



自监督异构双教师蒸馏的船舶点云目标检测方法研究

夏桂华, 柏隽, 李诗洋, 梁兆伟, 李玲亚

引用本文:

夏桂华, 柏隽, 李诗洋, 等. 自监督异构双教师蒸馏的船舶点云目标检测方法研究[J]. 智能系统学报, 2026, 21(5): 1268-1281.

XIA Guihua, BAI Jun, LI Shiyang, et al. Research on self-supervised heterogeneous dual-teacher distillation for ship detection in point clouds[J]. CAAI transactions on intelligent systems, 2026, 21(5): 1268-1281.

在线阅读 View online: <https://dx.doi.org/10.11992/tis.202606023>

您可能感兴趣的其他文章

自主移动机器人路径规划中的点云噪声处理

Point cloud noise processing in path planning of autonomous mobile robot

智能系统学报. 2021, 16(4): 699-706 <https://dx.doi.org/10.11992/tis.202007040>

改进Center-Net网络的自主喷涂机器人室内窗户检测

Indoor window detection of autonomous spraying robot based on improved CenterNet network

智能系统学报. 2021, 16(3): 425-432 <https://dx.doi.org/10.11992/tis.202005016>

基于深度学习的空间非合作目标特征检测与识别

Feature detection and recognition of spatial noncooperative objects based on deep learning

智能系统学报. 2020, 15(6): 1154-1162 <https://dx.doi.org/10.11992/tis.202006011>

生成对抗网络辅助学习的舰船目标精细识别

Fine-grained inshore ship recognition assisted by deep-learning generative adversarial networks

智能系统学报. 2020, 15(2): 296-301 <https://dx.doi.org/10.11992/tis.201901004>

基于反卷积和特征融合的SSD小目标检测算法

SSD small target detection algorithm based on deconvolution and feature fusion

智能系统学报. 2020, 15(2): 310-316 <https://dx.doi.org/10.11992/tis.201905035>

自主水面货船研究现状与展望

Research status and prospects of autonomous surface cargo ships

智能系统学报. 2019, 14(1): 57-70 <https://dx.doi.org/10.11992/tis.201806028>

DOI: 10.11992/tis.202606023

网络出版地址: <https://link.cnki.net/urlid/23.1538.tp.20260820.1648.002>

自监督异构双教师蒸馏的船舶点云目标检测方法研究

夏桂华¹, 柏隽¹, 李诗洋², 梁兆伟³, 李玲亚⁴

(1. 哈尔滨工程大学 智能科学与工程学院, 黑龙江 哈尔滨 150001; 2. 大连理工大学 船舶工程学院, 辽宁 大连 116024; 3. 武汉第二船舶设计研究所, 湖北 武汉 430060; 4. 亚太卫星宽带通信有限公司, 广东 深圳 518033)

摘要: 随着人工智能技术的快速发展, 无人艇自主航行与作业技术成为当前学术界的研究热点之一, 海上船舶目标检测与环境感知是保障无人艇自主决策与安全航行的关键技术。激光雷达可提供稳定的三维几何信息, 使船舶点云目标检测成为海上环境感知的重要技术路径。针对海上船舶点云目标检测中高质量数据集构建困难、远距离目标回波稀疏以及海面动态杂波干扰强等问题, 提出了一种自监督异构双教师对比网络指导的蒸馏方法, 以缓解船舶点云目标检测网络对船舶点云数据集的依赖, 并提升其在稀疏目标表征和复杂海面背景干扰条件下的检测鲁棒性。实验验证基于“海豚 1 号”智能船采集的海上船舶点云数据集开展, 结果表明, 该方法较基线方法平均精度均值 (mean average precision, mAP) 提升 6.4 百分点, 在稀疏点云和强杂波场景下保持较稳定的检测性能, 并具备面向无人艇船载感知平台部署与优化的应用潜力。

关键词: 无人艇; 船舶点云; 目标检测; 激光雷达; 自监督学习; 双教师蒸馏; 对比学习; 特征迁移

中图分类号: TP242 **文献标志码:** A **文章编号:** 1673-4785(2026)05-1268-14

中文引用格式: 夏桂华, 柏隽, 李诗洋, 等. 自监督异构双教师蒸馏的船舶点云目标检测方法研究 [J]. 智能系统学报, 2026, 21(5): 1268-1281.

英文引用格式: XIA Guihua, BAI Jun, LI Shiyang, et al. Research on self-supervised heterogeneous dual-teacher distillation for ship detection in point clouds[J]. CAAI transactions on intelligent systems, 2026, 21(5): 1268-1281.

Research on self-supervised heterogeneous dual-teacher distillation for ship detection in point clouds

XIA Guihua¹, BAI Jun¹, LI Shiyang², LIANG Zhaowei³, LI Lingya⁴

(1. College of Intelligent Systems Science and Engineering, Harbin Engineering University, Harbin 150001, China; 2. School of Naval Architecture and Ocean Engineering, Dalian University of Technology, Dalian 116024, China; 3. Wuhan Second Ship Design and Research Institute, Wuhan 430060, China; 4. APT Mobile SatCom Limited, Shenzhen 518033, China)

Abstract: With the rapid development of artificial intelligence, autonomous navigation and operation technologies for unmanned surface vehicle(USV) have become one of the current research focuses in academia. Maritime ship target detection and environmental perception are key technologies for ensuring autonomous decision-making and safe navigation of USVs. LiDAR can provide stable three-dimensional geometric information, which makes ship point-cloud target detection an important technical approach for maritime environmental perception. To address the problems of difficult construction of high-quality datasets, sparse echoes from long-range targets, and strong interference from dynamic sea clutter in maritime ship point-cloud target detection, this study proposed a distillation method guided by a self-supervised heterogeneous dual-teacher contrastive network. The method aims to reduce the dependence of ship point-cloud target detection networks on ship point-cloud datasets and improve their detection robustness under sparse target representation and complex maritime background interference. Experiments were conducted on a maritime ship point-cloud dataset collected by the “Dolphin-1” intelligent vessel. The results show that the proposed method improves the mean average precision (mAP) by 6.4 percentage points compared with the baseline method, maintains relatively stable detection performance in sparse point-cloud and strong-clutter scenes, and has application potential for deployment and optimization on onboard perception platforms of USVs.

Keywords: unmanned surface vehicle; ship point cloud; object detection; LiDAR; self-supervised learning; dual-teacher distillation; contrastive learning; feature transfer

收稿日期: 2026-06-12. 网络出版日期: 2026-08-21.

基金项目: 国家自然科学基金项目 (62403159, 52171302).

通信作者: 夏桂华. E-mail: xiaguihua@hrbeu.edu.cn.

人工智能推动了船舶智能控制与多源感知技术的快速发展, 为海洋装备智能化、自主化和无

人化提供了重要技术支撑, 也使无人艇 (unmanned surface vehicle, USV) 自主航行与海上作业成为当前学术界和工程应用领域的研究热点之一^[1-3]。无人艇作为海洋智能装备的重要代表, 因具备灵活部署、低成本运行和多任务适应等优势, 已广泛应用于海域巡逻、环境监测、目标探测和安防防控等场景^[4]。随着应用场景不断拓展, 无人艇在复杂海况下对环境信息获取与感知能力提出了更高要求, 环境感知成为支撑其自主航行、避碰决策与任务执行的关键技术^[5]。作为无人艇自主航行中的核心任务之一, 海上目标检测的准确性直接关系到任务执行质量与航行安全^[6]。相较于传统视觉传感器, 激光雷达能够更直接地获取目标三维几何信息, 因此点云目标检测逐渐成为海上环境感知的重要技术路径。点云目标检测在一定程度上缓解了视觉方法在弱光、强反光条件下感知稳定性不足以及目标三维结构表征不准确等问题。然而, 海上船舶点云通常具有远距离回波稀疏、船体边界不完整、海面动态杂波干扰强和船舶姿态变化明显等复杂特征, 对检测模型的特征表达、目标定位和鲁棒感知能力提出了更高要求。

在点云三维目标检测研究中, 现有国内外方法主要沿高效检测网络设计和自监督预训练两个方向发展^[7-8]。在通用三维目标检测领域, KITTI、nuScenes 等公开数据集推动了点云目标检测算法的发展^[9-10], PointPillars^[11]、SECOND^[12]、Voxel-Net^[13]、PointRCNN^[14]、PV-RCNN^[15]、3DSSD^[16] 和 CenterPoint^[17] 等方法分别从伪图像编码、稀疏卷积、体素特征提取、点级候选框生成、点体素融合、单阶段点结构检测和中心点预测等角度提升了三维目标检测性能。Chen 等^[18] 提出 VoxelNeXt, 直接基于稀疏体素特征完成目标预测, 推动了三维检测网络向全稀疏和高效化方向发展。Zhang 等^[19] 提出 SAFDNet, 通过自适应特征扩散增强全稀疏检测器的目标中心区域表征。

针对三维目标检测对大规模标注数据依赖较强的问题, MAE^[20]、PointContrast^[21]、Point-MAE^[22]、Point-M2AE^[23]、Occupancy-MAE^[24]、Voxel-MAE^[25]、BEV-MAE^[26] 和 R-MAE^[27] 等自监督方法通过无标注点云学习空间结构表征, 其中 Voxel-MAE、BEV-MAE 和 R-MAE 主要围绕掩码重构、点密度预测及空间占据关系进行预训练, 能够增强稀疏点云结构建模, 但总体仍以重构式表征学习为主。近年来, 相关研究进一步关注点云几何属性保持和稀疏结构建模, Nisar 等^[28] 提出 PSA-SSL, 通过姿态与尺度感知的自监督任务增强点云几何表征能

