

# 任务驱动下成像声呐水下目标识别方法综述

聂为之, 尹 斐, 苏毅珊

(天津大学 电气自动化与信息工程学院, 天津 300072)

**摘要:** 阐述了侧扫声呐、合成孔径声呐和前视声呐在图像分类、检测以及分割任务中的目标识别算法及其解决的主要问题。通过结合不同声呐的成像特点与应用场景, 分析总结上述成像声呐对应图像处理任务下的目标识别算法的优劣及仍需解决的关键问题, 并展望其未来发展方向。

**关键词:** 信息处理技术; 水下目标识别; 成像声呐; 图像处理

**中图分类号:** TN911.73; TP391.4 **文献标志码:** A **文章编号:** 1671-5497(2025)04-1163-13

**DOI:** 10.13229/j.cnki.jdxbgxb.20240002

## Review of task-driven imaging sonar for underwater target recognition approaches

NIE Wei-zhi, YIN Fei, SU Yi-shan

(School of Electrical Automation and Information Engineering, Tianjin University, Tianjin 300072, China)

**Abstract:** The target recognition algorithms and the major problems solved in image classification, detection, and segmentation tasks for side-scan sonar, synthetic aperture sonar, and forward-looking sonar were described. By combining the imaging characteristics and application scenarios of different sonars, the strengths and weaknesses of the target recognition algorithms under the corresponding image processing tasks of the above imaging sonar and the key issues that still need to be addressed were analyzed and summarized, and its future development direction was also looked forward to.

**Key words:** information processing technology; underwater target recognition; imaging sonar; image processing

## 0 引言

随着陆地资源日益枯竭, 各国逐渐将能源开采重心由陆地资源转移到海洋资源中。水下探测技术作为开发海洋资源、认识海洋环境、保护海洋生态以及拓展人类生存与发展空间的基础, 吸引

着越来越多的学者以此为中心开展研究。其中, 水下目标识别是水下探测领域的热门研究方向之一。由于电磁波在水中衰减大, 传播距离短, 因此, 水下探测领域通常使用传播损耗更小的声波作为传输介质<sup>[1]</sup>。

声呐 (Sound navigation and ranging, SO-

收稿日期: 2024-01-02.

基金项目: 国家自然科学基金项目 (62171310, 62272337).

作者简介: 聂为之 (1987-), 男, 教授, 博士. 研究方向: 计算机视觉. E-mail: weizhinie@tju.edu.cn

通信作者: 苏毅珊 (1985-), 男, 副教授, 博士. 研究方向: 水声通信与海洋立体信息网络. E-mail: yishan.su@tju.edu.cn

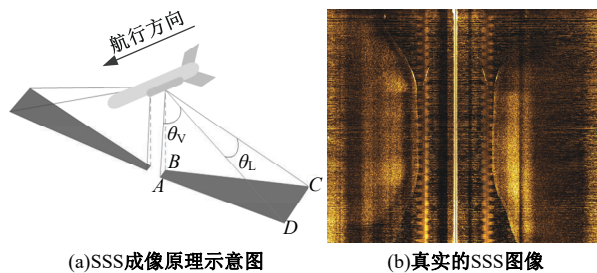
NAR)设备利用声波在水中的传播和反射特性,通过信息处理和电声转化对水下目标进行探测<sup>[2]</sup>,其中,成像声纳可以整合和处理接收返回波束信号的幅度和相位等信息,并采用图像重建技术实现对目标的成像,因此,在海洋生物监测、海洋垃圾检测、水下搜索与救援、海床测绘和海底管道探测等水下目标识别场景中得到广泛应用。目前普遍使用的成像声纳设备有3种<sup>[3]</sup>:侧扫声纳(Side-scan sonar, SSS)、合成孔径声纳(Synthetic aperture sonar, SAS)和前视声纳(Forward-looking sonar, FLS),它们的成像特点和应用场景各有不同。

图像处理领域的任务通常可以分为3种:图像分类、图像检测和图像分割。到目前为止,已经提出了多种基于成像声纳图像在分类、检测、分割这3种任务下的水下目标识别算法。分类任务中提出窄带 Chan-Vese (CV) 模型<sup>[4]</sup>、风格迁移 (Style transfer, ST)<sup>[5]</sup>等。检测算法有模板匹配 (Template matching, TM)<sup>[6]</sup>、改进经典网络模型 YOLO (You only look once)<sup>[7]</sup>等。分割任务中采用马尔科夫随机场 (Markov random field, MRF)<sup>[8]</sup>、水平集 (Level set, LS)<sup>[9]</sup>方法等。然而,上述算法并未能充分考虑声纳图像的采集过程,未能将不同成像声纳的成像特点与具体的图像处理任务充分结合以设计更有针对性的识别方法,从而限制了现有识别算法的识别率、鲁棒性等性能的进一步提高。

针对成像声纳图像识别存在的问题,本文回顾了3种常用成像声纳设备在3种图像处理任务中采用的目标识别算法,并根据不同成像声纳图像特点与应用场景,分析每种处理任务算法的优劣,最后展望成像声纳图像处理算法的未来发展。

# 1 侧扫声纳

侧扫声纳工作时,通常安装在载体的左、右两侧,并借助载体运动在航迹方向上成像。其工作原理如图1(a)所示,载体左右两侧的SSS分别发射扇形波束,垂直开角 $\theta_v$ 控制垂直平面的扫描范围,水平开角 $\theta_L$ 保证水平面上的扫描分辨率。扇形波束照射海床形成左右两片梯形区域,如图1(a)中航迹左侧梯形ABCD区域所示,受水平开角 $\theta_L$ 的影响,梯形区域靠近载体的底边AB长度小于底边CD,因此,SSS的图像分辨率会随着目标到声纳的距离增加而变差<sup>[10]</sup>。



(a)SSS成像原理示意图 (b)真实的SSS图像

图 1 SSS 成像示意图

Fig. 1 Imaging schematic of SSS

由于SSS成像依靠载体运动,探测范围大,所以SSS主要应用于大面积海洋地貌调查、测量与目标物搜索等场景。

## 1.1 水下图像分类算法

如图2所示,目前SSS图像分类任务中使用的算法可以分为3大类:统计模式识别、基于传统机器学习 (Machine learning, ML) 的分类器和人工神经网络 (Artificial neural networks, ANNs)。

TM是一种对感兴趣的目标提取特殊特征建模,通过计算测试图像与模板图像的相似度判别所属类别的模式识别算法。由SSS成像的先验知识可知,同一目标投射出的阴影形状可能有多

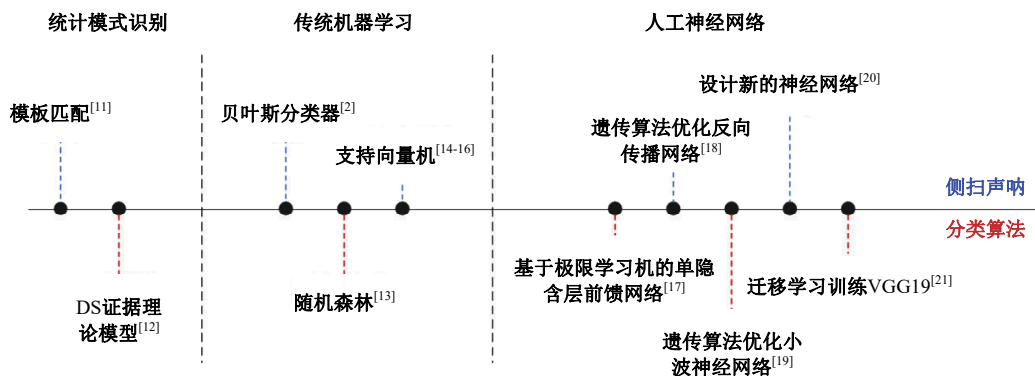


图 2 基于 SSS 图像在分类算法中的主要工作

Fig. 2 Major works in classification algorithms based on SSS images

种,Dura等<sup>[11]</sup>在目标阴影形状建模过程中通过引入一个可变参数生成不同形状的模板。但SSS通常应用在大面积搜寻场景,通过几何形态构建多种复杂形状目标的模板较为困难。其次,由于SSS图像分辨率较低,成像后的不同目标形状差距小,从而导致匹配错误率高,因此,TM实际应用有一定的局限性。

