

【引用格式】汤瑞, 陈依民, 高剑, 等. 基于深度学习的侧扫声呐图像目标检测[J]. 数字海洋与水下攻防, 2025, 8 (4): 472-480.

# 基于深度学习的侧扫声呐图像目标检测

汤 瑞, 陈依民\*, 高 剑, 郝邵文, 王亚周

(西北工业大学 航海学院, 陕西 西安 710072)

**摘要** 侧扫声呐图像目标精准检测技术在海洋资源勘探、水下考古研究及国防安全领域具有重要的应用价值。然而, 受侧扫声呐图像背景噪声、分辨率及目标尺度差异等因素制约, 现有自动检测方法仍面临检测精度低的问题。针对上述问题, 探究了一种基于深度学习的侧扫声呐图像目标检测新方法。该方法通过引入空频卷积替代基线模型主干网络中的普通卷积, 实现了对频率纹理特征的有效提取, 从而显著抑制侧扫声呐图像中散斑噪声对目标检测精度的影响。进一步, 利用多尺度特征融合模块改进基线模型中的快速空间金字塔池化, 以增强网络对小目标的检测能力。在沉船侧扫声呐图像数据集上的实验结果表明: 该方法的 mAP@50 和 mAP@50: 95 相比基线模型分别提高了 4.05% 和 5.61%。

**关键词** 侧扫声呐图像; 目标检测; 深度学习; 频率特征

**中图分类号** TN911.73    **文献标识码** A    **文章编号** 2096-5753(2025)04-0472-09

**DOI** 10.19838/j.issn.2096-5753.2025.04.010

## Object Detection in Side-Scan Sonar Images Based on Deep Learning

TANG Rui, CHEN Yimin\*, GAO Jian, HAO Shaowen, WANG Yazhou

( School of Marine Science and Technology, Northwestern Polytechnical University, Xi'an 710072, China )

**Abstract** Object detection in side-scan sonar images holds significant application value in marine resource exploration, underwater archaeological research, and national defense security. However, existing automatic detection methods still face the challenge of low detection accuracy due to factors such as background noise, resolution, and target scale variations in side-scan sonar images. To address these issues, a new target detection method for side-scan sonar images based on deep learning is explored in this paper. By replacing the conventional convolutions in backbone network of baseline model with space-frequency convolutions, this method effectively extracts frequency texture features, significantly suppressing the impact of speckle noise in side-scan sonar images on target detection accuracy. Furthermore, a multi-scale feature fusion module is used to improve the spatial pyramid pooling-fast in the baseline model, enhancing the ability to detect small targets. Experimental results on the shipwreck side-scan sonar dataset demonstrate that the proposed method achieves improvements of 4.05% and 5.61% in mAP@50 and mAP@50: 95, respectively, compared to the baseline model.

**Key words** side-scan sonar image; object detection; deep learning; frequency feature

收稿日期: 2025-07-15

作者简介: 汤瑞 (1997-), 男, 硕士, 主要从事水下航行器研究。

\*通信作者: 陈依民 (1990-), 男, 博士, 教授, 主要从事水下航行器研究。

基金项目: 国家自然科学基金“面向多源多方位探测的 AUV 集群分布式事件触发鲁棒协同控制研究”(52471347); “双一流”建设专项基金“师资队伍建设项目-国家级青年人才”(0206022GH0202); 西北工业大学博士论文创新基金“面向少样本侧扫声呐图像的域自适应目标检测方法研究”(CX2025046)。

## 0 引言

随着经济发展与国家安全等方面的需求不断增长,海洋战略价值开发已跃升为当代国际竞争的核心领域之一。党的十八大、十九大和二十大报告均提出“加快建设海洋强国”等一系列重大部署,并确定建设海洋强国是实现中华民族伟大复兴的重大战略任务。当前,我国在海洋领域已形成了众

多重要成果,但是在海洋核心装备和关键技术瓶颈方面,仍需要着力突破,其中包括无人智能平台和水下目标检测等技术<sup>[1]</sup>。水下目标检测技术作为海洋感知技术的“神经末梢”,其是海洋资源开发与维护、科研与生态保护、水下搜救与应急响应、渔业与海洋经济、军事与安防等多个领域中的重要一环(如图1所示),具有多维战略价值,也发挥着越来越关键的作用。

