3}} \right] \quad (4-5)$$

此外，通过逐元素点乘获取各分支加权特征，对于第 i 个分支，加权操作如下：

$$P'_i = P_i \cdot a'_i \quad (4-6)$$

最后，合并各分支加权特征得到输出特征 $P' = [P'_1, P'_2, P'_3]$ ，采用 1×1 卷积转换维度，并模拟 FPN 自上而下的层级间特征融合。经特征扩张提取、动态选择、加权融合，有效利用全局通道与尺度信息，实现动态尺度特征增强。

4.3.2 神经架构搜索头部

在 WS 分支的回归子网络与分类子网络中，头部采样池将 DSFEM 的输出特征映射至相应层级，执行目标定位和类别判定。在 SS 分支的角度子网络中，头部采样池则进行角度方向学习。然而，原始头部结构利用四个 3×3 标准卷积的固定拓扑结构，易导致目标预测或角度学习的性能陷入局部最优。

为探索头部的潜力并拓展顺序搜索空间以搜寻最佳头部结构，受到 NAS-FCOS^[68] 的启发，利用随机搜索策略，在 SS 分支和 WS 分支的三个子网络中集成 NAS-Head 模块。如图 4-4 所示，NAS-Head 由 NAS 搜索的头部采样池和预测输出组成。其中，头部采样池由独立搜索和共享搜索两部分构成。

首先，为简化搜索空间，需要在头部的不同层级间共享权重，将头部采样池结构所有批归一化层(Batch Normalization, BN)替换为组归一化层(Group Normalization, GN)；其次，为增加头部灵活性并了解权重共享效果，添加索引 i 作为头部共享权重的起始位。对于索引值 i 前的每一层，NAS-Head 模块为 DSFEM 的各输出层级创建独立的权重集，对于索引值 i 后的每一层，则使用全局权重进行共享；最后，为平衡 DSFEM 模块中各分支的工作量，共享搜索部分进行长度自适应，实现特定层级特征和跨层级特征的提取。

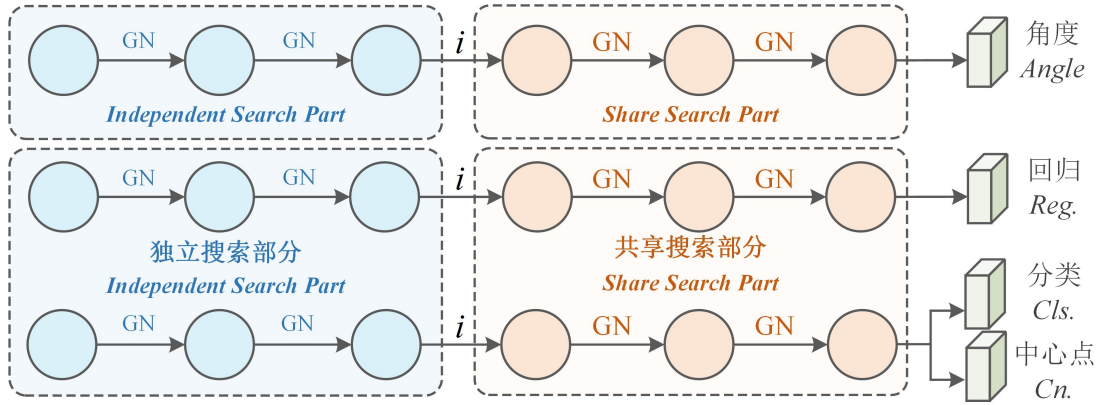


图 4-4 NAS-Head 的结构图

如图 4-5 所示，NAS 最佳采样池结构由可变形卷积(Deformable Convolution, DConv)、跳跃连接、标准卷积组成。其中， 3×3 DConv 在视图增强后采用动态卷积采样捕捉舰船的关键特征，跳跃连接是通过重新调整个别重要层级的权重，以确定共享权重层级与自动共享的层数，实现更为高效的特征融合和参数降低。 1×1 标准卷积和 3×3 标准卷积的作用是获得更好的前后性能比较。

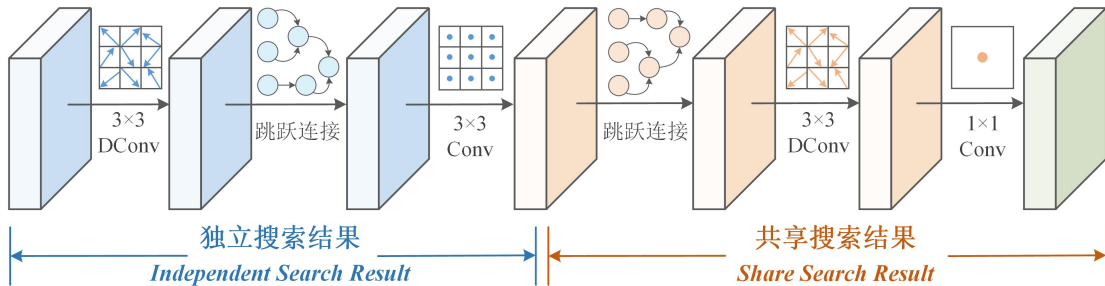


图 4-5 最佳头部采样池的结构图

与原始头部结构相比，NAS-Head 通过在 WS 分支和 SS 分支的头部引入 NAS 策略，从预定义的搜索空间中搜索最佳结构以自适应不同分支的任务。经 NAS 策略后，在头部结构中引入 3×3 可变形卷积与 1×1 标准卷积，使其在提高角度学习能力和舰船检测能力的同时实现计算资源占用的最小化。

4.3.3 外切 Scylla 交并比损失

损失函数作为训练惩罚措施，合理定义损失将为模型性能带来显著的提升。H2RBox-v2 的 WS 分支基于 IoU 设计 CircumIoU 损失，使其能对真实框进行旋转增强，进一步加速收敛，如图 4-6 所示。与全监督 RBox 方法不同，WS 任务的真实框是外切于预测框的 RBox。因此，CircumIoU 重新定义框回归损失：

$$L_{\text{CircumIoU}}(B_{\text{pred}}, B_{\text{gt}}) = -\ln \frac{B_{\text{proj}} \cap B_{\text{gt}}}{B_{\text{proj}} \cup B_{\text{gt}}} \quad (4-7)$$

式中， B_{pred} 表示预测框， B_{gt} 表示真实框， B_{proj} 是通过将 B_{pred} 映射到 B_{gt} 的方向所获得 RBox。CircumIoU 损失通过介质 B_{proj} 直接计算 B_{pred} 和 B_{gt} 的损失，使 B_{gt} 作为真实 RBox 外切至 B_{pred} ，避免先前利用 HBox 的 B_{gt} 计算 IoU 损失，在旋转视图增强中 B_{gt} 变成 RBox 的情况下，无法计算损失的问题。

然而，CircumIoU 损失在训练过程中存在些许问题，当 B_{pred} 与 B_{gt} 不相交时，即使 B_{proj} 映射至 B_{gt} 的方向，也会导致 CircumIoU 损失无法进行有效梯度回传和学习训练。同时，舰船目标方向变化大是非常常见的，当 B_{pred} 与 B_{gt} 的角度方向差异过大时， B_{proj} 的映射将存在偏差，导致预测边界框失真。