力; Abdelsamad 等^[29] 提出多尺度邻域占据掩码自编码方法, 通过多尺度占据重构提升稀疏 LiDAR 点云的自监督表征能力。

与此同时, 知识蒸馏逐渐应用于三维目标检测, PointDistille 通过局部结构蒸馏和关键点加权, 将教师网络的局部几何结构知识迁移至轻量化学生检测器, 以提高稀疏点云条件下的知识迁移效率^[30]。X-Ray Distillation 采用弱到强知识迁移策略, 利用目标结构更加完整的点云构建教师监督, 以缓解目标稀疏和遮挡造成的检测性能下降^[31]。CaKDP 进一步研究两阶段教师与单阶段学生之间的异构知识迁移, 通过类别感知蒸馏减小不同检测结构之间的预测差异, 并结合网络剪枝实现轻量化三维检测^[32]。上述方法分别从局部几何结构、完整目标监督和异构检测器对齐等角度提升了点云检测蒸馏效率, 但主要以单一教师检测器的特征或预测响应作为知识来源, 研究重点集中于模型压缩、检测知识迁移或教师—学生结构差异。总体来看, 现有方法在重构表征学习、稀疏结构建模和异构知识迁移方面取得了一定进展, 但对重构与判别表征的解耦学习、异构网络跨层级知识对齐以及极度稀疏点云局部结构不连续问题的协同处理仍显不足。针对海上船舶点云低标注、远距离回波稀疏和海面背景干扰等特点, 本研究从无标注点云中分别构建重构教师和判别教师, 使几何结构表征与目标判别表征进行解耦学习, 并通过异构特征对齐、注意力加权以及稀疏体素特征层与高层检测特征层的多粒度约束, 将两类互补知识协同迁移至 VoxelNeXt 学生网络, 从而增强稀疏船舶点云条件下的结构表征与目标判别能力。

在海上 LiDAR 感知领域, Ponzini 等^[33] 进一步开展了面向船舶态势感知的 LiDAR 目标检测与分类研究。相比道路交通等结构化场景, 海上船舶点云检测研究仍相对有限。现有海上目标检测研究既包括视觉图像^[34-35] 和 SAR 图像船舶检测方法^[36], 也逐渐扩展到激光雷达点云船舶目标检测任务。何芸倩等^[37] 构建了自航船模点云数据集, 并提出轻量型船舶点云检测网络 LASSD, 通过网络剪枝和局部注意力模块提高检测效率与精度。Zhang 等^[38] 提出双分支稀疏三维卷积和多尺度二维卷积结合的船舶点云检测网络, 以缓解大尺度船舶目标在体素化和特征压缩过程中的几何信息与高度信息丢失问题。Xie 等^[39] 面向自主水面艇提出激光雷达船舶目标检测与跟踪框架, 结合模块化检测网络和卡尔曼滤波方法, 提高了

繁忙水域中船舶运动、遮挡和漏检情况下的连续感知能力。总体来看,现有方法在高效检测、无标注预训练和海上目标感知方面取得了一定进展,但对海上船舶点云中远距离回波稀疏、船体边界不完整、海面动态杂波强以及低标注条件下特征学习不足等问题考虑仍不充分。因此,有必要结合海上船舶点云的场景特性,进一步研究面向船舶目标检测的自监督知识迁移方法。

针对上述问题,提出一种面向海上船舶点云目标检测的异构双教师自监督对比蒸馏框架。框架设置重构教师和判别教师两个功能互补的分支,分别通过掩码重构和双视图一致性学习船舶点云的几何结构知识与目标判别知识,使两类表征在独立分支中进行学习。在此基础上,利用投影映射完成教师网络与 VoxelNeXt 学生网络之间的异构特征对齐,并在稀疏体素特征层和高层检测特征层进行多层级知识迁移。同时,引入注意力加权机制强化船舶目标及关键结构区域的知识迁移,缓解远距离回波稀疏和局部结构缺失造成的特征表达不连续问题,从而提高有限标注条件下的检测鲁棒性。主要贡献如下:

1) 针对海上船舶点云缺乏稳定几何参照、目标回波稀疏和背景干扰强等特点,构建了异构双教师自监督学习框架。该框架分别设计重构教师网络和判别教师网络,通过掩码重构任务学习船舶目标的几何结构先验知识,通过双视图一致性与对比约束学习复杂扰动条件下稳定的目标判别表征,从而为下游船舶点云检测提供更加充分的先验知识。

2) 提出了面向学生检测器的协同蒸馏机制,将重构教师与判别教师获得的互补知识迁移至 VoxelNeXt 学生检测网络。与单一教师蒸馏方式相比,该方法能够同时增强学生网络对船舶局部几何结构、整体空间布局和目标-背景判别信息的表达能力,有助于缓解稀疏点云和有限标注条件下特征学习不足的问题。

3) 设计了注意力引导的多粒度知识迁移策略。该策略利用教师网络的空间响应生成注意力权重,使蒸馏监督更加关注船舶目标及其关键结构区域,降低海面杂波、尾流回波和非目标背景区域对特征学习的干扰。同时,从体素特征层和检测头层进行多粒度约束,进一步提升三维框定位精度和检测鲁棒性。

1 海上船舶多模态感知数据集构建

为研究海上船舶点云目标检测方法,利用哈

尔滨工程大学“海豚 1 号”智能船,搭建船载多模态感知采集平台,同步采集激光雷达点云、可见光视频和红外视频数据。采集设备安装位置如图 1 所示。



图 1 采集设备安装位置示意

Fig. 1 Schematic of the installation positions of the data acquisition equipment

激光雷达用于获取船舶目标的三维空间结构信息,可见光相机用于记录船舶外观、航行背景和环境变化,红外相机用于补充弱光、逆光和复杂照度条件下的目标热辐射信息。多源传感器在采集过程中进行时间同步和统一编号,使同一时刻的点云、可见光图像和红外图像能够对应于同一海上场景。

数据采集覆盖不同观测距离、航向角、目标姿态和海面扰动条件。经过无效帧剔除、时间同步校验和数据格式整理后,共获得 10 547 帧海上船舶样本。由于研究重点为基于激光雷达点云的船舶点云目标检测,后续实验主要使用其中的点云数据。

在数据标注方面,参照 KITTI 三维目标检测数据集的组织形式,对点云中的船舶目标进行三维包围框和类别标注。每个目标标注信息包括类别、三维框中心坐标、目标尺寸和航向角,可表示为

$$\mathbf{b}_i = (c_i, x_i, y_i, z_i, l_i, w_i, h_i, \theta_i)$$

式中: c_i 表示目标类别, (x_i, y_i, z_i) 表示三维框中心位置, (l_i, w_i, h_i) 表示目标长、宽、高, θ_i 表示目标航向角。针对海上船舶点云中远距离目标稀疏、船体边界不完整和海面杂波干扰明显等问题,标注过程中结合点云俯视图、三维视图以及同步可见光和红外图像进行复核,以提高类别判断和三维框标注的一致性。

数据集按照训练集、验证集和测试集进行划分, 比例约为 8:1:1, 对应 8 438、1 055 和 1 054 帧。为降低相邻帧相似性造成的数据泄露风险, 划分时避免同一连续航行片段同时出现在训练集和测试集中。自监督预训练阶段仅使用训练集中的无标注点云构造掩码重构任务和双视图一致性任务; 监督微调阶段使用训练集中的少量标注信息; 验证集用于参数选择和训练过程监控; 测试集仅用于最终性能评估。

2 自监督双教师蒸馏网络框架设计

2.1 异构双教师网络蒸馏框架设计

海上船舶点云目标检测与道路交通等结构化

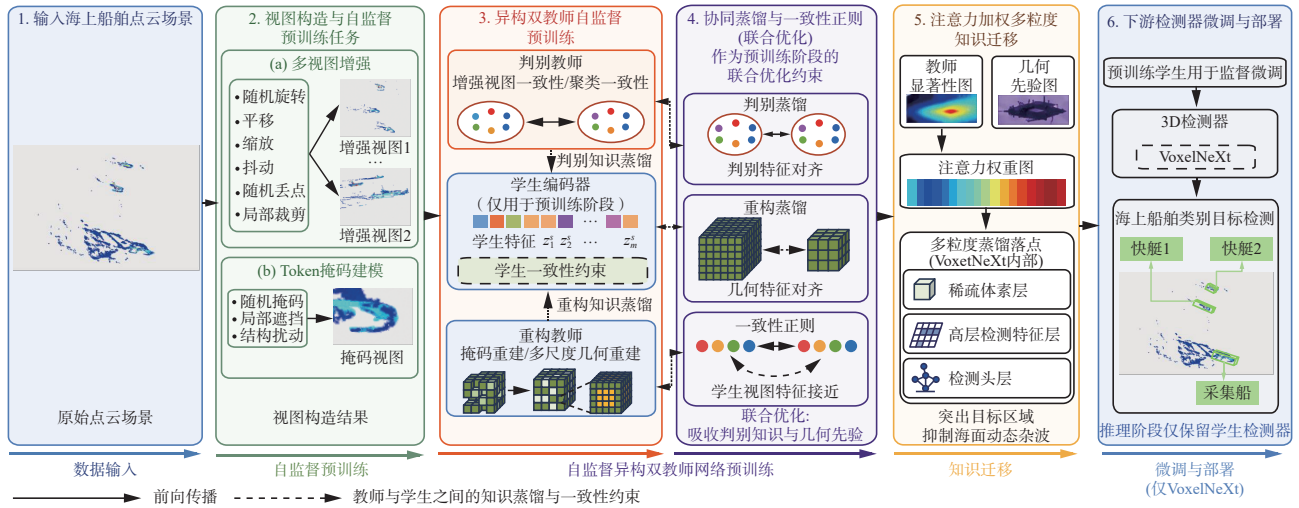


图 2 异构双教师指导自监督对比蒸馏网络

Fig. 2 Heterogeneous dual-teacher-guided self-supervised contrastive distillation network