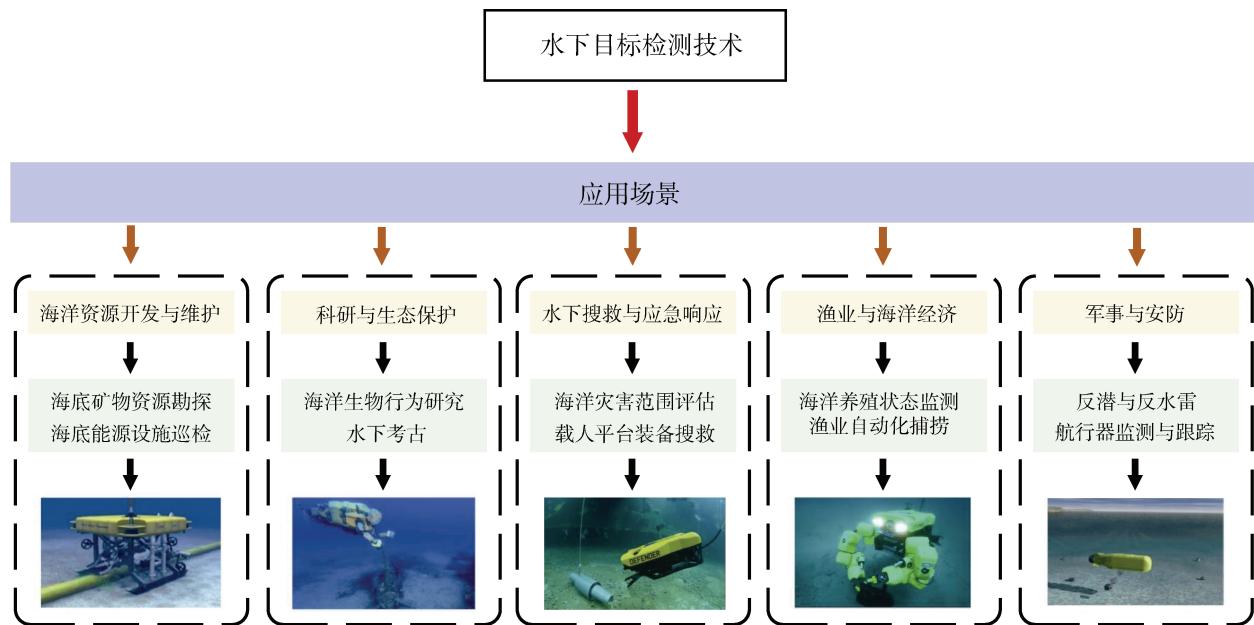


图1 水下目标检测技术应用场景  
Fig. 1 Application scenarios of underwater object detection

与光信号相比,声信号在水下具有更优异的传播性能,因此,基于船载或水下无人平台的声学观测技术已成为当前海洋探测领域的研究热点。其中,侧扫声呐凭借其灵活的搭载方式和获取高分辨率图像的能力,现已成为水下目标检测的核心装备之一。

与光学图像不同,侧扫声呐图像普遍分辨率低、轮廓模糊,这使得侧扫声呐图像目标检测的准确率不高,并存在很多误检。此外,人工目视判读方法不仅成本高昂、效率低,而且检测结果容易受到个人经验等主观因素的影响,难以保证检测的准确性与稳定性。基于此,学者们提出了各种方法。

传统的侧扫声呐图像水下目标检测算法主要利用空间域信息和频域信息来捕捉目标与背景在轮廓、灰度分布状态、局部对比度等几个表象上的

差异,从而将目标与背景区分开来。主流的声呐目标检测算法可以分为2类:基于特征域和基于像素。其中,基于特征域的算法可分为基于局部特征表示、主成分分析的检测算法。如YE等人<sup>[2]</sup>提出了2种新的水平集模型来解决声呐图像中的分割问题。基于像素的算法有帧差法、光流法、形态学滤波、恒虚警率等,如RAHNEMOONFAR等人<sup>[3]</sup>提出了一种基于顶帽变换的声呐图像增强框架,并利用形态学算子识别海草栖息地中的裸露斑块。这些算法的底层原理都比较简单,不需要提取目标的深层特征。同时,由于水下目标具有不确定性,人工获取先验信息的代价太大,大大限制了传统的声呐图像检测算法的应用场景和效果。

受深度学习技术发展和相关任务的驱使,当前基于深度学习的侧扫声呐图像目标检测已进行了

许多研究。针对侧扫声呐图像中目标稀疏和特征匮乏的问题, YU 等人<sup>[4]</sup>提出了一种集成 Transformer 模块、注意力模块和下采样模块的 TR-YOLOv5s 自动目标检测算法。HUANG 等人<sup>[5]</sup>通过提出了一种综合样本增强方法, 改善了侧扫声呐图像目标检测效果。XU 等人<sup>[6]</sup>通过一种多特征融合自注意力网络来生成新类别的侧扫声呐图像, 进而将生成的高质量的伪侧扫声呐图像用于网络的训练, 最终提高侧扫声呐图像识别的准确性。HUANG 等人<sup>[7]</sup>提出了一种基于空间自适应反归一化的 Seg2Sonar 网络。XI 等人<sup>[8]</sup>提出了一种基于 Transformer 和 YOLOv7 的声呐目标检测算法。同时, 引入了 2 种数据增强算法来提高检测器的性能。ZHANG 等人<sup>[9]</sup>在 YOLOv7 的基础上增加了针对小目标的检测层, 并在主干网络和颈部增加了 2 种注意力机制, 同时对 BiFPN 进行改进, 提高了侧扫声呐图像小目标检测效果。为了减小海底反射噪声和复杂背景对声呐图像目标检测的影响, WANG 等人<sup>[10]</sup>利用多尺度感受野提取块和自注意力块来增强模型的特征提取能力。同时, 利用双向特征金字塔网络和全局金字塔池化块来获得深层语义特征。在此基础上, 利用自适应特征融合块来融合不同尺度间的特征。