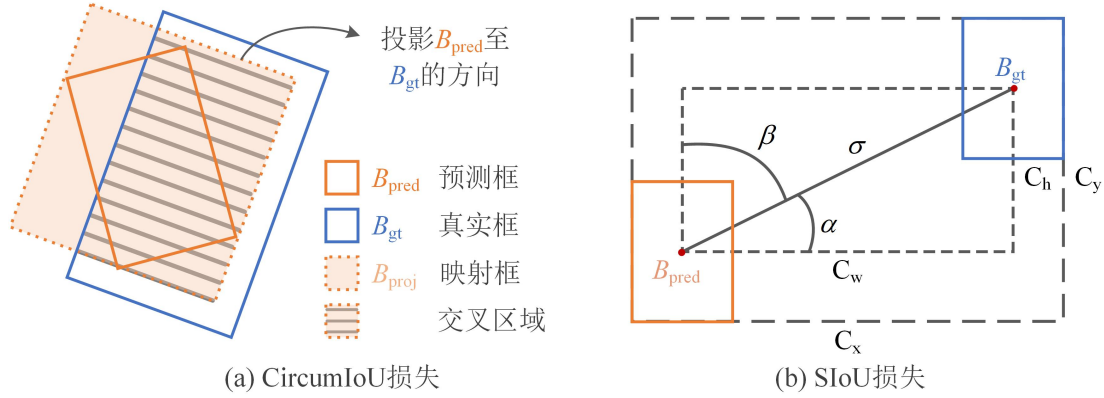


图 4-6 不同回归损失的结构图

为解决上述问题，引入 SIoU^[69]并设计 CircumSIoU 损失，SIoU 由角度损失、距离损失、形状损失组成，如图 4-6 所示。其中，角度损失旨在最小化 B_{pred} 与 B_{gt} 的角度偏差，在舰船识别中目标会以各角度出现，通过最小化角度差异，模型可准确预测舰船方向，解决 B_{proj} 投影偏差的问题，角度损失的定义如下：

$$\Lambda = 1 - 2 \sin^2(\arcsin(x)) - \frac{4}{\pi} \quad (4-8)$$

式中， Λ 表示角度损失， $x = \sin \alpha \cdot C_w$ ， C_w 表示 B_{pred} 与 B_{gt} 横坐标间的差异。

距离损失用于最小化 B_{pred} 与 B_{gt} 中心点间的距离，在复杂海洋场景下有助于模型准确定位舰船目标，距离损失的定义如下：

$$\Delta = \sum_{t=x,y} (1 - e^{-\tau P_t}) \quad (4-9)$$

式中， Δ 表示距离损失， t 表示 x 轴与 y 轴的选择因子， $\tau = 2 - \Lambda$ 表示角度损失的调整因子， P_t 表示 B_{pred} 和 B_{gt} 中心点的归一化距离损失。

形状损失致力匹配 B_{pred} 与 B_{gt} 的形状，如边界框的长宽比例。由于舰船形状不一，精确的形状匹配对提高识别精度至关重要。形状损失的定义如下：

$$\Omega = \sum_{t=w,h} (1 - e^{-w_t})^\theta \quad (4-10)$$

式中， Ω 表示形状损失， θ 用于控制对形状损失的关注程度，为避免过度关注，其取值范围为[2,6]，最优值为 4， w_t 表示 B_{pred} 和 B_{gt} 在宽度和高度间的相对误差。最终，结合 SIoU 损失与 CircumIoU 损失，得到 CircumSIoU 损失如下：

$$L_{CircumSIoU}(B_{pred}, B_{gt}) = 1 - L_{CircumIoU} + \frac{\Delta + \Omega}{2} \quad (4-11)$$

CircumSIoU 损失通过引入方向角度感知向量，使模型可以在旋转增强视图中 B_{gt} 成为真实 RBox 的同时实现随机旋转的回归损失计算，精确预测舰船方向、定位舰船目标以及匹配舰船形状，并进一步加速收敛速度。

4.4 实验结果分析

4.4.1 实验数据集

HRSC2016 数据集^[70]是西北工业大学收集的大型舰船数据集，图像大小为 300×300 到 1500×900 像素。HRSC2016 包含 1060 张图像和 2976 个实例，其中，训练集、验证集和测试集包含 436 张、181 张和 444 张图像。

DOSR 数据集^[59]是一个大型舰船检测数据集，包含 1066 张图像和 6127 个实例，舰船被划分为 20 个类别，训练集、验证集和测试集包含 523 张、223 张和 320 张图像。本章将 20 个舰船类别统一为单个“ship”类别。

SSDD 数据集^[71]是一个合成孔径雷达的舰船检测数据集，图像分辨率在 1m 到 15m 间。SSDD 覆盖大片海域和近岸地区，包含 1160 张 SAR 图像和 2456 个舰船实例，其中，训练集和测试集包含 928 张和 232 张图像。

光学遥感 HRSC2016 数据集、DOSR 数据集被用于评估模型的有向舰船检

测性能，SAR 遥感 SSDD 数据集被用于评估模型的泛化性能。其中，主要消融实验在光学遥感 HRSC2016 数据集、DOSR 数据集上实现。

4.4.2 实验配置与测评标准

实验硬件配置为 NVIDIA GeForce RTX 3060 12G 显卡，实验基于 PyTorch 1.11.0 的基准工具箱 MMRotate 1.0.0^[61]，采用权重衰减为 0.005 的 AdamW 作为优化器，批量大小为 2。对于 HRSC2016 与 DOSR，训练 72 个 epoch，在第 48 和 66 个 epoch 学习率降低 0.1 倍，输入大小为 800×800 。对于 SSDD，训练 36 个 epoch，在第 24 和 33 个 epoch 学习率降低 0.1 倍，输入大小为 512×512 。

本章选取经过 Pascal VOC2007 度量计算的平均准确率 AP 作为主要评价指标，AP 是基于准确率和召回率计算的，其中，不同 IoU 阈值的 AP (AP_{50} 、 AP_{75} 、 $AP_{50:95}$) 用于剖析整体检测性能，选取 FLOPs 反映模型整体复杂度。

4.4.3 对比实验

本节在通用的光学舰船检测数据集（HRSC2016 和 DOSR）上进行定量和定性分析，以验证模型有向舰船检测的先进性能。此外，在不同噪声水平的 SAR 舰船检测数据集（SSDD）上进行泛化分析，以验证其泛化性能。

(1) 光学舰船检测的定性实验

为验证模型在光学舰船检测上的有效性，采用 12 种主流目标检测器，同 DSENet 在 HRSC2016 和 DOSR 上进行定性比较分析，如表 4-1 所示。

表 4-1 HRSC2016 和 DOSR 的定性对比实验结果

方法	HRSC2016			DOSR		
	$AP_{50:95}(\%)$	$AP_{50}(\%)$	$AP_{75}(\%)$	$AP_{50:95}(\%)$	$AP_{50}(\%)$	$AP_{75}(\%)$
R ³ Det [24]	41.92	87.31	32.93	44.83	79.53	45.24
Faster R-CNN [16]	48.65	86.34	49.90	50.40	79.95	55.76
RetinaNet [6]	49.57	81.92	50.53	47.72	78.10	53.41
PSC [30]	51.95	84.22	59.24	49.34	85.64	53.46
GWD [31]	53.21	84.67	58.91	49.47	78.34	55.54
SASM [29]	55.33	81.95	65.19	51.62	85.76	62.56
CFA [72]	58.32	87.79	66.82	52.30	87.74	63.47
FCOS [63]	58.72	88.66	67.44	53.49	87.32	58.53
O-Reppoints [25]	61.29	89.30	74.42	53.45	88.76	64.76
S ² ANet [23]	61.58	90.07	75.50	54.06	88.77	60.65
H2RBox [41]	2.25	7.83	0.62	10.36	21.44	5.12
H2RBox-v2 [42]	57.70	89.64	63.35	46.52	80.26	54.95
DSENet (本章方法)	63.34	90.27	76.27	54.29	89.06	66.55

对于 HRSC2016，与全监督的有向目标检测器相比，DSENet 的 AP_{50} 、 AP_{75} 、 $AP_{50:95}$ 以 90.27%、76.27%、63.34% 的优势超越所有主流模型。DSENet 的 AP_{50} 、 AP_{75} 、 $AP_{50:95}$ 比 S^2ANet 高出 0.20%、0.77%、1.72%，而 S^2ANet 采用更复杂的特征对齐卷积对齐舰船对象。与基线相比，DSENet 的 AP_{50} 、 AP_{75} 、 $AP_{50:95}$ 提高 0.63%、12.92%、5.64%，反映无向监督有向条件下舰船检测的优越性；对于 DOSR，虽然全监督的有向检测器性能较好，但 DSENet 的 AP_{50} 、 AP_{75} 、 $AP_{50:95}$ 以 89.06%、66.55%、54.29% 的优势均脱颖而出。其中，DSENet 的 AP_{50} 、 AP_{75} 、 $AP_{50:95}$ 比 O-Reppoints 高出 0.30%、1.79%、0.84%，而 O-Reppoints 引入质量评估样本分配策略以实现高质量样本的训练。与基线相比，DSENet 的 AP_{50} 、 AP_{75} 、 $AP_{50:95}$ 提高 8.80%、11.60%、7.77%，验证了其舰船检测的优势。

根据基线方法和本章方法的检测结果，图 4-7 展示了 HRSC2016 和 DOSR 的混淆矩阵。与基线方法相比，本章方法在 HRSC2016 和 DOSR 数据集上的检测精度分别提高 2% 和 6%，而背景误检率也相应降低。