设单帧输入点云为

$$\mathbf{P} = \{\mathbf{p}_i\}_{i=1}^N, \mathbf{p}_i = [x_i, y_i, z_i, r_i]^T$$

式中: N 表示点数, (x_i, y_i, z_i) 点的三维坐标, r_i 表示反射强度等可选属性。为构建自监督学习样本, 对同一帧点云施加两组相互独立的随机增强, 生成两个扰动视图:

$$\mathbf{P}_a = \mathcal{T}_a(\mathbf{P}), \mathbf{P}_b = \mathcal{T}_b(\mathbf{P})$$

式中: $\mathcal{T}_a$ 和 $\mathcal{T}_b$ 表示随机旋转、平移、缩放、抖动、随机丢点和局部裁剪等增强操作。双扰动视图用于判别教师学习跨视图稳定的目标判别表示。与此同时, 为引入几何结构恢复约束, 将输入点云目标数据划分为若干局部 token, 并随机遮蔽其中一部分, 得到掩码视图 $\mathbf{P}_m = \mathcal{M}(\mathbf{P})$, $\mathcal{M}(\mathbf{P})$ 表示随机掩码操作。该视图用于重构教师学习船舶点目标的局部结构和空间布局。

框架包含重构教师网络、判别教师网络和 VoxelNeXt 学生网络 3 部分。重构教师网络以掩码视图为输入, 通过恢复被遮蔽区域的特征表示

场景不同, 海洋环境缺乏车道线、建筑物和道路边界等稳定几何参照, 难以为目标定位与背景分离提供可靠先验知识。在此基础上, 远距离回波稀疏、海面动态杂波强、船舶航向与姿态变化大以及数据集稀少等问题进一步增加了海上船舶点云检测难度, 容易造成目标结构表达不足、背景干扰增强和定位精度不稳定。针对上述挑战, 构建一种异构双教师对比蒸馏框架。如图 2 所示, 该框架包括点云视图构造、双教师自监督预训练、协同蒸馏、注意力加权多粒度知识迁移以及下游检测器微调与部署, 通过几何先验知识与判别知识的协同迁移提升 VoxelNeXt 学生网络的检测鲁棒性。

学习几何结构先验知识; 判别教师网络以双扰动视图为输入, 通过对比学习和聚类一致性约束学习目标判别知识; VoxelNeXt 学生网络在训练阶段接收双教师蒸馏知识监督, 并在监督微调阶段适配具体检测任务。部署阶段仅保留 VoxelNeXt 学生网络, 重构教师网络、判别教师网络和蒸馏分支不参与在线推理, 因此不会增加部署端网络结构和推理复杂度。

2.2 异构双教师自监督预训练与协同蒸馏

重构教师网络主要用于学习船舶点云的几何结构先验知识。设重构教师编码器为 $E_R(\cdot)$, 输入掩码视图 $\mathbf{P}_m$ 后得到多层特征:

$$\mathbf{F}_R^l = E_R^l(\mathbf{P}_m)$$

式中 l 表示特征层级。重构目标采用被遮蔽区域的中间 token 特征, 使网络在局部信息不完整的条件下恢复缺失结构。重构教师的训练损失可写为

$$\mathcal{L}_{\text{rec}} = \sum_{l=1}^L \left\| \hat{\mathbf{F}}_R^l - \mathbf{F}_{\text{gt}}^l \right\|_2^2$$

式中: $\hat{\mathbf{F}}_R^l$ 表示重构结果, $\mathbf{F}_{gt}^l$ 表示对应的目标特征。通过掩码重构任务, 重构教师能够学习船舶轮廓、局部形态和空间布局等几何结构信息, 为稀疏回波和局部缺失条件下的目标检测提供结构约束。

判别教师网络主要用于学习跨扰动视图稳定的目标判别知识。设判别教师编码器为 $E_D(\cdot)$, 投影头为 $g_D(\cdot)$, 则双视图的投影表示为

$$\mathbf{z}_a = g_D(E_D(\mathbf{P}_a)), \mathbf{z}_b = g_D(E_D(\mathbf{P}_b))$$

为使同一帧点云的不同扰动视图在特征空间中保持接近, 同时增强不同样本之间的可分性, 引入跨视图对比损失:

$$\mathcal{L}_{con} = -\log \frac{\exp(\text{sim}(\mathbf{z}_i^a, \mathbf{z}_j^b) / \tau)}{\sum_j \exp(\text{sim}(\mathbf{z}_i^a, \mathbf{z}_j^b) / \tau)}$$

式中: $\text{sim}(\cdot)$ 表示相似度函数, τ 为温度系数。考虑到海上船舶点云背景区域相似度高、目标回波稀疏, 仅依赖样本级对比学习可能导致监督不稳定, 因此进一步引入聚类一致性损失 $\mathcal{L}_{clu}$, 使相似结构区域在不同扰动视图下保持一致的原型分配。判别教师的训练目标为

$$\mathcal{L}_D = \mathcal{L}_{con} + \beta \mathcal{L}_{clu}$$

式中 β 为聚类一致性损失权重。完成双教师自监督预训练后, 固定教师网络参数, 并利用其指导 VoxelNeXt 学生网络学习。由于教师网络和 VoxelNeXt 学生网络可能具有不同的特征维度和空间分辨率, 直接匹配容易造成蒸馏不稳定。因此, 在蒸馏分支中引入轻量投影映射模块, 将教师特征和学生特征映射到统一表示空间:

$$\tilde{\mathbf{F}}_S^l = \phi_S^l(\mathbf{F}_S^l), \tilde{\mathbf{F}}_T^l = \phi_T^l(\mathbf{F}_T^l)$$

式中 $\phi_S^l(\cdot)$ 和 $\phi_T^l(\cdot)$ 分别表示学生端和教师端的投影映射。

重构蒸馏损失函数用于将重构教师的几何结构知识迁移至 VoxelNeXt 学生网络, 其损失定义为

$$\mathcal{L}_{rkd} = \sum_{l=1}^L \alpha_l \left\| \tilde{\mathbf{F}}_S^l - \tilde{\mathbf{F}}_R^l \right\|_2^2$$

式中 α_l 表示第 l 层蒸馏损失的权重。该损失约束 VoxelNeXt 学生网络学习船舶点云的局部形态、结构连续性和空间布局。

判别蒸馏损失函数用于将判别教师的目标判别知识迁移至 VoxelNeXt 学生网络, 其损失定义为

$$\mathcal{L}_{dkd} = \|g_S(E_S(\mathbf{P}_a)) - g_D(E_D(\mathbf{P}_a))\|_2^2 + \|g_S(E_S(\mathbf{P}_b)) - g_D(E_D(\mathbf{P}_b))\|_2^2$$

式中 $E_S(\cdot)$ 和 $g_S(\cdot)$ 分别表示学生编码器和学生投影

头。该损失促使 VoxelNeXt 学生网络学习对随机扰动、点云稀疏化和背景噪声更稳定的判别表征。

此外, 在 VoxelNeXt 学生网络引入一致性正则, 使其对同一帧点云的不同增强视图保持一致响应:

$$\mathcal{L}_{cons} = \|g_S(E_S(\mathbf{P}_a)) - g_S(E_S(\mathbf{P}_b))\|_2^2$$

因此, 自监督蒸馏阶段的总优化目标为

$$\mathcal{L}_{ssl} = \lambda_1 \mathcal{L}_{rkd} + \lambda_2 \mathcal{L}_{dkd} + \lambda_3 \mathcal{L}_{cons}$$

式中 λ_1 、 λ_2 和 λ_3 为损失权重。VoxelNeXt 学生网络能够在不使用人工三维框标注的条件下, 同时吸收几何结构先验知识和目标判别知识。

2.3 注意力引导的多粒度知识迁移机制设计

完成异构双教师自监督预训练与协同蒸馏后, 重构教师网络和判别教师网络分别获得了面向船舶点云目标的几何结构知识与目标判别知识。为进一步将双教师知识有效迁移至下游 VoxelNeXt 学生网络, 在检测微调阶段引入注意力引导的多粒度知识迁移机制。该机制在冻结双教师网络参数的基础上, 以 VoxelNeXt 学生网络作为唯一可训练网络, 将教师网络提供的中间层特征和空间注意力响应作为辅助监督信号, 与标注样本提供的检测监督共同约束 VoxelNeXt 学生网络优化, 整体流程如图 3 所示。

在微调阶段, 教师网络参数保持冻结, 仅作为特征监督源参与训练, 不参与梯度更新。VoxelNeXt 学生网络作为唯一可训练网络, 同时接受标注样本提供的检测监督和冻结教师网络提供的中间层知识监督。考虑到海上点云中船舶目标区域与海面背景区域的信息贡献存在显著差异, 直接对全部空间位置施加均匀蒸馏约束容易使 VoxelNeXt 学生网络受到海面动态杂波、稀疏噪声点以及非目标背景区域的干扰。因此, 引入注意力加权机制, 根据教师网络的空间响应强度自适应分配不同特征位置的蒸馏权重, 使知识迁移过程更加关注船舶目标及其关键结构区域。

