

中图法分类号: TN911.73 文献标识码: A 文章编号: 1006-8961(2026)05-1372-22

论文引用格式: Lin J, Chen W L, Xu X Y and Zhao T S. 2026. Review of the evaluation and enhancement of underwater sonar image quality. Journal of Image and Graphics, 31(5):1372-1393(林杰, 陈炜玲, 徐晓怡, 赵铁松. 2026. 声呐图像质量评价与增强研究综述. 中国图象图形学报, 31(5):1372-1393)[DOI:10.11834/jig.250421]

## 声呐图像质量评价与增强研究综述

林杰<sup>1,2</sup>, 陈炜玲<sup>1,2,3\*</sup>, 徐晓怡<sup>1,2</sup>, 赵铁松<sup>1,2</sup>

1. 福州大学物理与信息工程学院, 福州 350108; 2. 福建省媒体信息智能处理与无线传输重点实验室, 福州 350108;
3. 福建省模式识别与图像理解重点实验室, 厦门 361024

**摘要:** 声呐图像(Sonar Image, SI)是海洋信息的重要载体,然而,受水体散射、多径效应及声波波长限制等因素影响,声呐图像普遍存在分辨率低、斑点噪声强以及对比度弱等现象。这些问题不仅直接影响声呐图像的视觉效果,更会显著降低目标识别等下游任务的精度与可靠性。因此,发展有效的声呐图像质量评价(Image Quality Assessment, IQA)与增强技术,对保障海洋探测任务至关重要。尽管声呐图像处理研究近年来取得显著进展,但现有综述工作主要集中于水下光学图像,或仅聚焦于特定应用任务,对声呐图像质量评价体系及增强算法的系统性梳理、技术演进脉络的深入剖析尚显不足。为此,本文首次构建了声呐图像质量的“评价—增强”研究框架,旨在提供一份更系统、全面的介绍。系统梳理了声呐图像质量评价与质量增强领域的关键技术演进、代表性模型及最新研究进展;实验方面,针对质量评价领域,在公开数据集上对梳理出的主流声呐图像质量评价算法进行了性能对比实验,揭示了其在多种失真场景下的性能差异与适用性;针对质量增强领域,鉴于开源的声呐图像增强算法较少,重点评估了代表性通用超分辨率(Super-Resolution, SR)重建与去噪算法在声呐图像上的性能表现。综上,本文旨在为声呐图像处理领域的发展提供系统的理论参照与技术路线,通过实验揭示算法性能瓶颈与适用边界,最终为研究人员把握现状、理解挑战以及规划未来研究提供重要支撑。

**关键词:** 声呐图像(SI);图像质量评价(IQA);图像增强;超分辨率重建;图像去噪

## Review of the evaluation and enhancement of underwater sonar image quality

Lin Jie<sup>1,2</sup>, Chen Weiling<sup>1,2,3\*</sup>, Xu Xiaoyi<sup>1,2</sup>, Zhao Tiesong<sup>1,2</sup>

1. College of Physics and Information Engineering, Fuzhou University, Fuzhou 350108, China;
2. Fujian Key Laboratory for Intelligent Processing and Wireless Transmission of Media Information, Fuzhou 350108, China;
3. Fujian Key Laboratory for Pattern Recognition and Image Understanding, Xiamen 361024, China

**Abstract:** Sonar imaging serves as a foundational technology for deep-sea exploration, resource mapping, and underwater infrastructure inspection. However, inherent physical constraints (e.g., narrow bandwidth and acoustic wavelength limits) coupled with complex underwater propagation effects (including water scattering, multipath interference, and reverberation) inevitably degrade sonar image quality. Such quality degradation commonly manifests as low resolution, intense speckle noise, and weak contrast. These quality issues directly impair visual perception and considerably reduce the accu-

收稿日期: 2025-09-08; 修回日期: 2025-11-17; 预印本日期: 2025-11-25

\* 通信作者: 陈炜玲 weiling.chen@fzu.edu.cn

**基金项目:** 国家自然科学基金项目(62171134, 61901119); 福建省制造业技术创新重点攻关及产业化项目(2025G016); 福建省促进海洋与渔业产业高质量发展专项资金项目(FJHYF-L-2025-07-031);

**Supported by:** National Natural Science Foundation of China(62171134, 61901119); Fujian Technological Innovation Key Breakthrough and Industrialization Project(2025G016); Fujian Special Fund Project for High-Quality Development of Marine and Fishery Industrie(FJHYF-L-2025-07-031)

racy and reliability of downstream tasks, such as target recognition. Therefore, developing effective sonar image quality assessment and enhancement techniques is crucial for ensuring the efficacy of ocean exploration missions. Although much progress has been achieved by sonar image processing research in recent years, existing review works primarily focus on underwater optical images or specific application tasks. A systematic review of sonar image quality assessment frameworks and enhancement algorithms, along with an in-depth analysis of their technical evolution, remains insufficient. Furthermore, the articulation of key future challenges and breakthrough directions in this field lacks clarity. To address this issue, this study, for the first time, reviews and constructs a collaborative quality assessment-enhancement research framework for sonar images, aiming to provide a systematic and comprehensive survey. Single stimulus, double stimulus, and paired comparison: Centering on the core dimension of reference image availability, we systematically classify objective quality assessment methods and conduct an in-depth analysis of the latest developments in each category. Full-reference image quality assessment (IQA): we emphasize the analysis of local entropy-structure fusion models (local entropy backed sonar image (LESQP) and sonar image quality predictor (SIQP)) tailored for sonar statistical characteristics and reveal how they leverage edge sparsity and patch consistency to overcome gray-level nonuniformity. Concurrently, the gain boundaries of multiscale, frequency-domain, and perceptual weighting strategies in evaluating sonar compression artifacts are pointed out. Reduced-reference IQA: Three representative approaches are distilled: partial-reference sonar image quality predictor (PSIQP) based on information entropy comfort, sonar image utility quality assessment (SIUQA) oriented toward task utility, and TPSIQA, a meta-learning fusion framework for small-sample scenarios. task and perception oriented sonar image quality assessment (TPSIQA) has a three-tier framework for bandwidth-limited transmission and integrates application-specific features via selective ensemble learning. No-reference IQA: centered on attribute consistency theory, a new paradigm for cross-scenario no-reference models is constructed. This theory asserts that high-quality sonar images must maintain statistical consistency across four critical attributes, namely, regional distinctness, geometric integrity, detail preservation, and cleanliness, regardless of scene variation. Furthermore, no-reference contour degradation measurement (NRCDM), unified quality assessment method for sonar imaging and processing (UASIP), and perception-cognition aware sonar assessment (PCASS) models are analyzed. NRCDM quantifies contour degradation via feature ratios and support vector regression (SVR) regression. UASIP builds upon attribute consistency theory, enforcing the four task-critical attributes for cross-scenario generalization. PCASS evaluates super-resolution reconstruction (SR)-reconstructed images via hierarchical feature fusion, and SRIQA mimics ventral visual pathway processing. With regard to quality enhancement, we trace the evolution of SR from traditional interpolation and frequency-domain methods to deep learning paradigms. Early convolutional neural network (CNN)-based approaches (e. g. , ResNet blocks) and generative adversarial networks (e. g. , multi-branch generators) achieve 4× upscaling but struggle with texture fidelity. Recent breakthroughs integrate physical constraints and hybrid architectures. For example, STDPNet employs a dual-stream CNN-Transformer parallel structure to preserve global structures and local textures, overcoming the limitations of serial fusion. Its hybrid loss function, which combines mean absolute error, spectral loss, and local gradient-aware loss, effectively mitigates over-smoothing while enhancing edge integrity. Blind SR models address unknown degradation processes via dual-cross optimization, and MHGAN utilizes multiheaded adversarial learning to recover high-frequency details under data scarcity. Subsequently, the development trajectory of denoising techniques is reviewed. Early mean/median filters gave way to anisotropic diffusion, adaptive statistical modeling, and sparsity-driven optimization, whereas today's leading approaches embed anisotropic guidance into the kernel itself or couple dual-stage U-Nets with edge-aware losses. Specifically, dual-stage U-Net suppresses photometric noise while preserving textures, and anisotropic guided filtering introduces adaptive regularization and directional weighting, achieving a 6.3 dB peak signal-to-noise ratio gain and superior edge retention. Optimizes spatial pixel ranking for computational efficiency. Restoration techniques leverage multitask learning. For example, deep adaptive phase learning corrects synthetic aperture sonar (SAS) phase errors without iterative optimization, and recurrent neural networks/long short-term memory networks model reverberation dynamics. Furthermore, we investigate the emerging direction of multimodal fusion. RepDNet applies reparameterization to side-scan sonar (SSS) despeckling within a lightweight multibranch CNN, and FLSSNet integrates CNN-Transformer hybrids for cross-modal feature transformation to enhance saliency detection. In experimental section, systematic performance comparison experiments for sonar image quality assessment algo-

rithms are conducted on three public datasets (sonar image quality database(SIQD), super-resolution sonar image database (SRSID), and multi-scenario sonar image dataset (MSSID)). By quantitatively analyzing the performance of the three mainstream methods (i. e., full reference, reduced reference, and no reference) across different distortion scenarios, the experiments reveal substantial scene dependence in algorithm performance. For instance, PCASS exhibits prominent advantages in evaluating reconstructed images but struggles to adapt to other distortion types. SIQP is stable in traditional degradation evaluation but limited in reconstruction tasks. Moreover, validation experiments targeting general-purpose SR and denoising algorithms are performed to assess their applicability and performance boundaries on sonar images. The core value of this review lies in providing a systematic theoretical reference and technical roadmap for the advancement of sonar image processing; revealing algorithm performance bottlenecks and applicability boundaries through experiments; and offering crucial support for researchers to grasp the current situation, understand challenges, and plan future research.

**Key words:** sonar image (SI); image quality assessment (IQA); image enhancement; super-resolution reconstruction; image denoising

## 0 引言

声呐图像 (sonar image, SI) 作为深海探测信息的重要载体之一, 直接承载着海底地形、沉船目标、管线状态等关键信息(陈炜玲等, 2023), 是海洋资源勘探、环境监测和国防安全不可或缺的数据来源。然而, 受水体散射、多径效应、环境噪声及声波波长固有物理限制等因素影响, 声呐图像普遍存在分辨率低、对比度弱以及噪声复杂等退化问题(Xu等, 2025)。这些固有的质量缺陷使得图像中的关键信息难以被准确解析, 严重影响了目标识别、海底探测等后续分析任务的执行效果与可信度。因此, 建立一套科学、客观的声呐图像质量评价体系, 为海洋探测数据的有效性提供量化保障, 显得尤为迫切。

图像质量评价 (image quality assessment, IQA) 核心目标在于量化图像质量, 为算法优化和系统设计提供客观依据。在 IQA 体系中, 主观质量评价由人眼观察图像后, 根据个人感受对其质量好坏进行打分评价, 其独特价值在于能够捕捉复杂场景下难以建模的高阶感知信息。然而, 由于其成本昂贵且难以复现, 故需要与客观质量评价相辅相成, 共同建立完善的声呐 IQA 体系。而客观质量评价是通过构建计算模型, 提供自动化、可重复的质量度量。建立并完善的声呐 IQA 体系是优化声呐系统参数、量化算法性能以及保障海洋探测任务数据可信度的基础。

为应对声呐图像的退化问题, 质量增强技术迅速发展, 其核心目标在于提升关键目标的清晰度和信息保真度, 为后续任务提供可靠数据支撑。其中, 超分辨率 (super-resolution, SR) 重建技术致力于克服

声波波长限制, 恢复图像丢失的高频空间细节, 提升分辨率并锐化目标轮廓; 去噪技术则专注于抑制水体散射、多径效应和设备噪声引发的复杂噪声, 提高图像信噪比。此外, 图像恢复技术致力于解决特定的成像失真, 而多模态融合技术则通过整合声呐与其他传感器数据的互补信息, 在复杂环境下进一步提升图像质量和目标识别能力。

近年来, 在声呐图像处理领域的研究取得了显著成果。然而, 如表 1 汇总所示, 现有相关综述工作仍存在明显局限: 主要聚焦于水下光学图像, 或仅以特定任务为导向探讨声呐图像应用, 缺乏对声呐图像质量评价体系及增强算法的系统性梳理与深入剖析。

为此, 本文旨在建立一个贯通“质量评价”与“质量增强”的协同研究框架, 通过系统性回顾与分析, 为该领域贡献一份结构清晰、内容深入的综述, 主要贡献如下: 1) 详细整理主观质量评价方法, 重点阐述客观质量评价方法的分类及每个类别的代表性模型, 并在公开声呐图像数据集上对主流客观评价算法进行性能对比实验, 揭示不同失真场景下算法的性能差异, 为算法选择提供参考; 2) 系统梳理从传统方法到深度学习模型的声呐图像质量增强技术的研究进展与发展脉络, 在公开的声呐图像数据集上对代表性通用增强算法进行验证实验, 评估其在声呐图像上的适用性与性能边界, 为算法选型与改进提供依据; 3) 深入探讨声呐图像质量评价与增强领域当前面临的关键挑战, 理清核心问题, 并展望具有潜力的未来研究方向。