在文献[12]中,分类决策采用一种基于概率和集合的Dempster-shafer(DS)理论模型。DS证据理论以假设特征分布独立为前提,但SSS图像的多种特征具有相关性,导致该方法处理SSS图像效果并不理想。

分类SSS图像采用的传统ML分类器类型有3种:第1种分类器是概率模型贝叶斯(Bayes),目前已有一种新的贝叶斯网络(Bayes net)能将不可靠的阴影特征根据样本分布情况有效组合以实现可靠的目标分类<sup>[2]</sup>;第2种利用随机森林(Random forest)模型分类ReliefF算法优化的海床特征<sup>[13]</sup>;第3种分类器支持向量机(Support vector machine, SVM),相较于以上两种分类器,SVM与多种图像预处理方法结合取得了较高的分类准确率。SVM在分类前进行Sobel滤波处理,可将分类准确率提升到92.8%<sup>[14]</sup>。利用SVM对从声呐图像提取的纹理特征分类,最高准确率为95.6%<sup>[15]</sup>。以此为基础,结合图像的纹理特征和回波强度,准确率可提升1.05%<sup>[16]</sup>。

与手工提取特征相比,神经网络将图像本身作为输入以学习更深层、抽象、鲁棒性更高的特征,已经在光学图像识别领域取得了巨大的成功,因此,目前提出了多种基于神经网络的声呐图像识别算法。针对SSS图像的分类任务,文献[17]提出了一种基于极限学习机(Extreme learning machine, ELM)的单隐含层前馈神经网络,采用自适应增强算法优化,准确率可达93.56%。文献[18]中,遗传算法(Genetic algorithm, GA)优化简单的反向传播(Back propagation, BP)网络,准确率为92.5%。在文献[19]中,采用GA算法优化小波神经网络(Wavelet neural network, WNN),对海底沉积物的平均分类准确率为93.3%。针对类似的海底沉积物分类场景,文献[20]设计了两种不同的卷积神经网络(Convolutional neural network, CNN),实验证明在少样本的声呐数据集上训练,浅层CNN的分类性能优于

深层CNN。

由于声呐图像采集困难,且投入成本高,用于训练CNN的数据有限,容易造成模型过拟合,因此,将CNN与迁移学习结合的新方法被提出。模型在有大量高质量标注的数据集上预训练,再将模型迁移到目标数据集上进行微调训练。采用迁移学习在SSS数据集上训练VGG19网络的整体准确率为97.76%<sup>[21]</sup>。该方案目前在基于SSS图像的分类领域取得了更好的效果,且提高了该领域在大规模、多类型目标搜索时算法的泛化能力和鲁棒性。

## 1.2 水下图像检测算法

SSS应用于水下大规模目标检测时,对其处理算法不仅需要计算复杂度低以满足实时性,且具有抗水下复杂噪声干扰的能力。如图3所示,目前的解决方案有基于统计模式识别、传统ML和深度学习(Deep learning, DL)3类。

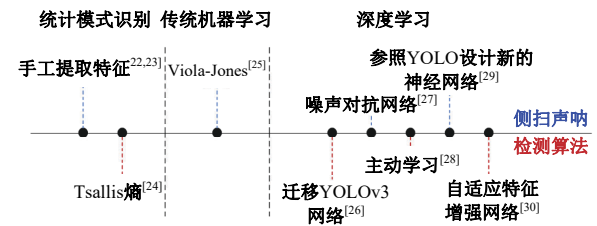


图3 基于SSS图像在检测算法中的主要工作

Fig. 3 Major works in detection algorithms based on SSS images

Petillot等<sup>[22]</sup>提取水下管道在SSS成像的阴影区域,对该区域进行1~3段不等的直线拟合,最后利用Bayes评估拟合效果的概率以实现检测。因为独立成分分析(Independent component analysis, ICA)能从混合信息中分离出与目标相关性更强的独立分量,所以文献[23]将图像纹理特征与ICA结合以检测目标。然而,上述工作手工提取浅层特征,在水下多样性地貌环境中鲁棒性较差。在文献[24]中,作者采用形态学变换占有率击中-击不中变换(Percentage occupancy hit-or-miss transform, POHMT)增强目标的高亮区域与阴影区域,再采用Tsallis熵自适应地确定POHMT的最佳阈值以检测目标。该方法对阈值的选择非常敏感,需要进一步优化阈值选择策略。

以上文献采用手工提取声呐图像特征,但目前主流算法是从图像数据中自主学习特征。Sawas等<sup>[25]</sup>将用于人脸检测的Viola-Jones引入SSS图像检测。但SSS图像中目标占据小部分,Viola-

la-Jones 检测器对区域选择没有针对性,导致窗口冗余,时间复杂度过高。文献[26]以 YOLOv3 作为基准模型,多尺度特征融合部分新增 4 倍和 2 倍下采样层学习阴影等浅层特征,且对该模型进行迁移训练,提高目标检测的精度和速度。在文献[27]中,为克服噪声扰动对声呐图像检测效果的影响,设计了噪声对抗网络(Noise adversarial network, NAN)与人为引入噪声的噪声块(Noise block, NB),提高检测时的噪声鲁棒性。由于标注 SSS 图像不仅耗时,还需要相关专业知识,因此,为降低标注成本, Jiang 等<sup>[28]</sup>提出基于主动学习(Active learning, AL)的检测算法。该框架使用 35% 的标注数据就能获得与使用全部标注数据的 CNN 相当的结果,但每次迭代都需要重新训练模型,计算成本高。

近些年,自动目标识别(Automatic target recognition, ATR)已经应用到成像声呐等领域。SSS 通常搭载在无人水下航行器(Unmanned underwater vehicle, UUV)进行实时目标识别。然而受硬件平台功率和计算能力等限制,文献[26-28]提出的大型 CNN 部署有困难。Topple 等<sup>[29]</sup>参考 YOLO 模型进行轻量级的网络设计,并尝试向受限硬件平台部署 CNN。文献[30]提出一种自适应特征增强网络 AGFE-Net,骨干网通过新增多尺度特征提取块和自注意机制块以提取更深层次的图像特征。其中,特征融合部分采用双向特征金字塔网络进行浅层和深层特征交互,全局金字塔池化块增强对全局特征的感知,自适应特征融合块实现声呐图像多尺度特征融合。该网络各类别平均检测精度(Mean average precision, mAP)为 96.83%,每秒可处理 67.28 帧数据,与目前其他方案相比,检测准确度更高、速度更快,可以部署在资源受限的移动平台中。且在多样性地貌环境中进行鲁棒性测试, mAP 保持在 95% 左右。

