**学校编码：10384**

**学 号：22320221151440**



**硕 士 学 位 论 文**

**基于自监督深度学习的海洋工程水下噪声提取研究**

**Research on Underwater Noise Extraction for Marine Engineering Based on Self-Supervised Deep Learning**

**杨承昊**

**指导教师姓名：陶毅 助理教授**

**专 业 名 称 ：海 洋 物 理**

**论文提交日期：2025 年 4 月**

**论文答辩时间：2025 年 5 月**

**学位授予日期：2025 年 月**

**2025 年 5 月**

**厦门大学学位论文原创性声明**

本人呈交的学位论文是本人在导师指导下,独立完成的研究成果。本人在论文写作中参考其他个人或集体已经发表的研究成果，均在文中以适当方式明确标明，并符合法律规范和《厦门大学研究生学术活动规范（试行）》。

另外，该学位论文为（ ）课题（组）的研究成果，获得（ ）课题（组）经费或实验室的资助，在（ ）实验室完成。（请在以上括号内填写课题或课题组负责人或实验室名称，未有此项声明内容的，可以不作特别声明。）

本人声明该学位论文不存在剽窃、抄袭等学术不端行为，并愿意承担因学术不端行为所带来的一切后果和法律责任。

声明人 （签名）：

指导教师（签名）：

年 月 日

**厦门大学学位论文著作权使用声明**

本人同意厦门大学根据《中华人民共和国学位条例暂行实施办法》等规定保留和使用此学位论文，并向主管部门或其指定机构送交学位论文（包括纸质版和电子版），允许学位论文进入厦门大学图书馆及其数据库被查阅、借阅。本人同意厦门大学将学位论文加入全国博士、硕士学位论文共建单位数据库进行检索，将学位论文的标题和摘要汇编出版，采用影印、缩印或者其它方式合理复制学位论文。

本学位论文属于：

（ ）1.经厦门大学保密委员会审查核定的涉密学位论文，于 年 月 日解密，解密后适用上述授权。

（ ）2.不涉密，适用上述授权。

（请在以上相应括号内打“√”或填上相应内容。涉密学位论文应是已经厦门大学保密委员会审定过的学位论文，未经厦门大学保密委员会审定的学位论文均为公开学位论文。此声明栏不填写的，默认为公开学位论文，均适用上述授权。）

声明人（签名）：

年 月 日

# 摘要

随着海洋工程建设活动的日益增多，海洋工程施工过程中的声学噪声问题引起了广泛关注。其中，海洋工程打桩作业以及桥梁结构振动是典型的水下声学信号源，其在工作过程中会产生显著的声波辐射。由于这些声波在远距离水下传播中经历显著衰减，导致接收信号信噪比较低，使得水下噪声的提取与监测面临巨大挑战。

远距离声学监测和海洋生态影响评估等应用场景对于采集和分析这些声源信号具有重要意义。然而，海洋环境中存在多种强背景噪声，使得微弱的冲击噪声中的脉冲成分容易被淹没。在强背景噪声环境中准确提取海洋工程施工产生的弱脉冲信号成为亟待解决的关键技术问题。为此，本研究围绕低信噪比情况下海洋工程水下噪声的提取与降噪展开，主要研究工作概况如下：

（1）搭建了水下噪声信号采集系统，在实际海洋施工现场获取了打桩噪声及海洋背景噪声数据，并对其时域和频域特性进行分析，明确了目标信号和背景噪声的特征差异，为模型设计提供支撑。结果显示，目标噪声的能量主要分布在100 Hz–1 kHz 的中低频段，且同时具有短时脉冲、高峰声压与复杂尾波等特征。上述特性使维纳滤波、小波变换等传统方法在尾波抑制和结构保真之间难以权衡。而水声数据的稀缺性也使得依赖纯净标签的监督式神经网络失去适用性，凸显了开发自监督式神经网络提取降噪技术的必要性。

（2）设计了自监督学习策略，通过构造训练信号对在无需纯净参考信号的情况下实现降噪模型的有效训练，有效克服了噪声连续性带来的伪标签污染问题。然后，开发了自适应多重聚焦降噪网络（AMFNet），采用编码与解码器架构，融合频谱动态聚焦、复数注意力、语义重校准和复数多尺度跳跃连接等模块，以充分挖掘水下噪声的复杂时频特征，实现对海洋工程水下噪声的高保真提取与降噪，从方法上突破了传统方法、监督式神经网络和现有自监督学习模型的局限。

（3）利用实地采集的水下打桩噪声和桥梁振动噪声混叠真实海洋背景噪声，构建了三类真实水下噪声数据集（包括累计能量段打桩噪声、原始打桩噪声和未参与训练的桥梁振动噪声）对AMFNet模型进行了性能验证。实验结果显示，AMFNet 在三类真实数据集上均取得领先性能。以混叠信噪比为 -10 dB 的低信噪比条件下的累计能量段打桩声实验为例，输出信噪比相较输入提高了 13.07 dB（从 -10.00 dB 提升至 3.07 dB），信号频谱结构从难以识别的状态被恢复到清晰可辨的结构。相较于传统降噪方法和自监督神经网络模型（如U-Net、N2N、NerNT、ONT等），AMFNet在各项指标上均表现出明显优势。

综上，本文方法为远距水下噪声监测、施工状态评估与海洋生态影响量化提供了新型海洋工程水下噪声提取技术路径，具有良好的工程推广价值。

**关键词：**水声信号处理；深度学习；自监督学习；水下打桩噪声；水下桥梁振动噪声；

# Abstract

With the increasing frequency of marine engineering construction activities, acoustic noise generated during these processes has become a significant concern. Underwater pile-driving operations in marine engineering and structural vibrations of bridges are typical sources of underwater acoustic signals, which generate significant acoustic radiation during operation. However, due to considerable attenuation during long-distance underwater propagation, the received signals often exhibit low signal-to-noise ratios, posing substantial challenges for the extraction and monitoring of underwater noise.

Long-distance acoustic monitoring and marine ecological impact assessments are critically dependent on accurately capturing and analyzing these noise sources. However, the marine environment is characterized by intense background noise, which often masks weak impulsive noise components. Thus, accurately extracting weak impulse signals generated by marine engineering activities in strong background noise conditions has become a critical technological challenge. Addressing this issue, this study focuses on underwater noise extraction and denoising under low-SNR conditions. The main research contributions are summarized as follows:

(1) An underwater acoustic signal acquisition system was developed to gather pile-driving noise and marine background noise from actual marine construction sites. Time-domain and frequency-domain analyses were conducted to clearly differentiate the characteristics of target signals from background noise, providing foundational insights for model design. The analysis indicated that target noise energy primarily concentrated in the medium-to-low frequency range (100 Hz – 1 kHz) and featured short-time impulses, high peak sound pressure levels, and complex tail-wave structures. These characteristics rendered traditional methods like Wiener filtering and wavelet transforms insufficient due to their struggle to balance tail-wave suppression and structural fidelity. Furthermore, the scarcity of clean underwater acoustic data limits the applicability of supervised neural networks reliant on clean labels, emphasizing the necessity for developing self-supervised neural network techniques.

(2) A self-supervised learning strategy was designed to enable effective training of the denoising model without the need for clean reference signals by constructing paired training samples. This approach effectively mitigates the pseudo-label contamination problem caused by the continuity of noise. Subsequently, an Adaptive Multi-Focus Denoising Network (AMFNet) was developed, employing an encoder-decoder architecture that integrates spectral dynamic focusing, complex-valued attention, semantic recalibration, and complex multi-scale skip connections. These modules are designed to fully exploit the complex time-frequency characteristics of underwater noise, enabling high-fidelity extraction and denoising of underwater acoustic signals in marine engineering scenarios. Methodologically, this framework overcomes the limitations of traditional approaches, supervised neural networks, and existing self-supervised learning models.

(3) By mixing field-recorded underwater pile-driving noise and bridge vibration noise with real ocean background noise, three types of realistic underwater noise datasets were constructed—namely, the cumulative energy segment pile-driving dataset, the original pile-driving dataset, and the unseen bridge vibration dataset not involved in training—to evaluate the performance of the proposed AMFNet model. Experimental results demonstrate that AMFNet achieves superior performance across all three datasets. For instance, in the cumulative energy segment pile-driving experiment under a low signal-to-noise ratio condition of -10 dB, the output SNR was improved by 13.07 dB (from -10.00 dB to 3.07 dB), with the spectral structure of the signal being recovered from an indistinguishable form to a clearly discernible pattern. Compared to traditional denoising methods and existing self-supervised neural network models (such as U-Net, N2N, NerNT, and ONT), AMFNet consistently outperforms in all evaluation metrics.

In conclusion, this study introduces a novel technical pathway for underwater noise extraction in marine engineering, supporting long-distance underwater noise monitoring, construction status evaluation, and quantitative assessment of marine ecological impacts, and demonstrates strong potential for engineering application and promotion.

**Keywords:** Underwater Acoustic Signal Processing; Deep Learning; Self-supervised Learning; Underwater Pile-driving Noise; Underwater Bridge Vibration Noise.

**目录**

[摘要 I](#_Toc196845387)

[Abstract III](#_Toc196845388)

[第一章 引言 1](#_Toc196845389)

[1.1 海洋工程水下噪声提取背景及意义 1](#_Toc196845390)

[1.2 国内外研究现状 4](#_Toc196845391)

[1.3 本文主要工作与技术路线 6](#_Toc196845392)

[第二章 海洋工程水下噪声特性分析 9](#_Toc196845393)

[2.1 海洋工程水下噪声评估参数 9](#_Toc196845394)

[2.1.1 声压级与峰值声压级 9](#_Toc196845395)

[2.1.2 均方根声压级 9](#_Toc196845396)

[2.1.3 声暴露级 10](#_Toc196845397)

[2.2 海洋背景噪声的来源与特性 11](#_Toc196845398)

[2.2.1 背景噪声的分类 11](#_Toc196845399)

[2.2.2 背景噪声特性分析 11](#_Toc196845400)

[2.3 海洋工程水下噪声信号采集系统 13](#_Toc196845401)

[2.3.1 软硬件系统架构 13](#_Toc196845402)

[2.3.2 水听器布放与数据采集 14](#_Toc196845403)

[2.4 打桩工程水下噪声特性分析 17](#_Toc196845404)

[2.4.1 采集环境与布置 17](#_Toc196845405)

[2.4.2 特性分析与讨论 18](#_Toc196845406)

[2.5 跨海桥梁水下噪声特性分析 20](#_Toc196845407)

[2.5.1 采集环境与布置 20](#_Toc196845408)

[2.5.2 特性分析与讨论 21](#_Toc196845409)

[2.6 本章小结 25](#_Toc196845410)

[第三章 自监督学习策略 26](#_Toc196845411)

[3.1 传统降噪方法能力分析 26](#_Toc196845412)

[3.1.1 高斯白噪声场景降噪 26](#_Toc196845413)

[3.1.2 海洋背景噪声场景降噪 28](#_Toc196845414)

[3.2 自监督学习机制 29](#_Toc196845415)

[3.2.1 监督方法的瓶颈 29](#_Toc196845416)

[3.2.2 自监督策略的演化 30](#_Toc196845417)

[3.3 自监督学习策略设计 32](#_Toc196845418)

[3.3.1 训练信号对构造 32](#_Toc196845419)

[3.3.2 频谱连续性缺失问题 34](#_Toc196845420)

[3.4 本章小结 37](#_Toc196845421)

[第四章 自适应多重聚焦降噪网络设计 38](#_Toc196845422)

[4.1 网络设计与训练策略 38](#_Toc196845423)

[4.1.1 AMFNet网络 38](#_Toc196845424)

[4.1.2 自监督训练方案 40](#_Toc196845425)

[4.2 AMFNet架构与关键模块 41](#_Toc196845426)

[4.2.1 编码与解码器结构 42](#_Toc196845427)