本文结构框架如图 1 所示, 共 6 部分: 1) 引言, 简要介绍声呐图像局限性及其质量评价与增强的意义, 并阐明综述主要贡献及文章框架结构; 2) 声呐成

表 1 相关综述汇总  
Table 1 Summary of relevant reviews

作者/年份	研究内容摘要	局限性
黄海宁等人 (2024)	系统归纳了声呐成像原理及目标识别方法	聚焦识别任务本身, 缺乏对图像质量评价及增强的探讨
焦文沛等人 (2025)	系统介绍了声呐图像智能感知算法, 并展望未来研究方向	未涉及声呐图像质量评价的指标及方法体系
周玲等人 (2025)	介绍了水下光学图像的退化机理及增强方法的最新进展	仅聚焦光学图像, 未涉及声呐图像
Zhang 等人 (2025)	系统介绍了声呐图像去噪方法, 指出目前所面临的挑战	聚焦去噪而非全面的质量评价与增强
Cai 等人 (2025b)	回顾了声呐图像分割的技术演进, 指明未来挑战与发展方向	仅聚焦分割任务, 未涉及图像质量评价与增强方法
陈炜玲等人 (2023)	综述了水下光学与声学图像的处理、质量评价及机器视觉应用	缺乏对声呐图像退化与增强技术的系统梳理
Xu 等人 (2025)	系统回顾了水下光学与声呐图像质量评价的方法	以质量评价为核心, 对声呐图像增强技术描述较少
庄培显等人 (2026)	系统构建了水下光场成像“理论—应用”框架	聚焦于水下光学成像范式, 未涵盖声学成像原理
郭继昌等人 (2017)	综述水下图像增强与复原方法, 并建立水下图像质量评测体系	缺乏对声呐图像质量评价标准及增强算法的探讨

像原理与特性, 介绍声呐成像的基本原理、关键特性及声呐图像质量的主要退化因素; 3) 声呐图像质量评价体系, 系统梳理主观质量评价方法, 并重点阐述全参考、半参考和无参考 3 类客观质量评价方法的原理、模型及最新进展; 4) 声呐图像质量增强技术, 详细介绍 SR 重建、去噪、图像恢复及基于数据融合

等增强技术的研究进展; 5) 性能对比与分析, 包含两部分实验: 一是在公开数据集上对主流声呐图像质量评价算法进行性能对比分析, 二是对代表性通用增强算法在声呐图像上的适用性进行验证实验; 6) 结语, 总结当前声呐图像质量评价与增强领域面临的核心挑战, 展望未来发展趋势。

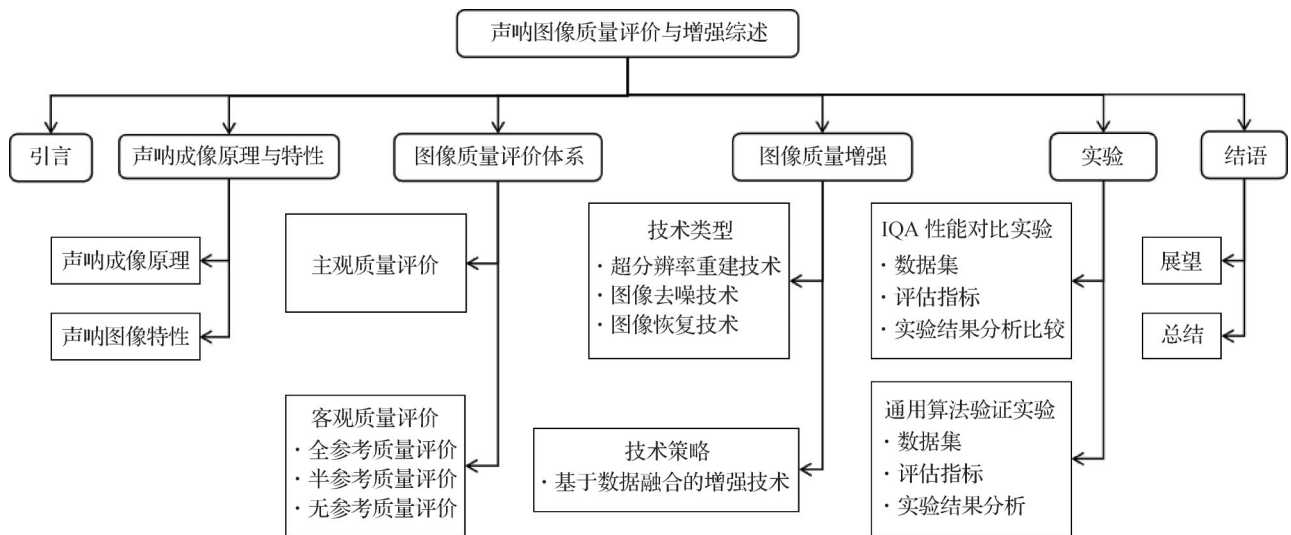


图 1 文章结构框架图

Fig. 1 Article structure framework diagram

## 1 声呐图像特性及成像原理

### 1.1 声呐成像原理

传统光学成像依赖可见光在直线传播路径下的几何投影原理,通过透镜聚焦和光电转换形成反映目标形态、色彩及纹理的图像。然而,这一机制导致其在水下面临根本性制约:水分子和悬浮颗粒对光线的强烈散射效应会严重模糊图像细节,同时水体对红光等波长的选择性吸收则导致色彩失真与信号快速衰减。相比之下,声呐成像主要是通过声呐发射器向水中发射一定频率和强度的声波,当声波遇到水中的物体或界面时,会产生反射、折射和散射等现象,从而改变声波的方向、频率和强度,然后声呐接收器接收这些变化后的声波,再经过信号处理和分析,最终得到水中物体或界面的位置、距离、速度、形状和性质等信息。声呐成像以声波为信息载体,其传播特性依赖水体的物理参数。声波在海水会受到水温、盐度以及水压梯度所形成的声速剖面影响显著,从而产生折射现象,导致声线曲化并引发目标定位的偏移。这种介质依赖性给予了声呐在浑浊水域中的穿透优势,可实现千米量级深水探测,但同时也成为其几何畸变的根源。而光学成像依赖波长 $400\sim 700\text{ nm}$ 的电磁波,其传播受水体吸收与散射效应严重制约,由于水分子对红光的选择性吸收致使在深度大于 $10\text{ m}$ 时会出现色彩失真,以及悬浮颗粒导致的前向散射会造成图像模糊(王柯俨等,2022)。因此,即使在清洁水体中,其有效探测深度通常不超过 $50\text{ m}$ (纪勋等,2024)。

声呐与光学成像的本质差异源于其信息载体与信息获取机制的根本不同:声呐依赖主动发射的声波,通过分析回波信号的时延与强度推断目标的空间位置与物理属性;而光学成像主要基于被动接收物体反射/发射的可见光,利用几何投影原理直接映射目标的形态与纹理细节。这种差异导致二者在介质依赖性、分辨率、对比度、边缘信息及噪声类型上形成鲜明对比。

### 1.2 声呐图像特性

声呐技术利用水中声波探测、定位和识别水下目标,如图2所示。然而,声波在水中传播时易受水体散射、海洋生物、环境噪声(如船舶干扰)及设备物理限制的影响,导致声呐图像普遍具有分辨率低、对

比度弱和噪声复杂等特征。具体表现为:1)噪声复杂:声波传播过程中受海床反射、水体悬浮物散射及设备自身电子噪声等多源干扰,形成耦合噪声;2)分辨率低:受限于声波波长,声呐图像分辨率远低于光学图像,导致目标边缘模糊并产生拖尾现象;3)对比度弱:水体中气泡和浮游生物的散射效应会引起低频信号畸变,产生类似“光学雾霾”的对比度衰减效应。声呐图像与传统光学图像的关键特征差异详见表2所示。值得注意的是,这些声呐图像特性往往呈现非线性耦合效应。例如,在浑浊港口水域,高浓度的悬浮颗粒物会同时加剧声波散射(增加噪声)和信号衰减(降低对比度),形成复合退化现象。由此可见,声呐图像的退化特征不仅造成图像信息损失,更会通过误差传导放大,引致系统性的决策风险。

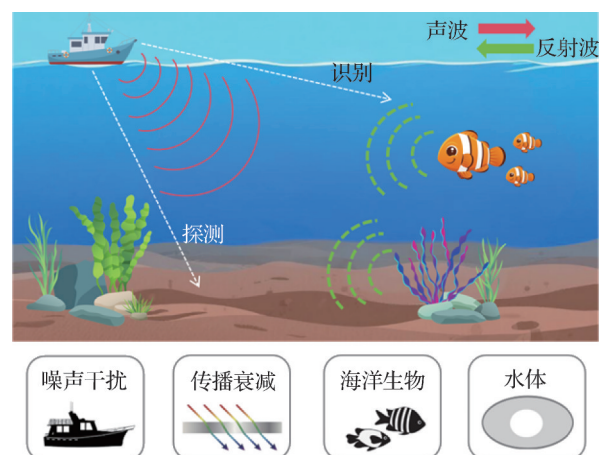


图2 声呐探测及退化因素示意图

Fig. 2 Schematic diagram of sonar detection and degradation factors

## 2 声呐图像质量评价

图像质量退化会造成信息损失和决策误差,因此需要科学可靠的质量评价体系提供量化支撑。水下探测领域对此需求更甚:声呐图像作为获取海洋数据的主要方式,其质量退化将直接影响目标识别、地形测绘和水下导航的精度,导致资源误判、导航偏差等风险。为此,建立专门的声呐图像质量评价框架是提升数据可信度、保障决策准确性的关键基础。图像质量评价根据主体和方法,主要分为主观质量评价与客观质量评价两类。

表2 图像特性差异

Table 2 Image characteristic differences

特性	声呐图像	光学图像
噪声类型	相干斑噪声、混响噪声	高斯噪声、椒盐噪声
分辨率	厘米—米级	毫米—微米级
边缘信息	模糊拖尾、几何畸变	锐利清晰、几何保真
对比度	整体偏低:远距离信号衰减、目标与背景灰度差异小	近距离时清洁水体中细节分明,深度大于10 m时骤降,红光吸收导致色彩失真
适用场景	深水、浑浊水域、大范围测绘	浅水、高细节识别

## 2.1 主观质量评价

主观质量评价是一种通过人类观察者的视觉感知和专业知识,对图像的可用性进行定性或半定量分析的方法体系。其主要依赖人类对图像内容的理解能力,然后结合特定应用场景的需求,以此评估图像在目标检测、目标分割和目标追踪等方面的表现。主观评价可以直接反映人类视觉系统(human visual system, HVS)对图像质量的综合感知,尤其在海洋这种复杂场景中,该评价方式也能够有效地捕捉到难以量化的部分。这使得主观质量评价成为声呐图像质量评估中不可替代的环节,尤其在一些对细节敏感度要求极高的领域中,更是不可或缺的一部分。

主观质量评价通常采用国际电信联盟(International Telecommunication Union, ITU)制定的5级评价尺度标准,该标准主要包含两种尺度:妨碍尺度与质量尺度。其中,妨碍尺度适用于专业领域人员,质量尺度适用于非专业人员。需要注意的是,与常见光学自然图像不同,声呐图像具有显著的任务导向性,面向应用场景的图像质量评估结果与脱离应用背景的评估结果往往不同。因此,声呐图像质量评价尺度标准往往需要根据任务及应用场景做出相应调整。Chen等人(2017b)根据ITU-R BT.500-13提出了针对声呐图像设计的5级评分妨碍尺度,如表3所示。

为确保主观评价结果的可靠性,实验前的准备工作必须严格遵循国际标准,具体要求如下:首先,组建包含领域专家和经过培训的非专业观察人员在内的团队,以保证评价视角的全面性;其次,实验环境需要严格依据国际标准化组织(International Organization for Standardization, ISO)的国际标准进行实验室环境的光学校准与把控,最大程度消除外部环

表3 5级评分尺度

Table 3 Five-level scoring scale

评分	质量尺度	妨碍尺度
5	优	目标清晰,信息保留完整
4	良	目标可见可辨认,有少量可感知失真
3	中	目标部分受损,失真可被明显感知
2	差	可辨别目标的轮廓,失真较严重
1	劣	无法辨认图像的内容,失真严重

境干扰,确保有效的观察条件。

质量评价流程通常分为训练与实测两个阶段:在训练阶段,需要通过培训使观察者掌握评分标准(表3),并借助涵盖不同质量等级的典型样本图像,使观察者熟悉评分尺度和失真类型;在实测阶段,需根据实验目标与资源条件,从单刺激法、双刺激法和成对比较法中选择一种方法进行图像质量评分。每幅声呐图像的呈现时间控制在10 s左右,然后观察者依据表3的5级尺度进行评分。最终通过计算所有观察者评分的算术平均值,得到每幅图像的平均意见得分(mean opinion score, MOS),表4详细介绍了3种方法的具体操作方法和评分依据。

主观质量评价方法能够捕捉人类视觉系统感知

表4 单刺激法、双刺激法和成对比较法介绍

Table 4 Introduction to single-stimulus method, double-stimulus method and pairwise comparison method