设第 l 层融合教师特征为

$$\mathbf{F}_T^l \in \mathbb{R}^{C \times H \times W}$$

式中 C 、 H 和 W 分别表示通道数、特征高度和特征宽度。这里的 $\mathbf{F}_T^l$ 由重构教师特征和判别教师特征共同构成。考虑到重构教师更侧重船舶点云的局部几何结构和空间连续性, 判别教师更侧重目标与背景之间的判别知识, 可将融合教师特征表示为

$$\mathbf{F}_T^l = \omega_l \tilde{\mathbf{F}}_R^l + (1 - \omega_l) \tilde{\mathbf{F}}_D^l$$

式中: $\tilde{\mathbf{F}}_R^l$ 和 $\tilde{\mathbf{F}}_D^l$ 分别表示经过投影映射或插值对齐后的重构教师特征和判别教师特征, ω_l 为第 l 层的

融合权重。对于浅层特征,可赋予重构教师更高权重,以突出船舶轮廓、局部形态和结构连续性;对于深层特征,可赋予判别教师更高权重,以增

强目标区域与海面背景之间的可分性。若某一检测器结构中仅存在单一适配层级,也可直接采用对应教师特征作为该层的蒸馏监督。

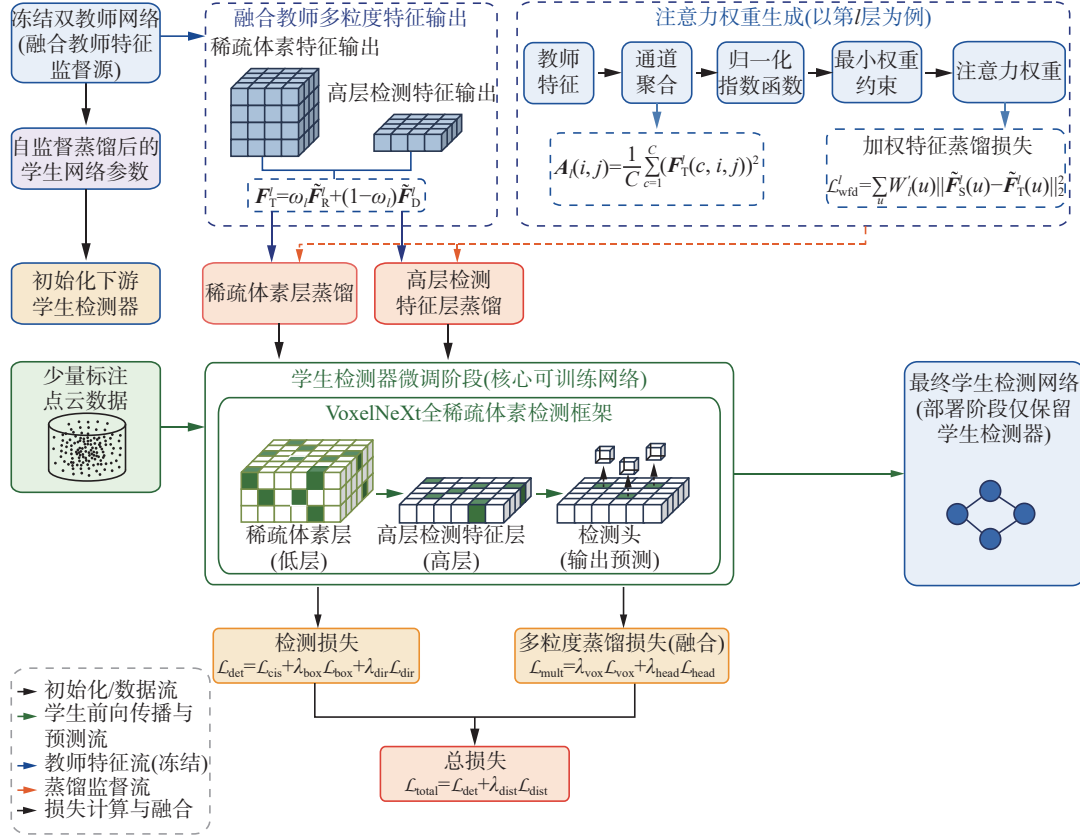


图 3 注意力引导的多粒度知识迁移机制

Fig. 3 Attention-guided multi-granularity knowledge transfer mechanism

首先通过通道聚合获得教师特征的空间响应图:

$$A_l(i, j) = \frac{1}{C} \sum_{c=1}^C (F_T^l(c, i, j))^2$$

式中 $A_l(i, j)$ 表示第 l 层特征在空间位置 (i, j) 处的响应强度。该响应图反映了教师网络在不同空间位置上的关注程度。平方操作能够增强高响应位置与低响应位置之间的差异,使船舶轮廓、船体边缘以及候选目标邻域等关键区域在后续蒸馏过程中获得更高关注度。

随后,引入注意力温度系数 τ_a 对空间响应进行平滑归一化:

$$W_l = \text{Softmax} \left(\frac{A_l}{\tau_a} \right)$$

式中 τ_a 用于调节注意力分布的平滑程度。当 τ_a 较小时,注意力分布更加集中,有利于突出高响应目标区域;当 τ_a 较大时,注意力分布更加平滑,有利于保持全局上下文信息。为便于不同层级特征之间的权重尺度保持一致,进一步将注意力权重归一化至 $[0, 1]$ 区间:

$$\bar{W}_l = \frac{W_l - \min(W_l)}{\max(W_l) - \min(W_l) + \varepsilon}$$

式中 ε 为防止分母为零的稳定项。由于海面背景区域中仍包含一定的上下文信息,完全抑制低响应区域可能导致 VoxelNeXt 学生网络对背景分布建模不足。为保持蒸馏过程的稳定性,设置最小权重下界 $a_{\min}$:

$$W'_l = \max(\bar{W}_l, a_{\min})$$

通过上述处理,最终得到的注意力权重 W'_l 既能够突出教师网络关注的目标相关区域,又能够保留背景区域的基础监督信号,从而在增强目标结构学习的同时避免训练过程过度偏置。

在注意力权重引导下,第 l 层的加权特征蒸馏损失定义为

$$\mathcal{L}_{\text{wfd}}^l = \sum_u W'_l(u) \left\| \tilde{F}_s^l(u) - \tilde{F}_T^l(u) \right\|_2^2$$

式中: u 表示空间位置或特征单元索引, $\tilde{F}_s^l$ 和 $\tilde{F}_T^l$ 分别表示经过特征对齐后的学生特征和融合教师特征。当教师网络与 VoxelNeXt 学生网络的通道维度或特征分辨率不一致时,通过轻量投影映射、

插值或稀疏特征索引对齐等操作, 将二者统一到相同表示空间后再计算蒸馏损失。该损失函数使 VoxelNeXt 学生网络在学习教师知识时更加关注高响应目标区域, 降低海面杂波、离散噪声和无效背景区域对特征优化的干扰。

结合 VoxelNeXt 全稀疏体素检测框架的结构特点, 将多粒度知识迁移作用于稀疏体素特征层和高层检测特征层两个层级。其中, 稀疏体素特征层蒸馏主要用于迁移船舶点云目标的局部几何结构知识, 增强 VoxelNeXt 学生网络对稀疏回波、局部缺失和船体轮廓变化的表达能力; 高层检测特征层蒸馏主要用于迁移目标区域的判别信息和空间上下文响应, 提升 VoxelNeXt 学生网络对船舶目标与海面背景杂波的区分能力。

总体多粒度蒸馏损失可表示为

$$\mathcal{L}_{\text{mult}} = \lambda_{\text{vox}} \mathcal{L}_{\text{vox}} + \lambda_{\text{head}} \mathcal{L}_{\text{head}}$$

式中: $\mathcal{L}_{\text{vox}}$ 表示稀疏体素特征层的注意力加权蒸馏损失, $\mathcal{L}_{\text{head}}$ 表示高层检测特征层的注意力加权蒸馏损失, λ_{vox} 和 λ_{head} 分别表示对应层级的损失权重。通过上述设计, VoxelNeXt 学生网络能够同时获得低层几何结构约束和高层目标判别约束, 从而增强其对稀疏船舶点云、复杂海面背景和多姿态船舶目标的适应能力。

下游检测任务的监督损失定义为

$$\mathcal{L}_{\text{det}} = \mathcal{L}_{\text{cls}} + \lambda_{\text{box}} \mathcal{L}_{\text{box}} + \lambda_{\text{dir}} \mathcal{L}_{\text{dir}}$$

式中: $\mathcal{L}_{\text{cls}}$ 表示船舶目标分类损失, $\mathcal{L}_{\text{box}}$ 表示三维边界框回归损失, $\mathcal{L}_{\text{dir}}$ 表示航向角估计损失, λ_{box} 和

λ_{dir} 分别为对应损失项的权重。结合注意力引导的多粒度知识迁移后, VoxelNeXt 学生网络微调阶段的总优化目标为

$$\mathcal{L}_{\text{total}} = \mathcal{L}_{\text{det}} + \lambda_{\text{dist}} \mathcal{L}_{\text{multi}}$$