[4.2.2 频谱动态聚焦单元（SDFU） 43](#_Toc196845428)

[4.2.3 复数注意力模块（CAM） 46](#_Toc196845429)

[4.2.4 语义重校准单元（SRU） 49](#_Toc196845430)

[4.2.5 复数多尺度跳跃连接单元（CMS-SCU） 53](#_Toc196845431)

[4.3 损失函数设计 55](#_Toc196845432)

[4.3.1 时频损失 55](#_Toc196845433)

[4.3.2 正则化损失 56](#_Toc196845434)

[4.4 本章小结 56](#_Toc196845435)

[第五章 降噪实验与结果分析 58](#_Toc196845436)

[5.1 实验设置 58](#_Toc196845437)

[5.1.1 实验设计内容 58](#_Toc196845438)

[5.1.2 数据集构建与来源 60](#_Toc196845439)

[5.1.3 实验环境与预处理流程 63](#_Toc196845440)

[5.1.4 评价指标 64](#_Toc196845441)

[5.2 实验结果 65](#_Toc196845442)

[5.2.1 累计能量段打桩声降噪实验 65](#_Toc196845443)

[5.2.2 原始打桩声信号降噪实验 70](#_Toc196845444)

[5.2.3 非训练类型声信号降噪实验 74](#_Toc196845445)

[5.3 模型有效性与降噪方法对比分析 76](#_Toc196845446)

[5.3.1 模型有效性对比分析 76](#_Toc196845447)

[5.3.2 模块消融与多方法对比分析 80](#_Toc196845448)

[5.4 本章小结 83](#_Toc196845449)

[第六章 总结与展望 85](#_Toc196845450)

[6.1 论文的主要工作与创新点 85](#_Toc196845451)

[6.2 未来工作展望 86](#_Toc196845452)

[参考文献 87](#_Toc196845453)

[攻读硕士学位期间获得的成果 95](#_Toc196845454)

[致谢 96](#_Toc196845455)

**Contents**

[Abstract in Chinese I](#_Toc196846938)

[Abstract in English III](#_Toc196846939)

[Chapter 1 Introduction​ 1](#_Toc196846940)

[1.1 Background and Significance of Underwater Noise Extraction in Marine Engineering 1](#_Toc196846941)

[1.2 Domestic and International Research Status 4](#_Toc196846942)

[1.3 Key Contributions and Technical Route of This Paper 6](#_Toc196846943)

[Chapter 2 Analysis of Underwater Noise Characteristics in Marine Engineering 9](#_Toc196846944)

[2.1 Evaluation Parameters for Underwater Noise in Marine Engineering 9](#_Toc196846945)

[2.1.1 Sound Pressure Level (SPL) and Peak Sound Pressure Level 9](#_Toc196846946)

[2.1.2 Root Mean Square Sound Pressure Level 9](#_Toc196846947)

[2.1.3 Sound Exposure Level (SEL) 10](#_Toc196846948)

[2.2 Sources and Characteristics of Marine Background Noise 11](#_Toc196846949)

[2.2.1 Classification of Background Noise 11](#_Toc196846950)

[2.2.2 Analysis of Background Noise Characteristics 11](#_Toc196846951)

[2.3 Underwater Noise Signal Acquisition System for Marine Engineering 13](#_Toc196846952)

[2.3.1 Software-Hardware System Architecture​ 13](#_Toc196846953)

[2.3.2 Hydrophone Deployment and Data Acquisition 14](#_Toc196846954)

[2.4 Analysis of Underwater Noise Characteristics in Pile-Driving Engineering 17](#_Toc196846955)

[2.4.1 Acquisition Environment and Configuration 17](#_Toc196846956)

[2.4.2 Characteristic Analysis and Discussion 18](#_Toc196846957)

[2.5 Analysis of Underwater Noise Characteristics in Cross-Sea Bridge Projects 20](#_Toc196846958)

[2.5.1 Acquisition Environment and Configuration 20](#_Toc196846959)

[2.5.2 Characteristic Analysis and Discussion 21](#_Toc196846960)

[2.6 Summary of This Chapter 25](#_Toc196846961)

[Chapter 3 Self-Supervised Learning Strategy 26](#_Toc196846962)

[3.1 Capability Analysis of Traditional Denoising Methods 26](#_Toc196846963)

[3.1.1 Denoising in Gaussian White Noise Scenarios 26](#_Toc196846964)

[3.1.2 Denoising in Marine Background Noise Scenarios 28](#_Toc196846965)

[3.2 Self-Supervised Learning Mechanism 29](#_Toc196846966)

[3.2.1 Limitations of Supervised Methods 29](#_Toc196846967)

[3.2.2 Evolution of Self-Supervised Strategies 30](#_Toc196846968)

[3.3 Design of Self-Supervised Learning Strategy 32](#_Toc196846969)

[3.3.1 Construction of Training Signal Pairs 32](#_Toc196846970)

[3.3.2 Spectral Continuity Missing Issue 34](#_Toc196846971)

[3.4 Summary of This Chapter 37](#_Toc196846972)

[Chapter 4 Design of Adaptive Multi-Focus Denoising Network 38](#_Toc196846973)

[4.1 Network Architecture and Training Strategy 38](#_Toc196846974)

[4.1.1 AMFNet Network 38](#_Toc196846975)

[4.1.2 Self-Supervised Training Scheme 40](#_Toc196846976)

[4.2 AMFNet Architecture and Key Modules 41](#_Toc196846977)

[4.2.1 Encoder and Decoder Structure 42](#_Toc196846978)

[4.2.2 Spectral Dynamic Focusing Unit (SDFU) 43](#_Toc196846979)

[4.2.3 Complex Attention Module (CAM) 46](#_Toc196846980)

[4.2.4 Semantic Re-Calibration Unit (SRU) 49](#_Toc196846981)

[4.2.5 Complex Multi-Scale Skip Connection Unit (CMS-SCU) 53](#_Toc196846982)

[4.3 Loss Function Design 55](#_Toc196846983)

[4.3.1 Time-Frequency Loss 55](#_Toc196846984)

[4.3.2 Regularization Loss 56](#_Toc196846985)

[4.4 Summary of This Chapter 56](#_Toc196846986)

[Chapter 5 Denoising Experiments and Results Analysis 58](#_Toc196846987)

[5.1 Experimental Setup 58](#_Toc196846988)

[5.1.1 Experimental Design Content 58](#_Toc196846989)

[5.1.2 Dataset Construction and Sources 60](#_Toc196846990)

[5.1.3 Experimental Environment and Preprocessing Workflow 63](#_Toc196846991)

[5.1.4 Evaluation Metrics 64](#_Toc196846992)

[5.2 Experimental Results 65](#_Toc196846993)

[5.2.1 Cumulative Energy Segment Pile-Driving Noise Denoising Experiment 65](#_Toc196846994)

[5.2.2 Raw Pile-Driving Noise Signal Denoising Experiment 70](#_Toc196846995)

[5.2.3 Denoising Experiment for Non-Trained Noise Types 74](#_Toc196846996)

[5.3 Model Validity and Comparative Analysis of Denoising Methods 76](#_Toc196846997)

[5.3.1 Model Validity Comparison 76](#_Toc196846998)

[5.3.2 Module Ablation and Multi-Method Comparative Analysis 80](#_Toc196846999)

[5.4 Summary of This Chapter 83](#_Toc196847000)

[Chapter 6 Conclusion and Future Work 85](#_Toc196847001)

[6.1 Key Contributions and Innovations of This Paper 85](#_Toc196847002)

[6.2 Future Research Directions 86](#_Toc196847003)

[References 87](#_Toc196847004)

[Achievements During Master’s Studies 95](#_Toc196847005)

[Acknowledgments 96](#_Toc196847006)

# 第一章 引言

## 1.1 海洋工程水下噪声提取背景及意义

近年来，随着近海基础设施建设的快速推进[1][2]，水下打桩、桥梁施工、隧道掘进等各类海洋工程活动在近岸海域频繁开展，产生了多种类型的水下噪声。其中，水下打桩作业释放的强脉冲型噪声，以及跨海桥梁钢制结构在车辆荷载作用下向水中辐射的弱脉冲型噪声，是其中代表性的两类海洋工程水下噪声声源。这两类噪声信号在特性上各具特点，在声场传播与环境干扰过程中均发挥着重要作用，因而成为海洋声学监测、水下施工状态评估及生态影响研究的重要对象[3]。海洋工程水下噪声通常具有能量密度高、传播距离远、结构特征显著等特点，不仅能够反映工程作业的动态过程，还对水下声场结构及周边海域生态环境产生影响[4]。

如图 1-1所示，在典型的海洋工程水下噪声场景中，近岸施工打桩会辐射出强烈的脉冲噪声（瞬间能量集中且持续时间短），而大型钢制跨海桥梁结构在车辆通行荷载作用下向水体辐射较弱的脉冲噪声，多种声源叠加形成复杂的水下声场。

|  |
| --- |
| 图 1-1 海洋工程水下声事件场景图 |

在现代水声工程应用中，准确提取海洋工程水下噪声信号的时频结构特征并实现对声源的精确识别，是开展声场建模及噪声传播特性分析等后续工作的关键基础。这一能力对于提高远场声传播仿真精度、增强水下目标监测能力和提升海洋生态风险评估水平具有重要意义[5]。图 1-2 直观揭示了海洋工程水下噪声传播过程中的能量递减趋势。只有当脉冲噪声信号的主体结构被清晰提取和正确识别，才能为相关模型的构建与参数拟合提供可靠依据，从而有效提升远程声场预测的准确性。

|  |
| --- |
|  |
| 图 1-2 海洋工程水下噪声衰减模型 |
|  |

在实际海域传播条件下，即便起始声压极高的打桩脉冲噪声，亦会随传播距离的增加而经历明显的几何扩散与频率相关吸收衰减，当其进一步叠加在多源海洋背景噪声之上时，信噪比将持续下降，最终仅以微弱脉冲形式抵达远场。而桥梁振动噪声信号本身幅度有限，属于先天低信噪比信号，更容易被风浪、船舶航迹及海洋生物声等广谱干扰所掩盖。与此同时，水声信号在传播过程中常伴随多径传输、海面与海底反射以及边界散射等复杂效应，这些效应一方面在时域上拉长尾波持续时间，另一方面在频域中引入显著能量扩散，导致原始脉冲结构被严重畸变，主体特征常被延迟尾波所淹没，对远程监测与能量评估带来不确定性[6][7]。

此外，海洋背景噪声本身成因多样，风浪破碎、船舶推进和海洋生物活动等声源叠加，其能量跨越低频至高频波段且随时间剧烈波动。这种宽带、非平稳背景不仅提高了目标信号与噪声的频谱重叠程度，进一步加大了海洋工程水下噪声信号提取与降噪难度[8]。

维纳滤波、小波变换等传统方法在理想条件下具有一定的降噪效果，但当它们应用于海洋工程水下噪声信号提取任务时，却常常出现信号结构被压缩、频谱失真以及尾波残留等问题。主要原因在于传统算法缺乏对水下非平稳信号特征的有效感知，难以聚焦于脉冲信号的主能量段并同时抑制复杂背景噪声的干扰[9][10][11]。

近年来，深度学习技术在语音增强与音频建模领域取得了显著进展，为复杂噪声环境下的信号处理提供了新的思路。然而，传统的有监督深度学习方法高度依赖高质量的纯净样本作为训练标签，而在实际海洋环境中几乎无法获取不含干扰的海洋工程水下噪声信号，这使得此类模型难以直接应用于真实场景[12]。此外，海洋工程水下噪声信号在时频结构上与人类语音等常见训练样本存在显著差异，现有模型难以有效泛化至这类特殊信号[13]。