方法	具体操作	评分依据
单刺激法	仅呈现待测图像	直接确定被评价图像的质量水平和分数
双刺激法	呈现参考图像和待测图像	根据参考图像的质量水平和分数来评估
成对比较法	呈现两幅待测图像比较	判断两图像中哪幅质量更好

到的高阶信息,这些信息往往是算法难以建模的,这使得它在理解图像感知质量方面具有不可替代的优势。然而,主观评价方法也存在显著的局限性:1)实施成本高昂,单次实验需投入设备校准、人员培训及数据采集等大量资源;2)过度依赖专家经验,导致评价结果难以复现;3)对动态场景适应性不足。因此,亟需发展有效、可靠的客观质量评价方法作为必要补充,与主观质量评价共同构建更完善的评价体系。

## 2.2 客观质量评价

声呐图像客观质量评价通过构建人类视觉感知模型,建立可计算的数学度量方法,在不需要人工干预的条件下量化图像质量表现。该方法为系统参数优化、图像处理算法及海洋探测任务提供了科学的依据。另外根据参考图像的可用性,客观质量评价可以分为三大类:全参考(full-reference, FR)评价、半参考(reduced-reference, RR)评价以及无参考(no-reference, NR)评价,如图3所示。

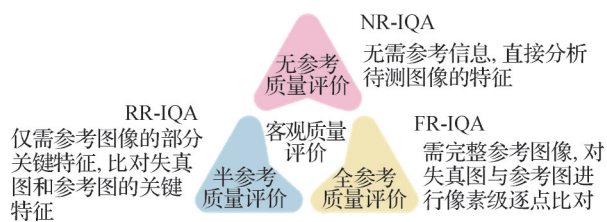


图3 基于参考图像可用性的声呐IQA分类

Fig. 3 The classification of sonar IQA based on the availability

### 2.2.1 全参考质量评价

全参考质量评价(full reference image quality assessment, FR-IQA)要求以完整的原始图像作为参考基准,通过比对退化图像与参考图像之间的差异来量化图像质量,该方法在参考图像易于获取的应用场景中具有显著优势,例如:海底特定目标识别、海底基建缺陷检测等。声呐成像通常能提供原始参考图像,故其FR-IQA研究较为成熟。

全参考质量评价算法还可以根据参考信息的利用方式进一步细分为两类:一类是基于失真度的方法,其依据HVS特性,通过比对原始图像定义待测图像的失真程度;另一类是基于相似度的方法,通过提取原始图像的结构特征,量化待测图像与参考特征的一致性。Chen等人(2017b)融合失真度与相似度双重视角,提出了一个基于局部熵的声呐图像质量预测器(local entropy backed sonar image quality

predictor, LESQP)。该模型通过局部熵映射量化失真导致的区域信息混乱度变化,其数学模型定义用表达式描述为

$$f_{LESQP} = \sum \sum S(x, y) \cdot I(x, y) \quad (1)$$

式中, $S(x, y)$ 表示局部熵映射相似度函数, $I(x, y)$ 表示归一化权重图。LESQP模型通过局部熵有效捕捉了失真导致的信息损失,适用于评估因压缩、传输等过程引起的全局信息退化。然而,该模型对局部结构失真敏感度不足,且高度依赖参考图像的质量。

为了克服LESQP模型对局部结构失真敏感度不足的局限性,Chen等人(2020)进一步提出了融合统计与结构信息的声呐图像质量预测器(sonar image quality predictor, SIQP)。该模型基于HVS特性,联合提取局部熵统计特征与边缘结构特征,通过二次多项式融合实现精准质量评估。其计算表达式为

$$f_{SIQP} = \sum_{i=1}^2 (r_{1i}s^i + r_{2i}e^i) + r_3se \quad (2)$$

式中, $r_{1i}, r_{2i} (i = 1, 2)$ 和 $r_3$ 均为二次多项式模型的参数, $s$ 表示统计信息特征, $e$ 表示结构信息特征。SIQP模型综合了统计与结构信息,在声呐图像压缩质量评估中表现优异,适用于对目标轮廓清晰度要求高的场景,如海底目标识别。其局限性在于计算复杂度较高,且同样严格依赖完整的参考图像,无法应用于参考图像缺失的场合。

FR-IQA方法虽然在参考图像可得场景中表现优异,但其性能高度依赖于参考图像的质量与完整性,在实际应用中常因参考图像难以获取或存在偏差而受限。

### 2.2.2 半参考质量评价

在复杂的海洋环境中,声呐图像的视觉质量受采集与传输限制影响显著,导致参考图像的获取十分困难。半参考质量评价(reduced reference image quality assessment, RR-IQA)作为全参考与无参考方法的平衡方案,其核心价值在于:通过比对失真前后的关键图像特征,实现客观质量量化。该方法专用于可保障少量参考信息无损传输的水下应用场景。

为应对参考图像缺失场景的质量评估挑战,Chen等人(2018a)提出了面向水下传输的部分参考声呐图像质量评测算法(partial-reference sonar image quality predictor, PSIQP)。该算法采用三阶段

渐进式框架:第1阶段为感知舒适度评估,通过信息熵提取计算感知舒适度指标;第2阶段为结构保真度验证,通过结构相似性(structural similarity, SSIM)度量原始图像与失真图像间的结构相似性;第3阶段为多特征融合预测,集成前两阶段的舒适度指标与结构相似性参数,构建最终质量预测模型。PSIQP用数学表达式描述为

$$f_{\text{PSIQP}} = \sum_{i=1}^4 p_i \cdot C_i \quad (3)$$

式中, $p_i, i \in (1, 4)$ 表示为信息指数、舒适度指数和结构相似度的权重, $C_i, i \in (1, 4)$ 表示信息与舒适度指数。PSIQP是带宽受限水下传输场景的理想选择,能够在参考信息不完整的条件下实现有效评估。但其性能高度依赖于所选取的少量特征,对于特征未能涵盖的复杂失真类型,其泛化能力可能受限。

此外,鉴于声呐图像质量评价需紧密结合特定应用场景,有研究者提出了任务驱动型的评价模型:Zheng等人(2020)提出声呐图像效用质量评测算法(sonar image utility quality assessment, SIUQA),融合感知特征与任务需求特征,提升质量指标与水下任务相关性,适合用于低带宽水下通信中的声呐图像质量评估,其局限性在于对全局结构变化敏感;Chen等人(2021b)基于声呐图像统计特性与任务背景,开发了面向任务和感知的声呐图像质量评测算法(task and perception oriented sonar image quality assessment, TPSIQA),针对声呐图像库小样本特性引入基学习器选择性集成机制,在保障精度同时有效抑制过拟合风险。值得注意的是,该模型凭借其具有任务和感知融合及选择性集成学习的特性,特别适合用于小样本声呐图像数据库和恶劣信道条件下的质量评估。然而,其计算复杂度高,且子带数量需要手动调优。

RR-IQA方法在参考信息有限的情况下实现了对声呐图像的质量评价,但当前方法多依赖于手工设计的特征,对声呐图像在不同传输、压缩和增强过程中的失真模式变化适应性不强,尤其在跨场景评估时易出现性能下降。

### 2.2.3 无参考质量评价

无参考质量评价(no reference image quality assessment, NR-IQA)与全参考方法不同,该方法能直接分析失真图像本身的质量特性,无需依赖任何原始参考信息。

Chen等人(2018b)提出了一种基于退化测量的无参考声呐图像质量评估方法(no-reference sonar image quality metric, NSIQM)。该方法通过量化测试图像与模糊退化版本间的轮廓退化程度评估质量,避免了对原始参考图像的需求。具体流程为:首先,从测试图像中提取轮廓稀疏特征,然后计算测试图像与模糊退化图像的特征比值,最后利用支持向量回归(support vector regression, SVR)模型将特征比值映射为质量分数。

为了在无参考条件下更精准地评估由失真导致的轮廓退化,Chen等人(2019)进一步提出了基于轮廓退化测量的声呐图像无参考质量评估模型(no-reference contour degradation measurement, NRCDM)。该方法通过分析失真导致的轮廓退化程度,并结合集成学习机制,在无需参考图像的前提下实现高精度质量评估,模型基本架构如图4所示。该模型的数学表达式为

$$f_{\text{NRCDM}} = \frac{1}{n} \cdot \sum_{i=1}^n S_i(t) \quad (4)$$

式中, $S_1(\cdot), S_2(\cdot), \dots, S_n(\cdot)$ 为SVR模型,其中, $S_n(\cdot)$ 是通过 $n$ 个自助样本拟合,并对输出结果的平均值进行整合。NRCDM模型具有基于轮廓退化测量和无需参考图像的特性,适合用于声呐图像在传输或压缩后的无参考质量评估。然而,该方法主要依赖宏观轮廓信息,可能忽略纹理细节等微观结构对质量的影响,且对模糊类失真的评估性能相对较弱,在失真类型复杂或轮廓不显著的场景下可能存在局限性。

与此同时,Zhang等人(2022)创新性地提出一种基于双路径深度神经网络(dual-path deep neural network, DPDNN)的无参考评估方法。该方法通过并行提取声呐图像的局部纹理和全局轮廓特征,并

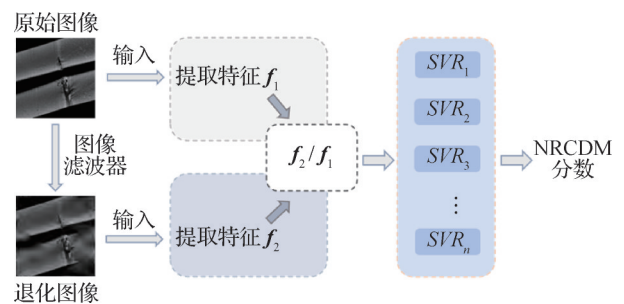


图4 NRCDM模型(Chen等人,2019)

Fig. 4 NRCDM model (Chen et al., 2019)

融合这些特征直接预测图像质量分数。沿着微观与宏观特征融合的思路, Tolie 等人(2025)进一步提出一种基于小波域微观纹理与宏观轮廓特征融合的无参考评估方法。该方法通过量化声呐图像在小波域的局部结构和整体轮廓特征退化程度评估其感知质量和实用质量。

由于声呐图像在采集、压缩、传输及增强过程中易受多重失真影响, 而且现有质量评估方法还普遍存在两大局限: 其一, 通用自然图像质量评价方法未能充分考虑声呐图像的灰度分布特性、任务导向性及独特失真模式; 其二, 现有声呐图像质量评价方法通常局限于单一场景, 缺乏跨场景泛化能力。针对此类现象, Cai 等人(2025a)提出属性一致性这一理论基础, 即高质量声呐图像无论场景如何变化, 都必须满足4个核心任务导向属性: 区域区分度、几何完整性、细节保留和洁净度。基于上述理论背景, Cai 等人(2025a)还提出了一种基于NR-IQA的声呐成像与处理统一质量评估方法(unified quality assessment method for sonar imaging and processing, UASIP), 旨在实现跨场景普适性的图像质量评估。需要注意的是, 该模型具有基于属性一致性原则和统一评估框架的特性, 适用于声呐图像在采集、传输和增强等多种处理场景下的无参考质量评估。然而, 该方法在特征设计上仍依赖人工设计的属性, 可能无法完全适应未知或极端复杂的失真类型, 另外其还涉及多属性提取与SVR回归, 在计算效率上相较于轻量级方法可能存在一定的局限。

为解决声呐图像因长距离成像导致的分辨率下降问题, SR算法通过从低分辨率图像重建高频细节, 生成细节增强的高分辨率图像。为了量化重建伪影并评估超分辨率声呐图像的质量, Chen 等人(2024)提出感知—认知协同驱动的质量评价方法(perception-cognition aware sonar assessment, PCASS), 该方法结合下游任务需求与SR重建失真特性, 通过分层特征融合实现无训练的质量预测。其计算模型的数学表达式为

$$f_{\text{PCASS}} = \sum_{i=1}^n \frac{(f_{1i} + f_{2i}) \times f_{3i}}{f_{1i}^2 + f_{2i}^2 + f_{3i}^2 + K} \quad (5)$$

式中,  $f_1$  表示低级特征,  $f_2$  表示中级特征,  $f_3$  表示高级特征,  $n$  表示尺度参数, 常数  $K$  设为 0.01, 以避免分母为零。PCASS 模型具有无参考、多层级特征融合和任务效用导向的特性, 是专门用于声呐图像 SR 重

建后的质量评估。然而, 其计算复杂度较高, 对图像轮廓结构的完整性较为依赖, 且适用场景单一, 仅针对 SR 重建图像, 不适用于其他类型的声呐图像处理。

此外, Feng 等人(2024)提出一种基于迁移学习的脑启发声呐超分辨率图像质量评估模型(super-resolution reconstructed image quality assessment, SRIQA), 通过模拟人脑视觉腹侧通路的层次化特征处理机制, 解决声呐超分辨率图像的复杂失真与任务效用评估难题。其计算模型的数学表达式为

$$f_{\text{SRIQA}} = FC(\mathbf{F}_{\text{fused}}) \quad (6)$$

$$\mathbf{F}_{\text{fused}} = \text{Concat}(\mathbf{F}^1 + \mathbf{F}^2, \mathbf{F}^3, \mathbf{F}^4, \mathbf{F}^5) \quad (7)$$