### 1.3 水下图像分割算法

图 4 展示了针对 SSS 图像分割领域的主要工作。聚类分析是一种标签信息未知,通过像素之间的相似度划分为若干类的无监督算法。文献[31]采用模糊  $k$  均值( $k$ -means)聚类算法将 SSS 图像分割为阴影区、高亮区和海底混响区。Celik 等<sup>[32]</sup>在  $k$ -means 算法分割之前,利用正反非抽取小波变换(Undecimated discrete wavelet trans-

form, UDWT)得到图像多层次的边缘信息,相较于单独采用  $k$ -means 算法,提升了分割边缘的准确度。MRF 分割算法类似于聚类算法,但 MRF 算法会考虑某一点像素的邻域像素标签,从而计算该点像素最大概率的标签。Mignotte 等<sup>[33]</sup>提出一种分层 MRF 模型,第一步采用迭代条件估计(Iterative conditional estimation, ICE)技术对噪声分布参数和 MRF 先验参数进行估计,第二步利用估计好的参数将 SSS 图像分割为阴影和海底混响区域。由于 SSS 图像分辨率较低,且受强度不均匀和散斑噪声干扰,导致上述算法分割目标边缘效果不理想。

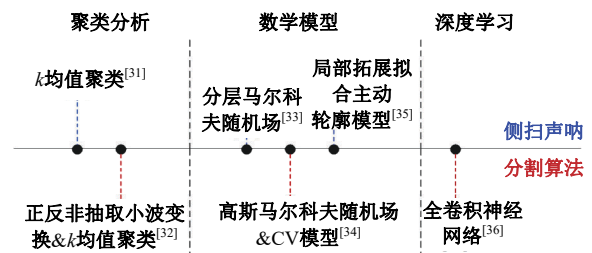


图 4 基于 SSS 图像在分割算法中的主要工作

Fig. 4 Major works in segmentation algorithms based on SSS images

为弥补上述算法的缺点,多种基于数学模型和几何推理的算法被引入 SSS 图像分割。文献[34]在传统 CV 能量函数中增加高斯马尔科夫随机场(Gauss-markov random field, GMRF)提取的纹理特征项,求解函数最小值以将曲线演化到目标区域实现分割。虽然文献[34]提出的方法可以改善分割边缘的效果,但时间复杂度过高。SSS 图像分割精度通常受到大量散斑噪声与强度分布不均匀的影响。针对这些挑战,文献[35]提出非局部均值散斑滤波(Nonlocal means-based speckle filtering, NLMSF)去除 SSS 图像的散斑噪声,并采用  $k$ -means 聚类与局部拓展拟合(Region-scalable fitting, RSF)主动轮廓模型结合的分割策略。

随着 DL 的兴起,文献[36]提出一种适合 SSS 图像分割的全卷积神经网络(Fully convolutional network, FCN)。该方法分割准确率为 91.62%,边缘分割效果优于传统方法,且对散斑噪声的鲁棒性强。此外,该网络只需要 18.7 MB 内存,预测时间为 27.9 ms。SSS 展开大规模目标搜索任务时,适合将该算法嵌入资源受限的硬件平台实时分割、识别目标,故 FCN 模型具有巨

大的应用优势。

## 2 合成孔径声呐

SSS 的图像分辨率受作用距离的影响,在实际应用中采集到的图像质量不稳定,不利于大规模水下地质勘探。合成孔径声呐(Synthetic aperture sonar, SAS)通过小孔径基阵的运动,在航迹方向合成大孔径,以提高图像的分辨率<sup>[37]</sup>。相较于一般 SSS, SAS 远距离成像的分辨率更高,适合在大面积海洋地质研究与目标搜索中实时作业。图 5 展示了 SAS 的工作原理。

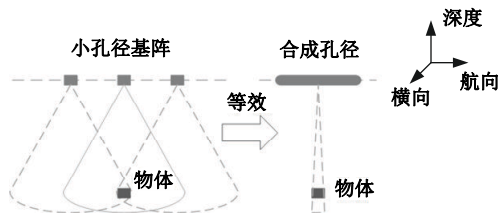


图 5 SAS 成像示意图

Fig. 5 Imaging schematic of SAS

### 2.1 水下图像分类算法

传统 ML 和 DL 的相关算法在 SAS 图像分类任务中得到广泛应用。图 6 展示了该领域的主要工作。

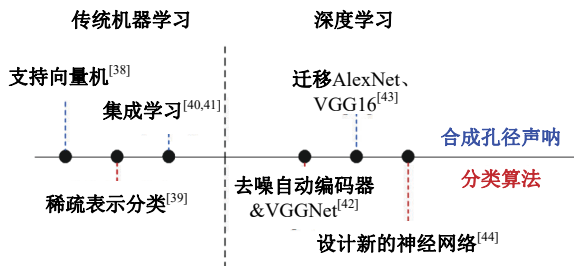


图 6 基于 SAS 图像在分类算法中的主要工作

Fig. 6 Major works in classification algorithms based on SAS images

基于 SAS 图像分类的传统 ML 分类器有多种类型。文献[38]采用卷积受限玻尔兹曼机(Convolutional restricted boltzmann machine, CRBM)提取图像特征,再采用 SVM 对该特征分类。稀疏表示分类(Sparse reconstruction-based classification, SRC)是创建一个超完备字典作为基函数,在字典中用较少元素线性表示输入对象,最后根据输入对象的线性表示系数进行分类的算法。文献[39]将 SRC 算法应用到 SAS 图像分类任务。Fei 等<sup>[40]</sup>提出一种基于 DS 模型的集成学习(Ensemble learning)方案,将不同分类器的结果合理融合以提高分类精度。

SAS 在多种海洋地质环境中进行大面积地质调查与目标搜索等工作中,不同的海洋地质具有不同的环境条件,导致成像有不同的特征分布。然而上述方法只在特定环境中具有较好性能,无法满足 SAS 对不同水下环境地质勘察的需求。为提高算法在不同环境中的分类泛化能力,文献[41]提出构建分类器集合。其中每个分类器与特定环境特征关联,然后根据不同分类环境自动调整每个分类器的权重。该方法在目前的方案中能分类多种环境下成像的目标,且具有很好的性能。

文献[42]采用去噪自动编码器(Denoising auto-encoder, DAE)与 VGGNet 联合的方案实现目标自动分类。DAE 是一种特殊的神经网络,能提取高维数据的有效特征,且在输入层添加随机噪声提高模型的鲁棒性。训练 CNN 模型需要大量数据,但目前采集大量具有地质环境多样性的 SAS 数据集有困难,导致模型易出现过拟合。迁移学习是解决搭建 CNN 模型数据不足的一种快速且有效的方法。文献[43]在少量 SAS 图像上,对 AlexNet 和 VGG16 尝试多组不同的微调方法。在文献[44]中,作者设计了 4 个小型 CNN 模型,分别对比从头训练和使用迁移学习训练的结果,后者的分类性能更好。然而,若源任务与目标任务相关性差,该方法分类多种环境成像目标的性能比构建分类器集合<sup>[41]</sup>差,且需要的训练样本更多。

### 2.2 水下图像检测算法

如图 7 所示,利用 SAS 成像进行水下目标检测的算法可分为 3 类。

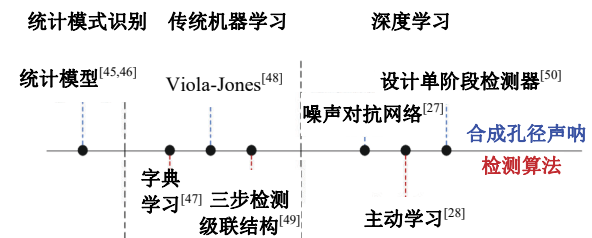


图 7 基于 SAS 图像在检测算法中的主要工作

Fig. 7 Major works in detection algorithms based on SAS images

第 1 类是统计模型算法。文献[45]提出一种基于拉普拉斯(Laplace)分布的多元统计模型,使用期望最大(Expectation maximization, EM)算法得到该模型概率密度函数的最大似然估计,最后利用似然比拟合测试图像。Abu 等<sup>[46]</sup>分别对目标高光区与阴影区建模,采用加权似然比检验法将