为突破对海量标注数据的依赖，自监督学习策略被引入信号提取与降噪领域。该类方法通过设计伪造的训练任务，从无标签数据中自驱动学习有意义的特征表示，已在图像复原和语音增强等任务中展现出初步成效[14]。自监督方法的优势在于具有良好的迁移能力和工程适应性，特别适用于处理海洋环境下水声信号数据缺乏的情况。但现有自监督降噪方法主要针对结构相对平稳的信号[15]，对于具有瞬态脉冲特征且伴随显著尾波混叠的海洋工程水下噪声信号而言，在保持频谱连续性和对齐信号主体结构方面仍存在明显不足[16]。

综上所述，在缺乏干净训练样本的背景下，亟需构建一种兼顾信号结构保持与频谱结构一致性的自监督降噪模型，以提升在低信噪比情况下海洋工程水下噪声信号的结构还原能力，并为复杂海洋环境下的水声信号处理提供关键技术支撑。例如，在跨海桥梁的施工与运营过程中，其结构振动向水中辐射的中低频噪声可作为声场建模与远程声源识别的重要依据。这类噪声有助于判断工程的运行状态和工况，分析声能传播情况，并为海洋生态环境影响评估提供重要参考。然而，此类弱脉冲信号往往被复杂背景噪声和尾波效应所淹没，传统方法难以恢复其完整的结构和能量特征[17]。目前，国内外对于上述打桩强脉冲和桥梁弱脉冲噪声信号的针对性提取研究仍然十分有限，在这一方向上的探索具有一定的创新性与开拓意义。为此，本研究提出了一种自监督式海洋工程水下噪声提取方法，能够在无需干净样本的条件下完成模型训练，实现对目标噪声信号的高保真提取，从而为跨海桥梁等水下工程场景的声场计算、工程状态监测和海洋生态影响评估提供有力支持。

## 1.2 国内外研究现状

随着海洋工程建设的推进及环保需求的增长，水声信号处理技术在水下监测、声源定位、生态评估等领域的应用日益深入，成为声学、人工智能与信号处理交叉的研究热点[18][19]。海洋工程水下噪声信号是人类海洋活动的重要声学产物，具有能量集中、瞬时性强、脉冲式等特点。但在复杂海洋环境下，传播路径不确定、水体扰动非平稳、边界反射增强以及多源背景噪声干扰等因素易使该类信号的时频结构特征被掩盖或扭曲[20][21]。如何在非理想声学环境下准确提取海洋工程水下噪声的主结构成分，已成为声源估计与声场计算等关键环节亟待突破的技术瓶颈[5][8]。

在早期研究中，传统信号处理方法如频域滤波、谱减法、时频重构等被广泛用于水下声信号的降噪处理。其中，维纳滤波作为经典方法，基于最小均方误差准则进行目标信号的功率谱估计，并利用统计特性构建滤波器实现信噪分离[22]。尽管该方法在稳定背景噪声环境下具有较好表现，但难以应对海洋工程水下噪声信号处理场景下的非平稳、多源海洋背景干扰。在频谱重叠、目标信号微弱或干扰强烈等情况下，维纳滤波容易削弱主频成分并丢失瞬态细节，进而影响结构完整性[23]。此外，小波变换等多尺度分析方法虽然具备一定的局部瞬态特征提取能力，但其基函数选取与频带划分存在刚性限制，难以灵活适配水下复杂场景中频谱结构的剧烈变化[24][25]。因此，在尾波混响和非结构冗余干扰普遍存在的情况下，传统方法难以同时保留信号主结构并抑制无效噪声成分，面临明显的性能瓶颈[26][27]。

近年来，深度学习在语音增强与音频建模领域取得突破性进展，尤其是在卷积神经网络（CNN）[28][29][30]、循环神经网络（RNN）[31][32]和Transformer[33][34]等模型的驱动下，实现了对语音信号的自适应建模与特征提取。在此背景下，一些研究开始尝试将深度学习方法引入水下声信号的降噪与识别任务中，通过将高信噪比信号近似为干净数据训练网络模型。例如，Zhou 等人[35]的噪声感知深度学习模型利用噪声先验实现了复杂水下噪声的自适应抑制；Zhou 等人[36]设计的DBSA-net利用双分支自注意力结构从不同尺度提取特征来有效提升水下声信号的降噪性能。此外，Song 等人[37]提出的双路径循环神经网络充分利用局部与全局信息，实现了对动态噪声的精准建模；同时，Zhou 等人[38]将注意力机制与多尺度卷积相结合，在复杂水下环境中显著改善了信号的降噪效果。

尽管上述方法在实验条件下取得了一定成果，但其普遍依赖于大量干净参考信号构建训练集，具有较强的监督学习依赖。对于自然语音信号，此类配对样本可以通过清晰语音叠加标准噪声合成获得，但在真实水下环境中，由于声源不稳定、传播路径复杂等因素，采集无背景干扰的水声信号几乎不具可行性，导致标注数据极度匮乏且获取成本高昂，严重限制了监督式模型的泛化能力与应用范围。

然而，当前公开的水下声学数据集也多集中于舰船声、环境背景声或生物声领域，难以涵盖水下工程应用所需的脉冲声信号。例如，ShipsEar 数据集[39]收录了多种类型舰船的水下辐射噪声，广泛应用于相关研究。通常，在降噪提取任务中，研究者将采集到的舰船辐射噪声视为信噪比较高、接近干净参考的目标信号，通过与海洋背景噪声混叠，构建用于模型训练的信号对[35][36][37][40]。此外，该数据集也被用于船舶目标识别等分类任务[38][41][42][43]，支持多场景下的水下声学研究；DeepShip 数据集[44]汇聚了大规模不同类型的舰船噪声样本，主要用于舰船目标的分类任务[38][41][42][43][45]；Watkins 海洋哺乳动物声库[46]收录了多种典型海洋哺乳动物的声信号，支持生物声学与生态监测；国内千岛 Ear22 数据集[47]则以多目标船舶噪声为主，支持噪声源分类检测与定位任务。

相比之下，水下打桩、桥梁振动等典型海洋工程水下噪声通常呈现典型的脉冲式结构特征，具有频谱宽、时域瞬态强、能量高度集中的特征，且易受多径反射、边界混响与远场折射影响，导致其结构在接收端高度畸变[48][49][50]。在实际采集过程中，此类信号的结构特征易被回波、边界反射及远场干扰所扭曲或淹没，高质量标注样本几乎无法获取。因此，现有公开数据集难以支撑海洋工程水下噪声领域的模型训练与降噪算法评估，使得传统监督模型难以开展有效训练。由此可见，水下工程信号的建模与降噪任务亟需发展摆脱对高质量标注数据依赖的其他方法。

在此背景下，自监督学习利用无标签数据构建训练目标，逐渐成为一种重要策略[51][52]。其核心思想是利用数据内部的关联构造伪监督任务，引导模型在无标签条件下学习具有语义或结构意义的特征表示[53]。对于能量集中且尾波混叠显著的水下脉冲信号来说，其独特的时频结构为自监督目标的构建提供了天然基础。目前已有一系列自监督音频降噪方法应用于语音增强场景，如NerNT[54]提出邻域扰动策略，增强信号结构对比学习能力；Noise2Noise(N2N)[55]方法通过两次独立噪声观测构建训练对，避免了对干净标签的依赖；ONT[56]则基于时序特征扰动，构造伪监督路径学习目标信号的短时动态特征。这些方法在无标签条件下取得了良好的语音清晰度提升效果。

但现有自监督方法多针对于人类语音语义建模，优化目标以听感提升为主，对于无语义、非周期的水下工程脉冲信号则面临诸多适应性问题。首先，工程信号在时频分布上具有高度瞬态性与脉冲主导性，易被子采样或差分处理后出现频谱连续性断裂，影响结构性建模；其次，尾波混响与冗余干扰较强，极易诱导网络过拟合噪声或丢失有效结构特征，现有方法在主结构保持与噪声抑制之间难以平衡。尤其在低信噪比条件下，传统自监督方法易导致特征弱化和频带压缩，亟需从结构感知、多尺度融合等方向进行机制优化。

综上，目前水下工程声信号降噪任务主要面临三大技术瓶颈：第一，复杂非平稳噪声场使传统静态滤波模型降噪性能大幅下降，难以适应动态场景中的冗余抑制需求；第二，受多径效应与传播介质等影响，结构性脉冲信号在传播中易遭破坏，信号主能量段与尾波部分难以准确区分与还原，严重影响降噪后信号的建模性；第三，缺乏针对该类信号的高质量标注数据集，限制了深度模型的训练稳定性与性能上限。因此，发展兼具结构感知能力与无标签建模能力的自监督网络模型，成为突破当前工程场景下水下噪声提取难题的重要方向。

## 1.3 本文主要工作与技术路线

针对水下工程声信号在复杂背景噪声干扰下，结构信息易被掩盖的问题，本文旨在提升信号的时频结构保持能力和复杂环境适应性，设计并实现了一种基于自监督学习的水下工程声信号降噪方法。研究工作从声信号特性分析入手，构建了涵盖理论分析、模型设计、系统实现与实验验证在内的完整技术路线，如图 1-3 所示。在这一技术路线指导下，本文围绕研究背景与意义阐述、噪声特性分析、自监督学习策略设计、神经网络模型构建及实验验证等方面展开了研究，主要工作如下：

第一章介绍了本研究的背景与意义，同时系统回顾了国内外水下声学降噪领域的研究进展。在总结现有方法在非平稳背景噪声干扰下对信号结构还原不足、工程应用泛化能力有限等问题的基础上，提出了以保持信号主要结构特征为核心的海洋工程水下噪声提取降噪思路，为后续研究奠定了理论基础。

第二章按照“指标体系—背景噪声—信号采集—特性分析”的逻辑顺序展开。首先，建立了包括峰值声压级、均方根声压级和声暴露级的量化指标体系，并分析了海洋背景噪声的来源与频谱特征。随后，本章设计并搭建了海洋工程水下噪声信号采集系统，采集高质量的声信号数据以支持后续实验。在此基础上，对典型的海洋工程水下噪声信号（如打桩声和桥梁振动声）进行了时域和频域结构特性的对比分析，为第三章评估经典降噪算法并引入自监督策略奠定了数据基础。

第三章围绕水声信号处理任务中存在的实际问题，系统设计了适应复杂海洋环境条件的自监督学习方案。针对缺乏干净参考信号、公开数据集稀少的现实困难，本研究首先评估了传统降噪方法在典型水下噪声场景下的性能表现，分析了现有监督学习策略在工程应用中的局限性，明确提出了引入自监督学习框架的必要性。在此基础上，进一步设计了基于含噪数据的训练信号对构造方法，并分析了降噪过程中子采样操作引起的频谱连续性缺失问题，为后续网络模型设计与优化提供了理论依据与实验支撑。

第四章完成了自适应多重聚焦降噪网络（AMFNet）的设计。基于第三章提出的自监督学习方案，本研究采用含噪数据子采样构造训练信号对，在不依赖干净参考样本的条件下完成了AMFNet网络的自监督训练。网络整体采用U型编码与解码主干结构，并在关键阶段引入频谱动态聚焦单元、复数域注意力模块与语义特征重校准单元，以强化对海洋工程水下噪声信号局部结构与全局特征的联合建模能力。上述网络结构与训练策略的设计，为后续章节的实验验证与方法对比提供了核心技术支撑。

第五章围绕所提方法的性能验证构建了相应的数据集并设计了多组实验。首先，根据声暴露级理论构造了“累计能量段”打桩声训练集，用于评估模型在结构约束条件下对信号的重构能力。随后，将训练好的模型应用于未经截断的原始打桩声和未参与训练的桥梁振动声样本，以测试模型面对非理想输入时的适应能力与稳健性。实验设置了多组不同噪声叠加条件和信噪比参数，并使用时域波形、时频谱图和帧级信噪比曲线等可视化手段对结果进行对比分析，同时以信噪比、峰值信噪比和均方根误差等指标对模型降噪性能进行定量评估。最后，本章还引入传统降噪方法（如维纳滤波、小波变换）以及其他自监督深度学习模型与所提方法进行了对比，并通过模块消融实验验证了所提模型在降噪效果和鲁棒性方面的有效性。