式中,  $\mathbf{F}^1, \mathbf{F}^2, \mathbf{F}^3, \mathbf{F}^4, \mathbf{F}^5$  分别表示声呐图像的特征向量,  $\mathbf{F}_{\text{fused}} = \text{Concat}(\cdot)$  表示 5 个特征的特征拼接,  $FC(\cdot)$  用于将多级特征映射到质量评分。SRIQA 模型具有专注于图像细节恢复和超分辨率引入失真的特性, 同样专门用于评估 SR 重建图像的质量。然而, 该模型未考虑任务背景, 没有针对特定应用领域的任务需求进行优化, 导致其在任务导向的质量评估中存在局限性, 且模型侧重于低层或中层特征, 而忽略高层语义信息。

需要注意的是, NR-IQA 方法虽然具备更强的实用性, 但其模型泛化能力仍是主要挑战。现有方法多基于特定数据集或失真类型训练, 对未见过的声呐场景或复合失真类型的评估能力较为有限。

## 3 声呐图像质量增强

### 3.1 基于图像处理的增强方法

#### 3.1.1 图像超分辨率重建

图像 SR 重建本质上是图像退化的逆过程, 其核心在于利用低分辨率图像本身所蕴含的结构与纹理等信息, 重建出对应的高分辨率图像(罗逸豪等, 2023)。该过程用数学表达式描述为

$$\mathbf{P}_{\text{HR}} = R(\mathbf{P}_{\text{LR}}, \theta) \quad (8)$$

式中,  $\mathbf{P}_{\text{HR}}$  表示高分辨率图像;  $\mathbf{P}_{\text{LR}}$  表示低分辨率图像;  $R$  和  $\theta$  分别表示 SR 重建模型及其参数。

声呐图像 SR 重建主要是针对低分辨率声呐图像进行算法处理, 以恢复其丢失的高频细节并提升空间分辨率。但由于受到水下声音信号衰减、窄带宽限制以及硬件采集过程中下采样等因素的影响, 传统的声呐系统获取的图像普遍存在三大缺陷: 其

一,空间分辨率显著低于光学成像系统;其二,细节呈现模糊化,关键结构易发生几何畸变;其三,重要目标特征常以低信噪比的块状伪影形式存在。这些缺陷不仅导致图像视觉信息严重损失,更会直接影响后续目标检测、分割等任务。因此,声呐图像SR重建技术应运而生,该技术融合图像处理与机器学习等方法,通过对低分辨率输入图像进行处理,克服上述物理和设备限制,最终生成细节更丰富、结构更清晰的高分辨率声呐图像。

本文系统梳理了声呐图像SR重建技术的发展历程,如图5所示。从回顾传统方法切入,进而聚焦当前主流的深度学习模型。传统的SR重建技术主要依赖插值优化、频域增强、多帧融合以及稀疏表示等方法。然而,这些方法受限于手工特征设计,重建

效果往往有限。得益于深度模型的发展,基于深度学习的SR重建技术取得了显著突破,当前主流架构以卷积神经网络(convolutional neural network, CNN)和Transformer为代表。为了突破传统手工特征方法重建效果的上限,Sung等人(2018b)提出了一种基于生成对抗网络(generative adversarial network, GAN)的框架,采用多分支生成器结合判别器结构,成功实现了4倍SR重建,该模型在视觉效果与峰值信噪比(peak signal-to-noise ratio, PSNR)指标上均超越了传统方法。同样为了提升性能但侧重于更稳健的架构,Sung等人(2018a)则利用卷积层与ResNet块构建的CNN模型,显著提升了重建性能,其效果优于传统插值法,PSNR指标更高且能有效保留声呐工作距离特性。

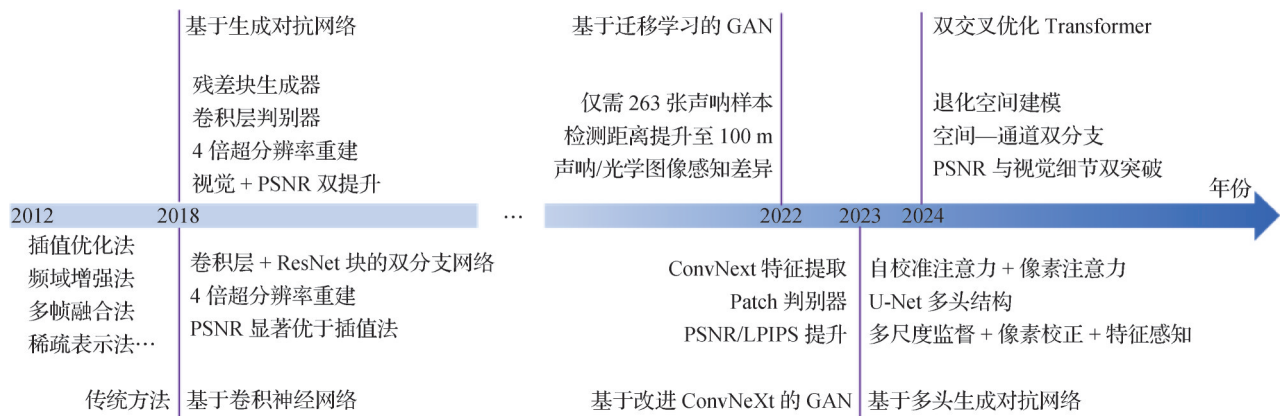


图5 声呐图像SR重建技术发展进程

Fig. 5 The development process of sonar image super-resolution reconstruction technology

近年来,该领域持续创新。针对声呐高质量数据稀缺难以训练深度模型挑战,Thomas等人(2022)提出的基于迁移学习的GAN模型,仅需263张声呐数据即可实现高质量重建,显著提升了关键参数覆盖率,并有效解决了声呐与光学图像间的感知差异问题。为了融合现代卷积架构的强大特征提取能力与GAN的细节生成优势,Ma等人(2023a)设计了一种融合改进ConvNeXt模块的GAN模型,结合ConvNeXt强大的特征提取能力与GAN的细节生成优势,大幅提升了声呐图像质量。为了解决实际应用中退化过程未知的“盲”超分问题,Rao等人(2024)针对盲声呐图像超分辨率问题,提出一种新型双交叉优化Transformer深度学习模型。同时,为了在数据匮乏和复杂背景的双重挑战下更好地恢复高频纹理,Ma等人(2023b)设计并构建了一种多头

生成对抗网络模型(multi-head generative adversarial network, MHGAN),该模型在缺乏大规模数据集且面临复杂水下背景纹理的挑战下,仍能有效恢复声呐图像的高频细节与真实纹理。

通过回顾近年声呐图像SR技术,发现现有方法存在以下局限:CNN因局部感受野的限制易导致全局结构失真;Transformer虽然有长程依赖优势,却常忽略高频纹理细节;传统平均绝对误差L1、均分误差L2损失易使重建结果过平滑,细节丢失严重;串行融合CNN与Transformer时,结构-纹理特征的相关性未得到充分挖掘。为了同时解决CNN因局部感受野限制导致的全局结构失真,以及Transformer易忽略高频纹理细节的问题,Wang等人(2025)提出了一种名为STDPNet(structure texture dual preserving network)的声呐图像SR方法。该方法基于结构

与纹理双重保持机制,通过并行整合 CNN 与 Transformer 模型,全面捕捉全局结构信息与局部纹理细节,显著提升重建质量。其核心工作原理用数学表达式描述为

$$P_{SR} = S \langle M_C \{P_{LR}, \theta_C\}, M_T \{P_{LR}, \theta_T\} \rangle : (F_h, F_l) \Rightarrow F_r \quad (9)$$

式中,  $P_{SR}$  表示 SR 图像,  $S \langle \cdot \rangle$  表示局部—全局特征交互策略,而  $\Rightarrow$  则是  $S \langle \cdot \rangle$  执行的细化操作,  $F_r$  代表细化后的特征表示,  $M_C$  和  $M_T$  分别表示基于 CNN 与 Transformer 的 SR 模型,  $\theta$  表示各自模型的参数,  $F_h$  和  $F_l$  分别表示高频和低频特征。

同时,损失函数的选择对重建质量也至关重要。当前主流方法多采用平均绝对误差(mean absolute error, MAE)损失,该损失能有效地减小高分辨率图像与低分辨率图像之间的平均像素差异。然而,仅依赖 MAE 损失会导致输出结果过于平滑,难以保持目标结构完整性。因此,Wang 等人(2025)引入了谱损失和局部梯度感知损失。前者指导全局结构重建;后者通过强调边缘与梯度特征有效捕捉局部形状。其损失函数定义用数学表达式描述为

$$L_{total} = L_{mae} + \alpha L_{spe} + \beta L_{ig} \quad (10)$$

式中,  $L_{mae}$ 、 $L_{spe}$  和  $L_{ig}$  分别代表 MAE 损失、谱损失和局部梯度感知损失,  $\alpha$  和  $\beta$  为权重参数。

Wang 等人(2025)提出的新型 SISR (sonar image super resolution) 框架,不仅性能优越且模型复杂度低,使重建图像更适配机器视觉任务。其设计的混合损失函数针对 SISR 模型优化具有高度适配性,显著提升了图像结构与纹理细节的重建质量。通过双流架构与混合损失的协同优化机制,该框架有效兼顾了结构—纹理特征,为声呐图像 SR 重建领域树立了新标杆。

除上述针对声呐特性设计的专用模型外,通用图像 SR 算法也广泛应用于声呐图像增强研究。代表性算法如 LBN (lightweight bimodal network) (Gao 等, 2022), 以及基于 Transformer 的 SRFormer (super-resolution former) (Zhou 等, 2023)、NGswin (N-Gram swin Transformer) (Choi 等, 2023)、ESRT (efficient super-resolution Transformer) 等,它们通过在声呐数据集上进行训练和验证,展现了对声呐图像 SR 重建任务的适应性。这些通用算法虽然非专为声呐设计,但其性能表现为评估声呐图像 SR 任务提供了重

要参考,并揭示了通用模型在处理声呐特定纹理和结构时的潜力与局限。

尽管 SR 重建技术显著提升了声呐图像的空间细节,但其重建质量仍受限于训练数据,尤其在数据稀缺的水下场景中容易出现过拟合或纹理失真的情况。此外,现有的模型多侧重于像素级保真度优化,对声呐图像中目标结构与边缘几何一致性的保持能力仍有待加强。

### 3.1.2 图像去噪

声呐图像的去噪技术旨在抑制声学成像中的噪声干扰,其核心是从受干扰的声呐信号中分离出目标的有效反射信息,并增强图像的视觉效果。由于声呐噪声主要是由多径效应引发的混响噪声、设备电子噪声以及水体中悬浮物散射噪声通过非线性方式耦合形成的。这类噪声不仅会掩盖沉船、管道等目标的边缘细节,还会影响后续的处理。去噪技术对声呐图像的重要性主要源于海洋探测任务对数据质量的严苛要求,有研究表明,噪声干扰导致的数据误判可能造成严重后果。因此,去噪是声呐数据处理中不可或缺的关键环节。鉴于其重要的研究价值与应用需求,本节将系统梳理并深入分析声呐图像去噪技术的发展进程及最新研究进展。

声呐图像去噪技术的发展经历了从基础方法到复杂模型、从单一处理到精细化优化的演进过程,如图 6 所示。早期的去噪技术主要依赖直接引导滤波、均值滤波以及中值滤波等传统方法,这些方法处理能力有限,尤其在保留边缘细节和抑制复杂噪声方面效果不佳。随后进入自适应建模阶段,为了克服早期基础方法在处理复杂的信号依赖性噪声和保留边缘细节方面的不足,Chen 等人(2017a)通过广义噪声建模和变分优化框架实现了对信号依赖性噪声的自适应去除,显著优于前馈模型;同时,为了更有效地解耦并处理由多源干扰形成的混合噪声,Wu 等人(2018)提出了一种基于分步稀疏表示的框架,创新性地利用对数变换解离线性噪声,有效处理混合噪声,同时提升了图像能量谱和结构信息的保留能力。

随着人工智能 (artificial intelligence, AI) 技术的迅猛发展,声呐图像去噪技术的研究重心正迅速转向深度学习与算法创新,并取得了显著成果。旨在结合物理噪声模型与数据驱动方法以同时提升去噪效果和边缘保持能力,Chen 等人(2021a)将实测固



图 6 声呐图像去噪技术发展进程

Fig. 6 The development process of sonar image denoising technology

定水域噪声建模与引导滤波相结合,在提升去噪效果和解决传统方法边缘模糊问题上取得突破,其中PSNR和结构相似性(SSIM)指标显著领先。针对传统去噪算法在计算效率、噪声去除和细节保留之间难以取得平衡点,Shi等人(2023)提出了一种优化空间像素排序算法(optimized spatial pixel ranking, OSPR),通过自适应筛选机制,在噪声去除、边缘保留和计算效率上全面超越传统方法。为了更有针对性地抑制特定噪声并充分利用深度网络的层次化特征,Liu等人(2024)提出了一种结合特征融合的双阶段U-Net去噪算法,该算法能高效抑制光度噪声,显著保留目标的细节纹理,并在评估指标上表现优异。与此同时,混合优化的去噪技术也在迅速发展。