阴影、高光检测与目标预期空间分布结合以实现目标检测。然而,统计建模法的检测性能与大量不同水下环境分布的先验知识有紧密联系。

第 2 类是基于传统 ML 的检测算法。文献[47]利用两个字典矩阵分别表示图像的目标和非目标片段,通过字典矩阵得到剩余误差功率比,随后在稀疏表示域进行目标检测匹配。文献[48]基于经典人脸检测器 Viola-Jones 做出改进以适应 SAS 成像的目标检测。然而,上述方法并未考虑不同海洋地质环境导致不同成像特征这一因素。Williams<sup>[49]</sup>采用阴影、波纹、高亮区域三步检测的级联结构和积分图像技术提出能嵌入 UUV 实时检测的算法。该方法准确率为 94.21%,在不同海床环境下试验结果均有一定的竞争力,且根据不同的环境特征自适应调整最佳观测方向,这对海洋地质中的沙纹调查具有重要意义。

第 3 类是基于 DL 的算法。文献[27]中设计的 NAN 和 NB 模块通过在 SAS 数据集训练,算法的噪声鲁棒性得到一定的提升。基于 AL 的目标检测算法<sup>[28]</sup>从未标记的 SAS 图像数据中按特定准则选择合适的图像添加注释,每次迭代联合新注释的图像与已经标注的数据重新训练模型,从而降低了 SAS 图像标注成本。但以上两种方案计算复杂度过高,无法部署到资源有限的硬件平台。文献[50]提出了一种基于 CNN 分类器的单阶段检测器,通过全连接层将得到的所有特征映射融合为一个特征映射输出。在输出的特征图中,每个像素代表该区域检测到目标的概率,通过选择最佳阈值确定该区域的目标存在性。然而,该方案并未考虑多样性海洋地质环境对算法稳定性的影响。

### 2.3 水下图像分割算法

文献[51]将由 SAS 图像得到的归一化直方图振幅数据映射到均值-标准差平面,分别计算各个轴上像素的熵,从熵值变化曲线中检测变化最明显的值作为最佳分割阈值。但该算法在边缘分割区域效果不理想。

如图 8 所示,模糊聚类算法同样被应用到 SAS 图像分割领域。文献[52]提出可能性模糊局部信息  $c$  均值(Possibilistic fuzzy local information  $c$ -means, PFLICM)算法,最小化 PFLICM 目标函数直至收敛得到分割结果。Abu 等<sup>[53]</sup>将模糊聚类算法与去噪算法结合,采用去噪算法对图像

预处理,随后于目标函数新增空间信息局部二阶矩和统计信息簇间项进行分割。然而,模糊聚类算法的目标函数易收敛到局部极值点甚至是鞍点,无法找到全局最优解。

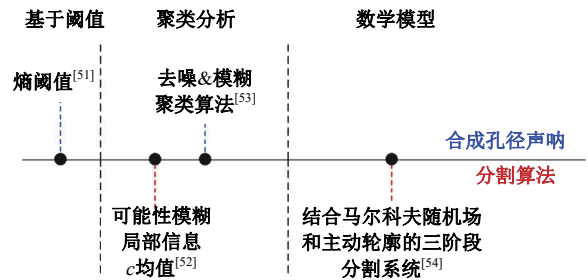


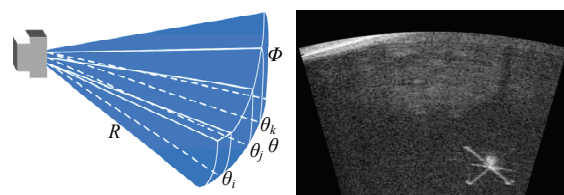
图 8 基于 SAS 图像在分割算法中的主要工作

Fig. 8 Major works in segmentation algorithms based on SAS images

在文献[54]中,作者提出一个 3 阶段的分割方案,第 1 阶段使用 MRF 将图像分割为目标高亮区、阴影区和背景混响区,第 2 阶段使用合作统计蛇(Cooperating statistical snake, CSS)主动轮廓模型进一步分割目标高亮区与阴影区,最后对分割好的区域分类。CSS 是该系统中极其重要的一部分,对散斑噪声的鲁棒性好,且能恢复前一阶段 MRF 没有理想分割的区域。然而,该算法目前无法满足 SAS 实时分割大面积海洋地质的需求。未来可以朝该方案继续探索适合 SAS 应用场景的更优方法,如设计新的优化算法使系统更快更准地收敛。

### 3 前视声呐

如图 9(a)所示,前视声呐(Forward-Looking sonar, FLS)以一定的水平开角  $\theta$  和垂直开角  $\varphi$  向前方发射声波,可到达最远探测距离  $R$ 。FLS 发射的波束可在水平面上划分为若干个固定垂直开角  $\varphi$  的扇形波束,在水平开角一定时,波束越多,图像的水平分辨率越高。此外,在水下探测中,可通过提高发射脉冲的带宽来提高接收的信噪比,同时提高系统的径向分辨率。



(a)FLS成像原理示意图

(b)真实的FLS图像

图 9 FLS 成像示意图

Fig. 9 Imaging schematic of FLS

FLS的探测距离有限,但成像分辨率较高,因此,目前在特定目标物检测、UUV导航避障等场景中得到广泛应用。

### 3.1 水下图像分类算法

基于FLS图像分类的相关研究在图10中展示。声呐图像的统计模式识别分类方法主要依靠目标阴影识别。但文献[55]利用声波垂直于目标界面切线会产生强烈的几何散射高光,且不同目标的几何散射高光特征不同这一发现,提出基于几何散射高光的二维分布近似目标特征的FLS图像分类方案。该方法有一定的局限性,FLS执行实时分类任务发射声波不一定与目标界面法线平行,导致产生的几何散射高光无法有效代表目标特征。



图10 基于FLS图像在分类算法中的主要工作

Fig. 10 Major works in classification algorithms based on FLS images

传统ML分类器也被应用到FLS图像。文献[56]评估了SVM、Random Forest和KNN分类器在该领域的性能。其中参数 $K=1$ 的KNN算法分类FLS图像的性能最优。

目前,基于DL的FLS图像分类算法研究非常广泛。在文献[57]中,作者参照AlexNet架构设计一个端到端模型,采用迁移学习的训练方法,准确率达到97.3%。文献[58]设计了一个轻量级CNN模型,相较于TM,可以用更少的参数实现更高精度。文献[59]采用经典的CNN模型AlexNet和GoogleNet对含有噪声的FLS图像分类。Matias Valdenegro-Toro针对能嵌入UUV低功耗系统的FLS图像分类算法做了大量的工作,通过实验证明最大池化(Max pooling, MP)层能减少参数、提高计算性能,且在小数据集上使用Dropout正则化与ADAM训练CNN性能更好。基于以上先验知识,文献[60]中,Matias Valdenegro-Toro设计了小型CNN,最高准确率可达99.7%,比TM、传统ML分类器和其他CNN模

型性能更好、需要的参数更少,且在树莓派2代上测试,时间仅用41~61 ms,准确率为98.8%~99.7%。

### 3.2 水下图像检测算法

如图11所示,目前基于FLS图像的检测算法可分为统计模式识别和DL两类。

在文献[61]中,作者将设计的目标模板与待测试图像对比进行相关性匹配验证。文献[62]针对目标被部分遮挡的现象,提出局部TM。然而,TM算法对噪声和强度变化的鲁棒性不高,且匹配效率低。