第六章对全文的研究工作进行了总结，并展望了未来的研究方向。

|  |
| --- |
|  |
| 图1-3 技术路线图 |

# 第二章 海洋工程水下噪声特性分析

## 2.1 海洋工程水下噪声评估参数

描述海洋工程水下噪声的声学参数主要包括：声压级、峰值声压级、均方根声压级以及声暴露级，对于来自水下打桩、桥梁振动声等这类宽带脉冲型海洋工程水下噪声，通常需要重点关注峰值声压级[3]、均方根声压级[4]与声暴露级[5]这三个表征参数。下面分别对这些关键评价参数进行定义和说明。

### 2.1.1 声压级与峰值声压级

声波在介质传播中引起微小压强变化，这种随时间波动的压力差称为声压（）。海洋环境中对声压的测量通常以帕斯卡（Pa）或微帕（μPa）为单位。声压的最大瞬时绝对值称为峰值声压（），它反映了声压波动的极值幅度，在强声源情况下通常可达数个数量级。为了便于量化和比较声信号的强度，常采用对数刻度定义的声压级（）进进行表征，其单位为分贝（dB），数学定义如下：

|  |  |  |
| --- | --- | --- |
|  |  | (2-1) |

其中， 为参考声压值，在水中通常取1 μPa，在空气中则取20 μPa。如果将峰值声压 代入该公式，即可得到峰值声压级（），用于描述声信号的瞬时最大声压。

### 2.1.2 均方根声压级

在声信号分析中，均方根声压级 是声信号均方根声压的对数表示，单位为分贝（dB），其数学定义如下：

|  |  |  |
| --- | --- | --- |
|  |  | (2-2) |

其中， 为声压的均方根值，定义如下：

|  |  |  |
| --- | --- | --- |
|  |  | (2-3) |

其中， 和 分别表示所取波形的起始和终止时间点，单位为秒（s）， 为参考声压值，通常取1 μPa（水中）。 的单位为dB。

在脉冲声的均方根声压级计算中，通常需要剔除信号两端能量较低部分，以更准确地反映信号的主要能量贡献。常用的方法是取信号总能量的90%进行统计，即剔除波形起始和末尾各约5%能量较低的部分。采用这种方式得到的均方根声压级能够更加准确地表征信号的有效声压水平，从而提高不同信号分析结果的稳定性和可比性[3][4][5]。

### 2.1.3 声暴露级

声暴露级（Sound Exposure Level, ）用于衡量单次或离散水下声事件的累计声能量，其数学定义如下：

|  |  |  |
| --- | --- | --- |
|  |  | (2-5) |

其中， 为参考时间（通常取1秒）， 和 为声暴露事件的起始和终止时刻， 为声压信号， 为参考声压（水下环境通常取1 μPa）。为了更准确地体现信号的实际能量分布，脉冲声的声暴露级计算常只考虑累计能量的主要持续时间段（例如90%累计能量所对应的持续时间）。通过这种方式，将不同时长的信号等效为相同累计能量的参考时间段信号，能够更客观地评估水下声事件对海洋环境可能产生的影响。在针对瞬态或间歇性高强度声源（如海上打桩作业等）的研究中，这一指标具有重要意义。

在环境声学与生态保护研究中，声压级、均方根声压级与声暴露级不仅用于衡量工程声源的物理强度特征，还为海洋生物声学影响评估提供了基础量化指标。尤其是在评估海洋工程活动对鲸类、鱼类等水生生物可能造成的行为干扰与生理伤害时， 常作为判定门槛之一[8]。例如，国际海洋哺乳动物保护组织制定的行为阈值标准中，即基于 值区分可接受暴露剂量与潜在致伤区间。通过准确计算短时或累积的声暴露级，可有效评估打桩声、振动声等工程活动是否超出敏感物种的听觉容忍范围，进而作为工程选址优化、施工时间避让、生物保护缓冲区划定等管控手段的科学依据[3][4]。因此，这些声学评估参数不仅在信号处理领域具有重要技术价值，更在推进水下工程生态友好性与环境监管制度化方面发挥着关键作用[5]。

## 2.2 海洋背景噪声的来源与特性

### 2.2.1 背景噪声的分类

海洋背景噪声广泛存在于水下探测与信号处理领域，其复杂特性给目标信号的提取与降噪带来了严峻挑战[57]。从产生机制角度，可将海洋背景噪声分为自然背景噪声和人为背景噪声两大类。

自然背景噪声源于海洋环境中的物理现象和生物活动，其中风浪噪声尤为常见。风作用在海面，形成波浪运动，并引发大量气泡形成与破裂；随着风速的增加，这类噪声能量显著提升，呈现连续且宽频带的分布特征[58]。此外，海洋生物的发声活动（如鲸类、鱼类、海豚及甲壳类动物等）也会在不同频段产生明显的噪声信号——这些生物声既是海洋生态系统内的重要交流方式，也使得水下声学环境更加复杂[59]。此外，海洋湍流、水体流动及海底地质活动也可产生背景噪声，进一步加剧了海洋噪声环境的多变性[57]。

人为背景噪声主要源自船舶活动产生的噪声，包括船体机械振动、发动机运转以及螺旋桨空化等，具有明显的周期性和结构化特征，对海洋声环境及水下探测系统造成持续且深远的影响[60]。总体而言，自然噪声和人为噪声在形成机制、空间分布及对环境因素的敏感度等方面存在显著差异：自然噪声广泛且相对均匀地充斥海洋，各种噪声源受气象、水文和季节等条件影响而呈现动态变化；而人为噪声则更具局域性，强度与人类活动的分布紧密相关，通常表现出更为明确的结构性和时空分布特征[61]。

### 2.2.2 背景噪声特性分析

如图 2-1 所示，海洋背景噪声在不同频率范围内表现出显著的频谱差异。这既与噪声来源的多样性密切相关，也深刻影响着水下声信号的传播机制和降噪策略的设计。通常根据频率范围可将海洋背景噪声划分为低频（<100 Hz）、中频（100 Hz 至 1 kHz）和高频（>1 kHz）三个典型频段，每个频段内的主导噪声来源及其与环境因素的耦合方式各不相同。

|  |
| --- |
| 图 2-1 厦门某海域背景噪声功率谱密度 |

在低频段，船舶噪声和海洋水体运动是主要成因，尤以螺旋桨空化现象产生的噪声最为突出，主频范围为30 Hz 至 100 Hz[62]。低频噪声具有衰减速率慢、传播距离远等特点，因此即使在远离人类活动区域的海洋环境中，仍能检测到明显的船舶航行噪声。Cui等[63]对曾母盆地的长期实测数据研究表明，在低于200 Hz 的频段内，船舶噪声始终占据主导地位；相比之下，风浪噪声在此频域内的能量贡献明显偏低。

中频段（100 Hz 至 1 kHz）中，风浪噪声与近距离的船舶噪声占据主导。于浅海环境中，海面风力形成的波浪和气泡运动会增强该频段内的风浪噪声水平。研究发现，随着频率的升高，船舶噪声的影响逐渐减弱，而风浪噪声显著提升，特别是在400 Hz 以上频率范围内，风浪噪声主导背景能量[63]。

高频段（>1 kHz）的背景噪声主要源于降雨、生物活动以及水体湍流等因素。其中，降雨噪声既具备频谱上的集中性，也呈现出明显的瞬态波动特征。Meng等人[64]在浅海环境中对10 kHz 至 35 kHz 范围的高频背景噪声进行了长时间观测，发现其能量波动特征与环境状态高度相关——例如降雨强度和风速的变化都会引起高频噪声能量的显著波动。此外，即便是在植被覆盖情况各异的浅水海域，高频背景噪声的整体频谱变化趋势仍表现出相对一致的规律，这进一步表明高频噪声在各类浅水生态系统中具有较强的普适性[65]。

## 2.3 海洋工程水下噪声信号采集系统

### 2.3.1 软硬件系统架构

本研究构建的水下噪声信号采集系统由水听器、数据采集卡及笔记本电脑组成，用于高精度获取和分析水下工程声信号。系统的核心部件包括：水听器用于声信号接收，数据采集卡用于模拟信号的模数转换，笔记本电脑用于数据记录和处理。

本系统选用丹麦 Brüel & Kjær（B&K）公司生产的8105型水听器，内置球形压电陶瓷敏感元件，具有优良的全指向性，涵盖了绝大多数水下工程声源的频段，各项技术参数均满足采集工作需求，具体规格详见表2-1。

表 2-1 B&K 8105型水听器主要参数

|  |  |
| --- | --- |
| 项目 | 参数 |
| 水听器型号 | 8105 |
| 声压灵敏度 | -85 dBV/Pa |
| 电容 | 7250 pF |
| 频率响应 | 0.1 Hz 至 160 kHz: +3.5/-10.0 dB |
| 垂直方向性 | 0°至270° @ 80 kHz ±2.5 dB @ 100 kHz |
| 灵敏度随温度变化值 | 0 to -0.03 dB/°C |
| 灵敏度随静压力变化值 | 0 to -3×10⁻³ dB/Pa (0 to -0.03 dB/atm.) |
| 最大静水压力 | 9.8 MPa (1000 m 水深) |

为确保水下声信号的高精度采集与分析，本系统采用美国国家仪器（NI）公司生产的NI USB-4431数据采集卡，专为高精度声音与振动测量设计，采样率范围为1 S/s 至102.4 kS/s，动态范围达100 dB，确保高分辨率信号获取。同时，其模拟输出通道可与输入通道同步，更新速率达96 kS/s，动态范围为89 dB。该设备的高性能保障了水下声信号采集的准确性与稳定性，为后续数据分析提供可靠支持。采集程序使用NI公司推出的LabView进行编写，实现数据的自动采集并实时观察采集信号波形，如图 2-2 所示。

|  |
| --- |
| 图 2-2 水下工程声信号采集系统 LabView 控制面板 |

### 2.3.2 水听器布放与数据采集

如图 2-3 所示，为多维度分析水下工程声信号的传播特性，采集工作依托调查船（或海上作业平台）作为载体，在不同水深位置垂直布设多层水听器，并通过高精度数据采集系统对水下声信号进行实时记录。图2-3（a）给出了数据采集系统的组成示意，图2-3（b）则显示了现场水听器的布放方式。

|  |
| --- |
| (a)    (b)  图 2-3 采集水听器布放示意图：（a）数据采集系统；（b）现场水听器布放 |

在实际数据采集过程中，调查船（或平台）搭载采集设备，负责水听器的布放与回收。为最大程度避免船只噪声干扰，采集期间关闭了作业船的机械设备并抛锚定位，以尽可能消除船舶动力系统产生的噪声干扰，使测量环境接近“静音”状态。水听器采用分层布设的方案，即在不同水深位置分别安放表层水听器、中层水听器和底层水听器，以同步获取不同深度处的水下声信号。所有水听器均通过电缆连接至船上的数据采集系统，实现信号的实时传输与记录。另外，如图 2-3（b）所示，为保证水听器阵列在水中的垂直稳定性，在阵列底端悬挂铅坠以抗衡水流扰动，确保水听器始终保持稳定姿态，避免因晃动导致的数据失真。

|  |  |
| --- | --- |
| (a) | (b) |
| (c) | 55f2e1911dd4ef7ebe19775141ceb28  (d) |
| 图 2-4 海洋工程水下噪声信号采集场景：（a） 打桩船；（b） 打桩平台；（c） 跨海桥梁；（d） 数据采集设备 | |

图 2-4 展示了海洋工程水下噪声现场采集的场景。在打桩噪声的采集过程中，通过采取一系列措施来保证数据质量和信噪比。首先，作业船在采集期间关闭动力装置并锚泊定位，从声源环境上减少了船舶噪声的干扰，以获取高信噪比、结构清晰的打桩声信号，同时有效避开机械振动干扰和信号传输畸变。对于桥梁振动噪声的采集，则选择在背景安静的时段进行，并确保在采集期间无其他船舶经过，最大程度降低了外界因素对数据采集的影响。