为了解决传统滤波方法在边缘保持与高强度噪声抑制之间的平衡难题,Wang等人(2022a)提出了一种名为各向异性引导滤波(anisotropic guided filtering, AnisoGF)的方法。该方法通过引入各向异性加权与自适应正则化机制改进了传统滤波技术,在边缘保持、高强度噪声抑制以及目标轮廓保留方面表现突出。实验结果表明,该方法实现了高达6.3 dB的PSNR提升,并在SSIM上取得了全面领先优势。各向异性引导滤波去噪流程如图7所示。

另外,通用图像去噪算法也可用于抑制声呐图像中的复杂噪声。如Restormer(Zamir等,2022)、Uformer(Wang等,2022b)、APD-Net(adaptive prior denoising network)(Jiang等,2022)、MSANet(multi-scale adaptive network)(Gou等,2022)等先进模型,这些模型强大的特征提取能力在处理声呐斑点噪声和混响干扰方面展现出一定效果。评估这些通用去

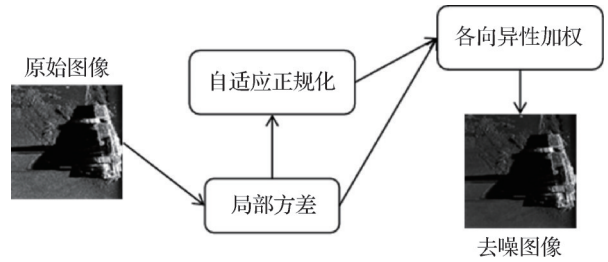


图 7 各向异性引导滤波去噪实例(Wang等,2022a)

Fig. 7 Illustration of the anisotropic guided filtering denoising (Wang et al., 2022)

噪算法的性能,有助于理解现有去噪技术在声呐领域的适用性边界,并为开发更契合声呐噪声特性的专用模型提供对比基准。

总体而言,当前去噪方法在抑制斑点噪声与混响干扰方面已取得显著进展,但仍普遍面临去噪强度与细节保留之间的权衡难题。深度学习模型虽然具有较强的噪声建模能力,但其对声呐图像中目标边缘、弱反射区域等关键信息的保护仍不足,且模型复杂度高,不利于实时或边缘计算场景部署。

### 3.1.3 图像恢复

声呐图像恢复技术是提升声呐成像系统在海洋探测、水下工程等领域获取的退化图像质量的关键技术。由于水下环境的复杂性,声波在传播过程中会受到介质吸收、多径反射、散射效应以及设备噪声的干扰,导致声呐图像常出现模糊、混响和对比度下降等问题。这种图像退化不仅影响后续的下游任务,还会掩盖微小目标,影响水下作业的安全性和可靠性。

近年来,随着计算机算力的提升与AI技术的发展,图像恢复技术也逐渐从传统信号处理方法向数

据驱动与物理模型融合的方向发展。在去模糊领域,为了解决传统迭代优化方法在校正声呐图像相位误差时存在的计算复杂、依赖人工调参且难以满足实时处理需求的问题,Gerg等人(2024)提出了一种名为深度自适应相位学习(deep adaptive phase learning, DAPL)的特殊去模糊技术,用于自动校正合成孔径声呐(synthetic aperture sonar, SAS)图像中的系统性相位误差,从而显著提升图像质量,DAPL通过数据驱动的深度网络替代传统迭代优化,实现了无需调参、快速鲁棒的SAS自聚焦,为水下自主平台的实时高清成像提供了新范式。

此外,深度学习还为去模糊技术提供了新的解决思路:CNN通过端到端的方式直接学习了退化图像与清晰图像的复杂非线性映射关系,然后利用残差学习、注意力机制等策略增强特征表达能力;GAN则通过对判别器与生成器的优化,生成感知质量更高的恢复结果,在处理运动模糊与离焦模糊交织的复杂退化场景中展现出明显优势。

去混响技术同样具有交叉学科特征。混响源于声波在水体界面、悬浮颗粒与目标表面间的多次反射,其干扰信号在时域和频域上与目标回波高度重叠。因此,传统方法通常通过自适应滤波、子带分解或统计建模来分离混响成分。然而,近年来深度学习的发展有力推动了去混响技术的革新:循环神经网络(recurrent neural network, RNN)与长短期记忆网络(long short-term memory, LSTM)通过时序相关性建模混响的动态衰减过程。

图像恢复技术虽然能有效校正特定类型的成像失真,但其多依赖于准确的退化模型或大量成对数据,在实际复杂水下环境中的适用性受限。此外,多数方法仍处于单任务优化阶段,缺乏对多类退化因素的联合建模能力,难以应对实际声呐图像中的复合退化问题。值得关注的是,当前技术发展正逐步突破单一任务的局限,向融合去模糊、去混响与去噪的协同恢复框架演进。这种多任务联合优化范式能够更准确地建模水下复杂声学环境对成像的耦合干扰,从而提升图像整体质量与信息保真度,为高精度水下探测与识别奠定更坚实的基础。

### 3.2 基于数据融合的增强方法

声呐图像多模态数据融合增强是一种通过整合声呐与其他传感模态(如光学图像、激光雷达、磁力

仪、惯性导航等)的异构数据,构建互补特征表达以提升图像质量与目标辨识能力的技术。其核心在于克服单一传感器受限于水下环境物理特性(如光学浑浊度、声波散射、磁场干扰)的固有局限,通过时空对齐、特征级或决策级融合,在多模态统一框架下实现信息转化与利用。例如,在浑浊水域中,光学相机因光线衰减难以获取目标细节,而侧扫声呐虽然能穿透悬浮物却易受混响干扰;融合声呐的穿透能力与光学图像的高分辨率纹理,通过跨模态特征互补可有效抑制噪声并修复细节。

得益于深度学习的突破,基于深度学习的多模态融合架构正推动该技术向端到端方向发展。第一种典型方案是多分支网络。例如,Li等人(2024)提出了一种名为RepDNet(despeckle denoising network)的新型轻量化卷积神经网络,首次将重参数化技术应用于侧扫声呐(side-scan sonar, SSS)图像去斑,并通过多分支网络结构融合,在保持边缘细节的同时高效去除斑点噪声。另一种方法则基于Transformer架构建模。Lei等人(2025)提出了一种名为FLSSNet(forward-looking sonar segmentation network)的混合架构,结合了CNN和Transformer骨干网络,并利用一个Transformer特征转换器模块进行特征转换与通道压缩,以此优化多模态融合。实验表明,该混合架构在25种先进显著性目标检测(salient object detection, SOD)方法中性能优异。

多模态融合方法虽然能有效互补各传感器优势,但其性能高度依赖于模态间数据的精确配准与对齐,在实际应用中常因时空同步误差而受限。此外,当前融合模型多基于监督学习,对跨域、跨设备数据的泛化能力较弱,且模型复杂度高,难以满足水下实时处理需求。因此,多模态数据融合仍面临诸多挑战,包括多源异构数据的标定难题、实时处理的算力约束及跨域泛化能力不足等。

## 4 实验结果与分析

为保障声呐图像质量评估指标对比的公平性与结果可现性,在公开数据集上对所有开源方法进行统一评估实验。在统一硬件、操作系统及软件环境下,采用相同数据划分,调用各算法的官方源码执行测试,所得结果即为该算法的基准成绩。鉴于算法实现及评估流程中固有的随机性,本次实验提供了

基准对照。对于无法获取公开源码的算法,本次对比实验未予纳入;其余 IQA 算法均直接调用其源代码完成评估。表 5 展示了实验数据集划分与配置总览。

表 5 实验数据集划分与配置总览

Table 5 Overview of the division and configuration of experimental datasets

实验类型	所用数据集	使用说明	硬件与软件配置
IQA 性能对比 (表 6—表 8)	SIQD, SRSID, MSSID	使用全部数据进行性能评估,无需训练;不划分训练/测试集。	软件平台:MATLAB R2024b 硬件:NVIDIA GeForce RTX 4060 GPU
通用 SR 算法验证 (表 9)	USOD	按 8:2 比例随机划分为训练集与测试集;实验结果取 5 次运行平均值。	深度学习框架:Python 3.7, PyTorch 硬件:NVIDIA GeForce RTX 4060 GPU
通用去噪算法验证 (表 10)	SIQD	按 8:2 比例随机划分为训练集与测试集;实验结果取 5 次运行平均值。	深度学习框架:Python 3.7, PyTorch 硬件:NVIDIA GeForce RTX 4060 GPU

#### 4.1 IQA 算法性能对比实验

##### 4.1.1 实验数据集

质量评价实验统一采用 3 种公开声呐图像数据集:SIQD (sonar image quality database) (Chen 等, 2017b)、SRSID (super-resolution sonar image database) (Zheng 等, 2022) 与 MSSID (multi-scenario sonar image dataset) (Cai 等, 2025a), 以全面评估算法在不同失真类型下的性能。具体信息如下:

1) SIQD 数据集共 840 幅声呐图像:参考图像 40 幅,由侧扫声呐和声学透镜声呐采集;失真图像 800 幅,内容涵盖游泳者、海洋生物、沉船及海底地形。失真类型分为 4 类,每类有 5 级强度。图像分辨率统一为  $320 \times 320$  像素,主观评分以 MOS 分数形式给出。

2) SRSID 数据集共 1 083 幅影像,包含 57 幅高

分辨参考图像及 1 026 幅失真图像。场景囊括飞机残骸、沉船、海底地形与散落碎片等典型水下目标。需要注意的是,图像分辨率并未统一,同时存在  $320 \times 320$  像素和  $480 \times 320$  像素两种规格的分辨率。

3) MSSID 数据集聚焦采集、传输和增强 3 大场景,共 2 131 幅影像,其中 61 幅参考图像通过高性能的前视声呐与侧扫声呐采集获得。数据集按场景划分,每个场景设置 4~6 级失真,其中采集场景因声呐回波斑点噪声显著而失真最重。另外, MSSID 与 SRSID 类似,同样存在  $320 \times 320$  像素和  $480 \times 320$  像素两种不同规格的分辨率。

图 8 展示了 SIQD、MSSID 和 SRSID 3 个数据集的样本图像及其对应的 MOS 分数。图 8(a)—(d) 来自 SRSID 数据集,图 8(f)—(i) 来自 MSSID 数据集,图 8(e) 和图 8(j) 则来自 SIQD 数据集。

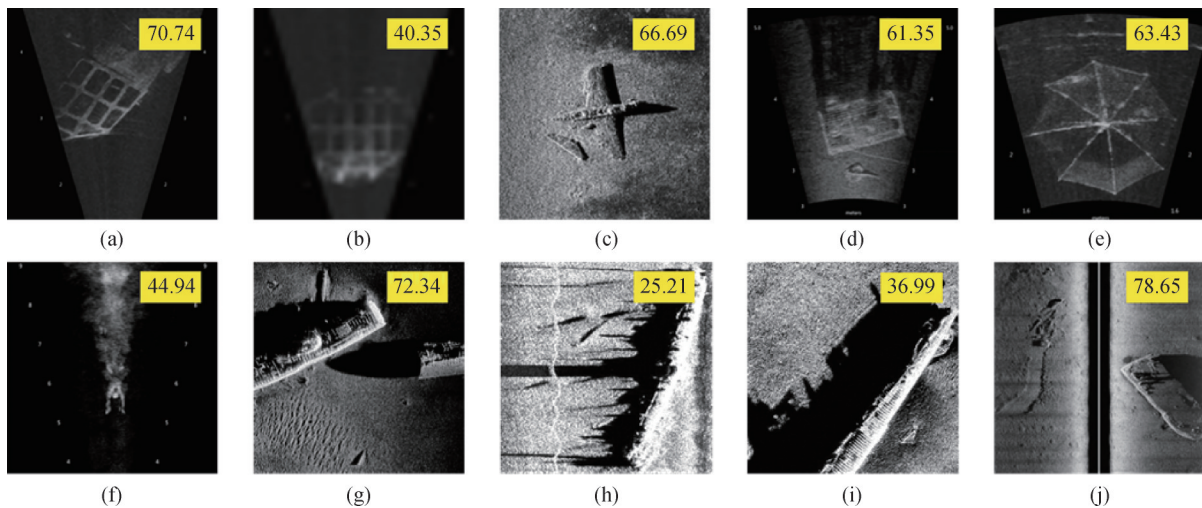


图 8 3 种数据集的示例图像和 MOS 分数

Fig. 8 Sample images and MOS scores of three datasets

## 4.1.2 评价指标与实施细节

实验对比了5种声呐IQA算法的性能。所有方法均在SIQD、MSSID和SRSID数据集上进行了全面评估。这些算法涵盖了不同的参考图像依赖级别,其核心特性与适用场景如下:

1)NRCMD。一种无参考方法,通过度量轮廓退化进行评估。其优势在于无需参考图像,适用于参考图像缺失的在线评估场景,但对纹理细节不显著的图像或复杂失真评估能力有限。参考已有论文的实验结果来看,其在传统失真和图像传输场景表现尚可,但在图像采集场景下的表现不佳,说明了其对严重斑点噪声的敏感性。