在迁移学习方面, XI 等人<sup>[11]</sup>基于侧扫声呐的成像机理和图像特性, 提出了一种目标到目标和背景到背景的改进风格迁移方法, 并在图像中加入噪声, 使得生成的侧扫声呐图像更加真实。CHENG 等人<sup>[12]</sup>利用扩散模型来生成水下声学数据。最后, 将增强的数据用于训练改进后的 YOLOv7 模型。LI 等人<sup>[13]</sup>设计了一种纹理特征去除网络来缩小光学图像与声呐图像之间的差异, 进而改善声呐图像的分类和检测精度。TANG 等人<sup>[14]</sup>基于迁移学习在 YOLOv3 模型的基础上, 通过改进其特征金字塔网络、目标锚框的参数以及损失函数, 进而提高了目标识别准确率。YE 等人<sup>[15]</sup>在 ImageNet 数据集上训练的 CNN 模型的基础上, 结合迁移学习并根据侧扫声呐图像的特点提出了一种识别方法。CHANDRASHEKAR 等人<sup>[16]</sup>将光学迁移到侧扫声呐图像上, 并对 EfficientNet 模型进行微调, 实现了沉积物分类。HUO 等人<sup>[17]</sup>基于深度迁移学习将光学图像合成为声呐图像, 并将生成样本和真实样

本混合用于深度卷积神经网络 (CNN) 的训练, 进而提高了水下目标分类精度。GE 等人<sup>[18]</sup>通过风格迁移网络将光学图像模拟为侧扫声呐图像, 并将其用于预训练的 CNN, 提高了分类精度。BAI 等人<sup>[19]</sup>提出了一种侧扫声呐图像风格迁移网络来合成伪 SSS 图像, 并设计了一种对比度损失算法来平衡光学与声学图像风格特征分布, 进而改善模型效果。

然而, 上述研究多聚焦于改进检测器的性能, 未充分考虑侧扫声呐图像中的散斑噪声对检测精度的影响, 仅将噪声归为低级视觉任务的处理范畴。同时, 这些工作大多未探究频率纹理特征在侧扫声呐图像目标检测中的潜在价值, 且当前检测器仍面临检测精度低的问题。因此, 本文探究了一种新的基于深度学习的侧扫声呐图像目标检测网络, 该网络基于 YOLOv8 算法<sup>[20]</sup>, 首先采用空频卷积<sup>[21]</sup>来替代主干网络中的普通卷积, 在保持轻量化的同时, 实现对频率纹理特征的有效提取, 从而显著抑制侧扫声呐图像中散斑噪声对目标检测精度的影响。进一步, 通过多尺度特征融合模块来改善基线模型中的快速金字塔池化模块, 以增强网络对小目标的检测能力。网络的整体结构如图 2 所示。

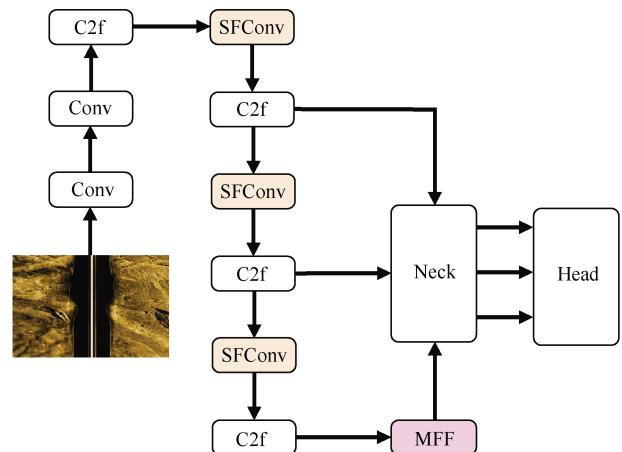


图 2 网络结构图  
Fig. 2 Overview of network structure

## 1 空频卷积

空频卷积 (SFConv) 主要由空间感知单元 (SPU) 和频率感知单元 (FPU) 2 部分组成, 其结构如图 3 所示。具体而言, 首先将输入特征图  $X$  按照  $\alpha$  的比例划分为 2 个部分, 分别记为空间特征  $X_s$  和频域特征  $X_f$ 。随后, 使用核为  $1 \times 1$  的卷积对其

进行调整, 并将其送入 SPU 和 FPU 组件来获得空间细化特征  $Y_s$  和频率细化特征  $Y_f$ 。然后, 对输出特征在通道维度进行联合拼接, 得到特征图  $Y_{sf}$ 。进一步, 使用全局平均池化 (GAP) 来收集全局的空间和频率信息, 并使用 Softmax 来生成特征权重。

进而, 将其与特征图  $Y_{sf}$  相乘来得到特征图  $Y_{sf}^+$ 。在此基础上, 将特征图  $Y_{sf}^+$  按照比例  $\alpha$  分成 2 部分, 并使用 add 操作将该 2 部分相加得到特征图  $Y^+$ 。随之, 使用核为  $1 \times 1$  的卷积对其进行调整, 并将调整后的特征图与输入特征图  $X$  相加得到最终输出  $Y$ 。