综上，得益于高规格的软硬件配置和研究的采集方案设计，本系统所采集的海洋工程水下噪声具有良好的稳定性，为后续的特性分析奠定了坚实基础。

## 2.4 打桩工程水下噪声特性分析

### 2.4.1 采集环境与布置

图 2-5 为厦门某海域的工程打桩现场。在该施工过程中采用直径为2.1 m 至3.0 m 的钢管复合桩。当钢管桩在遇到坚硬海底地层时无法继续贯入，需借助大型液压冲击锤进行冲击打桩，以继续贯入桩基。冲击打桩过程中产生的强冲击噪声是施工区域主要的声信号来源[66]，也是打桩工程水下噪声的最突出特征声源[67]。

|  |
| --- |
| 图 2-5 厦门某海域打桩现场图 |

如图 2-6 所示，打桩工程声信号使用固定式与走船式方法进行采集并根据水深分别布设表层、中层和底层的水听器：固定式采集是在打桩平台上选取固定点位采集；走船式采集则通过调查船在打桩过程中移动，并抛锚固定至不同距离处的进行声信号采集。

|  |
| --- |
| 图 2-6 打桩工程水下声信号采集布置 |

### 2.4.2 特性分析与讨论

为全面揭示打桩工程水下噪声的传播特性，本节选取距离打桩平台处的固定测点采集的声信号，从时域波形结构、时频能量分布特征及频带能量分布规律三方面进行了深入分析。

图 2-7 所示为信号的时域波形。信号具有典型瞬态脉冲特性：每次脉冲在极短时间内迅速达到峰值，随后快速衰减。主要脉冲之后伴随明显的尾波振荡现象，呈现明显的多径效应。

|  |
| --- |
| 图 2-7 打桩工程水下噪声信号时域图 |

图 2-8 给出了该信号的时频谱图。可以看出，打桩声在脉冲发生时刻的能量呈宽频带分布；在脉冲初始阶段，频谱中出现了明显的频率扩展现象，能量向高频方向延伸。随后，在较低频率范围内仍保持着时间较长的余振能量，对应于时域中观测到的尾波振荡和多径效应。

|  |
| --- |
|  |
| 图 2-8 打桩工程水下噪声信号时频谱图 |

为进一步分析打桩噪声的能量分布特性，我们记录了同一测点在未施工时段的环境背景噪声，并将其作为对照。图 2-9 展示了打桩噪声信号与海洋背景噪声的功率谱密度（PSD）与三分之一倍频程频带声压级图。

|  |  |
| --- | --- |
|  | |
| (a) | (b) |
| 图 2-9 打桩平台测得打桩声信号能量分布情况：（a）功率谱密度图；（b）三分之一倍频程频带声压级图 | |

如图 2-9（a）所示，打桩噪声信号在整个观测频率范围内（尤其是200 Hz 至1000 Hz 的中频段），打桩噪声信号的PSD显著高于背景噪声水平。其中在约400 Hz 处达到峰值，约为136 dB/Hz，体现出打桩噪声明显的中低频主导特性。随着频率的升高，打桩声与背景噪声的差异逐渐减小，说明打桩噪声的高频成分衰减较快。

为量化不同频带内的能量差异，图2-9（b）给出了打桩噪声与背景噪声的一组三分之一倍频程频带声压级比较结果。可以看出，打桩噪声的能量主要集中在4 kHz 以下频段，且在各个频带上均明显高于背景噪声。由于海洋背景噪声受水体运动的影响，其能量主要分布于中低频范围，因此在低频段打桩噪声与背景噪声的差异并不显著。而在约400 Hz 附近频带，二者差值达到最大值，约为40 dB；继续提高频率，二者差异趋于稳定，维持在约25 dB 左右。

综上所述，水下打桩噪声信号呈现出显著的瞬态脉冲特性和多径传播效应：时域上具有明显的尾波振荡，频域能量主要集中在中低频段，在中频范围（100 Hz～1000 Hz）与背景噪声水平存在显著差异。这些特征的深入分析为后续水下海洋工程水下噪声提取方法设计提供了重要的理论基础。

## 2.5 跨海桥梁水下噪声特性分析

### 2.5.1 采集环境与布置

本研究选取厦门某跨海大桥作为桥梁振动水下噪声信号采集的目标。如图2-10 所示，桥梁为多桥墩多跨结构，桥面设置有钢制伸缩结构。当车辆行驶在跨海大桥上时，桥面和主梁产生的振动会通过桥塔传递至水下，从而辐射出低频声信号[73]。当桥面或梁体上的钢制构件（例如伸缩缝、支座等）在车辆荷载作用下发生振动时，这些振动能量沿着主缆、吊索或塔柱等路径向下传递，最终汇集至水下的基础构件，在界面处引起周围水介质的振荡并向外辐射声波[74]。

|  |
| --- |
| 图 2-10 厦门某跨海大桥 |

如图 2-11（a）所示，采用走船式布放的采集方案，在桥梁附近水下1 m、4 m以及接近河床处（约7～8 m）的深度分别布放水听器进行同步采集。如图 2-11（b）所示，设置了A、B、C三个测点，分别距最近桥墩约32 m、34 m和32 m，其中B点位于桥梁钢结构件正下方。

|  |  |
| --- | --- |
| (a) | (b) |
| 图 2-11 采集布置与采集点位示意图：（a）走船式采集布置；（b）跨海大桥桥墩及采集点A、B、C位置 | |

### 2.5.2 特性分析与讨论

如表 2-2 所示，各测点在不同水深下测得的桥梁振动水下噪声峰值声压级、均方根声压级和声暴露级如下：测点A的三项指标平均值分别为155.2 dB、120.4 dB和125.5 dB；测点B的平均值分别为158.5 dB、127.5 dB和133.0 dB；测点C的平均值分别为157.7 dB、122.9 dB和126.3 dB。

表 2-2 各测点不同深度下三类声压级

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 测点 | 水深(m) | 峰值声压级(dB) | 均方根声压级(dB) | 声暴露级(dB) |
| A | 2 | 156.4 | 116.6 | 119.6 |
| 4 | 160.0 | **122.9** | 126.9 |
| 8 | 149.3 | 121.7 | **129.9** |
| B | 2 | 155.4 | 125.6 | 130.7 |
| 4 | 160.0 | **129.5** | **136.1** |
| 8 | 160.0 | 127.3 | 132.3 |
| C | 2 | 159.9 | 119.9 | 123.3 |
| 4 | 156.4 | 121.2 | 124.5 |
| 7 | 156.7 | **127.7** | **131.1** |

如图 2-12 给出了测点B（位于桥梁钢结构正下方，距桥墩约34 m）在4 m水深处桥梁振动产生的水下噪声信号时域波形。可以看出，该信号在时域上表现出显著的非平稳特性：总体背景噪声水平相对稳定，但在个别时刻出现了突出的短时脉冲峰值。在约58 s 处捕捉到数次显著的脉冲事件（红色方框突出显示了1 s时间窗内声压的变化情况），其瞬时声压峰值达到约28 Pa（约149 dB re 1 µPa）。相比之下，非脉冲信号在时域波形中往往难以分辨，其峰值声压不足5 Pa（约134 dB）。特别是在大型车辆通过时，引起的桥墩和地基振动会向水下辐射声波[76]。

|  |
| --- |
| 图 2-12 桥梁振动水下噪声信号时域图 |

如图 2-13 所示，该时段声信号的频谱进一步展示了其频域特征。图中红框部分对应于约58 s处多次强脉冲信号，引起了频谱的明显变化。这些脉冲信号在频域上主要集中于1 kHz 以下的中低频范围，呈现出典型的短脉冲“柱状”频谱特征。整体来看，环境背景噪声主要分布在中低频段，表现较为平稳，但其中叠加了间歇出现的高能短时脉冲成分。结合表 2-4 的数据可知，位于桥梁钢结构下方的B测点，其峰值声压级较均方根声压级高出约31 dB，进一步说明该段声信号中包含强烈的瞬态脉冲分量。

|  |
| --- |
|  |
| 图 2-13 桥梁振动水下噪声信号时频谱图 |

为进一步揭示桥梁振动水下噪声信号的频谱组成，我们在远离桥梁的海域（距离大桥约680 m，离岸最近约784 m，采集过程中无其它船舶经过）采集水下背景噪声，作为对比分析用的环境基线。图 2-14 展示了在4 m 水深处桥梁振动噪声信号与背景噪声的PSD曲线及三分之一倍频程声压级比较结果。

|  |  |
| --- | --- |
| (a) | (b) |
| 图 2-14 桥梁振动噪声信号能量分布情况：（a）功率谱密度图；（b）三分之一倍频程频带声压级图 | |

由图 2-14（a）可见，桥梁振动噪声信号的能量主要集中在1 kHz 以下的中低频段。在约100 Hz 左右的频率范围内，PSD曲线上出现了显著峰值，对应的声压级超过110 dB。随着频率升高，信号能量迅速衰减：当频率高于100 Hz 时，谱级已降至100 dB 以下；超过500 Hz 后进一步降至90 dB 以下；1 kHz 以上频率成分普遍低于85 dB，而4 kHz 以上频段信号强度已接近背景噪声水平（低于80 dB）。

如图 2-14（b）所示，为桥梁振动噪声信号与背景噪声的三分之一倍频程频带声压级对比结果。在1 kHz 以下的中低频段，两者声压级存在显著差异，尤其在约500 Hz 频带处达到最大差值，约25 dB。这一频谱特征表明，桥梁振动噪声信号的能量主要集中在中低频范围；随着频率进一步升高，其声压级逐渐趋近于周围海域的背景噪声水平。

桥梁振动噪声信号以中低频成分为主的特性与其传播机制密切相关。此类声源主要来自车辆荷载作用引发的桥梁结构振动，并通过钢制结构单元与水体的耦合辐射到水下，形成以低频为主的水下声场。尤其对于钢结构桥梁，由于其结构刚性强、传导性好，当车辆高速通过时，桥体结构中容易激发出强烈的低频结构波，并在与水体耦合的过程中转化为水下声波，形成持续或间歇的声辐射。

## 2.6 本章小结

本章从量化度量、环境噪声机理、数据采集体系以及典型海洋工程水下噪声信号结构四个维度，系统阐述了海洋工程水下噪声研究的基础问题，旨在为后续提取算法设计与实验评估奠定统一而可靠的技术基线。

首先，通过介绍了海洋工程水下噪声的重要量化评估指标，包括峰值声压级 、均方根声压级 与声暴露级 。这些评估参数分别从瞬时能量峰值、时均能量和累积能量的角度对海洋工程水下噪声信号进行了多维度衡量，为后续的信号分析与实验应用提供了重要的理论依据和工具支撑。

随后，结合海洋背景噪声的成因分析，揭示了低频段、中频段以及高频段不同的谱域规律，这不仅阐明了海洋工程水下噪声信号易被中低频背景掩蔽的本质原因，也为后续网络模型的设计提供了理论依据。

在数据获取层面，本章给出了由B&K 8105宽带水听器与NI USB‑4431高分辨率采集卡构成的海洋工程水下噪声信号采集系统，为后续的特性分析以及深度学习模型的训练提供了高质量数据源。

进一步的声信号特性分析表明，打桩噪声信号呈脉冲结构且集中于100 Hz至 1 kHz 中频段，多径效应突出；桥梁振动噪声信号则以20 Hz 至 1 kHz 中低频为主，表现为间歇脉冲与稳态混合模式。两类信号虽与海洋背景噪声存在频谱分离窗口，但尾波拖长和低频冗余或许会导致传统静态滤波难以完成有效信号提取，在下一章进行进一步尝试验证。