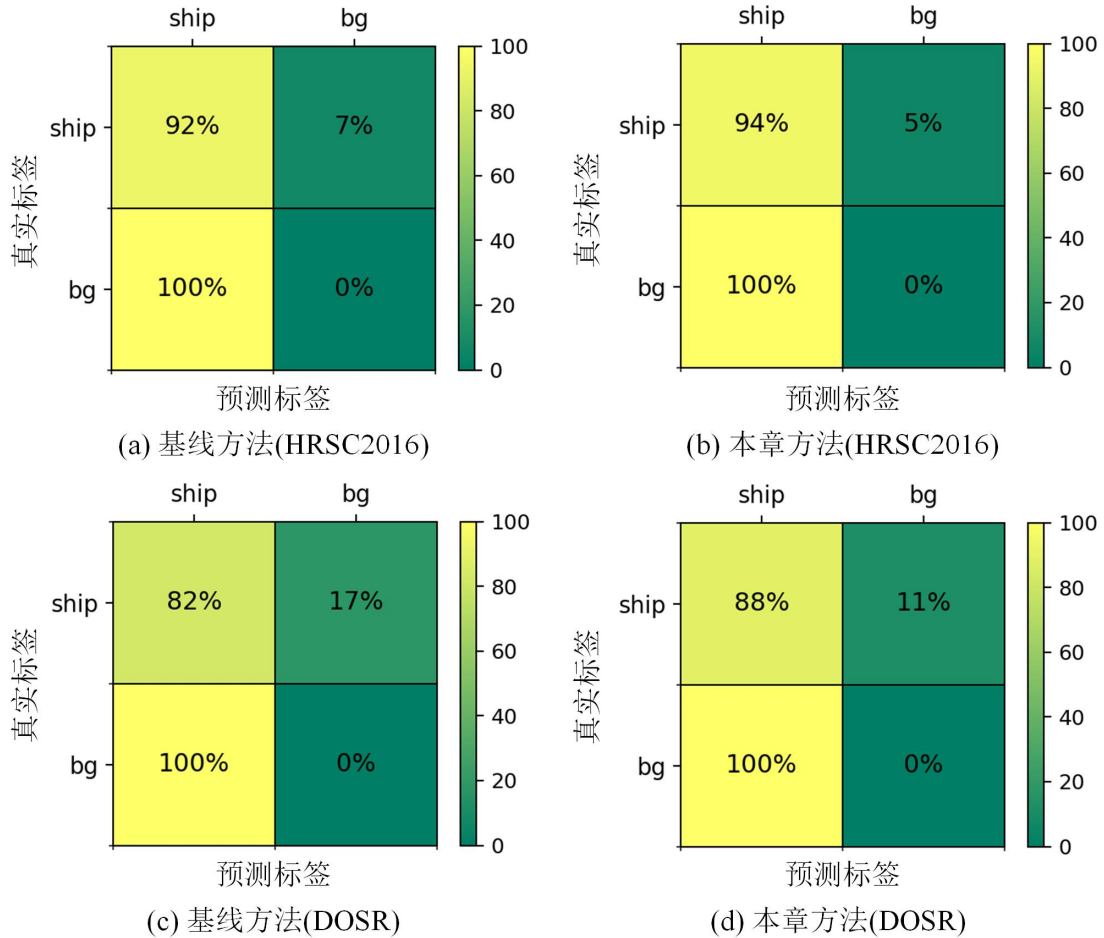


图 4-7 检测混淆矩阵的对比图

(2) 光学舰船检测的定量实验

为进一步直观展示本章方法在光学遥感舰船检测的有效性，图 4-8 与图 4-9 分别展示了 HRSC2016 与和 DOSR 的可视化定量对比结果。



图 4-8 HRSC2016 的定量对比图

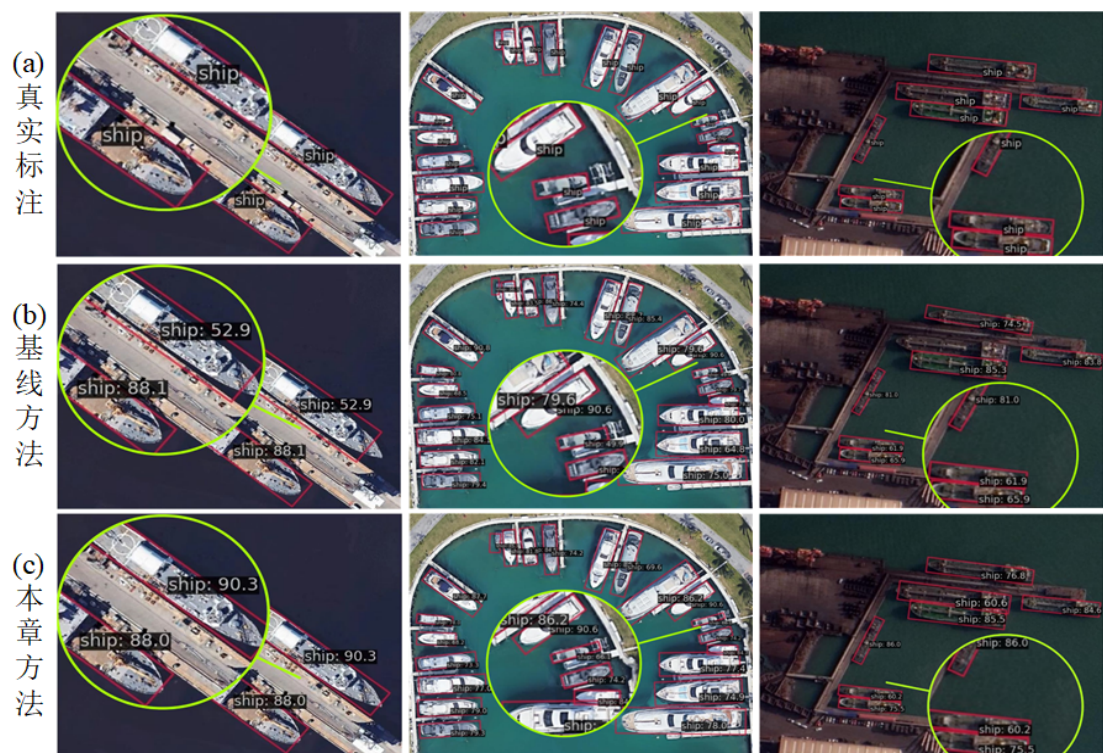


图 4-9 DOSR 的定量对比图

由图 4-8 可知，由于缺乏足够的多尺度特征信息，基线的 WS 分支检测效果不佳，且易出现漏检和误检情况，而 DSEFM 模块动态增强多尺度特征后，检测精度得到提高；NAS-Head 模块为 WS 分支和 SS 分支筛选最佳头部结构，误检情况得到改善；CircumSIoU 模块则使模型预测框更接近真实框。

由图 4-9 可知，与基线方法相比，本章方法在面向密集场景下的舰船检测准确率得到提高。此外，在检测尺度变化较大的舰船时，例如长宽比悬殊的舰船，基线方法更易出现漏检的情况，而本章方法避免此类情况。由此可知，本章方法在各种复杂光学遥感场景下的舰船检测性能均得到提升。

(3) SAR 舰船检测的泛化实验

为进一步验证模型在不同类型数据上舰船检测性能，以证明本章方法 DSENet 的泛化性能，选取用 11 种主流的目标检测器，同 DSENet 在 SAR 舰船检测数据集 SSDD 上进行比较分析，实验如表 4-2 所示。

表 4-2 SSDD 的泛化对比实验结果

方法	主干	SSDD		
		AP _{50:95} (%)	AP ₅₀ (%)	AP ₇₅ (%)
PSC [30]	R-50-FPN	39.56	79.20	35.77
SASM [29]	R-50-FPN	39.87	80.13	33.58
R ³ Det [24]	R-50-FPN	40.16	81.15	34.90
RetinaNet [6]	R-50-FPN	40.55	79.49	36.07
CFA [72]	R-50-FPN	43.28	85.35	35.67
GWD [31]	R-50-FPN	43.63	80.59	38.12
Faster R-CNN [16]	R-50-FPN	45.10	89.34	42.37
FCOS [63]	R-50-FPN	45.78	88.78	36.98
S ² ANet [23]	R-50-FPN	46.82	89.84	44.71
H2RBox [41]	R-50-FPN	7.30	15.37	4.10
H2RBox-v2 [42]	R-50-FPN	44.02	89.27	41.05
DSENet (本章方法)	R-50-DSFEM	47.26	90.12	47.43