式中 λ_{dist} 为蒸馏正则项权重, 用于平衡检测任务监督与双教师知识迁移之间的影响。通过该联合优化目标, VoxelNeXt 学生网络能够在有限标注样本条件下进一步吸收双教师网络提供的几何结构知识和目标判别知识, 提升其在稀疏点云、复杂海面背景和船舶姿态变化条件下的检测鲁棒性。训练完成后, 推理阶段仅保留 VoxelNeXt 学生网络, 重构教师网络、判别教师网络及蒸馏相关分支均被移除, 因此不会增加部署阶段的模型参数和网络结构复杂度。

2.4 船舶点云目标检测的 VoxelNeXt 学生网络构建

经过前述异构双教师自监督预训练、协同蒸馏以及注意力引导的多粒度知识迁移后, VoxelNeXt 学生网络已获得面向海上船舶点云的初始特征表征能力。为将该表征能力进一步转化为实际三维目标检测能力, 本节构建最终用于船舶点云目标检测的 VoxelNeXt 学生网络。需要说明的是, 本节并不对 VoxelNeXt 的基础检测结构进行重新设计, 而是将其作为学生网络的基本检测框架, 重点说明经过双教师蒸馏后的 VoxelNeXt 学生网络如何用于初始化下游检测器, 并在标注船舶点云数据上完成检测微调。最终 VoxelNeXt 学生网络结构如图 4 所示。

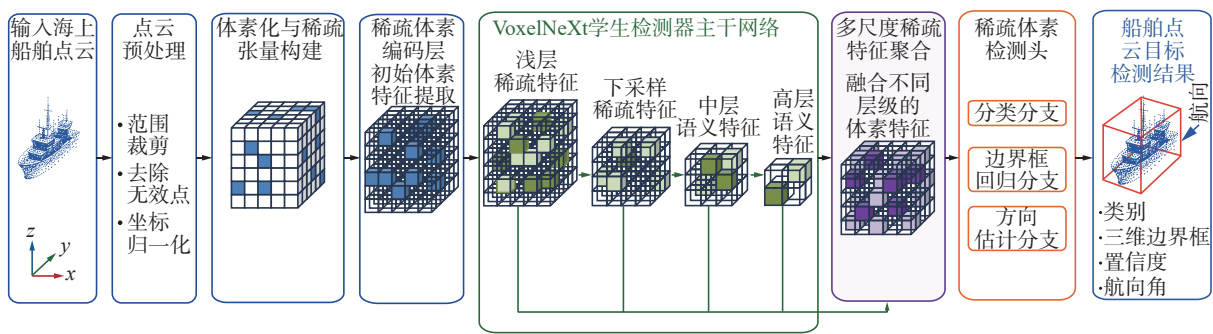


图 4 VoxelNeXt 学生网络架构

Fig. 4 VoxelNeXt student network architecture

输入单帧海上船舶点云目标后, 首先进行检测范围裁剪、无效点去除和坐标归一化等预处理操作, 以减少海面杂波、异常反射点及无效远距离点对检测结果的干扰。随后, 将预处理后的点云划分为规则体素, 并仅保留包含有效回波的非空体素, 构建稀疏张量输入。VoxelNeXt 学生网络主干采用 VoxelNeXt 的稀疏卷积编码结构, 对稀疏体素特征进行逐层提取, 获得由局部几何结

构到高层语义信息的层次化船舶点云目标表征, 为后续船舶类别判别、三维边界框回归和航向角估计提供特征基础。

VoxelNeXt 学生网络的参数由蒸馏参数初始化和检测微调两个步骤获得。首先, 将双教师蒸馏阶段获得的学生网络参数用于初始化 VoxelNeXt 学生网络, 并在标注样本上进行监督微调, 以适配海上船舶点云目标检测任务:

$$\theta_{\text{backbone}}^0 \leftarrow \theta_{\text{S}}^{\text{pre}}$$

式中 $\theta_{\text{S}}^{\text{pre}}$ 表示经过自监督预训练与知识蒸馏后得到的学生端参数, $\theta_{\text{backbone}}^0$ 表示检测微调阶段 VoxelNeXt 学生网络主干网络的初始参数。通过该初始化方式, VoxelNeXt 学生网络在进入有监督检测训练前已具备一定的船舶几何结构感知能力和目标区域判别能力, 有助于缓解标注样本有限条件下特征学习不足的问题。随后, 在标注船舶点云数据上对 VoxelNeXt 学生网络进行监督微调, 使其进一步适应船舶目标分类、三维空间定位和航向角估计任务。

训练完成后, 微调后的 VoxelNeXt 学生网络即作为最终的船舶点云检测部署模型。推理阶段仅保留该学生网络, 重构教师网络、判别教师网络、投影映射模块、注意力权重生成模块及蒸馏分支均不参与在线推理。输入单帧船舶点云目标后, 模型依次完成点云预处理、稀疏体素编码和 Voxel-to-Object 目标预测, 直接输出船舶目标的类别、三维边界框、置信度和航向角。因此, VoxelNeXt 学生网络能够在不增加推理阶段网络结构和计算负担的情况下, 利用前期双教师自监督蒸馏获得的几何结构先验知识和目标判别知识提升海上船舶点云目标检测性能。

3 实验结果与分析

3.1 实验设置

船舶多模态感知数据集包含 10 547 帧海上船舶多源目标数据, 实验主要使用其中的激光雷达点云数据。三维框标注仅用于下游检测器的监督微调、验证和测试评估; 自监督预训练阶段仅使用训练集点云数据, 不使用类别标签和三维边界框标注。下游检测器选用 VoxelNeXt 学生网络, 能够直接基于稀疏体素特征完成目标预测, 具有结构简洁、推理开销较低和部署友好的特点。因此, 本文以 VoxelNeXt 学生网络作为最终船舶点云检测网络, 用于验证所提异构双教师自监督蒸馏、注意力加权和多粒度知识迁移机制的有效性。

评价指标采用 IoU 阈值为 0.5 和 0.7 时的 3D mAP。对比方法包括未引入预训练策略的监督训练基线 Scratch, 以及 Voxel-MAE、BEV-MAE 和 R-MAE 等主流自监督预训练方法。为分析低标注条件下的检测性能, 设置 10%、30% 和 50% 3 种标注比例。低标注实验仅在训练集内部进行标注样本抽取, 采用固定随机种子对训练样本进行随机采样, 并按照 10%、30% 和 50% 的比例分别构建

标注子集, 其余训练样本不使用三维框标注, 仅作为无标注数据参与自监督预训练。不同标注比例实验均保持验证集和测试集不变, 以避免数据划分差异对实验结果造成影响。采用随机采样策略主要是为了在不人为引入目标距离、场景类型或船舶姿态等先验偏好的情况下模拟不同标注预算, 并保证不同方法在相同标注样本条件下进行公平比较。

为保证对比公平性, 所有方法采用相同的数据划分、数据增强策略、训练轮数、学习率设置、体素化参数和 NMS 阈值。除预训练或蒸馏策略不同外, 下游 VoxelNeXt 学生网络结构和推理阶段设置保持一致。损失函数中的权重超参数根据验证集检测性能确定。自监督蒸馏阶段, 重构教师蒸馏损失、判别教师蒸馏损失和学生一致性正则项的权重分别设置为 $\lambda_1 = 1.0$ 、 $\lambda_2 = 1.0$ 和 $\lambda_3 = 0.2$ 。检测微调阶段, 稀疏体素特征层和高层检测特征层的蒸馏权重分别设置为 $\lambda_{\text{vox}} = 1.0$ 和 $\lambda_{\text{head}} = 1.0$, 蒸馏正则项整体权重设置为 $\lambda_{\text{dist}} = 0.5$ 。

3.2 实验结果

表 1 给出了所提方法与主流自监督预训练方法在 VoxelNeXt 学生网络上的对比结果。由表 1 可知, 引入自监督预训练后, 各方法相较 Scratch 均取得一定提升, 说明无标注点云预训练能够改善船舶点云检测网络的特征初始化质量。其中, 所提方法在 VoxelNeXt 学生网络上取得最优结果, mAP@0.5 和 mAP@0.7 分别达到 77.2% 和 56.6%, 相比 Scratch 分别提升 6.4 和 7.0 百分点; 相比 R-MAE 分别提升 2.0 和 1.7 百分点。该结果表明, 异构双教师机制能够比单一掩码重构预训练更充分地利用无标注船舶点云目标中的几何结构先验知识与目标判别知识, 从而提升 VoxelNeXt 学生网络的特征表达能力。

表 1 与主流自监督预训练方法的对比结果

Table 1 Comparison results with mainstream self-supervised pre-training methods

方法	mAP@0.5/%	mAP@0.7/%	推理 参数量/ 10^6	推理速度/ (帧·s ⁻¹)
Scratch	70.8	49.6	3.4	3
BEV-MAE	74.4	53.7	3.4	3
Voxel-MAE	73.8	52.9	3.4	3
R-MAE	75.2	54.9	3.4	3
本文算法	77.2	56.6	3.4	3

从不同 IoU 阈值结果看, 所提方法在 mAP@0.7 上的提升更加明显。高 IoU 指标对目标中心位置、目标尺寸和航向角估计更加敏感, 因此该结

果说明所提方法不仅提升了船舶点云目标的检出能力,也改善了三维边界框定位精度。这主要得益于重构教师对船舶几何结构的建模,以及判别教师对目标区域与海面背景之间可分性的增强。

在模型复杂度方面,各方法推理阶段均采用相同的 VoxelNeXt 学生网络,模型参数量均为 3.4×10^6 。异构双教师网络及蒸馏相关分支仅在训练阶段参与知识迁移,训练完成后均被移除。因此,所提方法在提升检测精度的同时不会增加部署模型的参数量和网络结构复杂度。