综上，本章按照了“指标体系—背景噪声—信号采集—特性分析”的逻辑链条，系统刻画了海洋工程噪声信号在时域/频域/能量三维度的本质特征。下一章将在此基础上对经典降噪方法的局限性进行评估，并引入自监督深度学习思路，为面向频谱结构保持的海洋工程水下噪声提取算法研究奠定理论与实验基础。

# 第三章 自监督学习策略

## 3.1 传统降噪方法能力分析

为系统评估传统降噪方法在海洋工程水下噪声提取任务中的实际适用性与降噪提取能力，本节选取实际采集的水下打桩噪声信号作为参考信号（如图 3‑1 所示，信号采集自打桩平台，经幅值缩小处理）。该信号具备显著的瞬态脉冲特征与能量集中特性，同时尾波结构清晰，具备典型工程脉冲声信号特征。

将目标信号与噪声按其均方根能量归一化，使二者能量相等（对应SNR = 0 dB），后将处理后的高斯白噪声和实测海洋背景噪声分别叠加至目标信号，以构建两种含噪场景，模拟复杂水下环境中的实际干扰。随后，分别采用两种经典降噪方法进行对比处理：其一为基于最小均方误差准则构建的线性估计方法——维纳滤波，适用于平稳背景干扰条件下的频谱增强；其二为具备多尺度分析能力的小波变换方法，适用于非平稳条件下瞬态结构的提取与恢复。两者结合为后续算法性能验证提供了对比基准与参考框架。

|  |
| --- |
| 图 3-1 水下打桩噪声信号 |

### 3.1.1 高斯白噪声场景降噪

图 3-2 展示了在信噪比SNR = 0条件下，参考信号混叠高斯白噪声后的降噪结果。其中，图 3-2（a）为叠加高斯白噪声后的波形，原始信号结构受到显著干扰；图 3-2（b）为维纳滤波处理结果。作为基于统计特性的自适应滤波器，维纳滤波能够有效建模平稳噪声，主脉冲区域得到初步恢复，但尾波干扰未能充分抑制，仍存在结构模糊；图 3-2（c）为小波变换的降噪结果。通过多尺度分解与阈值去噪策略，小波变换在抑制背景噪声方面表现出一定优势，波形整体趋于平滑，主脉冲结构得以保留，但尾波部分残留依旧明显。

综合上述分析，传统方法在面对高斯白噪声这一典型平稳噪声时，仍具备一定降噪能力，但在结构完整性和尾波恢复方面存在不足。

|  |
| --- |
|  |
| (a) |
| (b) |
| (c)  图 3-2 混叠高斯白噪声（SNR = 0）水下打桩声信号降噪情况：（a）含噪信号；（b）维纳滤波结果；（c）小波变换结果 |

### 3.1.2 海洋背景噪声场景降噪

图 3-3 进一步展示了在信噪比SNR = 0条件下，参考信号混叠真实海洋背景噪声后的降噪效果。

|  |
| --- |
| (a) |
| (b) |
| (c) |
| 图 3-3 混叠真实海洋背景噪声（SNR=0）水下打桩声信号降噪情况：（a）含噪信号；（b）维纳滤波结果；（c）小波变换结果 |

图 3-3（a）所示为参考信号叠加海洋背景噪声后的时域波形。从图中可见，背景噪声具有显著的非平稳性与高度随机性，不仅导致主脉冲结构的清晰度显著下降，尾波部分亦被严重掩盖，整体信号建模难度显著增加。特别是在约0.9 s 处（红色箭头标注位置），背景中出现明显的突发干扰成分，进一步体现出真实海洋背景噪声的复杂性。图 3-3（b）为维纳滤波的处理结果，由于该方法依赖于背景噪声的频域统计建模，当背景噪声具有强非平稳特性时，其噪声估计能力受限，导致降噪效果不理想。尽管主脉冲能量略有恢复，但红色箭头所示区域的高幅值干扰仍未有效抑制，表明线性滤波方法在处理局部突发性噪声方面的局限性。图 3-3（c）为小波变换处理结果，尽管小波分析理论上具备较强的非平稳信号处理能力，能够在时频域进行多尺度分解，但在实际处理过程中仍受到噪声非结构性和多源扰动特征的影响，降噪后信号中仍残留大量背景干扰，主脉冲与尾波结构的恢复程度有限，红色箭头标注区域的非平稳干扰同样未能有效剔除。

通过对比图 3-2 和图 3-3 可以发现，传统降噪方法在不同噪声条件下展现出明显的性能差异。尽管维纳滤波和小波变换在应对高斯白噪声等平稳噪声方面具备一定优势，能够在一定程度上恢复信号主脉冲的能量，但在信号结构完整性和细节保真度方面仍存在明显不足。当面对频谱复杂、结构性弱且高度非平稳的海洋背景噪声时，这两种方法的结构保持能力均显著不足，尾波等细节难以有效重建。具体而言，维纳滤波由于高度依赖噪声频谱的精确估计，当噪声模型与实际环境不匹配时，容易过度抑制有效信号成分；小波变换虽然具有多尺度分析能力，但在复杂尾波结构和多源背景干扰叠加条件下，仍难以稳定提取信号的主结构特征。这些在信号结构重建、尾波抑制以及鲁棒性方面的不足，限制了传统方法在实际海洋工程水下噪声提取中的应用效果，难以满足高保真重构的任务需求。上述局限性进一步凸显了引入具备结构感知和非线性建模能力的深度学习方法的必要性。

## 3.2 自监督学习机制

### 3.2.1 监督方法的瓶颈

近年来，监督式的深度学习方法在音频增强和音频建模领域展现出优异的性能，尤其在特征提取建模能力和非线性表达能力方面明显优于传统算法。然而，主流的监督式学习方法往往依赖于大量准确标注的干净信号作为训练目标。在真实海洋环境下，获取完全不含背景干扰的目标声信号几乎不可行，干净样本极度匮乏已成为限制监督学习应用推广的核心瓶颈。尤其是对于突发性强、结构复杂的水下工程脉冲信号而言，样本采集难度高、场景变化显著、数据分布不均等问题更加突出，导致现有监督式降噪模型在泛化能力和工程实用性方面存在明显限制。

监督式的深度学习方法通常依赖明确标注的干净信号作为监督信息，以训练深度神经网络实现信号增强或降噪。这类方法被称为“噪声到干净”策略（Noise-to-Clean, N2C）[77]。在N2C策略中，网络模型通过最小化预测输出 与真实干净信号 之间的均方误差，来学习从含噪信号中提取干净信号的映射关系：

|  |  |  |
| --- | --- | --- |
|  |  | (3-1) |

其中， 为神经网络对含噪音频输入 的估计输出， 为对应地干净参考信号。含噪音频信号可表示为 ，而 表示叠加的噪声信号成分。然而，在真实的海洋环境中，大量干净信号数据往往难以获取，甚至在某些情景下无法获得，这导致N2C方法在实际应用中面临瓶颈。为了克服这种限制，近年来，自监督降噪方法逐渐引起研究人员关注。自监督降噪的优势在于仅依赖无标注的含噪音频数据进行训练，使模型在没有干净参考数据的条件下也能学习有效提取干净信号特征[15]。因此，自监督降噪方法在复杂海洋环境中具有广阔的应用前景和重要的现实价值。

### 3.2.2 自监督策略的演化

自监督方法通过构建伪标签或基于信号自身结构关系设计训练任务，使模型能够利用未标注的数据自主学习有效特征表示，具有显著的数据适应性和任务迁移能力。相较于传统方法对噪声统计分布等先验假设的依赖，自监督策略更贴近真实噪声环境下信号的结构规律，特别适用于具有明显主结构特征且尾波扰动复杂的脉冲类信号建模需求。该策略不仅有效缓解了缺乏干净样本的数据困境，还可以在训练目标中融入对信号结构保持的约束，自适应增强信号的主能量部分并动态抑制冗余干扰成分，从而有望实现更高保真的水下声信号重构。

有学者率先提出自监督方法Noise2Noise(N2N)，其利用同一干净信号的两种不同含噪观测数据进行模型训练，以避免对干净信号数据的直接需求[55]。在该策略中，每对训练数据由同一干净信号与两个相互独立的噪声分量叠加生成，即训练数据对为 和 ，其中 和 表示互不相关的零均值随机噪声变量。

具体而言，N2N方法在训练过程中，以含噪数据 作为输入，另一个含噪数据 网络训练目标，通过最小化以下损失函数来优化网络：

|  |  |  |
| --- | --- | --- |
|  |  | (3-2) |

在N2N条件下，由于 是独立从零均值分布采样，可以得到 的期望值等于 的方差与其期望平方之和。利用该原理，可以对公式（3-2）中的第三项进行展开。样本分布的方差 等于总体方差除以样本大小，因此随着音频训练数据集规模的增大，公式（3-2）中的第二项和第三项趋近于零。可以得出，最终N2N的 损失值接近于N2C的 损失值，即：

|  |  |  |
| --- | --- | --- |
|  |  | (3-3) |

理论上，当训练样本趋于无限大时，N2N策略的性能能够逼近传统监督式N2C策略[55]。然而，在实际有限数据条件下，由于训练数据规模限制，N2N策略的性能仍略逊于N2C策略。此外，N2N策略的实际应用还面临更大的局限性——难以在真实环境下获得独立噪声叠加的同源信号对。尤其在复杂多变的水下场景中，环境噪声往往随时间快速变化，很难满足严格独立条件的数据对采集需求。

针对N2N方法在数据配对条件上的局限性，研究人员提出了Neighbor2Neighbor(Nbr2Nbr)[78]策略，通过子采样的方式从单个含噪数据中构造一对近似独立的伪噪声样本，以满足自监督训练的需求。这种方法无需严格的配对噪声信号，极大提升了训练数据构建的灵活性。在Nbr2Nbr基础上，Wu 等人[56]将该策略进一步扩展到音频降噪领域，通过子采样构造伪训练目标对，实现了音频领域的高效自监督降噪训练。

具体来说，使用子采样器 从单个含噪音频数据 中生成一对音频训练数据 和 。与N2N方法不同的是，这里抽取得到的两段带噪音频 和 在期望意义上的真实值并不完全相同，因此产生了一个偏差项 ：

|  |  |  |
| --- | --- | --- |
|  |  | (3-4) |

通过引入偏差项的差值 ，其中 是噪声项，并假设 ，可以得到：

|  |  |  |
| --- | --- | --- |
|  |  | (3-5) |

考虑到理想的降噪网络 是用干净的音频数据训练并采用MSE损失，它可以确保在给定含噪输入 ，对于任意的噪声项 的情况下，对于 ，满足 且 。因此，在理想降噪网络 条件下，以下结论成立：

|  |  |  |
| --- | --- | --- |
|  |  | (3-6) |

换言之，网络在训练过程中被迫忽略随机噪声成分，仅捕捉信号的结构性成分。自监督策略的优势正在于此：无需真实的无噪材料，仅通过单一含噪数据的子采样构造即可实现模型的有效训练。该策略利用了噪声信号中潜在的结构一致性，促使网络仅关注信号内在特征结构，自动忽略随机噪声部分，在实际应用中具有重要价值。

## 3.3 自监督学习策略设计

### 3.3.1 训练信号对构造

本研究采用自监督学习方式训练模型，无法直接获取干净的目标信号，因此需要利用噪声信号自身构造伪“干净”目标，以提供有效的训练监督信号。本文针对单通道水下噪声数据设计了一种结构化随机子采样机制，用于生成自监督训练所需的信号对。