2)PCASS。专为评估SR重建图像设计的无参考方法,采用感知—认知协同驱动框架。其优势在于紧密结合SR重建任务的失真特性,在评估重建图像质量方面具有显著的效果。参考相关研究的实验结果,表明其在SR重建数据集和图像增强场景中性能领先,但在其他失真类型上表现较差,场景适用较为单一。

3)PSIQP。一种半参考方法,通过比对失真前后的部分特征进行质量评估,适用于带宽受限但可传输少量参考信息的水下通信场景。在文献报道的实验中,其在声呐数据集上表现良好,在图像传输场景中也具竞争力,但在其他类型的数据集上表现较为一般。

4)SIQP。一种全参考方法,融合图像的统计与结构信息。其优势在于充分利用参考图像信息,评估精度高,尤其适用于有高质量参考图像或特定目标比对的场景。参考相关实验结果,表明其在模拟实际声呐环境的数据集及图像采集、传输场景中均表现最优,但在图像增强场景表现较弱,反映了全参考方法对参考图像与待测图像内容一致性具有高度的依赖。

5)TPSIQA。一种面向任务和感知的半参考方法,针对小样本问题引入了选择性集成学习机制。适合用于样本量有限且需考虑任务效用的复杂水下环境。通过参考已有论文的实验结果,可以看出其性能通常优于无参考方法,在多个场景下表现稳定,说明了半参考方法在有限参考信息下具有平衡精度与实用性的优势。

性能定量分析采用4种指标:斯皮尔曼等级相关系数(Spearman rank-order correlation coefficient,

SRCC)、全称肯德尔等级相关系数(Kendall rank-order correlation coefficient, KRCC)、均方根误差(root mean square error, RMSE)以及皮尔逊线性相关系数(Pearson linear correlation coefficient, PLCC)。

1)SRCC。用于衡量IQA模型预测结果的单调性,其取值范围为0~1,性能值等于1时,表明两组数据完全一致,其计算式为

$$f_{\text{SRCC}} = 1 - \frac{6 \sum_{i=1}^N d_i^2}{N(N^2 - 1)} \quad (11)$$

式中, $N$ 表示样本数量, $d_i$ 表示第 $i$ 幅图像主观质量分数排名与客观质量分数排名的差值。

2)KRCC。与SRCC一样,用于衡量IQA模型预测结果的单调性。KRCC的数值越大,说明两个信号数据之间的相关性越好,值越小说明相关性越差。其计算式为

$$f_{\text{KRCC}} = \frac{2(N_c - N_d)}{N(N - 1)} \quad (12)$$

式中, $N$ 表示样本数量, $N_c$ 是数据集中的一致对的个数, $N_d$ 是数据集中的不一致对个数。

3)RMSE。用于评估IQA模型预测的一致性,其计算式为

$$f_{\text{RMSE}} = \sqrt{\frac{1}{n} \sum_{i=1}^N (s_i - p_i)^2} \quad (13)$$

式中, $n$ 表示样本数量, $s_i$ 和 $p_i$ 分别表示第 $i$ 幅图像的主观质量分数及客观质量分数。

4)PLCC。描述两组数据之间的线性相关性,其取值范围为-1~1。当PLCC的值为零时,表示两组数据完全不相关;当PLCC值为1或-1时,表明两组数据完全相关。另外,在计算PLCC前,需对客观分数和主观分数进行非线性回归操作。用于非线性回归的logistic函数计算表达式以及PLCC计算表达式分别为

$$p(Q) = \beta_1 \left[ \frac{1}{2} - \frac{1}{1 + e^{(\beta_2(Q - \beta_3))}} \right] + \beta_4 Q + \beta_5 \quad (14)$$

$$f_{\text{PLCC}} = \frac{\sum_{i=1}^N (s_i - \bar{s})(p_i - \bar{p})}{\sqrt{\sum_{i=1}^N (s_i - \bar{s})^2 \sum_{i=1}^N (p_i - \bar{p})^2}} \quad (15)$$

式中, $Q$ 表示原始的客观质量分数, $\beta_1, \beta_2, \beta_3, \beta_4, \beta_5$ 为模型参数, $p$ 是经过回归操作的客观质量分数。 $s_i$ 和 $p_i$ 分别表示第 $i$ 幅图像的主观质量分数和客观

质量分数,  $\bar{s}$  和  $\bar{p}$  分别表示主观质量分数平均值和客观质量分数平均值。

实验中使用的算法均为非学习算法, 评估结果通过计算整个数据集图像的质量分数与其主观评分之间的统计相关性获得。计算工作基于配备单个 NVIDIA 4060 GPU 的 Windows 11 电脑, 在 MATLAB R2024b 平台上执行。

4.1.3 实验结果对比分析

声呐图像质量评价算法在 SIQD、SRSID 以及 MSSID 数据集上的性能对比结果如表 6—表 8 所示。可以看出, 在 SIQD 数据集上, SIQP 算法性能最优; 在 SRSID 数据集上, PCASS 与 SRSIQA 算法性能显著优于其他算法; 在 MSSID 数据集上, SIQP 和 TPSIQA 在采集与传输场景表现优异, 但在增强场景中, PCASS 表现最佳, NRCDM 次之, 而 SIQP 则表现最弱。

表 6 声呐 IQA 算法在 SIQD 数据集上的性能比较  
Table 6 Performance comparison of sonar IQA algorithm on SIQD dataset

算法	PLCC ↑	SRCC ↑	KRCC ↑	RMSE ↓
NRCDM	0.734 0	0.709 1	0.513 8	9.497 9
PCASS	0.525 6	0.161 5	0.122 4	11.897 4
PSIQP	0.764 0	<u>0.757 7</u>	0.562 2	9.022 7
SIQP	<b>0.810 7</b>	<b>0.800 2</b>	<b>0.610 3</b>	<b>8.187 5</b>
TPSIQA	<u>0.767 6</u>	0.756 0	<u>0.562 3</u>	<u>8.963 5</u>

注: 加粗、下划线字体分别表示各列最优、次优结果, “↑”表示值越高越好, “↓”表示值越低越好。

从实验结果可以得出以下结论:

1) 不同数据集模拟的失真类型和场景差异会影响算法的性能表现。SIQD 数据集模拟水下声学信

表 7 声呐 IQA 算法在 SRSID 数据集上的性能比较  
Table 7 Performance comparison of sonar IQA algorithm on SRSID dataset

算法	PLCC ↑	SRCC ↑	KRCC ↑	RMSE ↓
NRCDM	<u>0.699 4</u>	<u>0.699 4</u>	<u>0.458 2</u>	<u>9.561 2</u>
PCASS	<b>0.775 7</b>	<b>0.763 0</b>	<b>0.568 9</b>	<b>8.442 6</b>
PSIQP	0.474 2	0.446 2	0.318 7	11.777 9
SIQP	0.494 4	0.449 3	0.306 1	11.181 4
TPSIQA	0.579 7	0.514 1	0.347 5	10.901 2

注: 加粗、下划线字体分别表示各列最优、次优结果, “↑”表示值越高越好, “↓”表示值越低越好。

道失真, 更贴近实际声呐应用环境; SRSID 数据集聚焦于 SR 重建图像, 主要反映重建过程的特定失真; MSSID 数据集则涵盖了采集、增强和传输 3 种典型场景, 提供了更全面的评估视角。

2) NRCDM、SIQP、TPSIQA 及 PSIQP 主要针对常规声呐图像质量评估设计, 在常见失真特征提取与评估方面具备优势; SIQP 作为一种 FR 声呐 IQA 方法, 利用原始图像参考信息, 在模拟实际声呐环境的 SIQD 数据集上表现突出; TPSIQA 和 PSIQP 作为 RR 声呐 IQA 方法, 其性能通常优于 NRCDM 等无参考方法, 尤其在 MSSID 的采集和传输场景中表现良好。

3) PCASS 专为声呐图像重建任务设计, 充分融入了声呐图像的特有属性和重建伪影。这使得其在专注于重建失真的 SRSID 数据集上表现优异, 在 MSSID 的增强场景中也拔得头筹。然而, 这种与声呐重建特性的深度绑定也限制了通用性: 在模拟信道失真的 SIQD 数据集上表现不佳, 在 MSSID 的采集和传输场景表现也较弱甚至最差。

4) MSSID 数据集的结果揭示了声呐 IQA 的场景

表 8 声呐 IQA 算法在 MSSID 数据集上的性能比较  
Table 8 Performance comparison of sonar IQA algorithm on MSSID dataset

算法	采集场景				增强场景				传输场景			
	PLCC ↑	SRCC ↑	KRCC ↑	RMSE ↓	PLCC ↑	SRCC ↑	KRCC ↑	RMSE ↓	PLCC ↑	SRCC ↑	KRCC ↑	RMSE ↓
NRCDM	0.305 6	0.257 6	0.174 5	16.712 5	<u>0.699 4</u>	<u>0.652 3</u>	<u>0.458 2</u>	<u>9.561 2</u>	0.763 4	0.738 7	0.534 7	8.032 9
PCASS	0.387 7	0.125 6	0.069 8	16.187 2	<b>0.775 7</b>	<b>0.763 0</b>	<b>0.568 9</b>	<b>8.442 7</b>	0.560 3	0.215 5	0.161 2	10.301 3
PSIQP	0.528 6	0.401 1	0.277 9	14.899 4	0.478 4	0.452 7	0.323 6	11.747 8	0.800 9	0.806 7	0.616 2	7.446 7
SIQP	<b>0.580 9</b>	<b>0.612 4</b>	<b>0.433 8</b>	<b>14.286 9</b>	0.474 1	0.429 4	0.292 0	11.779 1	<b>0.895 9</b>	<b>0.880 9</b>	<b>0.702 9</b>	<b>5.525 2</b>
TPSIQA	<u>0.549 5</u>	<u>0.524 7</u>	<u>0.363 1</u>	<u>14.664 8</u>	0.586 7	0.519 8	0.351 4	10.833 9	<u>0.817 1</u>	<u>0.821 6</u>	<u>0.625 6</u>	<u>7.169 7</u>

注: 加粗、下划线字体分别表示各列最优、次优结果。

依赖性。没有任何单一算法在所有失真场景下表现最优。SIQP在采集和传输场景领先,却在增强场景垫底;PCASS在增强场景最优,却在其他场景表现较差。这表明声呐IQA算法的选择和设计必须紧密结合目标应用所面临的主要失真类型。

## 4.2 通用增强算法验证实验

### 4.2.1 实验数据集

验证实验采用以下数据集:SIQD(Chen等,2017b)、USOD(underwater salient object detection)(Wang等,2024)。SIQD数据集用于通用去噪实验,USOD数据集用于通用SR重建实验。USOD数据集共包含1602幅声呐图像。该数据集提供了485幅具有代表性的高分辨率图像作为参考图像,内容涵盖飞机、橡胶人、沉船等典型目标。所有图像分辨率均统一为 $256 \times 256$ 像素。在此基础上,每幅参考图像还提供了对应的2倍和4倍下采样版本。

### 4.2.2 评价指标与实施细节

SR重建验证实验采用4种通用SR算法:LBNNet、ESRT、NGswin和SRFormer,并选用3种指标进行定量分析:PSNR与SSIM作为SR领域广泛采用的基础评价指标;PCASS则是最新提出的声呐图像专用质量评价指标,可显著提升评估的可靠性。4种通用SR算法的详细介绍如下:

1)LBNNet。通过对称的CNN分支和递归Transformer分支相结合,其优势在于以轻量级的架构实现局部特征提取与依赖建模的平衡。在自然图像上的表现,表明了算法在计算效率和重建质量间取得了良好权衡,尤其适合移动端或边缘设备部署。

2)ESRT。一种轻量级Transformer模型,专注于在保持性能的同时降低计算复杂度。相关文献报告其在多个基准数据集上能以更少的参数达到很好的性能,适用于对实时性有要求的场景。

3)NGswin。在Swin Transformer基础上引入N-Gram机制以增强局部建模能力。在自然图像SR重建实验中,其性能与ESRT相近,改进的局部注意力机制对图像纹理恢复有一定作用。

4)SRFormer。采用置换自注意力机制的Transformer模型,该机制能更灵活地建模图像块之间的依赖关系,从而更好地恢复图像的全局结构。实验表明,其在多项自然图像超分基础指标上均领先,说明其重建图像与参考图像在像素级相似度上极高。

去噪验证实验同样采用4种通用算法:

Restormer、Uformer、APD-Net和MSANet,以及3种指标:PSNR与SSIM作为通用的图像质量基础指标;NRCDM专为声呐图像设计的质量评估指标,该指标以高精度、高效性以及声呐图像关键特征的侧重而著称。4种通用去噪算法的详细信息如下:

1)Restormer。一种基于Transformer的高性能恢复模型,采用多头注意力和门控机制。其优势在于强大的全局信息整合与复杂噪声建模能力。在自然图像去噪任务中,其展现出领先于其他去噪算法的性能表现。

2)Uformer。采用U型架构与局部-全局自注意力机制的网络。其设计有利于在去噪的同时保持边缘结构。其在低光图像增强和去噪等任务中表现优异,尤其在保持边缘结构方面有显著优势。