为进一步验证异构双教师蒸馏机制对不同三维目标检测网络的适用性,在 VoxelNeXt 之外,选取 PointPillars、SECOND 和 PV-RCNN 作为学生检测器开展迁移实验,结果如表 2 所示。引入所提方法后,PointPillars、SECOND 和 PV-RCNN 在 mAP@0.7 上分别提升 5.9、6.3 和 6.5 百分点,VoxelNeXt 提升 7.0 百分点。不同结构的学生检测器均获得稳定性能增益,说明重构教师提供的几何结构知识和判别教师提供的目标判别知识能够通过特征对齐迁移至不同三维检测网络,所提异构双教师蒸馏机制具有较好的跨检测器适用性。

表 2 不同学生检测器上的迁移实验结果
Table 2 Transfer results with different student detectors %

学生检测	训练方式	mAP@0.5	mAP@0.7
PointPillars	Scratch	65.7	43.8
PointPillars	本文算法	71.2	49.7
SECOND	Scratch	67.9	46.2
SECOND	本文算法	73.6	52.5
PV-RCNN	Scratch	69.6	48.3
PV-RCNN	本文算法	75.5	54.8
VoxelNeXt	Scratch	70.8	49.6
VoxelNeXt	本文算法	77.2	56.6

表 3 给出了不同标注比例下各方法的 mAP@0.7 结果。随着标注比例由 50% 降低至 10%,所有方法的检测性能均出现下降,说明高质量三维标注样本数量对海上船舶点云目标检测具有重要影响。相比其他预训练方法,所提方法在不同标注比例下均取得最优结果。特别是在 10% 标注比例下,所提方法的 mAP@0.7 达到 47.9%,相比 Scratch 提升 11.5 百分点,相比 R-MAE 提升 5.7 百分点,表明该方法在低标注条件下具有更明显的性能优势。

表 3 不同标注比例下各预训练方法的检测性能对比
Table 3 Detection performance comparison of different pretraining methods under different annotation ratios %

预训练方法	10%标注	30%标注	50%标注
Scratch	36.4	44.8	47.8
Voxel-MAE	40.1	47.8	50.4
BEV-MAE	41.0	48.7	51.3
R-MAE	42.2	50.1	52.6
本文算法	47.9	54.6	56.9

该结果表明,异构双教师蒸馏能够利用无标注船舶点云中的几何结构信息与判别一致性信息,为 VoxelNeXt 学生网络提供有效补充监督。随着标注比例增加,监督信息更加充分,方法优势虽有所缩小,但仍保持稳定领先,说明其在低标注和常规标注条件下均能提升船舶点云检测性能。

为分析各核心模块对检测性能的贡献,本文以 VoxelNeXt 作为学生网络的基础检测框架开展消融实验,实验结果如表 4 所示。

表 4 消融实验结果对比
Table 4 Comparison of ablation experiment results %

实验 编号	重构 教师	判别 教师	注意力 加权	多粒度 蒸馏	mAP@0.5	mAP@0.7
A0	—	—	—	—	70.8	49.6
A1	√	—	—	—	74.0	53.5
A2	—	√	—	—	74.6	54.2
A3	√	√	—	—	76.0	54.9
A4	√	√	√	—	76.6	55.6
A5	√	√	√	√	77.2	56.6

由表 4 可知,在 VoxelNeXt 监督训练基线的基础上,单独引入重构教师后,mAP@0.7 由 49.6% 提升至 53.5%,说明几何结构知识能够增强学生网络对船舶轮廓、局部形态和空间布局的建模能力。单独引入判别教师后,mAP@0.7 提升至 54.2%,说明目标判别知识有助于提高不同船舶类别以及船舶目标与海面背景之间的特征可分性。当重构教师和判别教师同时使用时,mAP@0.7 进一步提升至 54.9%,相比单独采用判别教师提高 0.7 百分点。重构教师与判别教师学习的知识存在一定共享,但侧重点不同,其中重构教师更侧重船舶几何结构和空间布局信息,判别教师更侧重类别判别及目标与背景之间的特征可分性,因此双教师联合后的性能增益并非两个单教师增益的简单叠加。在此基础上加入注意力加权机制后,mAP@0.7 提升至 55.6%,说明注意力权重能够引导蒸馏监督更加关注船舶目标区域及其关键

结构位置,降低海面杂波和无效背景区域对特征学习的干扰。进一步加入多粒度蒸馏后,完整方法的 $mAP@0.7$ 达到 56.6%, 相比基础方法提升 7.0 百分点。结果表明,稀疏体素层和高层检测特征层的联合知识迁移能够同时增强 VoxelNeXt 学生网络的低层几何表达能力和高层目标判别能力,从而提升船舶点云检测的定位精度和鲁棒性。

为进一步分析重构教师与判别教师之间的知识互补性,对不同蒸馏方式获得的特征进行 t-SNE 降维可视化,如图 5 所示。图 5(a) 中,经过重构教师蒸馏后,两类船舶特征已呈现一定的聚集趋势,但不同船舶类别之间以及船舶目标与背景之间仍存在较明显的特征交叠;图 5(b) 中,经过判别教师蒸馏后,两类船舶及背景特征之间的分离更加明显,表明双视图一致性与对比约束能够增强特征的类别可分性;图 5(c) 中,双教师联合蒸馏后的特征保持了较好的类别分离趋势,同时类内分布更加集中,类别交叠和离群特征进一步减少。可见,重构教师提供的几何结构信息能够对判别教师获得的判别知识形成补充,使学生网络在保持类别可分性的同时获得更加稳定的特征表示,也进一步说明了 A3 相较 A2 取得增益的原因。

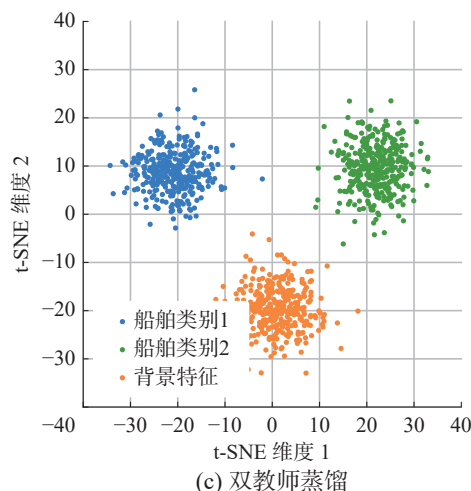
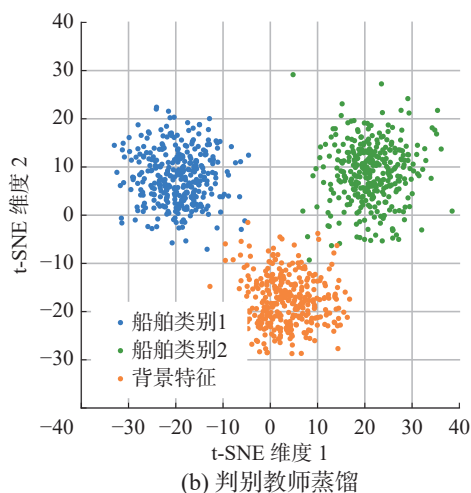
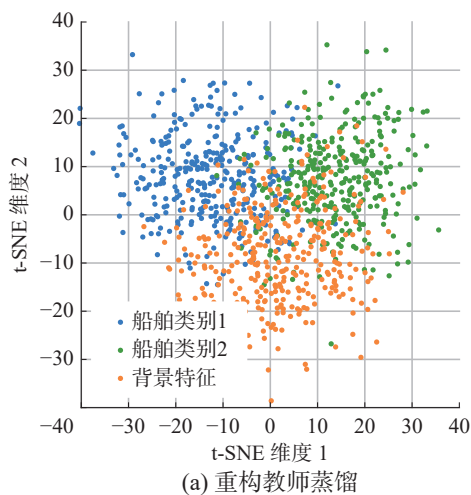


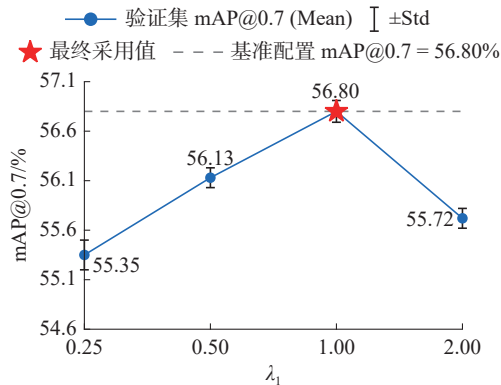
图 5 不同蒸馏方式下的特征分布可视化

Fig. 5 Feature distribution visualization under different distillation strategies

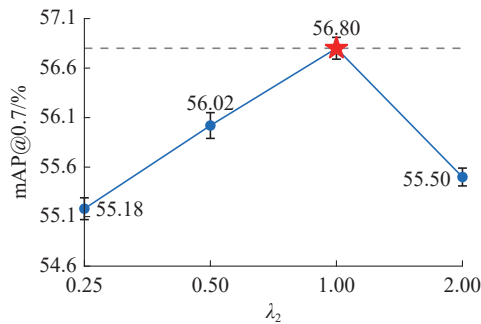
为分析不同损失权重对模型性能的影响,并进一步验证参数设置的稳定性,对 λ_1 、 λ_2 、 λ_3 、 λ_{vox} 、 λ_{head} 和 λ_{dist} 进行超参数敏感性分析。实验采用单因素控制变量法,每次仅调整一个待分析参数,其余参数固定为基准配置,并采用 3 个随机种子进行独立训练,以验证集 $mAP@0.7$ 的均值和标准差评价不同参数设置下的检测性能,实验结果如图 6 所示。由图 6 可知,各损失权重在基准配置附近均取得较优检测性能。当权重过小时,蒸馏监督或一致性约束作用不足;随着权重进一步增大,过强的辅助约束会对检测任务优化产生一定干扰。其中, λ_{dist} 的变化对性能影响相对更明显,而其余参数在合理范围内均表现出较好的稳定性。各组实验标准差整体较小,说明模型对随机初始化具有较好的鲁棒性。综合检测精度与训练稳定性,最终设置 $\lambda_1 = 1.0$ 、 $\lambda_2 = 1.0$ 、 $\lambda_3 = 0.20$ 、 $\lambda_{vox} = 1.0$ 、 $\lambda_{head} = 1.0$ 和 $\lambda_{dist} = 0.50$ 。