设原始噪声信号为 ，信号采样点数为dim，可将其视为一维张量（单通道）。为了实现随机子采样，设定子采样步长 ，将信号按连续的两个采样点分组。设分组后信号的组数为dim1，计算如下：

|  |  |  |
| --- | --- | --- |
|  |  | (3-7) |

其中 为边界安全额度，以避免索引越界。为确保随机子采样的均匀性和数据覆盖率，在每一组数据中生成随机二值变量 ，其形状为 ，并定义两个布尔掩码：

|  |  |  |
| --- | --- | --- |
|  |  | (3-8) |

同时，设每组的起始索引序列为：

|  |  |  |
| --- | --- | --- |
|  |  | (3-9) |

则每一组的子采样信号可分别表示为：

|  |  |  |
| --- | --- | --- |
|  |  | (3-10) |
|  |  | (3-11) |

最后，将每组结果重塑为 的矩阵，得到两个子采样信号：

|  |  |  |
| --- | --- | --- |
|  |  | (3-12) |
|  | | |
| 图 3-4 子采样示意图 | | |

如图 3-4 所示，相较于Nbr2Nbr[78]，该子采样机制优势在于保证每一组采样点被完全利用且无冗余或遗漏，通过组内交叉掩码实现了局部差异性与整体一致性并存的伪标签构造策略。这不仅提升了训练样本的表达效率，也有助于保持背景噪声的全局能量一致性，避免因随机丢弃而引入的频谱能量偏差。

此外，该策略在实现层面采用矢量化操作，支持 GPU 加速，显著提升了训练效率与扩展性，特别适用于大规模水下噪声数据的训练任务。接下来，将子采样信号 和 转换为频谱表示，分别得到 和 ，并构成训练对 。在每次训练迭代中， 作为模型输入，经过网络得到估计输出 ，然后计算 与参考目标 之间的损失。

这一训练机制的理论依据在于第3.2节介绍的噪声随机无关性假设。假设背景噪声是零均值的随机过程，则 和 之间的差异主要由噪声的随机波动部分构成，而它们的公共成分即为信号的稳定成分（即非随机海洋工程水下噪声信号）。因此，通过最小化 和 之间的损失，模型倾向于还原出 中 缺失的部分，这相当于在训练过程中主动补全子采样导致的频谱缺失信息。

### 3.3.2 频谱连续性缺失问题

引入子采样器的自监督策略虽然有效规避了对连续干净数据的依赖，成功构建了无标签条件下的训练样本，但同时也带来了新的建模挑战，即频谱连续性缺失问题。具体而言，子采样的预处理操作虽有助于提升训练数据的多样性与局部扰动鲁棒性，但也不可避免地在时间维度上引入了随机扰动，破坏了原始信号在频谱域中的结构完整性，导致时频谱图中出现跳变与缝隙，影响信号的连续性建模与结构特征提取。这种频谱结构的不连续性会削弱网络对关键脉冲成分的感知能力，进而影响整体的降噪重建质量。

需要指出的是，当前音频增强与降噪网络普遍采用时频联合建模策略，通常以短时傅里叶变换（Short-Time Fourier Transform, STFT）将一维时间序列映射为二维复数时频谱图作为输入。该时频谱图不仅包含信号的幅度与相位信息，还保留了时间分辨率，便于模型在频谱空间中建模信号结构与噪声扰动之间的差异。STFT通过选取窗函数将长时域信号分割成多个短时帧，并对每个短帧进行傅里叶变换，从而获得信号在时间和频率两个维度上的联合表示。在STFT中，每个时间帧之间的间隔由连续窗口之间的采样点数 （即帧移长度）决定：

|  |  |  |
| --- | --- | --- |
|  |  | (3-13) |

其中， 为信号的采样频率。当使用子采样技术对信号进行预处理时，其等效采样率降低为原来的一半。此时，相邻STFT帧之间的时间间隔将被拉长为原来的两倍：

|  |  |  |
| --- | --- | --- |
|  |  | (3-14) |

|  |
| --- |
| （a）    (b)  图 3-5 不同频谱场景下卷积核采样点分布情况：（a）原始信号频谱卷积核采样点分布情况；（b）子采样信号频谱卷积核采样点分布情况； |

这一变化对时频谱图的影响在图 3-5 有所体现。图 3-5（a）展示了原始音频信号的STFT时频谱图，其中红色方框突出显示了沿时间轴分布的一个柱状频谱结构（例如桥梁振动引起的一系列瞬时脉冲），该红框覆盖了约20个连续的时间帧。而在图 3-5（b）中，针对同一信号应用子采样处理后得到的时频谱图，对应的柱状频谱结构被“压缩”到仅约10个时间帧内。原本在高时间分辨率下沿时间轴连续分布的能量特征，被挤压到更少的帧中呈现，造成频谱中细节信息与结构连续性的严重破坏。许多原本沿时间连续变化的谱纹理被割裂开来。这样的信息损失尤其影响对于具有短暂脉冲特征信号的建模和学习。例如，第二章分析的海洋工程水下噪声信号的瞬态宽带脉冲部分，在高时间分辨率的频谱上表现为清晰的短时柱状能量分布；但经过子采样压缩后，这些脉冲在频谱上的表征变得模糊、拉伸，能量分布不再集中，干扰了网络对目标信号完整特征的准确学习，同时也降低了后续频谱重建的质量和准确性。

除了时间分辨率下降，子采样对频率分辨率也有明显影响。在STFT时频谱图中，每个垂直点代表某一时间窗内的频率成分，展示该时刻下各频率对应的能量或功率分布。时频谱图中频率点之间的间距即为频率分辨率 ，由采样率 和FFT点数 共同决定，计算公式如下：

|  |  |  |
| --- | --- | --- |
|  |  | (3-15) |

由公式（3-9）可知，当FFT点数 保持不变而采样率 减半时，频率分辨率 也降低一半，同时可显示的最高频率（奈奎斯特频率）变为原来的一半。这意味着时频谱图的频率轴被压缩，频率范围变窄：原本分布在多个频率点上的谱能量将集中到更少的频点上，相邻频点间距增大，不同频率成分之间的区分度降低，削弱了频谱的细节解析能力。此外，那些原本高于新奈奎斯特频率上限的信号成分可能因欠采样产生的频谱混叠效应而折叠至低频段，进一步破坏时频谱图的完整性与准确性。

综上所述，子采样操作虽然成就了自监督训练的数据构造，但伴随而来的频谱结构连续性受损不容忽视。第二章中我们了解到，海洋工程水下噪声信号在频域上通常表现为低频连续噪声叠加瞬时宽带脉冲的形式，要求模型既关注持续背景噪声又关注瞬态脉冲特征。然而子采样预处理使得频谱的局部纹理和整体结构遭到破坏，时间分辨率和频率分辨率的降低使网络难以有效捕捉瞬时宽带脉冲等关键结构信息。这种频谱连续性缺失削弱了网络的特征提取能力，也降低了降噪后信号的自然性和保真度。因此，如何在自监督框架下弥补因子采样引起的频谱信息损失，成为了提升模型降噪性能的关键问题。

## 3.4 本章小结

本章按照了“传统方法适用性验证—自监督需求提出—自监督策略设计—频谱连续性问题”的逻辑思路，为下一章深入阐述降噪网络结构与自监督训练细节奠定了坚实的理论基础与技术动机。

首先以两类典型噪声场景——高斯白噪声与实测海洋背景噪声，为实验对象，系统评估了维纳滤波与小波变换在海洋工程水下噪声信号降噪提取任务中的表现。实验结果显示：当噪声呈平稳特性时，两种传统方法虽能在主脉冲区域实现初步能量恢复，但在尾波结构与细节保真方面仍存在明显缺陷；当噪声高度非平稳且成分复杂时，两者对信号结构的保持能力急剧下降，已无法满足工程应用对高保真重构的要求。该结论充分暴露了传统线性与半线性算法在真实海域环境中的适用性边界。

随后，第3.2节阐述了以自监督学习机制，以及其取代监督式训练的可行性。自监督策略通过“用数据自身监督数据”的思想，避免了对干净训练数据的依赖，可显著提升模型在稀缺海洋声信号处理场景下的泛化鲁棒性，尤其适用于瞬态脉冲与连续尾波混叠的海洋工程水下噪声信号。

最后，针对上述问题，第3.3节提出了本研究所采用的自监督学习策略，设计了具备结构一致性与局部互补性的子采样策略，以从单一含噪音频中构造训练样本对。同时指出，该策略虽突破了数据配对难题，但在频谱变换过程中会引发频率连续性缺失的问题，进一步明确了在自监督框架下弥补因子采样引起的频谱信息损失，成为了提升模型降噪性能的关键问题。

综上所述，本章按照“传统方法适用性验证—自监督需求提出—自监督策略设计—频谱连续性问题”的逻辑链条，为下一章深入阐述降噪网络结构与自监督训练细节奠定了坚实的理论基础与技术动机。

# 第四章 自适应多重聚焦降噪网络设计

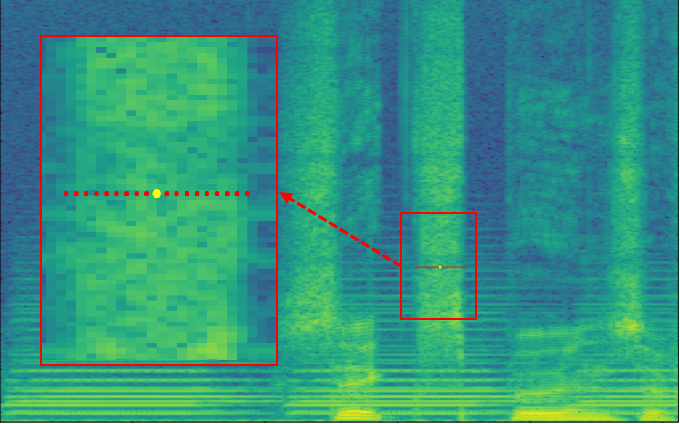
本章的核心目标是设计基于自监督深度学习的降噪网络，通过一系列结构与模块设计弥补因子采样引起的频谱信息损失，以提升海洋工程水下噪声信号（如打桩、桥梁振动信号）的降噪提取能力。具体内容包括网络架构、自监督方案、模块设计与损失函数。

## 4.1 网络设计与训练策略

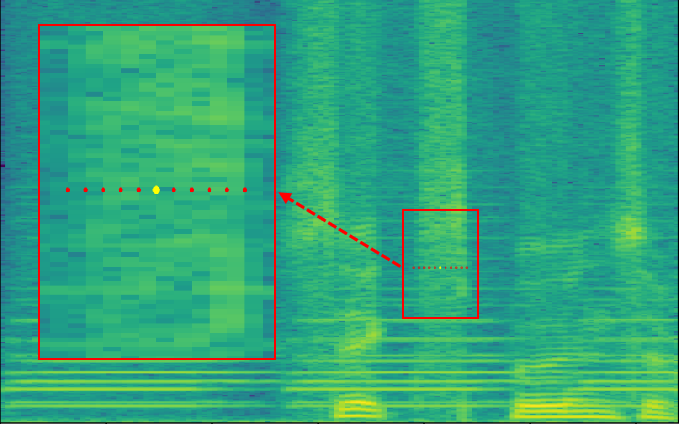
### 4.1.1 AMFNet网络

针对第三章总结的自监督策略中频谱信息丢失的不足，本研究提出了一种新型的自适应多重聚焦降噪网络（Adaptive Multi-Focus Denoising Network, AMFNet），以弥补子采样造成的频谱连续性缺失，增强网络对短时宽带脉冲结构的学习能力。