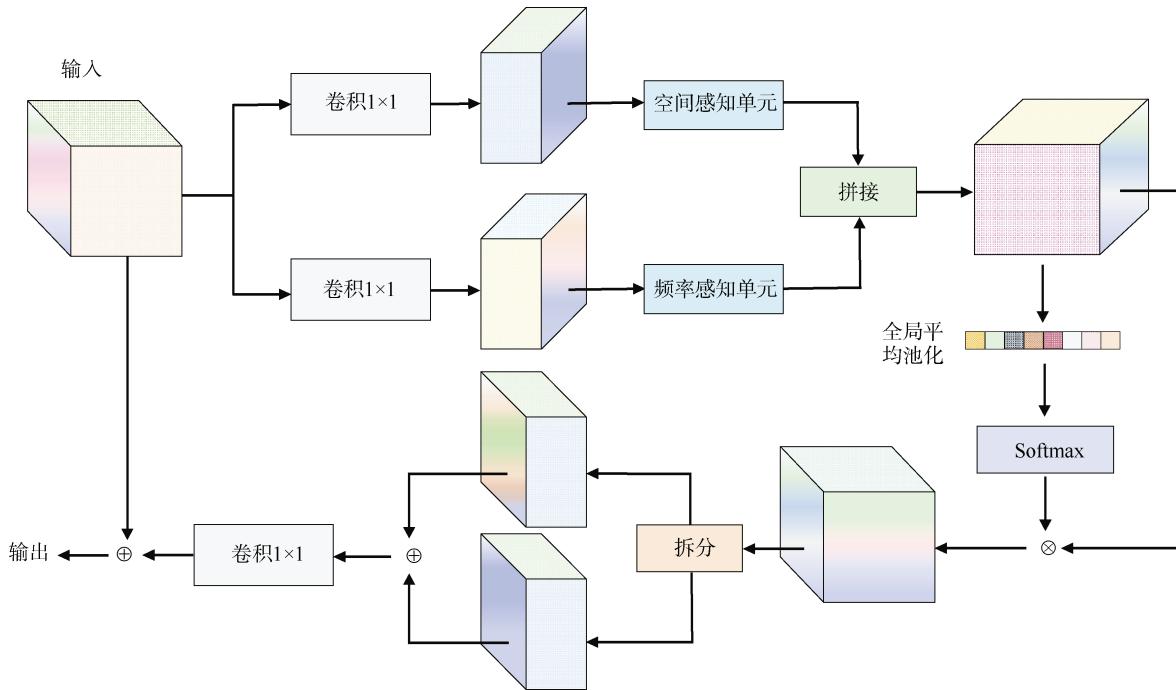


图 3 空频卷积结构图

Fig. 3 Structure diagram of space-frequency convolution

## 1.1 空间感知单元

空间感知单元主要用于对不同尺度的上下文信息进行动态建模。具体来说, 首先将特征图  $X^s$  在通道维度均匀地划分为  $n$  个特征图组  $X_g^s$ ,  $g \in \{1, 2, \dots, n\}$ 。对于每个特征图  $X_g^s + Y_{g-1}^s$ , 设置一个对应大小为  $k_g$  的卷积核,  $k_{g+1} = k_g + 2$ 。当  $g=1$  的时候, 特征图仅为  $X_g^s$ , 卷积核设置为  $k_1=3$ 。

为了更好地融合不同尺度的信息, 进一步将所有组沿通道维度进行拼接, 并通过一个核大小为  $1 \times 1$  的卷积, 进而得到空间感知单元的输出, 即空间细化特征  $Y_s$ 。

## 1.2 频率感知单元

为了使同一卷积层内的标准卷积核能够捕捉不同尺度和方向的纹理特征, 采用分数阶 Gabor

变换 (FrGT) 滤波器来代替普通卷积核, 进而构建频率感知单元。

具体来说, 通过将输入特征  $X^f$  划分为  $V$  组, 每组记为  $X_v^f$ , 并对每个  $X_v^f$  设置  $N = C/VU$  个卷积核来生成相应的频率特征  $Y_v^f$ 。进一步, 将所有的  $Y_v^f$  进行拼接, 并通过核大小为  $1 \times 1$  的卷积可以得到频率感知单元输出, 即频率细化特征  $Y_f$ 。

## 2 多尺度特征融合模块

多尺度特征融合 (MFF) 模块主要由核大小为  $1 \times 1$  的卷积和参数共享的核大小为  $3 \times 3$  空洞卷积组成, 其可嵌入主干网络中。该模块可以通过空洞卷积序列以及特征重组策略实现跨尺度特征的高效融合, 结构图如图 4 所示。