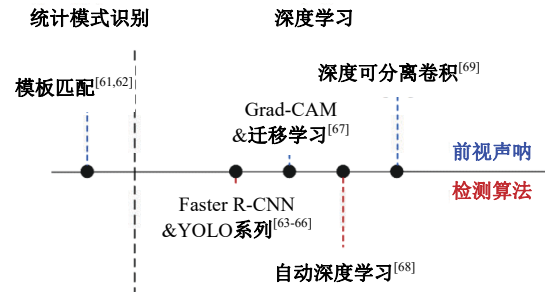


图11 基于FLS图像在检测算法中的主要工作

Fig. 11 Major works in detection algorithms based on FLS images

随着DL的兴起,CNN模型在FLS图像检测任务中得到广泛应用。Faster R-CNN和YOLO系列等经典模型应用到该领域<sup>[63-66]</sup>,检测精度都具有一定的竞争力。文献[67]采用Grad-CAM提高定位能力,且使用迁移学习提高模型提取特征的能力。在文献[68]中,作者引入自动深度学习(AutoDL)法,自主设计最优检测器且自训练各层权重,不仅实现了高检测精度,更加快了检测速度。AutoDL算法包括神经架构搜索与参数优化两部分,神经架构搜索就是根据搜索策略在确定的搜索空间寻找合适的结构,并用评估策略评估找到的所有架构以得到最优检测器。Qin等<sup>[69]</sup>探索能嵌入低功耗平台的FLS图像检测方法,提出了采用深度可分离卷积降低参数数量与操作成本的MRF-Net (Multiple receptive field network)模型。

不同于SSS大规模检测多样性目标,FLS通常应用于特定目标物实时检测。虽然CNN模型取得了不错的检测结果,但参数冗余,且并非FLS检测某种特定目标物的最优架构。采用AutoDL<sup>[68]</sup>自主设计检测器架构且自训练该检测器,可以有效控制模型复杂度和计算复杂度,能在检

测速度和精度之间实现更好的权衡,适合嵌入水下机器人实时检测FLS成像的目标。

### 3.3 水下图像分割算法

通常,UUV等水下勘探设备依靠FLS成像快速精准地识别障碍物,因此,应用于FLS成像领域的分割算法需要提高在非目标高亮区和目标高亮区交织的复杂区域上的精准度,且计算复杂度低。目前该领域的相关研究可分为两类,在图12中展示。

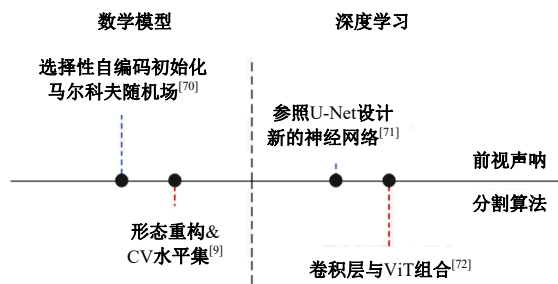


图 12 基于 FLS 图像在分割算法中的主要工作

Fig. 12 Major works in segmentation algorithms based on FLS images

第1类是基于数学模型和几何推理的分割算法。MRF选择理想的初始标签有困难,导致无法找到全局最优解。文献[70]提出利用选择性自编码(Selective autoencoding, SAE)初始化MRF模型,针对不同区域选择不同的激活函数自适应地保留每部分的本质特征。该方法虽然可以选择较理想的初始状态,但迭代条件模式(Iterated conditional mode, ICM)对细小目标识别率低。为提高小目标的轮廓分割精度,Zhang等<sup>[9]</sup>提出一种结合形态重构与水平集的新型CV水平集分割方法MRLSM,但其演化效果依赖于能量函数的能量惩罚项和时间步长项的初始设置。

第2类是基于CNN模型的分割方案。文献[71]在U-Net基准框架下,嵌入新设计的注意力模块和空洞卷积金字塔(Atrous pyramid, AP)模块,取得了较高的分割准确率。在文献[72]中,作者为降低CNN的计算开销,将ViT(Vision Transformer)思想引入FLS图像分割领域,构建卷积层和ViT组合的U型网络MiTU-Net,极大降低了模型的复杂度,且具有很强的鲁棒性。MiTU-Net的分割精度略低于文献[71]中的模型精度,但模型复杂度由124.46 GFLOPs(Floating Point Operations)降低到9.86 GFLOPs,适合嵌入资源有限的平台开展实时目标分割。因此,文

献[72]提出的MiTU-Net更能满足FLS图像分割任务的应用需求。

## 4 本文总结

本文综述了基于SSS、SAS和FLS在图像分类、检测以及分割任务中主要的水下目标识别算法。根据图13中总结的这3类声呐的成像特点与应用场景,分析总结不同成像声呐对应图像处理任务下目标识别算法的特点以及目前性能较好的方案,如图14所示。

## 5 未来展望

(1)用于科学研究的无任何偏好、多种类以及多数量的开源成像声呐数据集稀缺。未来需要:①开展成像声呐图像生成模拟器的研究。目前已经存在相关方法,如调整光学图像不同区域的强度分布模拟声呐图像的半合成数据方法<sup>[1]</sup>,采用生成对抗网络(Generative adversarial network, GAN)<sup>[73]</sup>以及基于风格迁移的端到端网络<sup>[74]</sup>等声呐图像生成方法。进一步结合不同成像声呐的成像特点,设计更加逼真的成像声呐图像生成模拟器是应对缺乏声呐图像挑战的趋势;②不局限于文献[72]采用的线性拉伸、旋转和颜色变换等图像增强方法进行成像声呐数据扩增,还可以利用Curvelet变换等仿人眼视觉方案增强声呐图像。

(2)由于水下环境的复杂性,成像声呐的成像质量受大量散斑噪声干扰,导致目标识别精度低。目前已有递归最小二乘(Recursive least squares, RLS)法<sup>[75]</sup>、非局部均值散斑滤波<sup>[76]</sup>以及基于CNN的自编码器<sup>[30]</sup>等算法在声呐图像去噪领域取得了一定的进展。未来如何结合不同成像声呐图像特点探索最大限度保留细节的去噪方法是极具前瞻性和挑战性的课题。

(3)声呐图像受声波成像原理的影响存在大量的非目标高亮区,相较于自然图像<sup>[77]</sup>和医学图像<sup>[78]</sup>的目标识别方法,现有的成像声呐水下目标识别方法在目标高亮区与非目标高亮区交织的复杂区域误报率(False positive rate, FPR)较高。在面向复杂场景的光学图像处理领域,文献[79]提出的CGNet取得了较好的效果。因此,可通过借鉴CGNet在网络各个阶段收集上下文信息的思路,提高水下目标识别算法融合局部特征、周围上下文特征以及全局上下文信息的能力以有效抑制