由表可知，在检测 SAR 舰船目标时，DSENet 的 AP₅₀、AP₇₅、AP_{50:95} 分别以 90.12%、47.43%、47.26% 的优势超越全监督目标检测器。在相同实验设置下，与弱监督检测器 H2RBox 相比，DSENet 的 AP₅₀、AP₇₅、AP_{50:95} 分别高出 74.75%、43.33%、39.96%，与基线 H2RBox-v2 相比，AP₅₀、AP₇₅、AP_{50:95} 分别以 0.85%、6.38%、3.24% 得到提高，证明本章方法在不同传感器数据下优异泛化性。

为更清晰地可视化本章方法 DSENet 的泛化性能，分别列出基线方法与本章方法在 SSDD 数据集上的可视化结果，如图 4-10 所示。

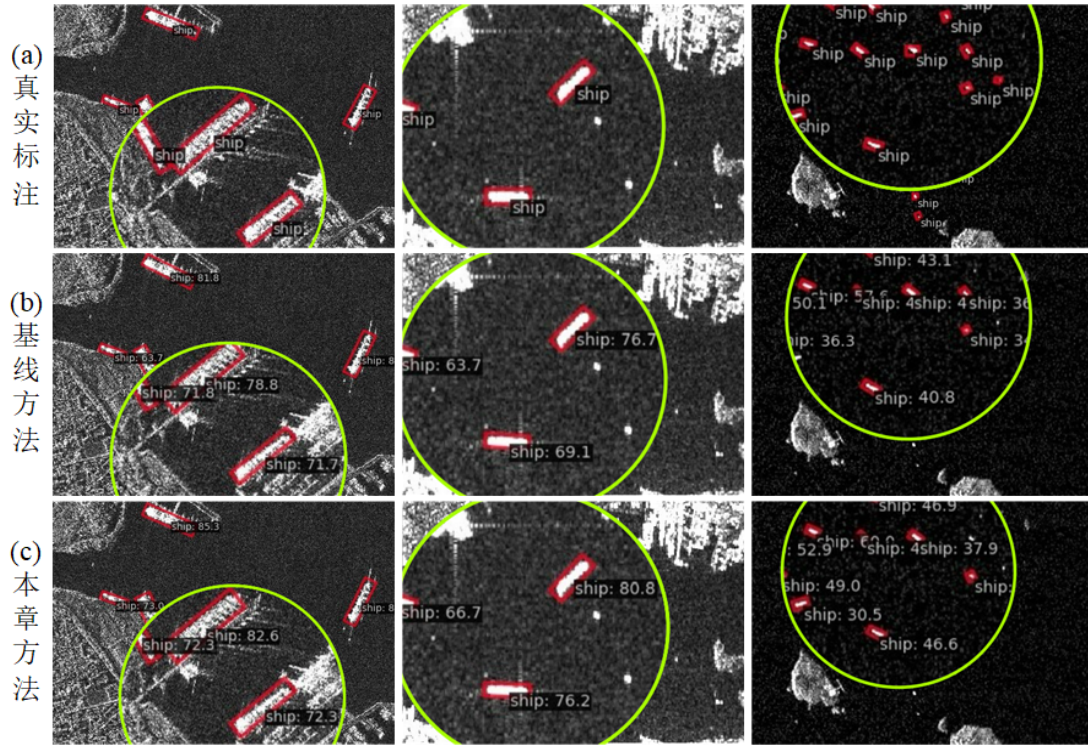


图 4-10 SSDD 的泛化对比图

由图可知，即使在噪声杂波较高的极化 SAR 环境中，本章方法的舰船检测性能也明显优于基线。具体地，在检测小型 SAR 舰船时，基线方法易出现漏检，本章方法则避免此类情况，更深入验证本章方法出色的泛化性。

4.4.4 消融实验

(1) 各模块的有效性消融实验

为彻底分析 DSENet 中三个模块的有效性，执行组件式消融实验，如表 4-3 所示，DSFEM、NAS-Head、CircumSIoU 模块对各数据集性能均有益。

表 4-3 各模块有效性的消融实验结果

数据集	DSFEM	NAS-Head	CircumSIoU	AP _{50:95} (%)	AP ₅₀ (%)	AP ₇₅ (%)	FLOPs(G)
HRSC2016				57.70	89.64	63.35	126.04
	√			60.04	89.80	73.90	120.88
	√	√		61.25	90.08	74.03	82.05
	√	√	√	63.34	90.27	76.27	82.05
DOSR				46.52	80.26	54.95	126.04
	√			49.53	82.42	62.88	120.88
	√	√		51.84	85.82	63.27	82.05
	√	√	√	54.29	89.06	66.55	82.05

在 HRSC2016 和 DOSR 上，首先，借助 DSFEM 模块，DSENet 准确利用不同尺度的特征表示，为 WS 分支和 SS 分支提供精确的舰船特征，AP₅₀ 分别提高

0.16%、2.16%， AP_{75} 提高 10.55%、7.93%， $AP_{50:95}$ 提高 2.34%、3.01%，同时 FLOPs 均降低 5.16；其次，在 DSFEM 的基础上应用 NAS-Head，为 WS 和 SS 分支搜寻最佳头部结构，使模型更好地进行舰船预测和角度学习， AP_{50} 分别提高 0.44%、5.56%， AP_{75} 提高 10.68%、8.32%， $AP_{50:95}$ 提高 3.55%、5.32%，同时 FLOPs 均降低 43.99；最后，在 DSFEM 和 NAS-Head 的基础上应用 CircumSIoU，使模型精确预测舰船方向、定位目标、匹配形状， AP_{50} 分别提高 0.63%、8.80%， AP_{75} 提高 12.92%、11.60%， $AP_{50:95}$ 提高 5.64%、7.77%。

(2) DSFEM 的消融实验

在 DSFEM 中，采用空洞卷积提取尺度特征，而膨胀率的变化将生成对应的膨胀卷积，即尺度层级分支，尺度分支数量对特征融合和动态增强有不同影响。因此，对 DSFEM 同一层级内不同膨胀率进行消融分析，如表 4-4 所示。

表 4-4 DSFEM 膨胀率的消融实验结果

膨胀率	HRSC2016			DOSR		
	$AP_{50:95}(\%)$	$AP_{50}(\%)$	$AP_{75}(\%)$	$AP_{50:95}(\%)$	$AP_{50}(\%)$	$AP_{75}(\%)$
-	57.70	89.64	63.35	46.52	80.26	54.95
(1,2)	59.65	89.77	73.68	48.80	81.37	61.15
(1,2,3)	60.04	89.99	73.90	49.53	82.42	62.88
(1,2,3,4)	59.39	89.54	73.36	47.97	81.23	60.71

与基线相比，无论膨胀率取值如何变化，DSFEM 均能为模型带来增益。在 HRSC2016 和 DOSR 上，(1,2) 的膨胀率设置使 AP_{50} 提高 0.13%、1.11%， AP_{75} 提高 10.33%、6.20%， $AP_{50:95}$ 提高 1.95%、2.28%；(1,2,3) 的膨胀率设置分别使 AP_{50} 提高 0.35%、2.16%， AP_{75} 提高 10.55%、7.93%， $AP_{50:95}$ 提高 2.34%、3.01%；(1,2,3,4) 的膨胀率设置使 AP_{75} 提高 10.01%、7.93%， $AP_{50:95}$ 提高 1.69%、1.45%，在 HRSC2016 上 AP_{50} 下降 0.1%，即感受野过度扩大引入冗余噪声。DSFEM 模块的引入使 AP_{75} 大幅提高，表示动态尺度特征增强使模型对舰船尺度变化更敏感，加强目标定位能力。实验表明，(1,2,3) 的膨胀率为最佳设置。

为直观地展示特征增强模块在捕捉多尺度特征信息的有效性，分别从 FPN 模块和 DSFEM 模块中提取可视化特征图，并生成热力图，如图 4-11 所示。与原始 FPN 相比，DSFEM 模块更准确地提取与融合多尺度舰船特征，并动态增强舰船的重要尺度细节，缓解不同尺度特征的稀释，聚焦更广感受野。综上所述，进一步说明 DSFEM 模块在舰船检测中的有效性。