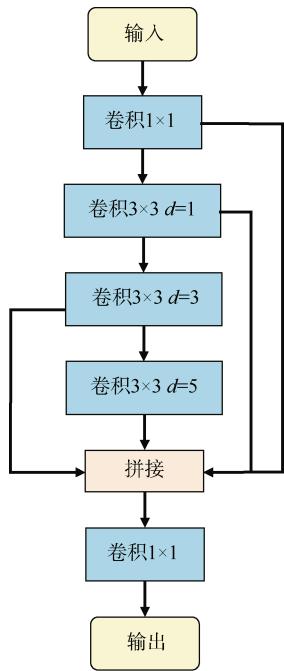


图 4 MFF 模块结构图

Fig. 4 Structure diagram of MFF module

该模块首先经过核大小为  $1 \times 1$  卷积进行通道维度对齐，进一步，将其送入多尺度提取核心层，通过连续 3 组核大小为  $3 \times 3$  的空洞卷积来构建渐进式感受野。其中，第 1 级空洞率  $d = 1$  可捕获局部细节，用于增强对小目标的高敏感感知。第 2 级空洞率  $d = 3$  用于提取区域级特征。第 3 级空洞率  $d = 5$  可覆盖全局空间关系，增强对遮挡目标的强鲁棒性。该设计以指数级扩展感受野，避免传统池化造成的分辨率损失。在此基础上，将卷积输出特征在通道维度进行联合拼接，使细粒度到粗粒度的多尺度信息整合为统一特征张量。最后，利用核大小为  $1 \times 1$  的卷积对拼接结果实施特征压缩与重组，在降低通道复杂度的同时强化跨尺度特征交互。此外，空洞卷积参数共享可显著降低计算成本。

### 3 网络结构

在目标检测领域，YOLOv8 算法凭借其卓越的检测精度，且已受到广泛的应用测试，现已成为当前最具竞争力的检测架构之一。基于此，本研究选用 YOLOv8 作为侧扫声呐图像目标检测的基准模型，重点解决因背景噪声、分辨率及目标尺度差异等因素导致的检测精度下降问题，开展基于深度学习的侧扫声呐图像目标检测研究。

具体而言，本文首先引入空频卷积替代 YOLOv8 主干网络中  $P_3$ 、 $P_4$  和  $P_5$  层的普通卷积。与仅聚焦空间特征的普通卷积不同，引入空频卷积可以适配侧扫声呐图像对空间轮廓与频率纹理的双重需求。通过将输入特征划分为空间与频域 2 个分支进行细化处理来提高特征鲁棒性。其中，空间感知单元采用多尺度卷积核动态建模目标的空间上下文来解决目标尺度差异大的问题，频率感知单元则通过分数阶 Gabor 变换捕获频域纹理来增强含噪侧扫声呐图像中提取到特征的判别性。

此外，根据侧扫声呐图像的噪声特性，本文调整了空频卷积的特征划分比例和频率滤波器方向，使其更侧重声呐图像频率纹理提取。同时，通过全局平均池化与权重融合机制，在保持轻量化的同时实现空频特征的有效融合。

其次，本文采用多尺度特征融合模块替代 YOLOv8 原有的快速空间金字塔池化（SPPF）模块。传统 SPPF 模块采用池化操作进行特征融合，容易丢失小目标的细节信息，而 MFF 模块通过采用空洞率为 1, 3, 5 的  $3 \times 3$  空洞卷积构建渐进式感受野，有效避免了传统池化操作导致的特征损失。

## 4 数据集与实验

### 4.1 数据集

本文使用沉船侧扫声呐图像数据集<sup>[22]</sup>来进行性能验证，其包含 851 幅图像，具有不同背景颜色、分辨率大小和目标区域大小的图像数据，如图 5 所示。该数据集主要包含 7 种目标类型：小目标、模糊目标、需要依靠阴影识别的目标、大部分被泥沙掩埋的目标、多目标、瀑布图中间存在的弱目标、被瀑布图切成两半的目标。

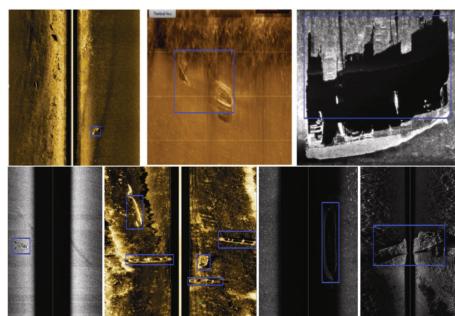


图 5 沉船侧扫声呐图像数据集中典型目标

Fig. 5 Typical objects in shipwreck side-scan sonar image dataset

## 4.2 评价指标

本文选取平均精度均值 (mAP) 作为模型的性能评价指标, 计算公式如下:

$$P = \frac{TP}{TP + FP}$$

$$R = \frac{TP}{TP + FN}$$

$$AP = \int_0^1 P(R)dR$$

$$mAP = \frac{\sum AP}{m}$$

式中: TP 表示预测样本和实际样本均为正例的情况; FP 表示预测样本为正例, 但实际样本为负例的情况; FN 表示预测样本为负例, 但实际样本为正例的情况。

## 4.3 实验配置

所有实验均在配备 inter (R) i5-13490F 处理器、32G 内存、NVIDIA RTX4070 12.0G 显卡, 使用 Ubuntu20.04 系统以及 Pytorch2.0.0 + cu118 的台式机上进行。