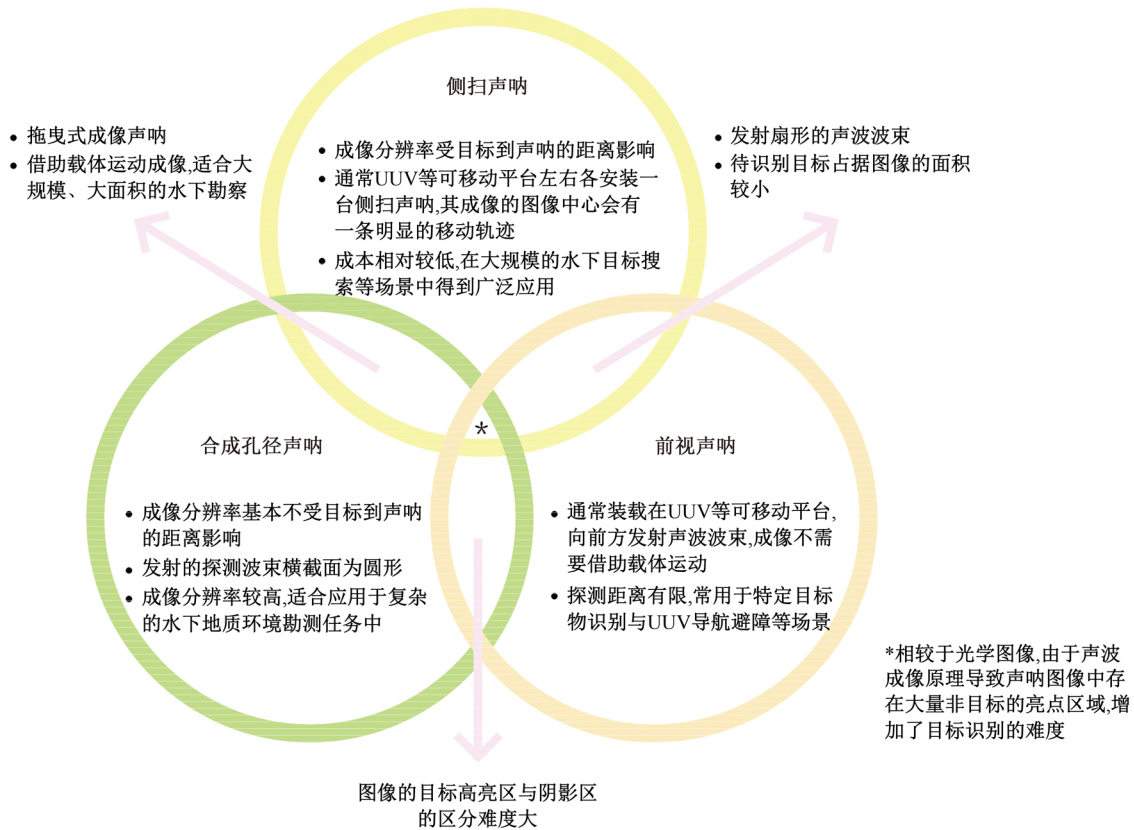


图 13 三类成像声呐成像特点与应用场景的异同比较

Fig. 13 Comparison of similarities and differences in imaging characteristics and application scenarios of three types of imaging sonar

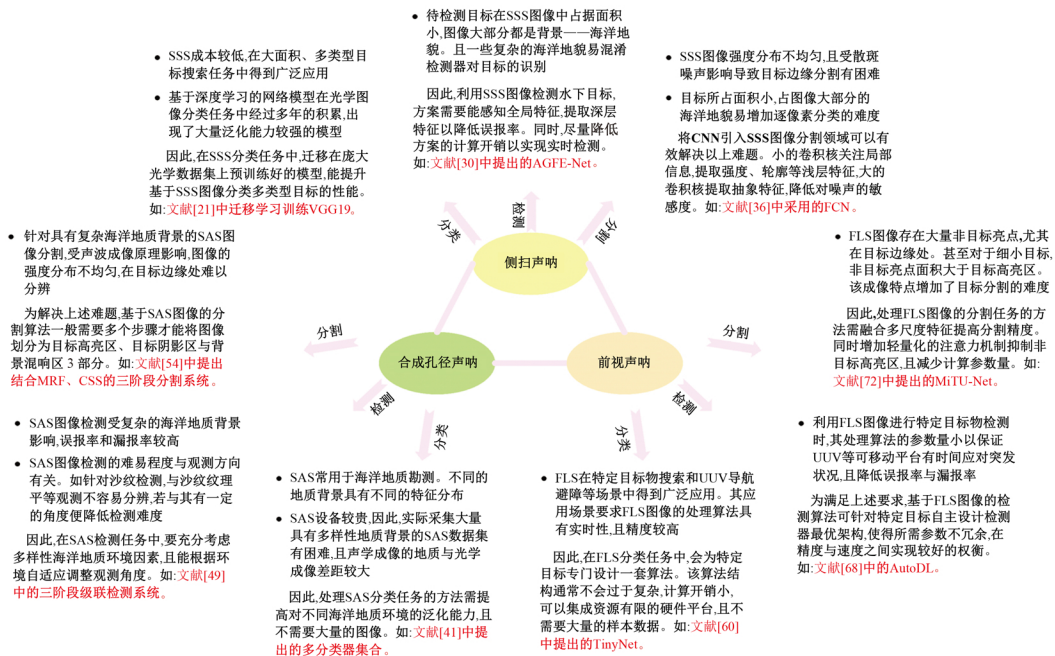


图 14 三类成像声呐对应图像处理任务水下目标识别算法的特点总结

Fig. 14 Summary of characteristics of underwater target recognition algorithms for three types of imaging sonar corresponding to image processing tasks

非目标高亮区干扰是未来重要的研究方向之一。  
(4)随着UUV逐渐呈现智能化发展趋势,这

将对水下机器人通过成像声呐自主勘探、识别目标的技术提出更高要求。目前,文献[60]提出的

FLS 图像分类算法已成功嵌入树莓派 2 代。未来可以进一步探索减少计算参数并节省内存空间的方法,使其能运行在资源有限的硬件系统进行智能化水下作业。

#### 参考文献:

- [1] Kim H G, Seo J, Kim S M. Underwater optical-sonar image fusion system[J]. *Sensors*, 2022, 22(21): 8445-8457.
- [2] Kumar N, Mitra U, Narayanan S S. Robust object classification in underwater sidescan sonar images by using reliability-aware fusion of shadow features[J]. *IEEE Journal of Oceanic Engineering*, 2015, 40(3): 592-606.
- [3] 罗逸豪, 刘奇, 张吟, 等. 基于深度学习的水下图像目标检测综述[J]. *电子与信息学报*, 2023, 45(6): 1-15.  
Luo Yi-hao, Liu Qi, Zhang Yin, et al. Review of underwater image object detection based on deep learning[J]. *Journal of Electronics & Information Technology*, 2023, 45(6): 1-15.
- [4] Chen H Y, Zhang H Q. Combined fuzzy clustering and Chan Vese model for multicolor gradual ochotona curzoniae image segmentation[J]. *Mathematical Problems in Engineering*, 2021, 2021(1): 1-9.
- [5] Ge Q, Ruan F, Qiao B, et al. Side-scan sonar image classification based on style transfer and pre-trained convolutional neural networks[J]. *Electronics*, 2021, 10(15): 1823-1838.
- [6] Oliveira A J, Ferreira B M, Diamant R, et al. Sonar-based cable detection for in-situ calibration of marine sensors[C]//*IEEE OES Autonomous Underwater Vehicles*, Singapore, 2022:1-6.
- [7] Chen R Y, Chen Y. Improved convolutional neural network YOLOv5 for underwater target detection based on autonomous underwater helicopter[J]. *Journal of Marine Science and Engineering*, 2023, 11(5): 989-1002.
- [8] Mignotte M. Saliency map estimation using a pixel-pairwise-based unsupervised markov random field model[J]. *Mathematics*, 2023, 11(4): 986-997.
- [9] Zhang B M, Zhong T, Shi Z F, et al. An underwater small target boundary segmentation method in forward-looking sonar images[J]. *Applied Acoustics*, 2023, 207(2): 109341-109350.
- [10] Yang Z W, Zhao J H, Zhang H M, et al. A side-scan sonar image synthesis method based on a diffusion model[J]. *Journal of Marine Science and Engineering*, 2023, 11(6): 1103-1115.
- [11] Dura E, Bell J, Lane D. Superellipse fitting for the recovery and classification of mine-like shapes in side-scan sonar images[J]. *IEEE Journal of Oceanic Engineering*, 2008, 33(4): 434-444.
- [12] Reed S, Petilot Y, Bell J. A model based approach to mine detection and classification in sidescan sonar [C]//*Oceans*, San Diego, USA, 2003: 1402-1407.
- [13] Sun J, Fan M, Cui X D, et al. A multibeam seafloor classification method combining ReliefF and random forest model[J]. *Marine Science Bulletin*, 2022, 41: 131-139.
- [14] Rhinelander J. Feature extraction and target classification of side-scan sonar images[C]//*IEEE Symposium Series on Computational Intelligence (SSCI)*, Athens, Greece, 2016: 1-6.
- [15] 熊明宽, 吴自银, 李守军, 等. 基于 SVM 的海底声纳图像底质识别[J]. *海洋通报*, 2012, 2012(4): 409-414.  
Xiong Ming-kuan, Wu Zi-yin, Li Shou-jun, et al. Seafloor sonar sediment image recognition with the SVM[J]. *Marine Science Bulletin*, 2012, 2012(4): 409-414.
- [16] 郭军, 马金凤, 王爱学. 基于 SVM 算法和 GLCM 的侧扫声纳影像分类研究[J]. *测绘与空间地理信息*, 2015, 38(3): 4-13.  
Guo Jun, Ma Jin-feng, Wang Ai-xue. Study of side scan sonar image classification based on SVM and GLCM[J]. *Geomatics & Spatial Information Technology*, 2015, 38(3): 4-13.
- [17] Wang J C, Wu Z Y, Wang M W, et al. ELM-AdaBoost method of acoustic seabed sediment classification[J]. *Acta Oceanologica Sinica*, 2021, 43: 144-151.
- [18] Yang F L, Liu J N. Seabed classification using BP Neural Network based on GA[J]. *Acta Oceanologica Sinica*, 2006, 4: 523-531.
- [19] 熊明宽, 吴自银, 李守军, 等. 基于遗传小波神经网络的海底声学底质识别分类[J]. *海洋学报*, 2014, 36(5): 90-97.  
Xiong Ming-kuan, Wu Zi-yin, Li Shou-jun, et al. Wavelet neural network identification and classification of sediment seabed sonar images based on genetic algorithms[J]. *Acta Oceanologica Sinica*, 2014, 36(5): 90-97.
- [20] Luo X, Qin X, Wu Z, et al. Sediment classification of small-size seabed acoustic images using convolutional neural networks[J]. *IEEE Access*, 2019, 7:

- 98331-98339.
- [21] Huo G, Wu Z, Li J. Underwater object classification in sidescan sonar images using deep transfer learning and semisynthetic training data[J]. *IEEE Access*, 2020, 8: 47407-47418.
- [22] Petillot Y R, Reed S R, Bell J M. Real time AUV pipeline detection and tracking using side scan sonar and multi-beam echo-sounder[C]//*Oceans*, Biloxi, USA, 2002:217-222.
- [23] Fakiris E, Papatheodorou G, Geraga M, et al. An automatic target detection algorithm for swath sonar backscatter imagery, using image texture and independent component analysis[J]. *Remote Sensing*, 2016, 8(5): 373-380.
- [24] Zheng L Y, Tian K. Detection of small objects in sidescan sonar images based on POHMT and Tsallis entropy[J]. *Signal Processing the Official Publication of the European Association for Signal Processing*, 2018, 142: 168-177.
- [25] Sawas J, Petillot Y, Pailhas Y. Cascade of boosted classifiers for rapid detection of underwater objects [C]//*The 10th European Conference on Underwater Acoustics*, Istanbul, Turkey, 2010:1507-1516.
- [26] 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.
- [27] Ma Q, Jiang L, Yu W, et al. Training with noise adversarial network: a generalization method for object detection on sonar image[C]//*Workshop on Applications of Computer Vision*, Snowmass, USA, 2020: 188-196.
- [28] Jiang L, Cai T, Ma Q, et al. Active object detection in sonar images[J]. *IEEE Access*, 2020, 8: 102540-102553.
- [29] Topple J M, Fawcett J A. MiNet: efficient deep learning automatic target recognition for small autonomous vehicles[J]. *IEEE Geoscience and Remote Sensing Letters*, 2021, 18(6): 1014-1018.
- [30] Wang Z, Zhang S, Huang W, et al. Sonar image target detection based on adaptive global feature enhancement network[J]. *IEEE Sensors Journal*, 2022, 22(2): 1509-1530.
- [31] Tang X Y, Zhang X W, Xu X L, et al. Methods for underwater sonar image processing in objection detection[C]//*International Conference on Computer Systems, Electronics and Control(ICCSEC)*, Dalian, China, 2017:941-944.
- [32] Celik T, Tjahjadi T. A novel method for sidescan sonar image segmentation[J]. *IEEE Journal of Oceanic Engineering*, 2011, 36(2): 186-194.
- [33] Mignotte M, Collet C, Pérez P, et al. Sonar image segmentation using an unsupervised hierarchical MRF model[J]. *IEEE Transactions on Image Processing*, 2000, 9(7): 1216-1231.
- [34] Ye X F, Zhang R H, Liu R X, et al. Sonar image segmentation based on GMRF and level-set models [J]. *Ocean Engineering*, 2010, 37(10): 891-901.
- [35] Huo G, Yang S X, Li Q, et al. A robust and fast method for sidescan sonar image segmentation using nonlocal despeckling and active contour model[J]. *IEEE Transactions on Cybernetics*, 2016, 47(4): 1-18.
- [36] Wu M, Wang Q, Rigall E, et al. ECNet: efficient convolutional networks for side scan sonar image segmentation[J]. *Sensors*, 2019, 19(9): 2009-2010.
- [37] Wei B, He C L, Xing S Y, et al. Accelerated deconvolved imaging algorithm for 2D multibeam synthetic aperture sonar[J]. *Sensors*, 2022, 22(16): 6016-6025.
- [38] Hollesen P, Connors W A, Trappenberg T P. Comparison of learned versus engineered features for classification of mine like objects from raw sonar images [C]//*Canadian Conference on Advances in Artificial Intelligence*, Hollensen, Canada, 2011: 174-185.
- [39] McKay J, Monga V, Raj R G. Robust sonar ATR through bayesian pose corrected sparse classification [J]. *IEEE Transactions on Geoscience & Remote Sensing*, 2017, 55(10): 5563-5576.
- [40] Fei T, Kraus D, Zoubir A M. Contributions to automatic target recognition systems for underwater mine classification[J]. *IEEE Transactions on Geoscience & Remote Sensing*, 2014, 53(1): 505-518.
- [41] Williams D P, Fakiris E. Exploiting environmental information for improved underwater target classification in sonar imagery[J]. *IEEE Transactions on Geoscience & Remote Sensing*, 2014, 52(10): 6284-6297.
- [42] Denos K, Ravaut M, Fagette A, et al. Deep learning applied to underwater mine warfare[C]//*Oceans*, Aberdeen, UK, 2017:1-7.
- [43] Warakagoda N, Midtgaard I. Transfer-learning with deep neural networks for mine recognition in sonar images[C]//*Synthetic Aperture Sonar, Synthetic Aperture Radar*, Lerici, Italy, 2018:1-5.
- [44] Williams D P. Transfer learning with SAS-image