3)APD-Net。一种基于自适应先验的去噪网络,通过学习自适应先验来指导去噪过程,以此获得更好的泛化能力。但在自然图像去噪实验中,相关研究指出其性能在不同噪声类型和水平下可能存在波动。

4)MSANet。通过并行多尺度通路和自适应特征选择机制,旨在处理不同尺度的噪声和图像结构。相关实验表明,其在处理具有复杂纹理和大小不一目标的自然图像时具有鲁棒性。

PSNR通过计算重建图像与参考图像之间的均方误差(mean squared error, MSE)评估图像质量。其单位为dB,值越高表明图像失真越小。在图像压缩领域,典型的PSNR值范围约为20 dB至40 dB,且数值越大代表质量越好。

SSIM同样是一种全参考型图像质量评价指标。它通过分别度量亮度、对比度和结构3个方面的相似性评估图像质量。相较于PSNR,SSIM在模拟人眼对图像结构的感知方面表现更优,因为它不仅考虑像素间的差异,还融入了图像的局部结构、亮度及对比度信息。

实验采用的算法均为基于学习的方法。数据集按8:2的比例划分为训练集与测试集。所有训练及测试均在配备单张NVIDIA 4060 GPU的Windows 11系统上完成,以确保评估公平性。实验代码基于Python 3.7实现。

### 4.2.3 实验结果分析

通用SR算法在 $\times 4$ 重建任务中的实验结果如表9所示。可以看出,在 $\times 4$ 重建任务中,通用SR算法

SRFormer 在基础评价指标 PSNR 和 SSIM 上均取得最优值。其原因在于 SRFormer 采用了置换自注意力机制,该机制使 Transformer 模型能够有效建模图像块之间的长程依赖关系,从而能更加精准地恢复目标的全局结构信息。相比之下,LBNet 则在声呐图像专用评价指标 PCASS 上表现最优。这是由于 PCASS 指标侧重于评估图像在视觉感知质量和任务实用性方面的表现。而 LBNet 作为一种轻量级双模网络,采用了对称 CNN 与递归 Transformer 的架构。其中对称 CNN 分支有效提取局部特征,递归 Transformer 分支则通过参数共享机制在轻量级前提下建模长程依赖。这种设计使 LBNet 在保持模型轻量化的同时,较好地平衡了局部细节与全局结构的重建,其输出在纹理自然度方面可能更符合主观感知,因此在强调实用性的 PCASS 指标上表现突出。另外,ESRT 与 NGswin 作为轻量级超分模型,其性能介于 SRFormer 和 LBNet 之间。ESRT 通过精心设计的轻量 Transformer 架构实现效率与性能的平衡,而 NGswin 则在 Swin Transformer 基础上引入 N-Gram 机制以增强局部结构建模。两者都体现了轻量级设计思想,在计算效率与重建质量间取得了较好权衡。

表9 通用SR算法在×4重建任务中的实验结果

Table 9 Experimental results of the general SR algorithm in the ×4 reconstruction task

算法	PSNR/dB ↑	SSIM ↑	PCASS ↑
LBNet	23.292 8	0.594 4	<b>6.364 5</b>
ESRT	<u>23.711 6</u>	<u>0.613 0</u>	6.282 4
NGswin	23.580 7	0.609 9	6.296 4
SRFormer	<b>23.835 2</b>	<b>0.618 7</b>	<u>6.333 1</u>

注:加粗、下划线字体分别表示各列最优、次优结果,“↑”表示值越高越好。

通用去噪算法实验结果如表 10 所示。可以看出,在去噪任务中,通用去噪算法 Restormer 在 PSNR 和 SSIM 上展现出显著优势。这是由于 Restormer 是一种基于 Transformer 的先进恢复模型,其核心是多头 Transformer 块与门控机制,具有强大的全局信息整合与复杂噪声建模能力,能有效抑制声呐图像中广泛存在的相干斑噪声,从而在像素级相似度上达到很高水平。但值得注意的是,在声呐图像专用指标 NRCDM 上,Uformer 表现最优,MSANet 紧随其

后。NRCDM 指标侧重于通过度量轮廓退化来评估质量,对边缘和结构的保持能力较为敏感。而 Uformer 采用 U 型架构与局部—全局自注意力,这种编码器—解码器结构及其窗口注意力机制在有效去噪的同时,能更好地保留目标的边缘锐利度和局部结构。同理,MSANet 的多尺度自适应机制同样有助于在去除噪声时保护不同尺度的细节。相比之下,Restormer 虽然去噪能力强,但其过于强大的平滑效应可能导致边缘模糊化,从而在强调轮廓保真度的 NRCDM 指标上略有失分。另外,APD-Net 在所有指标上均表现最弱。表明该模型所依赖的先验知识或网络结构与声呐图像的噪声特性之间存在显著的不匹配,导致其在水下场景的泛化能力不足。

表10 通用去噪算法实验结果

Table 10 Experimental results of the general denoising algorithms

算法	PSNR/dB ↑	SSIM ↑	NRCDM ↑
Restormer	<b>35.611 0</b>	<b>0.874 5</b>	48.527 2
Uformer	34.036 8	0.862 4	<b>49.033 7</b>
APD-Net	32.680 3	0.851 5	47.826 2
MSANet	<u>34.274 9</u>	<u>0.867 7</u>	<u>49.027 5</u>

注:加粗、下划线字体分别表示各列最优、次优结果,“↑”表示值越高越好。

## 5 展 望

鉴于声呐图像在海洋探测、水下目标识别等应用中的重要性日益凸显,对其质量进行精准评价与有效优化的需求持续增长。在这一背景下,未来该领域的研究可重点关注以下方向:

1) 声呐图像质量评价应深度融合其成像机理与应用背景。潜在路径:未来的研究可探索构建物理机理嵌入的深度学习模型。例如,将声学波动方程、水体衰减模型、传感器点扩散函数等物理先验作为神经网络的物理约束层或可微分模块。同时,开发高保真的声呐成像仿真平台同样具有意义,仿真平台可用于生成与真实物理过程一致的合成数据,以便进一步进行模型预训练或数据增强。核心挑战:如何对复杂多变的海洋环境进行可计算且足够精确的数学建模?如何简化物理模型?如何降低其与真

实场景之间的差距?如何实现物理模型与数据驱动模型的高效、稳定耦合?都是需要攻克的核心问题。

2)发展面向小样本甚至零样本学习的质量评价与优化方法。潜在路径:重点发展面向声呐的自监督与生成式预训练策略。利用无标注声呐图像,通过设计如拼图、去噪和掩膜图像重建等前置任务,让模型学习其通用表征。对于零样本场景,可借鉴对比语言—图像预训练模型的范式,构建一个共享嵌入空间,将声呐图像特征与语义信号或其他信息进行对齐,这些信息包括但不限于成像参数(如频率、入射角)、物理特征(如混响强度)或任务标签(如“沉船”、“平坦底质”)。核心挑战:声呐图像与自然图像存在巨大域差异,使得从自然图像迁移而来的方案在声呐图像上效果受限。同时,声呐领域缺乏定义清晰的“语义描述”,如何构建有效、量化的替代监督信号,是实现零样本学习面临的首要难题。

3)确立以效用为导向的质量评价体系。潜在路径:推动“任务感知”的质量度量,如以下游任务(如目标检测器的召回率、地貌分类器的准确率)作为优化目标的评价指标。在方法上,可构建基于注意力的可解释性模型,通过分析下游任务模型(如分类器)在图像上的关注区域,反推哪些图像特征对任务成功至关重要,从而将这些特征的质量作为评价依据。核心挑战:如何量化定义并平衡“图像保真度”与“任务效用”之间的关系是一大难题。这种任务驱动的质量度量可能缺乏通用性,为一个任务优化的图像可能不利于另一个任务。此外,联合训练的收敛性和稳定性也需要深入研究。

4)构建感知—优化一体化的嵌入型系统。潜在路径:研究可微分的质量评价模块作为损失函数或强化学习中的奖励函数,直接引导图像增强、超分辨率重建等优化算法。如,可训练一个质量评价网络作为生成对抗网络中判别器的补充,或者作为强化学习智能体的奖励,从而实现评价与优化的闭环反馈与联合迭代。核心挑战:如何保证评价网络本身的准确性和可靠性是确保系统有效的基础。更重要的是,这种闭环系统可能存在稳定性问题,如陷入局部最优或模式崩溃。此外,这类复杂模型的计算开销巨大,与水下场景的“轻量化”需求存在天然矛盾,需要在算法设计和效率之间做出权衡。

基于上述方向,未来研究还可进一步探索:多模态数据(如光学与声学图像)的融合质量评价,关键

在于解决不同模态间特征空间与尺度不一致、信息互补性等挑战;发展针对恶劣成像条件(如强混响、低信噪比)的鲁棒性增强方法,其核心是提高模型对复杂噪声和干扰的泛化能力;以及进行轻量化模型设计,通过神经网络剪枝、量化和知识蒸馏等技术,以适应边缘计算设备的实时处理需求。通过这种多学科、多层次的交叉合作,将有望推动声呐图像质量技术从理论创新走向系统化、实用化的发展,最终提升其在水下探测与应用中的实际效能。

## 6 结 语

随着海洋探测需求的日益增长,声呐成像技术的重要性愈发凸显。为克服水下复杂环境导致的图像退化问题并充分挖掘数据价值,声呐图像质量增强与评估技术的研究至关重要。质量增强技术已从传统滤波演进至深度学习方法,有效应对了信道衰减、混响噪声等挑战,显著提升了图像清晰度与细节表现。与此同时,质量评估技术为衡量增强效果提供了客观的量化依据。二者相辅相成,共同构成了声呐图像高价值信息提取的核心环节。未来研究需着力突破复杂场景适应性、评估指标体系完善以及计算效率优化等关键方向。

## 参考文献(References)

- Cai B Q, Chen W L, Zhang J H, Junejo N U R and Zhao T S. 2025a. Unified no-reference quality assessment for sonar imaging and processing. *IEEE Transactions on Geoscience and Remote Sensing*, 63: #5902711 [DOI: 10.1109/TGRS.2024.3524835]
- Cai W Y, Zhu J F and Zhang M Y. 2025b. From classical approach to deep-learning: a review on underwater target segmentation with sonar image. *Neurocomputing*, 637: #130087 [DOI: 10.1016/j.neucom.2025.130087]
- Chen D S, Chu X M, Ma F and Teng X X. 2017a. A variational approach for adaptive underwater sonar image denoising//*Proceedings of the 4th International Conference on Transportation Information and Safety (ICTIS)*. Banff, Canada: IEEE: 1177-1181 [DOI: 10.1109/ICTIS.2017.8047920]
- Chen M, Li L, Li Z J and Xie X M. 2021a. Research on sonar image denoising method based on fixed water area noise model//*Proceedings of 2021 IEEE International Conference on Mechatronics and Automation (ICMA)*. Takamatsu, Japan: IEEE: 231-235 [DOI: 10.1109/ICMA52036.2021.9512575]