在训练过程中, 本文模型和对比模型的输入图像大小固定为  $640 \times 640$ , 优化器采用 AdamW<sup>[23]</sup>, 批量大小为 32, 训练轮数为 300, 并采用在线马赛克数据增强方法。空频卷积中的超参数  $\alpha$  设置为 2,  $n$  设置为 3,  $u$  设置为 [0, 45, 90, 135]。



## 4.4 实验结果

为了全面验证我们的模型对侧扫声呐图像的性能, 在沉船侧扫声呐图像和公共声学图像数据集上进行了一系列广泛的实验。此外, 为了进一步说明本文模型的性能, 选择了深度学习中一些知名的目标检测模型, 如 YOLOv5<sup>[24]</sup>、YOLOv6<sup>[25]</sup> 和 YOLOv7<sup>[26]</sup> 进行实验结果比较。以上模型都有不同尺度的多个版本。本文选择在同一数量级上与本文模型进行比较, 具体模型与测试集对比实验结果如表 1 所示。测试集中多个模型的可视化结果如图 6 所示, 其中红色和绿色的方框是预测的目标位置, 方框上方显示的依次是预测的类别和对应的置信度。本文模型与基线模型的 mAP@50 和 mAP@50:95 对比如图 7 所示, 与基线模型训练过程中的 loss 变化对比如图 8 所示。

表 1 沉船侧扫声呐数据集上的比较结果

Table 1 Comparison results on shipwreck side-scan sonar image dataset

模型	$F_1$	mAP@50%	mAP@50:95%
YOLOv5s	0.69	72.1	37.9
YOLOv6n	0.76	80.6	54.1
YOLOv7t	0.80	84.4	49.3
基线模型	0.80	83.9	53.5
本文模型	0.82	87.3	56.5

图 6 多个检测模型结果的可视化

Fig. 6 Visualization of detection results for multiple models

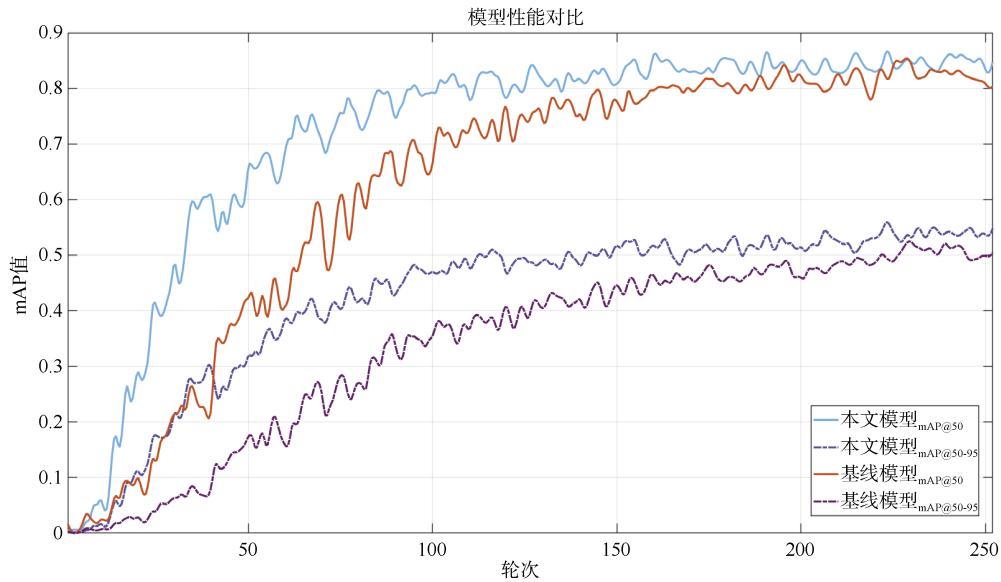


图 7 与基线模型的 mAP@50 和 mAP@50:95 对比

Fig. 7 Comparison of mAP@50 and mAP@50:95 of our method with that of baseline model