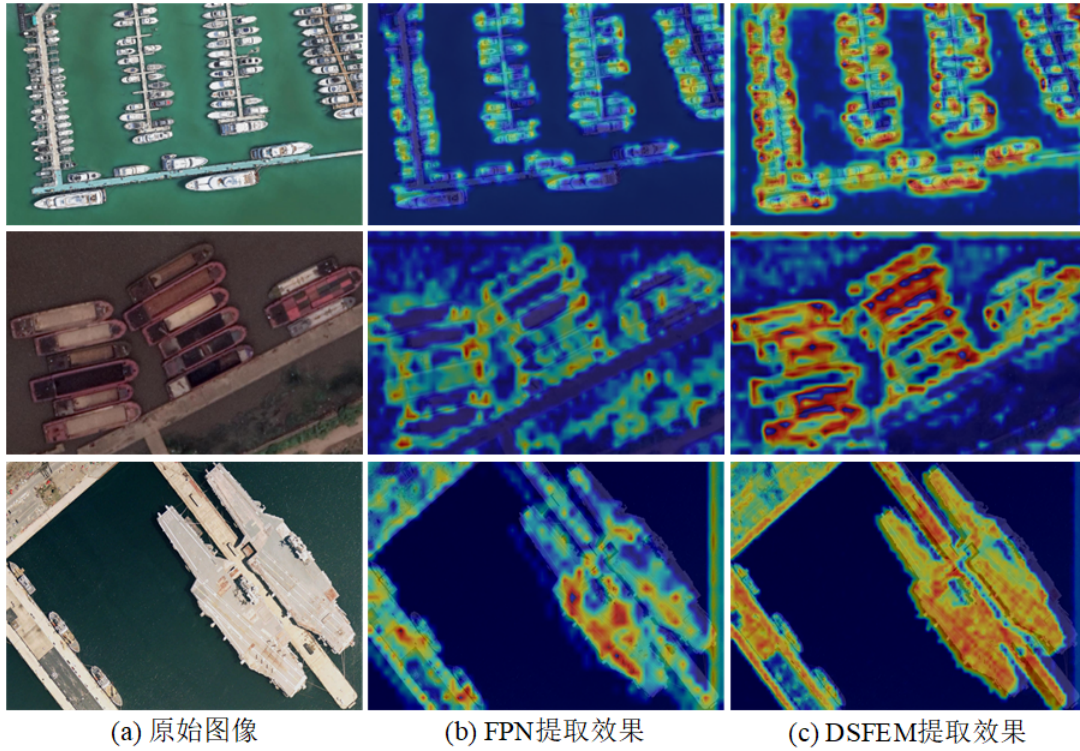


图 4-11 特征提取和特征融合效果的对比图

(3) NAS-Head 的消融实验

与固定拓扑结构相比，NAS-Head 采用 3×3 DConvs 与 1×1 卷积的方式，在适应分支间不同任务的同时最小化资源占用。为探索 NAS-Head 应用不同卷积的效果，在应用 DSFEM 基础上进行消融实验，如表 4-5 所示。

表 4-5 NAS-Head 卷积的消融实验结果

数据集	1×1 Conv	3×3 DConv	AP _{50:95} (%)	AP ₅₀ (%)	AP ₇₅ (%)	FLOPs(G)
HRSC2016	√	√	60.04	89.80	73.90	120.88
			60.14	89.89	73.97	106.89
	√	√	60.80	90.03	73.46	96.04
			61.25	90.08	74.03	82.05
DOSR	√	√	49.53	82.42	62.88	120.88
			50.30	83.56	62.95	106.89
	√	√	50.91	84.49	62.39	96.04
			51.84	85.82	63.27	82.05

在 HRSC2016 和 DOSR 上，单独应用 3×3 DConvs 与 1×1 卷积均可提高模型检测能力，并降低复杂度。然而，当同时应用 3×3 DConvs 与 1×1 卷积时，准确率增益达到最大化，分别使 AP₅₀ 提高 0.28%、3.40%，AP₇₅ 提高 0.13%、0.39%，AP_{50:95} 提高 1.21%、2.31%，而 FLOPs 均降低 38.83G，有效地证明经 NAS 策略所得的 NAS-Head 结构可获得最优收益。

(4) CircumSIoU 的消融实验

在 CircumSIoU 中, 采用 SIoU 结合 CircumIoU 以弥补角度、距离以及形状偏差过大的缺陷。然而, CircumIoU 可结合其他 IoU 损失, 为确定最佳回归损失, 将 EIoU、DIoU、GIoU、SIoU 同 CircumIoU 结合, 如表 4-6 所示。

表 4-6 CircumSIoU 损失的消融实验结果

IoU 类型	HRSC2016			DOSR		
	AP _{50:95} (%)	AP ₅₀ (%)	AP ₇₅ (%)	AP _{50:95} (%)	AP ₅₀ (%)	AP ₇₅ (%)
CircumIoU	57.70	89.64	63.35	46.52	80.26	54.95
+EIoU	55.40	89.37	62.10	45.56	79.24	53.46
+DIoU	56.44	89.15	62.40	45.78	80.20	54.68
+GIoU	57.54	89.10	62.21	47.66	80.34	55.71
+SIoU	58.55	89.86	64.87	49.34	80.47	56.49

在 HRSC2016 和 DOSR 中, EIoU 致力于解决样本不平衡的问题, 而模型过度关注高质量样本, 结合 CircumIoU 后准确率明显降低; DIoU 在预测框和真实框的相对位置变化时出现较大波动, 导致 AP₅₀ 下降 0.49%、0.06%, AP₇₅ 下降 0.95%、0.27%, AP_{50:95} 下降 1.26%、0.74%; GIoU 改善了相对位置变化产生的波动, 提高在 DOSR 上密集检测能力, 但缺乏角度差异考虑, 面对 HRSC2016 的角度变化, 检测能力有所降低; SIoU 充分考虑角度、距离以及形状差异, 结合 CircumIoU 后引导精准边框回归, 使 AP₅₀ 提高 0.22%、0.21%, AP₇₅ 提高 1.52%、1.54%, AP_{50:95} 提高 0.85%、2.82%, 证明了 CircumSIoU 有效性。

由于真实框与预测框的形状匹配至关重要, 而 CircumSIoU 的超参数 θ 控制形状成本的关注度。为探索超参数 θ 的最优值, 表 4-7 在区间 (2,6) 对 θ 执行消融实验。若 θ 值为 1 时模型会立刻优化形状, 将损害形状的自由移动。

表 4-7 CircumSIoU 超参数 θ 的消融实验结果

θ 值	HRSC2016			DOSR		
	AP _{50:95} (%)	AP ₅₀ (%)	AP ₇₅ (%)	AP _{50:95} (%)	AP ₅₀ (%)	AP ₇₅ (%)
2	57.59	89.27	63.34	47.31	80.28	54.28
3	58.27	89.63	64.07	48.59	80.22	55.96
4	58.55	89.86	64.87	49.34	80.47	56.49
5	57.88	89.44	63.84	49.08	80.24	56.35
6	57.39	89.38	64.05	47.72	80.28	55.32

当 θ 值较小时 (如 $\theta=2$), 模型对形状成本的关注过少, 使预测框与真实框之间执行较少的形状匹配便过早优化形状。当 θ 值较大时 (如 $\theta=6$), 模型将过

度关注形状成本，降低角度成本关注，使预测框和真实框之间的角度偏差过大，而上述情况的结果均不佳。当 θ 值为4时，平衡形状成本和角度成本的关注，在HRSC2016和DOSR上获取最佳结果，使 AP_{50} 达到89.86%、80.47%， AP_{75} 达到64.87%、56.49%， $AP_{50:95}$ 达到58.55%、49.34%。因此， θ 的最佳取值为4。

图4-12展示HRSC2016和DOSR上的CircumIoU和CircumSIoU的边界框回归损失。与原始CircumIoU相比，CircumSIoU损失收敛更优异。