为进一步直观分析所提方法在实际海上环境中的船舶点云目标检测效果,选取远距离稀疏目标、多船目标、海面杂波以及不同尺度船舶等代表性场景,并将同步采集的可见光真实图像与对应点云检测结果进行对照展示,如图 7 所示。其中,可见光图像用于展示实际海上场景及船舶目标分布,点云检测结果中红色框表示预测结果,绿色框表示真实标注框。由可视化结果可以看出,预测框与真实框在空间位置、目标尺度和航向角上整体具有较高一致性,能够较准确地覆盖船舶点云主体区域,表明所提方法具备较好的三维定位与姿态估计能力。在多目标分布、点云回波稀疏以及海面杂波干扰等复杂场景下,模型仍

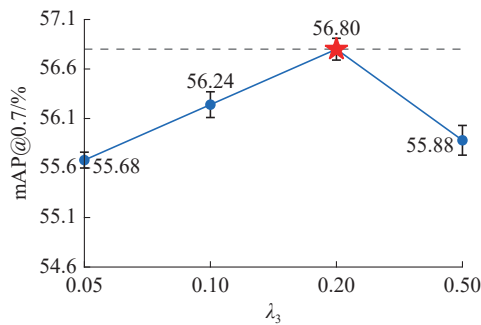
能够有效识别船舶目标,并保持预测框与真实框之间较好的匹配关系,说明该方法能够在一定程度上抑制海面背景噪声对检测结果的影响。同时,对于不同尺度的船舶目标,模型均能获得较稳定的检测结果,体现出较好的多尺度适应能力。综上,检测结果可视化进一步验证了所提方法在复杂海上点云场景下的检测鲁棒性和三维框定位精度。



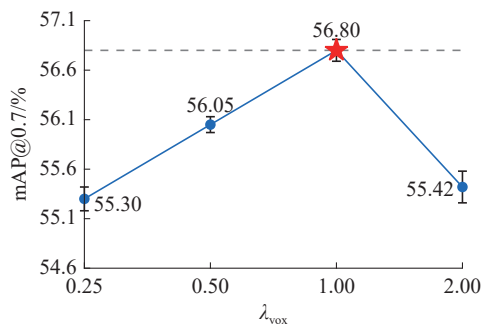
(a) 重构教师蒸馏权重



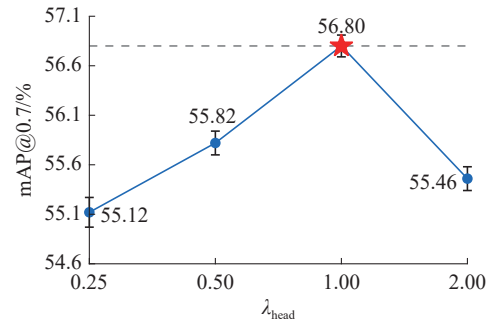
(b) 判别教师蒸馏权重



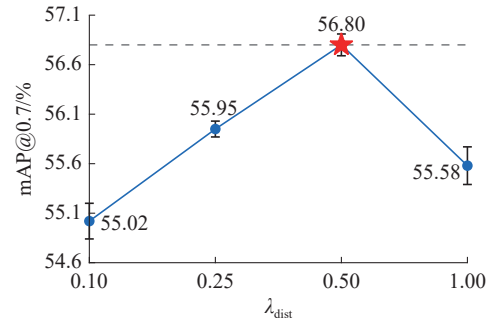
(c) 一致性正则权重



(d) 体素层蒸馏权重



(e) 高层检测特征蒸馏权重



(f) 总体蒸馏正则权重

图 6 损失权重超参数敏感性分析

Fig. 6 Sensitivity analysis of loss-weight hyperparameters

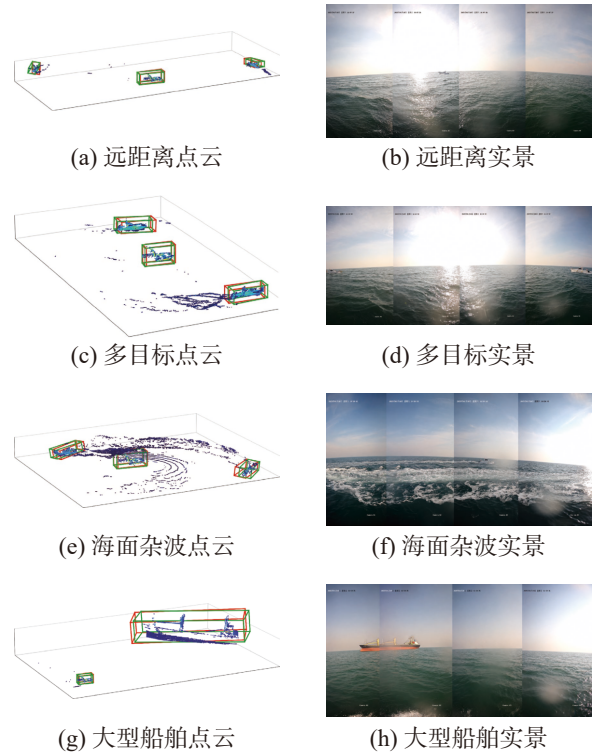


图 7 真实海上场景及对应船舶点云目标检测结果可视化

Fig. 7 Visualization of real maritime scenes and corresponding ship point-cloud detection results

4 结束语

针对海上船舶点云目标检测中回波稀疏、海面动态杂波强、船舶姿态变化大以及数据集构建成本高等问题,提出的融合异构双教师自监督蒸

馏与注意力加权知识迁移的船舶点云目标检测方法, 能够有效提升 VoxelNeXt 学生网络的检测性能。与从头训练的 Scratch 方法相比, $mAP@0.5$ 由 70.8% 提升至 77.2%, 提升 6.4 百分点; $mAP@0.7$ 由 49.6% 提升至 56.6%, 提升 7.0 百分点。与主流自监督预训练方法相比, 所提方法同样取得最优结果, 相较于 R-MAE 在 $mAP@0.5$ 和 $mAP@0.7$ 上分别提升 2.0 和 1.7 百分点, 表明异构双教师自监督蒸馏能够更有效地增强稀疏船舶点云目标的特征表达能力。消融实验进一步验证了双教师蒸馏、注意力加权和多粒度知识迁移等关键设计的有效性。由于教师网络和蒸馏分支仅在训练阶段参与参数优化, 部署推理阶段仅保留 VoxelNeXt 学生网络, 因此该方法在提升检测精度的同时不会增加额外模型参数, 具备面向无人艇船载感知平台部署的应用潜力。

当前研究仍存在一定局限, 数据集主要覆盖典型海况下的船舶目标, 对极端海况、复杂航道和更多目标类型的覆盖仍需进一步扩展; 同时, 当前方法主要基于单帧点云建模, 尚未充分利用跨帧时序信息。后续研究将扩展数据集覆盖范围, 并探索轻量级时序融合和多模态船舶目标检测方法, 以提升复杂海上船舶点云检测的鲁棒性和泛化能力。

参考文献:

- [1] LIU Zhixiang, ZHANG Youmin, YU Xiang, et al. Unmanned surface vehicles: an overview of developments and challenges[J]. *Annual reviews in control*, 2016, 41: 71–93.
- [2] BAI Xiangen, LI Bohan, XU Xiaofeng, et al. A review of current research and advances in unmanned surface vehicles[J]. *Journal of marine science and application*, 2022, 21(2): 47–58.
- [3] QIAO Yuanyuan, YIN Jiaxin, WANG Wei, et al. Survey of deep learning for autonomous surface vehicles in marine environments[J]. *IEEE transactions on intelligent transportation systems*, 2023, 24(4): 3678–3701.
- [4] CHENG Yuwei, JIANG Mengxin, ZHU Jiannan, et al. Are we ready for unmanned surface vehicles in inland waterways? the USVInland multisensor dataset and benchmark[J]. *IEEE robotics and automation letters*, 2021, 6(2): 3964–3970.
- [5] ZHANG Wei, JIANG Feng, YANG Chifu, et al. Research on unmanned surface vehicles environment perception based on the fusion of vision and lidar[J]. *IEEE access*, 2021, 9: 63107–63121.
- [6] 李小毛, 张鑫, 王文涛, 等. 基于 3D 激光雷达的无人水面艇海上目标检测[J]. *上海大学学报 (自然科学版)*, 2017, 23(1): 27–36.
- LI Xiaomao, ZHANG Xin, WANG Wentao, et al. 3D lidar-based marine object detection for USV[J]. *Journal of Shanghai University (natural science edition)*, 2017, 23(1): 27–36.
- [7] 武淑文, 李燕焱, 张少琛, 等. 基于深度学习的 3D 点云目标检测研究综述[J]. *遥测遥控*, 2024, 45(5): 1–18.
- WU Shuwen, LI Yanxi, ZHANG Shaochen, et al. 3D object detection methods based on point cloud with deep learning: a survey[J]. *Journal of telemetry, tracking and command*, 2024, 45(5): 1–18.
- [8] 鲁斌, 孙洋, 杨振宇. 融合体素图注意力的三维目标检测算法[J]. *智能系统学报*, 2024, 19(3): 598–609.
- LU Bin, SUN Yang, YANG Zhenyu. 3D object detection algorithm with voxel graph attention[J]. *CAAI transactions on intelligent systems*, 2024, 19(3): 598–609.
- [9] GEIGER A, LENZ P, URTASUN R. Are we ready for autonomous driving? The KITTI vision benchmark suite[C]//2012 IEEE Conference on Computer Vision and Pattern Recognition. Providence: IEEE, 2012: 3354–3361.
- [10] CAESAR H, BANKITI V, LANG A H, et al. nuScenes: a multimodal dataset for autonomous driving[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle: IEEE, 2020: 11618–11628.
- [11] LANG A H, VORA S, CAESAR H, et al. PointPillars: fast encoders for object detection from point clouds[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach: IEEE, 2019: 12689–12697.
- [12] YAN Yan, MAO Yuxing, LI Bo. SECOND: sparsely embedded convolutional detection[J]. *Sensors*, 2018, 18(10): 3337.
- [13] ZHOU Yin, TUZEL O. VoxelNet: end-to-end learning for point cloud based 3D object detection[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018: 4490–4499.
- [14] SHI Shaoshuai, WANG Xiaogang, LI Hongsheng. PointRCNN: 3D object proposal generation and detection from point cloud[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach: IEEE, 2019: 770–779.
- [15] SHI Shaoshuai, GUO Chaoxu, JIANG Li, et al. PV-RCNN: point-voxel feature set abstraction for 3D object detection[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle: IEEE, 2020: 10526–10535.

- [16] YANG Zetong, SUN Yanan, LIU Shu, et al. 3DSSD: point-based 3D single stage object detector[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle: IEEE, 2020: 11037–11045.
- [17] YIN Tianwei, ZHOU Xingyi, KRAHENBUHL P. Center-based 3D object detection and tracking[C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville: IEEE, 2021: 11779–11788.
- [18] CHEN Yukang, LIU Jianhui, ZHANG Xiangyu, et al. VoxelNeXt: fully sparse VoxelNet for 3D object detection and tracking[C]//2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Vancouver: IEEE, 2023: 21674–21683.
- [19] ZHANG Gang, CHEN Junnan, GAO Guohuan, et al. SAFDNet: a simple and effective network for fully sparse 3D object detection[C]//2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle: IEEE, 2024: 14477–14486.
- [20] HE Kaiming, CHEN Xinlei, XIE Saining, et al. Masked autoencoders are scalable vision learners[C]//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New Orleans: IEEE, 2022: 15979–15988.
- [21] XIE Saining, GU Jiatao, GUO Demi, et al. PointContrast: unsupervised pre-training for 3D point cloud understanding[C]//Computer Vision—ECCV 2020. Cham: Springer, 2020: 574–591.
- [22] PANG Yatian, WANG Wenxiao, TAY F E H, et al. Masked autoencoders for Point cloud self-supervised learning[C]//Computer Vision—ECCV 2022. Cham: Springer, 2022: 604–621.
- [23] ZHANG Renrui, GUO Ziyu, GAO Peng, et al. Point-M2AE: multi-scale masked autoencoders for hierarchical point cloud pre-training[C]//Advances in Neural Information Processing Systems 35. New Orleans: NeurIPS, 2022: 27061–27074.
- [24] MIN Chen, XIAO Liang, ZHAO Dawei, et al. Occupancy-MAE: self-supervised pre-training large-scale LiDAR point clouds with masked occupancy autoencoders[J]. *IEEE transactions on intelligent vehicles*, 2024, 9(7): 5150–5162.
- [25] HESS G, JAXING J, SVENSSON E, et al. Masked autoencoder for self-supervised pre-training on lidar point clouds[C]//2023 IEEE/CVF Winter Conference on Applications of Computer Vision Workshops. Waikoloa: IEEE, 2023: 350–359.
- [26] LIN Z, WANG Y, QI S, et al. BEV-MAE: Bird's eye view masked autoencoders for point cloud pre-training in autonomous driving scenarios[C]//Proceedings of the AAAI Conference on Artificial Intelligence. 2024, 38(4): 3443–3451.
- [26] LIN Zhiwei, WANG Yongtao, QI Shengxiang, et al. BEV-MAE: bird's eye view masked autoencoders for point cloud pre-training in autonomous driving scenarios[C]//Proceedings of the AAAI Conference on Artificial Intelligence. Vancouver: AAAI, 2024: 3204–3212.
- [27] TAYEBATI S, TULABANDHULA T, TRIVEDI A R. Sense less, generate more: pre-training LiDAR perception with masked autoencoders for ultra-efficient 3D sensing[EB/OL]. (2024-06-12)[2026-06-12]. <https://doi.org/10.48550/arXiv.2406.07833>.
- [28] NISAR B, WASLANDER S L. PSA-SSL: pose and size-aware self-supervised learning on LiDAR point clouds[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Music City Center: IEEE, 2025: 6402–6411.
- [29] ABDELSAMAD M, ULRICH M, GLASER C, et al. Multi-scale neighborhood occupancy masked autoencoder for self-supervised learning in LiDAR point clouds [C]//2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville: IEEE, 2025: 22234–22243.
- [30] ZHANG Linfeng, DONG Runpei, TAI H S, et al. Point-Distiller: structured knowledge distillation towards efficient and compact 3D detection[C]//2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Vancouver: IEEE, 2023: 21791–21801.
- [31] GAMBASHIDZE A, DADUKIN A, GOLYADKIN M, et al. Weak-to-strong 3D object detection with X-ray distillation[C]//2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle: IEEE, 2024: 15055–15064.
- [32] ZHANG Haonan, LIU Longjun, HUANG Yuqi, et al. CaKDP: category-aware knowledge distillation and pruning framework for lightweight 3D object detection[C]//2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle: IEEE, 2024: 15331–15341.
- [33] PONZINI F, ZACCONE R, MARTELLI M. LiDAR target detection and classification for ship situational awareness: a hybrid learning approach[J]. *Applied ocean research*, 2025, 158: 104552.
- [34] 姚婷婷, 张波, 柳晓鸣. 特征增强全卷积网络下的船舶检测[J]. *计算机辅助设计与图形学学报*, 2022, 34(7): 1028–1036.
- YAO Tingting, ZHANG Bo, LIU Xiaoming. Feature enhanced fully convolutional network for ship detection[J]. *Journal of computer-aided design & computer graphics*, 2022, 34(7): 1028–1036.
- [35] 李忠智, 尹航, 左剑凯, 等. 基于 UNet++网络与多边形

- 出融合策略的船舶检测模型[J]. [计算机工程](#), 2022, 48(4): 276–283.
- LI Zhongzhi, YIN Hang, ZUO Jiankai, et al. Ship detection model based on UNet++Network and multiple side-output fusion strategy[J]. [Computer engineering](#), 2022, 48(4): 276–283.
- [36] 袁国文, 张彩霞, 杨阳, 等. 复杂场景下深度表示的 SAR 船舶目标检测算法[J]. [计算机工程与应用](#), 2022, 58(2): 289–294.
- YUAN Guowen, ZHANG Caixia, YANG Yang, et al. SAR target detection algorithm for depth representation in complex scenes[J]. [Computer engineering and applications](#), 2022, 58(2): 289–294.
- [37] 何芸倩, 夏桂华, 冯鸿超, 等. 自航船模点云数据集的海上船舶检测[J]. [哈尔滨工程大学学报](#), 2022, 43(8): 1156–1162, 1168.
- HE Yunqian, XIA Guihua, FENG Hongchao, et al. Marine ship detection on the point cloud dataset of autonomous navigation ship models[J]. [Journal of Harbin engineering university](#), 2022, 43(8): 1156–1162, 1168.
- [38] ZHANG Qiuyu, WANG Lipeng, MENG Hao, et al. Ship detection in maritime scenes under adverse weather conditions[J]. [Remote sensing](#), 2024, 16(9): 1567.
- [39] XIE Yongchang, NANLAL C, LIU Yuanchang. Reliable

LiDAR-based ship detection and tracking for Autonomous Surface Vehicles in busy maritime environments[J]. [Ocean engineering](#), 2024, 312: 119288.

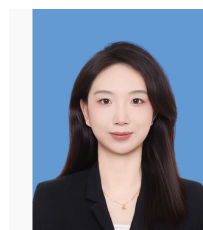
作者简介:



夏桂华, 教授, 博士生导师, 主要研究方向为智能船舶与海洋装备、海洋环境感知及人工智能技术。享受国务院政府特殊津贴, 获省部级以上奖励 10 余项, 发表学术论文 90 余篇。E-mail: xiaguihua@hrbeu.edu.cn。



柏隼, 博士研究生, 主要研究方向为海上船舶点云目标检测、计算机视觉及海上环境感知。E-mail: baijun19920816@hrbeu.edu.cn。



李诗洋, 博士后研究员, 主要研究方向为海上船舶环境感知、船舶数据驱动建模。E-mail: lishiyang@dlut.edu.cn。