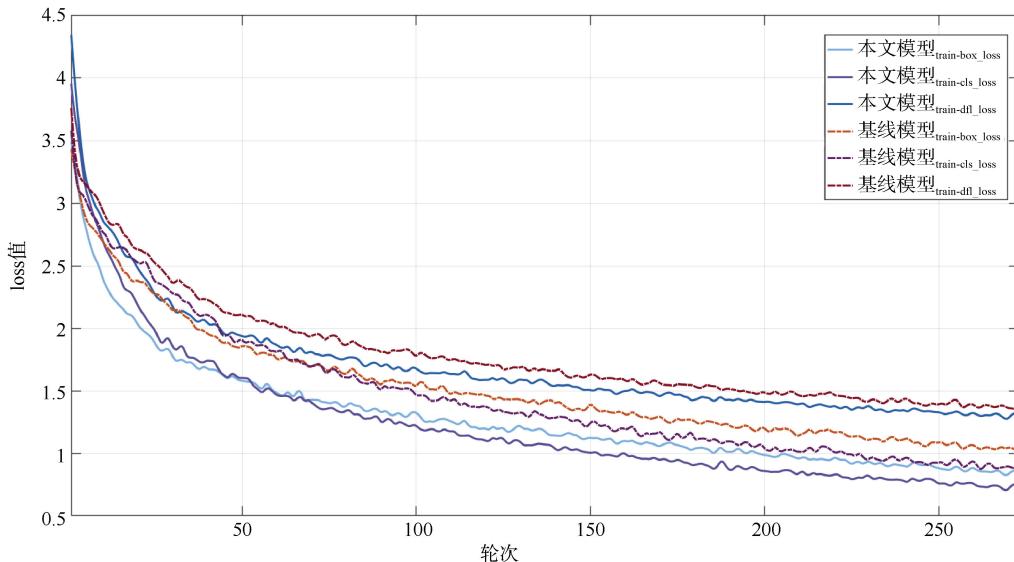


图 8 与基线模型的 loss 变化对比

Fig. 8 Comparison of loss value of our method with that of baseline model

#### 4.5 消融实验

为了验证每个设计和改进模块的有效性和重要性，进行了充分的消融实验。消融实验主要包括以下几个方面：1) 只包含空频卷积；2) 只包含多尺度特征融合模块；3) 包含空频卷积和多尺度特征融合模块。

消融实验的结果如表 2 所示，其中“√”表示包含该模块，“×”表示不包含该模块。结果表明，本文使用的空频卷积和多尺度特征模块是有效的。

表 2 消融实验结果

Table 2 Results of ablation experiment

SFConv	MFF	mAP@50/%	mAP@50:95/%
√	×	86.1	55.1
×	√	84.7	54.4
√	√	87.3	56.5

#### 4.6 实验分析

从表 1 可以看出，与基线模型和其他模型相比，本文模型具有更高的精度。与基线模型相比，

本文模型  $F_1$  值提高了 2.50%, mAP@50 值提高了 4.05%, mAP@50:95 值提高了 5.61%。

从图 6 可以看出, 在案例 1 中, 仅本文模型准确识别出了 2 只沉船; 在案例 2 中, 本文模型对于损毁程度较大沉船的识别具有最高置信度; 在案例 3 中, 本文模型识别到了最多数量的沉船目标; 在案例 4 中, 本文模型对于大部分被泥沙掩埋的目标具有高置信度。

从表 2 可以看出, 通过对比不同模块组合对目标检测模型性能的影响, 验证了 SFConv 和 MFF 模块的有效性。

## 5 讨论

### 5.1 数据集

本文选择沉船侧扫声呐图像数据集进行实验验证, 虽然该数据集的样本规模相比光学图像数据集偏小, 但是其包含有 7 种典型目标类型(如小目标、模糊目标、被泥沙掩埋的目标等), 能够覆盖侧扫声呐图像检测中常见的难点场景, 具有较高的可靠性。

需要说明的是, 受限于侧扫声呐图像数据采集环境、设备成本及应用场景限制, 公开可用的高质量标注数据集较少。此外, 由于私密性原因, 该领域多数专业数据集暂无法公开获取。因此, 本文优先选择该数据集进行验证。

后续研究中, 将通过多场景协同采集并构建可公开的侧扫声呐数据, 进一步扩充数据集规模, 开展对比实验, 以更全面地证明本文方法的有效性。

### 5.2 空频卷积数量

针对 SFConv 模块的插入数量问题, 本文在设计时已结合网络特性进行针对性分析, 具体如下:

在主干网络的第 P1 和 P2 层中, 未采用 SFConv 替换普通卷积主要原因在于这两层需快速压缩输入图像的空间维度, 普通卷积在降维效率上更具优势, 可减少冗余计算。随着网络加深, 特征图逐渐抽象, 需要同时强化空间上下文与频域纹理的融合能力。因此, 将 P<sub>3</sub>, P<sub>4</sub> 和 P<sub>5</sub> 层的普通卷积替换为 SFConv。若将 SFConv 减少至 1 个或 2 个, 会导致频率纹理特征丢失, 造成目标检测精度下降; 若增加至 4 个及以上, 虽特征提取更充分, 但计算量相较基线模型会有较大增加。

## 6 结束语

本文针对当前侧扫声呐图像自动目标检测方法面临的精度不足问题, 探究了一种基于深度学习的侧扫声呐图像目标检测新方法。在沉船侧扫声呐图像数据集上的实验表明, 本文模型相比基线模型和其他模型具有更高的检测精度。同时, 消融实验验证了本文引入的各模块的有效性。