- Chen W L, Cai B Q, Zheng S M, Zhao T S and Gu K. 2024. Perception-and-cognition-inspired quality assessment for sonar image super-resolution. *IEEE Transactions on Multimedia*, 26: 6398-6410 [DOI: 10.1109/TMM.2024.3349929]
- Chen W L, Gu K, Lin W S, Xia Z F, Le Callet P and Cheng E. 2019. Reference-free quality assessment of sonar images via contour degradation measurement. *IEEE Transactions on Image Processing*, 28(11): 5336-5351 [DOI: 10.1109/TIP.2019.2910666]
- Chen W L, Gu K, Lin W S, Yuan F and Cheng E. 2020. Statistical and structural information backed full-reference quality measure of compressed sonar images. *IEEE Transactions on Circuits and Systems for Video Technology*, 30(2): 334-348 [DOI: 10.1109/TCSVT.2019.2890878]
- Chen W L, Gu K, Min X K, Yuan F and Cheng E and Zhang W J. 2018a. Partial-reference sonar image quality assessment for underwater transmission. *IEEE Transactions on Aerospace and Electronic Systems*, 54(6): 2776-2787 [DOI: 10.1109/TAES.2018.2829378]
- Chen W L, Gu K, Zhao T S, Jiang G Y and Callet P L. 2021b. Semi-reference sonar image quality assessment based on task and visual perception. *IEEE Transactions on Multimedia*, 23: 1008-1020 [DOI: 10.1109/TMM.2020.2991546]
- Chen W L, Qiu Y L, Zhao T S, Wei H and Cheng E. 2023. Advances in ocean-oriented underwater image processing and vision technology. *Journal of Signal Processing*, 39(10): 1748-1763 (陈炜玲, 邱艳玲, 赵铁松, 魏宏安, 程恩. 2023. 面向海洋的水下图像处理与视觉技术进展. *信号处理*, 39(10): 1748-1763) [DOI: 10.16798/j.issn.1003-0530.2023.10.003]
- Chen W L, Yuan F, Cheng E and Lin W S. 2017b. Subjective and objective quality evaluation of sonar images for underwater acoustic transmission//*Proceedings of 2017 IEEE International Conference on Image Processing (ICIP)*. Beijing, China: IEEE: 176-180 [DOI: 10.1109/ICIP.2017.8296266]
- Chen W L N, Yuan F, Cheng E and Gao C X. 2018b. Sonar image quality assessment based on degradation measurement//*Proceedings of 2018 OCEANS-MTS/IEEE Kobe Techno-Oceans (OTO)*. Kobe, Japan: IEEE: 1-5 [DOI: 10.1109/OCEANSKOB.2018.8559343]
- Choi H, Lee J and Yang J. 2023. N-Gram in swin transformers for efficient lightweight image super-resolution//*Proceedings of 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Vancouver, Canada: IEEE: 2071-2081 [DOI: 10.1109/CVPR52729.2023.00206]
- Feng Q X, Zheng S M, Zhang K K and Wei H A. 2024. A brain-inspired quality assessment model for sonar image super-resolution. *Displays*, 82: #102620 [DOI: 10.1016/j.displa.2023.102620]
- Gao G W, Wang Z X, Li J C, Li W J, Yu Y and Zeng T Y. 2022. Lightweight bimodal network for single-image super-resolution via symmetric CNN and recursive transformer//*Proceedings of the 31st International Joint Conference on Artificial Intelligence*. Vienna, Austria: IJCAI: 913-919
- Gerg I D, Cook D A and Monga V. 2024. Deep adaptive phase learning: enhancing synthetic aperture sonar imagery through learned coherent autofocus. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 17: 9517-9532 [DOI: 10.1109/JSTARS.2024.3393139]
- Gou, Y B, Hu P, Lv J C, Zhou J T and Peng X. 2022. Multi-scale adaptive network for single image denoising//*Proceedings of the 36th International Conference on Neural Information Processing Systems*. New Orleans, USA: Curran Associates Inc.: 14099-14112
- Guo J C, Li C Y, Guo C L and Chen S J. 2017. Research progress of underwater image enhancement and restoration methods. *Journal of Image and Graphics*, 22(3): 273-287 (郭继昌, 李重仪, 郭春乐, 陈善继. 2017. 水下图像增强和复原方法研究进展. *中国图象图形学报*, 22(3): 273-287) [DOI: 10.11834/jig.20170301]
- Huang H N, Li B Q, Liu J Y, Liu Z J, Wei L Z and Zhao S. 2024. Sonar image underwater target recognition: a comprehensive overview and prospects. *Journal of Electronics and Information Technology*, 46(5): 1742-1760 (黄海宁, 李宝奇, 刘纪元, 刘正君, 韦琳哲, 赵爽. 2024. 声呐图像水下目标识别综述与展望. *电子与信息学报*, 46(5): 1742-1760) [DOI: 10.11999/JEIT231207]
- Ji X, Leng N and Guo H. 2024. Underwater image enhancement and restoration techniques: state-of-the-art methods and prospects. *Journal of Computer-Aided Design and Computer Graphics*, 36(6): 805-830 (纪勋, 冷娜, 郭慧. 2024. 水下图像增强与复原技术进展与展望. *计算机辅助设计与图形学学报*, 36(6): 805-830) [DOI: 10.3724/SP.J.1089.2024.20116]
- Jiang B, Lu Y, Wang J H, Lu G M and Zhang D. 2022. Deep image denoising with adaptive priors. *IEEE Transactions on Circuits and Systems for Video Technology*, 32(8): 5124-5136 [DOI: 10.1109/TCSVT.2022.3149518]
- Jiao W P, Li J, Zhang C Y, Xie G M, Xiao W D and Zhang J L. 2025. Intelligent perception algorithms for sonar images: a review. *Journal of Unmanned Undersea Systems*, 33(3): 559-572 (焦文沛, 李杰, 张春燕, 谢广明, 肖文栋, 张建磊. 2025. 声呐图像智能感知算法综述. *水下无人系统学报*, 33(3): 559-572) [DOI: 10.11993/j.issn.2096-3920.2024-0127]
- Lei J, Wang H G, Lei Z L, Li J Y and Rong S W. 2025. CNN-transformer hybrid architecture for underwater sonar image segmentation. *Remote Sensing*, 17(4): #707 [DOI: 10.3390/rs17040707]
- Li Z Y, Wang Z S, Chen D S, Yip T L and Teixeira A P. 2024. RepD-Net: a re-parameterization despeckling network for autonomous underwater side-scan sonar imaging with prior-knowledge customized convolution. *Defence Technology*, 35: 259-274 [DOI: 10.1016/j.dt.2023.12.007]
- Liu W J, Yang X H, Li J W and Yang H X. 2024. Two-stage network architecture for sonar image denoising based on U-Net//*Proceedings of the 13th International Conference on Communications, Circuits and Systems (ICCCAS)*. Xiamen, China: IEEE: 548-553 [DOI:

- 10.1109/ICCCAS62034.2024.10652637]
- Luo Y H, Cao X, Zhang J T, Wang T J and Feng Q. 2023. A review of underwater optical image super-resolution reconstruction based on deep learning. *Digital Ocean and Underwater Warfare*, 6(1): 17-33 (罗逸豪, 曹翔, 张钧陶, 王天江, 冯琪. 2023. 基于深度学习的水下光学图像超分辨率重建综述. *数字海洋与水下攻防*, 6(1): 17-33) [DOI: 10.19838/j.issn.2096-5753.2023.01.003]
- Ma Z D, Ding J, Li S S and Zou B B. 2023a. Super-resolution of underwater sonar image based on generative adversarial network//Proceedings of the 49th IECON Annual Conference of the IEEE Industrial Electronics Society. Singapore, Singapore: IEEE: 1-6 [DOI: 10.1109/IECON51785.2023.10312168]
- Ma Z D, Li S S, Ding J and Zou B B. 2023b. MHGAN: a multi-headed generative adversarial network for underwater sonar image super-resolution. *IEEE Transactions on Geoscience and Remote Sensing*, 61: #4209416 [DOI: 10.1109/TGRS.2023.3327045]
- Rao J H, Peng Y N, Chen J and Tian X. 2024. Various degradation: dual cross-refinement transformer for blind sonar image super-resolution. *IEEE Transactions on Geoscience and Remote Sensing*, 62: #4206114 [DOI: 10.1109/TGRS.2024.3398188]
- Shi B, Cao T Y, Wang Z T, Ge Q Q and Guo W B. 2023. Research on sonar image denoising based on optimized spatial pixel ranking algorithm//Proceedings of the 8th International Conference on Signal and Image Processing (ICSIP). Wuxi, China: IEEE: 249-253 [DOI: 10.1109/ICSIP57908.2023.10271087]
- Sung M, Joe H, Kim J and Yu S C. 2018a. Convolutional neural network based resolution enhancement of underwater sonar image without losing working range of sonar sensors//Proceedings of 2018 OCEANS — MTS/IEEE Kobe Techno-Oceans (OTO). Kobe, Japan: IEEE: 1-6 [DOI: 10.1109/OCEANSKOBE.2018.8559421]
- Sung M, Kim J and Yu S C. 2018b. Image-based super resolution of underwater sonar images using generative adversarial network//Proceedings of 2018 TENCON IEEE Region 10 Conference. Jeju, Korea (South): IEEE: 0457-0461 [DOI: 10.1109/TENCON.2018.8650176]
- Thomas C T, Nambiar A M and Mittal A. 2022. A GAN-based super resolution model for efficient image enhancement in underwater sonar images//Proceedings of OCEANS 2022. Chennai, India: IEEE: 1-8 [DOI: 10.1109/OCEANSChennai45887.2022.9775508]
- Tolie H F, Ren J C, Chen R J, Zhao H M and Elyan E. 2025. Blind sonar image quality assessment via machine learning: leveraging micro- and macro-scale texture and contour features in the wavelet domain. *Engineering Applications of Artificial Intelligence*, 141: #109730 [DOI: 10.1016/j.engappai.2024.109730]
- Wang C J, Shen L R, Fan Y S, Chen T S and Tan X Y. 2022a. Sonar image denoising based on anisotropic guided filtering//Proceedings of the 5th International Conference on Intelligent Autonomous Systems (ICoIAS). Dalian, China: IEEE: 54-59 [DOI: 10.1109/ICoIAS56028.2022.9931211]
- Wang K Y, Huang S R and Li Y S. 2022. An optical reconstruction based underwater image analysis. *Journal of Image and Graphics*, 27(5): 1337-1358 (王柯俨, 黄诗芮, 李云松. 2022. 水下光学图像重建方法研究进展. *中国图象图形学报*, 27(5): 1337-1358) [DOI: 10.11834/jig.210819]
- Wang M J, Chen W L, Lan F Q, Ur Rehman Junejo N and Zhao T S. 2025. Sonar image super-resolution based on structure-texture dual preservation. *IEEE Transactions on Geoscience and Remote Sensing*, 63: #5913711 [DOI: 10.1109/TGRS.2025.3570835]
- Wang M J, Zhang K K, Wei H A, Chen W L and Zhao T S. 2024. Underwater image quality optimization: researches, challenges, and future trends. *Image and Vision Computing*, 146: #104995 [DOI: 10.1016/J.IMAVIS.2024.104995]
- Wang Z D, Cun X D, Bao J M, Zhou W G, Liu J Z and Li H Q. 2022b. Uformer: a general u-shaped transformer for image restoration//Proceedings of 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New Orleans, USA: IEEE: 17662-17672 [DOI: 10.1109/CVPR52688.2022.01716]
- Wu D, Du X and Wang K Y. 2018. An effective approach for underwater sonar image denoising based on sparse representation//Proceedings of the 3rd IEEE International Conference on Image, Vision and Computing (ICIVC). Chongqing, China: IEEE: 389-393 [DOI: 10.1109/ICIVC.2018.8492877]
- Xu X Y, Cai H, Wang M J, Chen W L, Zhang R X and Zhao T S. 2025. Exploring underwater image quality: a review of current methodologies and emerging trends. *Image and Vision Computing*, 154: #105389 [DOI: 10.1016/j.imavis.2024.105389]
- Zamir S W, Arora A, Khan S, Hayat M, Khan F S and Yang M H. 2022. Restormer: efficient transformer for high-resolution image restoration//Proceedings of 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New Orleans, USA: IEEE: 5718-5729 [DOI: 10.1109/CVPR52688.2022.00564]
- Zhang C, Chai Y Q, Wang X Y and Chen Y Y. 2025. Enhancing sonar image quality for underwater object recognition in marine environments: a review. *Ocean Engineering*, 336: #121862 [DOI: 10.1016/j.oceaneng.2025.121862]
- Zhang H Q, Li S, Li D H, Wang Z C, Zhou Q X and You Q X. 2022. Sonar image quality evaluation using deep neural network. *IET Image Processing*, 16(4): 992-999 [DOI: 10.1049/ipr2.12199]
- Zheng S M, Chen W L, Zhao T S, Wei H A and Lin L Q. 2022. Utility-oriented quality assessment of sonar image super-resolution//Proceedings of OCEANS 2022, Hampton Roads. Hampton Roads, USA: IEEE: 1-5 [DOI: 10.1109/OCEANS47191.2022.9977189]
- Zheng Y N, Chen W L, Zhao T S, Gu K and Xia Z F. 2020. Semi-reference utility quality assessment of sonar images//Proceedings of Global Oceans 2020; Singapore — U.S. Gulf Coast. Biloxi, USA: IEEE: 1-4 [DOI: 10.1109/IEEECONF38699.2020.9389367]
- Zhou L, Liu Q M, Jin K J, Zhao W Y and Zhang W D. 2025. Research progress of underwater image restoration and enhancement meth-

ods. *Journal of Image and Graphics*, 30(1): 51-65 (周玲, 刘庆敏, 金凯杰, 赵文义, 张卫东. 2025. 水下图像复原和增强方法研究进展. *中国图象图形学报*, 30(1): 51-65) [DOI: 10.11834/jig.240050]

Zhou Y P, Li Z, Guo C L, Bai S, Cheng M M and Hou Q B. 2023. SRFormer: permuted self-attention for single image super-resolution//*Proceedings of 2023 IEEE/CVF International Conference on Computer Vision*. Paris, France: IEEE: 12734-12745 [DOI: 10.1109/ICCV51070.2023.01174]

Zhuang P X, Wang Y H, Zhang X H, Liu F, Tong J J and Fu Z Q. 2026. A survey of underwater light field imaging: theory and applications. *Journal of Image and Graphics*, 31(5): 1503-1522 (庄培显, 王一航, 张新恒, 刘飞, 童俊杰, 富振奇. 2026. 水下光场成像综述: 理论与应用. *中国图象图形学报*, 31(5): 1503-1522

[DOI: 10.11834/jig.250255]

## 作者简介

林杰,男,硕士研究生,主要研究方向为图像处理与计算机视觉。E-mail:241127032@fzu.edu.cn

陈炜玲,通信作者,女,副教授,主要研究方向为智慧海洋、图像处理与计算机视觉、机器学习。

E-mail:weiling.chen@fzu.edu.cn

徐晓怡,女,硕士研究生,主要研究方向为图像处理与计算机视觉。E-mail:231120110@fzu.edu.cn

赵铁松,男,教授,主要研究方向为人工智能与多媒体通信。

E-mail:t.zhao@fzu.edu.cn