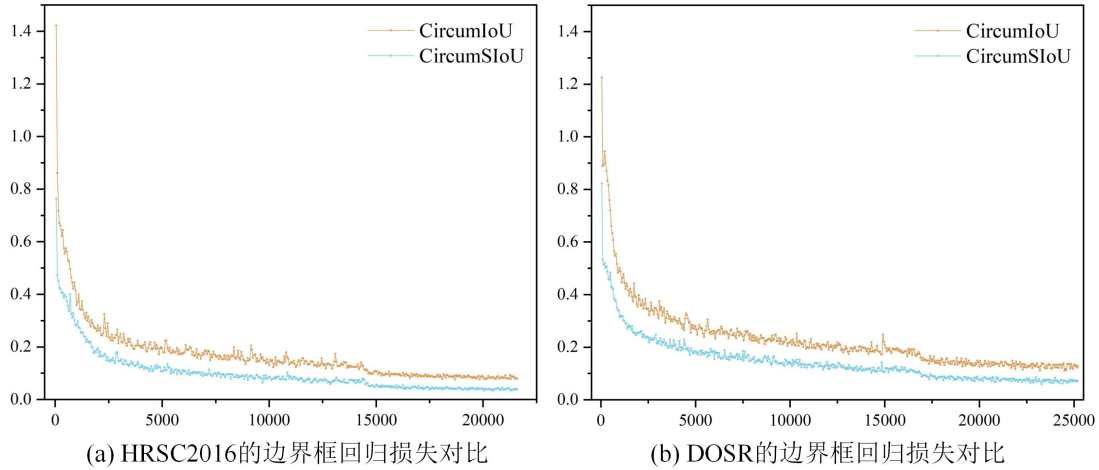


图4-12 回归损失曲线的对比图

4.5 本章小结

为解决光学遥感图像中有向舰船检测的旋转框标注成本高、注释少的问题，本章探究了基于动态尺度增强的无向监督有向舰船检测网络DSENet。其中，模型的主要贡献在三个方面：①动态增强多尺度舰船特征（DSFEM模块）；②集成平衡精度和轻量级头部（NAS-Head模块）；③引导模型精准定位舰船并加速训练（CircumSIoU模块）。实验结果表明，在三个通用舰船数据集上执行消融实验和对比实验，与其他主流的全监督检测器相比，无向监督有向的DSENet模型在有向舰船检测中表现更优异且更具泛化性。

第 5 章 总结与展望

5.1 总结

舰船作为浩瀚海洋的主要运输工具，围绕遥感舰船识别技术备受世界各国关注。本文基于深度学习的方法，解决无向、有向、无向监督有向舰船检测的现存问题。首先，针对遥感无向检测中小型舰船不易检测的问题，探讨了一种基于改进 YOLOv8s 的无向小型舰船检测算法；其次，针对遥感有向检测中细粒度舰船检测特征混淆的问题，研究了一种基于候选框增强的有向细粒度舰船检测算法；最后，针对遥感无向监督有向检测中舰船检测准确率低的问题，探究了一种基于动态尺度增强的无向监督有向舰船检测算法。研究内容如下：

(1) 探讨了一种基于改进 YOLOv8s 的无向小型舰船检测算法。首先，优化模型结构，均衡浅层空间特征与深层语义特征的权重分布，提升小型舰船的定位与尺度判别能力；其次，采用 C2f-FE 模块，分流特征通道，聚焦跨通道特征交互和跨空间信息融合，增强特征提取能力；最终，引入 Dyhead 模块，动态增强特征融合后的小型舰船特征，加强模型的分类与定位能力。经过消融实验和对比实验，改进模型在准确性和复杂度上具备优越性。

(2) 研究了一种基于候选框增强的有向细粒度舰船检测算法。首先，设计横向尺度均衡特征金字塔网络 LSE-FPN，为候选框生成提供均衡的尺度特征，以便后续生成更精确的候选框；其次，设计多通道融合网络 MCFN，以提取候选框中具有鉴别力的独立特征表示；最终，设计定向动态检测网络 ODPN，采用动态训练策略为训练提供指导，有效利用高质量候选框对舰船进行识别和定位。根据定量和定性的实验分析，改进方法实现先进的 FGSD 性能。

(3) 探究了一种基于动态尺度增强的无向监督有向舰船检测算法。首先，设计动态尺度特征增强 DSFEM 模块，经多尺度特征提取、融合及动态增强，为自监督分支和弱监督分支提供精准舰船尺度特征；其次，引入 NAS 策略封装头部，搜索头部结构最优解；最终，设计外切 Scylla 交并比 CircumSIoU 损失，通过方向角度感知向量引导模型准确预测舰船目标，并加速训练。实验表明，所探究的模型在无向监督有向的条件下实现先进的舰船检测性能。

5.2 展望

本文从无向舰船检测、有向舰船检测、无向监督有向舰船检测的角度出发，解决了光学遥感舰船检测领域的现存问题。虽取得一定成效，但尚有值得改进之处需要深入剖析探讨，具体如下：

(1) 在无向舰船检测中，重叠排列或遮挡场景下舰船目标区分困难，传统水平框难以实现多方向舰船的高效定位，且光照变化、天气干扰显著影响检测鲁棒性，应多方面提升不同困难遥感场景下的无向舰船检测性能。

(2) 在有向舰船检测中，针对极端长宽比舰船（如货轮与快艇）的旋转定位精度不足。角度分布不均衡削弱模型泛化性，密集舰船因角度重叠难以精准分离，导致角度歧义问题（如 180° 与 0° 等效性）影响模型收敛。

(3) 在无向监督有向舰船检测中，在生成旋转视图与翻转视图后，输入至网络特征提取时，由于旋转和翻转导致特征位置变换，使输入至自监督与弱监督的学习与预测产生偏差，应设计旋转等变特征提取以平衡特征。

参考文献

- [1] Zhao T, Wang Y, Li Z, et al. Ship Detection with Deep Learning in Optical Remote-Sensing Images: A Survey of Challenges and Advances [J]. Remote Sensing, 2024, 16(07): 1145.
- [2] Li K, Wan G, Cheng G, et al. Object Detection in Optical Remote Sensing Images: A Survey and A New Benchmark [J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2020, 159(09): 296-307.
- [3] Amani M, Ghorbanian A, Asgarimehr M, et al. Remote Sensing Systems for Ocean: A Review (Part 1: Passive Systems) [J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2022, 15: 210-234.
- [4] 赵泉华,王肖,李玉,等.基于散射梯度矢量的 PolSAR 数据舰船检测方法[J].仪器仪表学报,2020,41(03):39-48.
- [5] 陈艳浩.基于全局背景减法滤波器与多形状特征的红外舰船检测算法[J].电子测量与仪器学报,2018,32(10):138-146.
- [6] Lin T, Goyal P, Girshick R, et al. Focal Loss for Dense Object Detection [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 42(02): 318-327.
- [7] Redmon J, Divvala S, Girshick R, et al. You Only Look Once: Unified, Real-Time Object Detection [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, Las Vegas, NV, USA, 27-30 June, 2016: 779-788.
- [8] Lv W, Zhao Y, Xu S. DETRs Beat YOLOs on Real-time Object Detection [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, Seattle, WA, USA, 16-22 June, 2024: 16965-16974.
- [9] Zhang X, Liu Q, Chang H, et al. High-Resolution Network with Transformer Embedding Parallel Detection for Small Object Detection in Optical Remote Sensing Images [J]. Remote Sensing, 2023, 15(18): 4497.
- [10] Zhang X, Yuan S, Luan F, et al. Similarity Mask Mixed Attention for YOLOv5 Small Ship Detection of Optical Remote Sensing Images [C]//Proceedings of the WRC Symposium on Advanced Robotics and Automation. IEEE, Beijing, China, 20-20 August, 2022: 263-268.
- [11] Mishra S, Shalu P, Soman D, et al. Advanced Ship Detection System Using YOLOv7 [C]//Proceedings of the IEEE Conference on Information and Communication Technology (CICT). IEEE, Jabalpur, India, 15-17 December, 2023: 1-6.
- [12] Kim M, Jeong J, Kim S. ECAP-YOLO: Efficient Channel Attention Pyramid YOLO for Small Object Detection in Aerial Image [J]. Remote Sensing, 2021, 13(23): 4851.
- [13] 刘树东,任慧娟,张众维.基于特征融合和注意力机制的遥感目标检测[J].遥感信息,2023,38(05):1-7.
- [14] Fu J, Sun X, Wang Z, et al. An Anchor-Free Method Based on Feature Balancing and Refinement Network for Multiscale Ship Detection in SAR Images [J]. IEEE Transactions on Geoscience and Remote Sensing, 2021, 59(02): 1331-1344.
- [15] Girshick R, Donahue J, Darrel T. Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, Columbus, OH, USA, 23-28 June, 2014: 580-587.
- [16] Ren S, He K, 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(06): 1137-1149.