- convolutional neural networks for improved underwater target classification[C]//IEEE International Geoscience and Remote Sensing Symposium, Yokohama, Japan, 2019: 78-81.
- [45] Klausner N, Azimi-Sadjadi M R. Non-Gaussian target detection in sonar imagery using the multivariate laplace distribution[J]. *IEEE Journal of Oceanic Engineering*, 2015, 40(2): 452-464.
- [46] Abu A, Diamant R. A statistically-based method for the detection of underwater objects in sonar imagery [J]. *Sensors Journal*, 2019, 19(16): 6858-6871.
- [47] Azimi-Sadjadi M R, Klausner N, Kopacz J. Detection of underwater targets using a subspace-based method with learning[J]. *IEEE Journal of Oceanic Engineering*, 2017, 42(4): 869-879.
- [48] Sawas J, Petillot Y. Cascade of boosted classifiers for automatic target recognition in synthetic aperture sonar imagery[C]//Meetings on Acoustics, Edinburgh, Scotland, 2012: 1-11.
- [49] Williams D P. Fast target detection in synthetic aperture sonar imagery: a new algorithm and large-scale performance analysis[J]. *IEEE Journal of Oceanic Engineering*, 2015, 40(1): 71-92.
- [50] Berthomier T, Williams D P, Dugelay S. Target localization in synthetic aperture sonar imagery using convolutional neural networks[C]//OCEANS MTS/IEEE SEATTLE, Seattle, USA, 2019: 1-9.
- [51] Maussang F, Chanussot J, Hetet A, et al. Mean - standard deviation representation of sonar images for echo detection: application to SAS images[J]. *IEEE Journal of Oceanic Engineering*, 2008, 32(4): 956-970.
- [52] Peeples J, Cook M, Suen D, et al. Comparison of possibilistic fuzzy local information c-means and possibilistic k-nearest neighbors for synthetic aperture sonar image segmentation[J]. *Detection and Sensing of Mines, Explosive Objects, and Obscured Targets XXIV*, 2019, 19(4): 1904-1914.
- [53] Abu A, Diamant R. Enhanced fuzzy-based local information algorithm for sonar image segmentation[J]. *IEEE Transactions on Image Processing*, 2019, 29: 445-460.
- [54] Bell J M, Petillot Y R, Lebart K, et al. Target recognition in synthetic aperture and high resolution side-scan sonar[C]//IET Seminar on High Resolution Imaging and Target Classification, London, UK, 2006: 99-106.
- [55] Lu Y C, Sang E F. Underwater target's size/shape dynamic analysis for fast target recognition using sonar images[C]//Proceedings of International Symposium on Underwater Technology, Tokyo, Japan, 1998:257-262.
- [56] Dos Santos M, Ribeiro P O, Núñez P, et al. Object classification in semi structured environment using forward-looking Sonar[J]. *Sensors*, 2017, 17(10): 2235-2251.
- [57] Jin L, Liang H, Ang C. Accurate underwater ATR in forward-looking sonar imagery using deep convolutional neural networks[J]. *IEEE Access*, 2019, 7: 125522-125531.
- [58] Valdenegro-Toro M. Object recognition in forward-looking sonar images with convolutional neural networks[C]//OCEANS MTS/IEEE Monterey, Monterey, USA, 2016: 1-6.
- [59] Nguyen H T, Lee E H, Lee S. Study on the classification performance of underwater sonar image classification based on convolutional neural networks for detecting a submerged human body[J]. *Sensors(Basel, Switzerland)*, 2020, 20(1): 94-100.
- [60] Valdenegro-Toro M. Deep neural networks for marine debris detection in sonar images[D]. Edinburgh: Engineering & Physical Sciences, Heriot-Watt University, 2019.
- [61] Wang W W, Cheng B B, Chen Y. A real-time object recognition for forward looking sonar[C]//The 2nd International Conference on Image, Vision and Computing (ICIVC), Chengdu, China, 2017: 58-61.
- [62] Hurtós N, Palomeras N, Nagappa S, et al. Automatic detection of underwater chain links using a forward-looking sonar[C]//MTS/IEEE OCEANS, Bergen, Norway, 2013: 1-7.
- [63] Zacchini L, Ridolfi A, Topini A, et al. Deep learning for on-board AUV automatic target recognition for optical and acoustic imagery[J]. *IFAC*, 2020, 53(2): 14589-14594.
- [64] Wu Y C. Sonar image target detection and recognition based on convolution neural network[J]. *Mobile Information Systems*, 2021, 2021(6): 1-8.
- [65] Kvasić I, Mišković N, Vukić Z. Convolutional neural network architectures for sonar-based diver detection and tracking[C]//Oceans, Marseille, France, 2019: 1-6.
- [66] Zacchini L, Topini A, Franchi M, et al. Autonomous underwater environment perceiving and modeling: an experimental campaign with FeelHippo AUV for forward looking sonar-based automatic target rec-

- ognition and data association[J]. *IEEE Journal of Oceanic Engineering*, 2022, 48: 277-296.
- [67] Lou G, Zheng R, Liu M, et al. Automatic target recognition in forward-looking sonar images using transfer learning[C]//*Global Oceans*, Biloxi, USA, 2020: 1-6.
- [68] Zhang P, Tang J, Zhong H, et al. Self-trained target detection of radar and sonar images using automatic deep learning[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2022, 60: 1-14.
- [69] Qin R, Zhao X, Zhu W, et al. Multiple receptive field network (MRF-Net) for autonomous underwater vehicle fishing net detection using forward-looking sonar images[J]. *Sensors*, 2021, 21(6): 1933-1945.
- [70] Song S, Si B, Feng X, et al. Label field initialization for MRF-based sonar image segmentation by selective autoencoding[C]//*Oceans*, Shanghai, China, 2016: 1-5.
- [71] Liang Y, Zhu X, Zhang J. Maanu-Net: multi-level attention and atrous pyramid nested U-Net for wrecked objects segmentation in forward-looking sonar images[C]//*IEEE International Conference on Image Processing (ICIP)*, Bordeaux, France, 2022: 736-740.
- [72] Liang Y, Zhu X, Zhang J. MiTU-Net: an efficient mix transformer U-like network for forward-looking sonar image segmentation[C]//*IEEE 2nd International Conference on Computer Communication and Artificial Intelligence (CCAI)*, Beijing, China, 2022: 149-154.
- [73] Jegorova M, Karjalainen A I, Vazquez J, et al. R2D2-GAN: unlimited resolution image generation for acoustic data[J]. *Marine Technology Society Journal*, 2021, 55(4): 99-107.
- [74] Chungath T T, Nambiar A M, Mittal A. Transfer learning and few-shot learning based deep neural network models for underwater sonar image classification with a few samples[J]. *IEEE Journal of Oceanic Engineering*, 2022, 49(1): 294-310.
- [75] Muthuraman D L, Santhanam S M. Contrast improvement on side scan sonar images using retinex based edge preserved technique[J]. *Mar Geophys Res*, 2022, 43(17): 1-19.
- [76] Jiao W, Zhang J. Sonar images classification while facing long-tail and few-shot[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2022, 60: 1-20.
- [77] 薛雅丽, 俞潼安, 崔闪, 等. 基于级联嵌套U-Net的红外小目标检测[J/OL]. [2024-01-13]. *吉林大学学报:工学版*, <https://doi.org/10.13229/j.cnki.jdxbgxb.20230785>.
- Xue Ya-li, Yu Tong-an, Cui Shan, et al. Infrared small target detection based on cascaded nested U-Net[J/OL]. [2025-01-13]. *Journal of Jilin University (Engineering and Technology Edition)*, <https://doi.org/10.13229/j.cnki.jdxbgxb.20230785>.
- [78] 刘萍萍, 商文理, 解小宇, 等. 基于细粒度分析的不均衡图像分类算法[J/OL]. [2025-01-13]. *吉林大学学报:工学版*, <https://doi.org/10.13229/j.cnki.jdxbgxb.20230991>.
- Liu Ping-ping, Shang Wen-li, Xie Xiao-yu, et al. An unbalanced image classification algorithm based on fine-grained analysis[J/OL]. [2025-01-13]. *Journal of Jilin University (Engineering and Technology Edition)*, <https://doi.org/10.13229/j.cnki.jdxbgxb.20230991>.
- [79] Wu T, Tang S, Zhang R, et al. CGNet: a light-weight context guided network for semantic segmentation[J]. *IEEE Transactions on Image Processing*, 2021, 30: 1169-1179.