未来将重点开展更为广泛类型的侧扫声呐图像数据的采集与获取, 同时针对侧扫声呐图像预处理以及本文引入的侧扫声呐图像频域特征进行更深入的研究。

## 参考文献

- [1] 科学网. 自然资源部副部长、国家海洋局局长: 持续深耕建设海洋强国关键领域[EB/OL]. ( 2022-10-21 ) [2025-06-05]. <https://news.science.net.cn/html/news/2022/10/488075.shtml>.
- [2] YE X F, ZHANG Z H, LIU P X, et al. Sonar image segmentation based on GMRF and level-set models[J]. Ocean Engineering, 2010, 37 ( 10 ): 891-901.
- [3] RAHNEMOONFAR M, RAHMAN A F, KLINE R J, et al. Automatic seagrass disturbance pattern identification on sonar images[J]. IEEE Journal of Oceanic Engineering, 2018, 44 ( 1 ): 132-141.
- [4] YU Y C, ZHAO J H, GONG Q H, et al. Real-time underwater maritime object detection in side-scan sonar images based on transformer-YOLOv5[J]. Remote Sensing, 2021, 13 ( 18 ): 3555.
- [5] HUANG C, ZHAO J H, YU Y C, et al. Comprehensive sample augmentation by fully considering SSS imaging mechanism and environment for shipwreck detection under zero real samples[J]. IEEE Transactions on Geoscience and Remote Sensing, 2021, 60: 1-14.
- [6] XU H L, BAI Z Y, ZHANG X Y, et al. MFSANet: Zero-shot side-scan sonar image recognition based on style transfer[J]. IEEE Geoscience and Remote Sensing Letters, 2023, 20: 1-5.
- [7] HUANG C, ZHAO J H, ZHANG H M, et al. Seg2Sonar : A full-class sample synthesis method applied to underwater sonar image target detection, recognition, and segmentation tasks[J]. IEEE Transactions on Geoscience and Remote Sensing, 2024, 62: 1-19.
- [8] XI J, YE X F. Sonar image target detection based on simulated stain-like noise and shadow enhancement in optical images under zero-shot learning[J]. Journal of

- Marine Science and Engineering, 2024, 12 ( 2 ): 12020352.
- [9] ZHANG F H, ZHANG W, CHENG C S, et al. Detection of small objects in side-scan sonar images using an enhanced YOLOv7-based approach[J]. Journal of Marine Science and Engineering, 2023, 11 ( 11 ): 2155
- [10] WANG Z, ZHANG S W, HUANG W Z, et al. Sonar image target detection based on adaptive global feature enhancement network[J]. IEEE Sensors Journal, 2021, 22 ( 2 ): 1509-1530.
- [11] XI Z, ZHAO J H, ZHU W Q. Side-scan sonar image simulation considering imaging mechanism and marine environment for zero-shot shipwreck detection[J]. IEEE Transactions on Geoscience and Remote Sensing, 2023, 61: 1-13.
- [12] CHENG C S, HOU X J, WEN X, et al. Small-sample underwater target detection: a Joint Approach utilizing diffusion and YOLOv7 model[J]. Remote Sensing, 2023, 15 ( 19 ): 15194772.
- [13] LI C L, YE X F, XI J, et al. A texture feature removal network for sonar image classification and detection[J]. Remote Sensing, 2023, 15 ( 3 ): 15030616.
- [14] TANG Y L, JIN S H, BIAN G, et al. Shipwreck target recognition in side-scan sonar images by improved YOLOv3 model based on transfer learning[J]. IEEE Access, 2020, 8: 173450-173460.
- [15] Y X F, LI C L, ZHANG S Y, et al. Research on side-scan sonar image target classification method based on transfer learning[C]// OCEANS 2018. Charleston: IEEE, 2018.
- [16] CHANDRASHEKAR G, RAAZAA, RAJENDRAN V, et al. Side scan sonar image augmentation for sediment classification using deep learning based transfer learning approach[J]. Materials Today: Proceedings, 2023, 80: 3263-3273.
- [17] HUO G Y, WU Z Y, LI J B. Underwater object classification in sidescan sonar images using deep transfer learning and semisynthetic training data[J]. IEEE Access, 2020, 8: 47407-47418.
- [18] GE Q, RUAN F X, QIAO B J, et al. Side-scan sonar image classification based on style transfer and pre-trained convolutional neural networks[J]. Electronics, 2021, 10 ( 15 ): 1823.
- [19] BAI Z Y, XU H L, DING Q C, et al. Side-scan sonar image classification with zero-shot and style transfer[J]. IEEE Transactions on Instrumentation and Measurement, 2024, 73: 3352693.
- [20] ULTRALYTICS. YOLOv8[EB/OL]. [2025-06-05]. <https://docs.ultralytics.com/yolov8>.
- [21] LI K, WANG D, HU Z Y, et al. Unleashing channel potential: Space-frequency selection convolution for SAR object detection[C]// 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition ( CVPR ). Seattle: IEEE, 2024.
- [22] TANG R, CHEN Y M, GAO J, et al. Towards real-time detection of underwater target with pruning lightweight deep learning method in side-scan sonar images[J]. Neurocomputing, 2025, 620: 129254.
- [23] ZHOU P, XIE X Y, LIN Z C, et al. Towards understanding convergence and generalization of AdamW[J]. IEEE Transactions on Pattern Analysis And Machine Intelligence, 2024., 46 ( 9 ): 6486-6493.
- [24] ULTRALYTICS. YOLOv5[EB/OL]. [2025-06-05]. <https://docs.ultralytics.com/yolov5>.
- [25] LI C Y, LI L L, JIANG H L, et al. YOLOv6: A single-stage object detection framework for industrial applications[EB/OL]. [2022-09-07]. <https://openreview.net/pdf?id=7c3ZOKGQ6s>.
- [26] WANG C Y, BOCHKOVSKIY A, LIAO H Y M. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[C]// 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition ( CVPR ). Vancouver: IEEE, 2023.

(责任编辑: 曹晓霖)