- [17] Li Q, Mou L, Liu Q, et al. HSF-Net: Multiscale Deep Feature Embedding for Ship Detection in Optical Remote Sensing Imagery [J]. IEEE Transactions on Geoscience and Remote Sensing, 2018, 56(12): 7147-7161.
- [18] Wang P, Liu J, Zhang Y, et al. A Novel Cargo Ship Detection and Directional Discrimination Method for Remote Sensing Image Based on Lightweight Network [J]. Journal of Marine Science and Engineering, 2021, 09(09): 932.
- [19] 方梦梁, 黄刚. 一种光学遥感图像船舶目标检测技术 [J]. 计算机技术与发展, 2019, 29(08): 136-141.
- [20] Yao Y, Jiang Z, Zhang H. Ship Detection in Optical Remote Sensing Images Based on Deep Convolutional Neural Networks [J]. Journal of Applied Remote Sensing, 2017, 11(04): 042611.
- [21] Nie S, Jiang Z, Zhang H. Inshore Ship Detection Based on Mask R-CNN [C]//Proceedings of the IEEE International Geoscience and Remote Sensing Symposium (IGARSS). IEEE, Valencia, Spain, 22-27 July, 2018: 693-696.
- [22] Yang F, Xu Q, Li B. Ship Detection from Thermal Remote Sensing Imagery Through Region-based Deep Forest [J]. IEEE Geoscience and Remote Sensing Letters, 2018, 15(03): 449-453.
- [23] Han J, Ding J, Li J, et al. Align Deep Features for Oriented Object Detection [J]. IEEE Transactions on Geoscience and Remote Sensing, 2022, 60: 1-11.
- [24] Yang X, Liu Q, Yan J, et al. R3Det Refined Single-Stage Detector with Feature Refinement for Rotating Object [C]//Proceedings of the AAAI Conference on Artificial Intelligence (AAAI). AAAI, Virtual Event, 02-09 February, 2021: 3163-3171.
- [25] Li W, Chen Y, Hu K, et al. Oriented RepPoints for Aerial Object Detection [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, New Orleans, LA, USA, 18-24 June, 2022: 1819-1828.
- [26] Zhuang Y, Liu Y, Zhang T, et al. Contour Modeling Arbitrary-Oriented Ship Detection from Very High-Resolution Optical Remote Sensing Imagery [J]. IEEE Geoscience and Remote Sensing Letters, 2023, 20: 1-5.
- [27] Ge J, Tang Y, Guo K, et al. KeyShip: Towards High-Precision Oriented SAR Ship Detection Using Key Points [J]. Remote Sensing, 2023, 15(08): 2035.
- [28] Chen Y, Wang J, Zhang Y, et al. Arbitrary-Oriented Ship Detection Based on Kullback-Leibler Divergence Regression in Remote Sensing Images [J]. Earth Science Informatics, 2023, 16(04): 3243-3255.
- [29] Hou L, Lu K, Xue J, et al. Shape-Adaptive Selection and Measurement for Oriented Object Detection [C]//Proceedings of the AAAI Conference on Artificial Intelligence (AAAI). AAAI, Vancouver, BC, Canada, 22 February-01 March, 2022: 923-932.
- [30] Yu Y, Da F. Phase-Shifting Coder: Predicting Accurate Orientation in Oriented Object Detection [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, Vancouver, BC, Canada, 17-24 June, 2023: 13354-13363.
- [31] Yang X, Yan J, Ming Q, et al. Rethinking Rotated Object Detection with Gaussian Wasserstein Distance Loss [C]//Proceedings of the International Conference on Machine Learning (ICML). PMLR, Virtual Event, 18-24 July, 2021: 11830-11841.
- [32] Xie X, Cheng G, Wang J, et al. Oriented R-CNN for Object Detection [C]//Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV). IEEE, Montreal, QC, Canada, 10-17 October, 2021: 3500-3509.

- [33] Han J, Ding J, Xue N, et al. ReDet: A Rotation-equivariant Detector for Aerial Object Detection [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, Nashville, TN, USA, 20-25 June, 2021: 2785-2794.
- [34] Ding J, Xue N, Long Y, 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 (CVPR). IEEE, Long Beach, CA, USA, 15-20 June, 2019: 2844-2853.
- [35] 周国庆,黄亮,孙乔.改进 Oriented R-CNN 的遥感舰船目标细粒度检测方法[J].计算机工程与应用,2024,60(15):307-317.
- [36] 张磊,陈文,王岳环.用于遥感舰船细粒度检测与识别的关键子区域融合网络[J].中国图象图形学报,2023,28(09):2940-2955.
- [37] 刘万军,高健康,曲海成,等.多尺度特征增强的遥感图像舰船目标检测[J].自然资源遥感,2021,33(03):97-106.
- [38] 徐志京,丁莹.自适应旋转区域生成网络的遥感图像舰船目标检测[J].激光与光电子学进展,2020,57(24):408-415.
- [39] Ouyang L, Fang L, Ji X. Multigranularity Self-Attention Network for Fine-Grained Ship Detection in Remote Sensing Images [J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2022, 15: 9722-9732.
- [40] Yan Z, Li Z, Xie Y, et al. ReBiDet: An Enhanced Ship Detection Model Utilizing ReDet and Bi-Directional Feature Fusion [J]. Applied Sciences, 2023, 13(12): 7080.
- [41] Yang X, Zhang G, Li W, et al. H2RBox: Horizontal Box Annotation is All You Need for Oriented Object Detection [C]//Proceedings of the International Conference on Learning Representations (ICLR). ICLR, Kigali, Rwanda, 01-05 May, 2023: 1-15.
- [42] Yu Y, Yang X, Li Q, et al. H2RBox-v2: Incorporating Symmetry for Boosting Horizontal Box Supervised Oriented Object Detection [C]//Proceedings of the Conference on Neural Information Processing Systems (NeurIPS). Curran Associates, New Orleans, LA, USA, 10-16 December, 2023: 59137-59150.
- [43] 任洋,陈绪君,王磊.基于跨空间多尺度的弱监督有向目标检测算法研究[J].激光杂志,2024,45(07):63-70.
- [44] Yue T, Zhang Y, Wang J. A Weak Supervision Learning Paradigm for Oriented Ship Detection in SAR Image [J]. IEEE Transactions on Geoscience and Remote Sensing, 2024, 62: 1-12.
- [45] Chen J, Kao S, He H, et al. Run, Don't Walk: Chasing Higher FLOPS for Faster Neural Networks [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, Vancouver, BC, Canada, 17-24 June, 2023: 12021-12031.
- [46] Ouyang D, He S, Zhang J, et al. Efficient Multi-Scale Attention Module with Cross-Spatial Learning [C]//Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, Rhodes Island, Greece, 04-10 June, 2023: 1-5.
- [47] Dai X, Chen Y, Xiao B, et al. Dynamic Head: Unifying Object Detection Heads with Attentions [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, Nashville, TN, USA, 20-25 June, 2021: 7369-7378.
- [48] Alashhab S, Gallego A, Pertusa A, et al. Precise Ship Location with CNN Filter Selection from Optical Aerial Images [J]. IEEE Access, 2019, 07: 96567-96582.
- [49] Xia G, Bai X, Ding J, et al. DOTA: A Large-Scale Dataset for Object Detection in Aerial Images [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, Salt Lake City, UT, USA, 18-23 June, 2018: 3974-3983.

- [50] Kim M, Jeong J, Kim S. ECAP-YOLO: Efficient Channel Attention Pyramid YOLO for Small Object Detection in Aerial Image [J]. Remote Sensing, 2021, 13(23): 4851.
- [51] 刘树东,任慧娟,张众维.基于特征融合和注意力机制的遥感目标检测[J].遥感信息,2023,38(05):1-7.
- [52] Li W, Zhao D, Yuan B, et al. PETDet: Proposal Enhancement for Two-Stage Fine-Grained Object Detection [J]. IEEE Transactions on Geoscience and Remote Sensing, 2024, 62: 1-14.
- [53] Wang X, Zhang S, Yu Z, et al. Scale-Equalizing Pyramid Convolution for Object Detection [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, Seattle, WA, USA, 13-19 June, 2020: 13356-13365.
- [54] Kim J, On K, Lim W, et al. Hadamard Product for Low-rank Bilinear Pooling [C]//Proceedings of the International Conference on Learning Representations (ICLR). ICLR, Toulon, France, 24-26 April, 2017: 1-14.
- [55] Dauphin Y, Fan A, Auli M, et al. Language Modeling with Gated Convolutional Networks [C]//Proceedings of the International Conference on Machine Learning (ICML). JMLR, Sydney, NSW, Australia, 06-11 August, 2017: 933-941.
- [56] Zhang H, Ghang H, Ma B, et al. Dynamic R-CNN: Towards High Quality Object Detection via Dynamic Training[C]//Proceedings of the European Conference on Computer Vision (ECCV). Springer, Glasgow, SEC, Scotland, 23-28 August, 2020: 260-275.
- [57] Zhang Z, Zhang L, Wang Y, et al. ShipRSImageNet: A Large-Scale Fine-Grained Dataset for Ship Detection in High-Resolution Optical Remote Sensing Images [J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2021, 14: 8458-8472.
- [58] Zhang F, Wang X, Zhou S, et al. Arbitrary-Oriented Ship Detection Through Center-Head Point Extraction [J]. IEEE Transactions on Geoscience and Remote Sensing, 2022, 60: 1-14.
- [59] Han Y, Yang X, Pu T, et al. Fine-Grained Recognition for Oriented Ship Against Complex Scenes in Optical Remote Sensing Images [J]. IEEE Transactions on Geoscience and Remote Sensing, 2022, 60: 1-18.
- [60] 禹文奇,程堃,王美君,等.MAR20: 遥感图像军用飞机目标识别数据集[J].遥感学报,2023,27(12):2688-2696.
- [61] Zhou Y, Yang X, Zhang G, et al. MMRotate: A Rotated Object Detection Benchmark Using PyTorch [C]//Proceedings of the ACM International Conference on Multimedia (ACM MM). ACM, Lisbon, Portugal, 10-14 October, 2022: 7331-7334.
- [62] Liu Z, Mao H, Wu C, et al. A ConvNet for the 2020s [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, New Orleans, LA, USA, 18-24 June, 2022: 11966-11976.
- [63] Tian Z, Shen C, Chen H, et al. FCOS: Fully Convolutional One-Stage Object Detection [C]//Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV). IEEE, Seoul, Korea (South), 27-02 November, 2019: 9626-9635.
- [64] Xi Y, Fu M, Wang Q, et al. Gliding Vertex on the Horizontal Bounding Box for Multi-Oriented Object Detection [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2019, 43(04): 1452-1459.
- [65] Li Y, Hou Q, Zheng Z, et al. Large Selective Kernel Network for Remote Sensing Object Detection [C]//Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV). IEEE, Paris, France, 01-06 October, 2023: 16748-16759.
- [66] Yu F, Koltun V. Multi-Scale Context Aggregation by Dilated Convolutions [C]//Proceedings

- of the International Conference on Learning Representations (ICLR). ICLR, San Juan, PR, USA, 02-04 May, 2016: 1-13.
- [67] Li X, Wang W, Hu X, et al. Selective Kernel Networks [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, Long Beach, CA, USA, 15-20 June, 2019: 510-519.
- [68] Wang N, Gao Y, Chen H, et al. NAS-FCOS: Fast Neural Architecture Search for Object Detection [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, Seattle, WA, USA, 13-19 June, 2020: 11940-11948.
- [69] Gevorgyan Z. SIOU Loss: More Powerful Learning for Bounding Box Regression [J]. ArXiv, 2022, abs/2205.12740.
- [70] Liu Z, Wang H, Weng L, et al. Ship Rotated Bounding Box Space for Ship Extraction from High-Resolution Optical Satellite Images with Complex Backgrounds [J]. IEEE Geoscience and Remote Sensing Letters, 2016, 13(08): 1074-1088.
- [71] Zhang T, Zhang X, Li J, et al. SAR Ship Detection Dataset (SSDD): Official Release and Comprehensive Data Analysis [J]. Remote Sensing, 2021, 13(18): 3690.
- [72] Guo Z, Liu C, Zhang X, et al. Beyond Bounding-Box: Convex-hull Feature Adaptation for Oriented and Densely Packed Object Detection [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, Nashville, TN, USA, 20-25 June, 2021: 8788-8797.

攻读学位期间发表的学术论文

- [1] **Shengbo Zhu**, Lisheng Wei. PEDNet: A Proposal Enhancement Dynamic Network for Fine-Grained Ship Detection in Optical Remote Sensing Images [J]. IEEE Access, 2024, 12: 129813-129825. (**SCI** 中科院 3 区)
- [2] **Shengbo Zhu**, Lisheng Wei, Bowen Zheng, et al. Weakly Supervised Arbitrary-Oriented Ship Detection Network Based on Dynamic Scale Feature Enhancement [J]. Journal of Electronic Imaging, 2025, 34(01): 013047. (**SCI** 中科院 4 区)
- [3] **Shengbo Zhu**, Lisheng Wei. Optical Remote Sensing Ship Detection Based on Improved YOLOv8 [C]//Proceedings of the International Conference on Computer, Control and Robotics. IEEE, Shanghai, China, 19-21 April, 2024: 40-45. (**EI** 会议)
- [4] **Shengbo Zhu**, Gang Gao, Lisheng Wei. Classification Detection of Packaging Bag Defects Based on the Improved MobileNetV3 [C]//Proceedings of the Asian Control Conference. IEEE, Dalian, China, 05-08 July, 2024: 1206-1210. (**EI** 会议)
- [5] **朱圣博**, 魏利胜, 高港, 等. 基于改进 YOLOv8s 的光学遥感小型船舶目标检测算法 [J]. 电子测量与仪器学报, 2024, 38(10): 48-57. (**CSCD**)
- [6] **朱圣博**, 魏利胜. 光学遥感船舶目标检测综述 [J]. 北部湾大学学报, 2024, 39(06): 17-25. (**AMI** 入库期刊)

攻读学位期间获得的奖励

- [1] 获得 2024 年硕士研究生**国家奖学金**
- [2] 获得 2025 年安徽省普通高等学校**优秀毕业生**
- [3] 获得 2025 年安徽工程大学研究生**优秀毕业生**
- [4] 获得 2024~2025 年安徽工程大学研究生**一等奖学金**
- [5] 获得 2023~2024 年安徽工程大学研究生**一等奖学金**
- [6] 获得 2022~2023 年安徽工程大学研究生**二等奖学金**
- [7] 获得 2024 年中国(国际)传感器创新创业大赛(华东赛区)**二等奖**
- [8] 获得 2023 年中国(国际)传感器创新创业大赛(华东赛区)**二等奖**
- [9] 获得 2023 年中国(国际)传感器创新创业大赛(华东赛区)**三等奖**

致谢

在硕士研究生短短三年时光里，我发现并学会了接受自己的平庸。诚然科研自带的不确定性以及伴随的曲折与一些累计的失败不断加深我对自己的负面认知，阻碍我的自洽。但我坚信人生有梦，我仍然会为科研带给我的巨大成就感而感到慰藉，仍然感谢这段人生经历，特别是那些宝贵而又温暖的帮助。

首先，感谢导师魏利胜教授和校外导师上海欧朔智能包装科技有限公司刘小琿老师。三年以来，魏老师一直是律人先律己。在生活上，叮嘱我们身体是革命的本钱，每当遇到生活中问题时，总尽心竭力帮助我们解决问题；在学术上，为我们树立求真务实、科学严谨的学习态度，教会我们跟踪学术前沿、总结学术文献的方式。衷心祝愿魏老师在生活上健康顺遂，学术上自由广阔。

其次，感谢 509 实验室所有同学，感谢高港、汪瑞、唐绍语、项诗雨等师兄师姐给予的学习指导；感谢郑泊文、陶新杰、李明、王勤同门的督促；感谢金莉莉、邵自强、武涛等师弟师妹的陪伴，同舟共济是你我最真实的写照。

最后，感谢沈丹捷女士、朱建平先生，人间一场，他们让我自降生一刻起，便收获世间最真挚无私的爱，并小心翼翼呵护包容我至今，父母大恩，无以为报，唯有督促自己时时勤勉，刻刻努力，做一个正直勇敢、澄澈清白的人。

但远山长，云山乱，晓山青，未来路远，我始终在路上。

朱圣博

安徽工程大学电气工程学院

二〇二五年五月

尚德敏学 唯实惟新