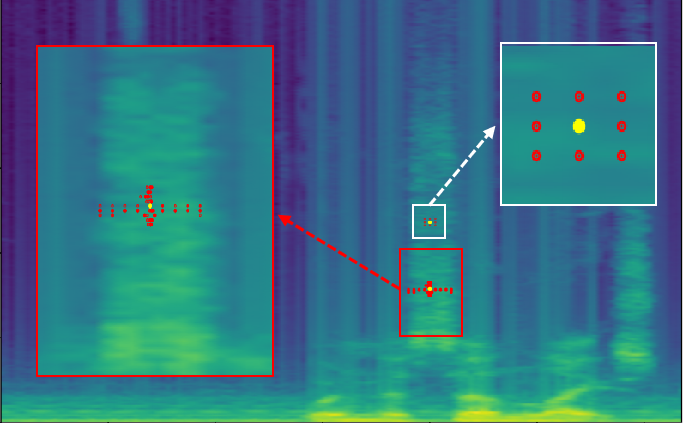
AMFNet在网络设计对卷积特征提取过程进行了改进，引入自适应的可变卷积核设计（详见第4.2.2节）使模型能够更有效地聚焦目标信号的时频特征。如图 4-1 所示，传统卷积在处理子采样后的“压缩”频谱时，卷积核的采样点在频谱平面上仍是固定均匀分布的。图 4-1(c)中白色虚线方框示意了这种传统卷积采样：无论频谱结构如何变化，卷积核在横向（时间轴）和纵向（频率轴）都按固定间隔取样，导致无法充分捕捉频谱随时间和频率变化的细节特征。而图 4-1(c)中的红色方框突出展示了AMFNet的作用效果：在横向（时间维），卷积核根据目标信号的频谱结构动态调整采样位置，使其能够准确对齐并抓取沿时间轴排列的脉冲频谱结构；在纵向（频率维），AMFNet显著提高卷积核在局部频带内的采样密度，以补偿子采样后频谱压缩所导致的信息空缺。通过这种自适应的可变卷积核设计，网络在提取特征时相当于同时具有“广角”关注整体趋势和“聚焦”关注细节的能力，对短时瞬态脉冲和细微谱变化更加敏感。此外，AMFNet在网络的各级卷积层中均配备了专门的多尺度特征提取模块（详见第4.2节），逐层融合不同尺度的谱信息，主动弥补子采样带来的信息损失。借助上述设计，AMFNet能够更有效地学习并提取海洋工程水下噪声信号中的瞬态脉冲频谱特征，同时抑制海洋环境中的连续背景噪声，显著提升复杂噪声环境下的整体降噪性能。



(a)



(b)



(c)

图 4-1 不同频谱场景下卷积核采样点分布情况：（a）原始信号频谱卷积核采样点分布情况；（b）子采样信号频谱卷积核采样点分布情况；（c）AMFNet网络处理频谱卷积核采样点分布情况

### 4.1.2 自监督训练方案

在建立了如上具有自适应特征提取能力的网络结构后，我们构建了相应的自监督训练方案流程，如图 4-2 所示。训练阶段，AMFNet完全基于含噪音频数据进行训练。具体来说，首先将原始含噪音频信号 通过子采样器生成一对含噪音频训练数据 。其中，子采样后的信号 经过STFT得到频谱表示，作为训练网络的输入；网络输出降噪后的频谱经过逆短时傅里叶变换（Inverse STFT, ISTFT）得到 与 计算损失，以更新网络权重。

|  |
| --- |
| 图 4-2 本研究提出的自监督策略概述 |

在训练过程，网络被迫学习两个子采样片段共同的信号结构，忽略所含噪声的差异部分。这意味着AMFNet掌握海洋工程水下噪声信号典型的脉冲频谱结构，同时将随机的背景噪声视作可忽略的扰动加以抑制。训练完成后的AMFNet在推理阶段只需将原始含噪音频经过STFT得到的频谱输入已经训练好的模型，网络便会直接输出对应的干净频谱估计 。最后，对 执行ISTFT重构回时域，即可得到降噪后的音频信号 。整个自监督降噪方案在训练不需要任何外部干净参考信号，完全依赖含噪数据自身的内部关联来驱动模型学习，充分体现了自监督学习在海洋工程水下噪声信号提取任务中的价值。

## 4.2 AMFNet架构与关键模块

AMFNet采用编码与解码器的U型网络整体结构，并融合多种特定设计的模块单元，有针对性地对海洋工程水下噪声信号的频谱特性进行高效特征提取和重构降噪。

|  |
| --- |
| 图 4-3 本研究提出的AMFNet网络架构 |

如图 4-3 所示，AMFNet的左半部分为编码器，用于逐级压缩并提取输入复数频谱的主要特征；右半部分为解码器，用于逐步重建并输出去噪后的细粒度频谱。在编码器和解码器对应层之间，引入了改进的跨层跳跃连接机制，以替代传统 U-Net 直接拼接的方式，实现不同尺度特征的深度融合；同时在网络的瓶颈层（编码器与解码器交汇处）设计了全局语义校准单元，以充分整合最高层次的特征信息。通过这些设计，AMFNet能够在无干净数据的情况下完成模型训练，充分捕获水下脉冲信号的瞬态频谱脉冲特征并抑制宽带连续背景噪声，为后续高质量重构提供保障。下面将分别介绍AMFNet整体架构以及各关键模块的设计原理与优势。

### 4.2.1 编码与解码器结构

如图 4-3 所示，AMFNet总体采用对称的编码与解码器架构，以典型U-Net结构为基础骨干。整个网络由对称的编码端和解码端组成：编码端逐层提取并压缩输入信号的频谱特征，解码端逐步恢复出目标信号的细节频谱。编码器包括4个逐级串联的编码模块，对应4个不同的特征尺度；解码器同样包含4个逐级堆叠的解码模块，与编码器层次一一对应。通过这种由“粗到细”再“细到精”的分层结构，网络能够逐步从原始复杂频谱中抽取多尺度特征表示，并最终重建出净化后的目标信号频谱。

在编码端，每一级编码器首先接收来自上一层（或直接来自输入）的复数频谱特征图。为突出目标脉冲信号的显著特征并减少下采样导致的关键信息丢失，每个编码器层输入处都引入了频谱动态聚焦单元，其根据输入频谱的时频局部结构，动态调整卷积核的采样形状与位置，使卷积运算能够自适应地“聚焦”于频谱中能量显著的目标区域（详细机制见4.2.2节）。每级编码器模块由复数注意力模块和下采样操作组成。复数注意力模块旨在复数域内高效提取频谱特征，同时建模幅度与相位信息的交互（详细机制见4.2.3节）。

编码器中的下采样模块主要实现频谱特征维度的压缩。本研究使用了专门的复数卷积层来完成下采样提取特征。复数卷积的计算可表示为对输入复数向量 与复数滤波器 进行卷积运算，即：

|  |  |  |
| --- | --- | --- |
|  |  | (4-1) |

复数卷积层具体由两个步长为2、大小为3×3的实值卷积核实现，分别处理输入频谱的实部和虚部。在下采样的同时，这种卷积操作高效提取频谱特征并压缩特征空间维度。此外，在复数卷积之后，我们配备了复数批量归一化层（Complex Batch Normalization, CBN）和复数ReLU激活函数（Complex ReLU, CReLU），这些复数域专用操作被证明能更稳定地处理复杂频谱特征[79]，保证网络在复杂水下噪声环境中的训练收敛性和特征表达能力。

与编码端对称，解码端同样由4个逐层对应的解码模块组成，用于逐步恢复编码压缩过程中丢失的频谱细节。每一级解码器接收来自上一层解码器的上采样特征，并融合来自编码端的跨层特征信息后，再经卷积生成更高分辨率的特征图。不同于经典U-Net直接将对应编码器特征与解码器特征在通道维度拼接，本模型在每层解码前引入复数多尺度跳跃连接单元来替代直接跳跃连接，其通过对编码端不同尺度的特征进行对齐融合（详细机制见4.2.5节），实现编码特征与当前解码特征的深度整合。同时，解码器中的上采样模块采用复数转置卷积等操作将特征图逐步放大回原始尺寸。

综上，AMFNet通过编码端的逐级特征压缩与解码端的逐级特征重建，形成一个从频谱粗略结构到精细细节的完整处理流程。尤其借助SDFU、CAM、SRU和CMS-SCU等创新模块的有机配合，网络能够有效捕获频谱中突出的脉冲结构和连续的背景噪声成分，实现多尺度特征的自适应融合。在复杂海洋工程水下噪声提取与降噪任务中，该架构有效克服了自监督方法子采样引起的频谱连续性缺失问题。

### 4.2.2 频谱动态聚焦单元（SDFU）

频谱动态聚焦单元（Spectrum Dynamic Focusing Unit, SDFU）是AMFNet中为增强脉冲信号细节提取而设计的核心模块之一。其设计动机在于：编码器的下采样虽可提取抽象特征，但同时也会导致时频谱图中部分细节和连续性丢失。当水下脉冲信号与背景噪声在频谱上高度叠加时，传统固定形状的卷积核难以同时兼顾全局结构和局部瞬时变化，可能遗漏短时突发的局部能量峰值。为了解决这一问题，SDFU引入了动态形变卷积（Deformable Convolution, DSConv）的思想[80]，通过可学习的卷积核偏移量来自适应调整卷积采样位置和如如区域，从而令卷积操作针对每个输入样本的频谱结构进行精准捕捉。

如图4-4，SDFU在标准二维卷积的基础上增加了形变机制：卷积核的采样位置不再是固定规则网格，而是通过学习一系列位置偏移量 来动态调整。对于中心坐标为 的3×3卷积核 ，其采样位置定义为规则网格：

|  |  |  |
| --- | --- | --- |
|  |  | (4-2) |

|  |
| --- |
| 图4-4 SDFU组成结构与效果示意图 |

如图4-4，SDFU在标准二维卷积的基础上增加了形变机制：卷积核的采样位置不再是固定规则网格，而是通过学习一系列位置偏移量 来动态调整。对于中心坐标为 的3×3卷积核 ，其采样位置定义为规则网格：

|  |  |  |
| --- | --- | --- |
|  |  | (4-2) |

受到可变形卷积[80]启发，而在SDFU的DSConv中，卷积核 沿着水平轴（时间轴）的采样位置定义变为：

|  |  |  |
| --- | --- | --- |
|  |  | (4-3) |

同样的，沿着垂直轴（频率轴）方向，卷积核 位置定义变为：

|  |  |  |
| --- | --- | --- |
|  |  | (4-4) |

上述公式中学习到的偏移量 和 通常为非整数，因此在实际实现中，引入了双线性插值方法计算非整数位置处的卷积核采样值。对于偏移后的位置 ，其卷积采样值通过周围四个整数栅格点：

|  |  |  |
| --- | --- | --- |
|  |  | (4-5) |

其中 表示由公式(4-12)和(4-13)计算出的非整数位置，而 表示其周围所有整数坐标位置, 表示的是卷积核 在整数坐标位置 处的取值。插值内核 由两个一维线性插值核 和 的张量积构成：

|  |  |  |
| --- | --- | --- |
|  |  | (4-6) |

在实现上，SDFU模块的最终输出由三个并行卷积分支的特征构成：沿时间轴（频谱横轴）的形变卷积特征 、沿频率轴（频谱纵轴）的形变卷积特征 和常规卷积特征 。这三个特征在通道维度拼接后，通过一个1×1卷积层进行初步融合，再经过组归一化（Group Normalization, GN）和挤压-激发注意力机制（Squeeze-and-Excitation Attention, SE）[81]进一步优化。其中，GN层对特征分布进行归一化处理，提升网络训练稳定性；SE机制则根据每个通道对输出结果的贡献，自适应增强与目标信号强相关的通道权重，降低与目标信号弱相关的冗余通道权重。对于输入复数频谱特征 而言，SDFU整个计算流程可表示为：

|  |  |  |
| --- | --- | --- |
|  |  | (4-7) |

其中符号 表示通道维度的拼接操作，函数 的具体定义如下：

|  |  |  |
| --- | --- | --- |
|  |  | (4-8) |

其中 表示1×1卷积，GN和SE分别表示组归一化和挤压-激发机制。SDFU引入的形变卷积机制显著提升了模型对频谱细节的捕捉能力。如图4-1（c）所示，对于缓变背景噪声中夹杂的瞬态脉冲信号，SDFU能够根据时频谱图中能量分布的变化自动调整卷积感受野：在时间轴方向增加对瞬时突发事件的采样密度，在频率轴方向关注脉冲能量所在的频带范围。例如，对于桥梁振动产生的短促脉冲，其频谱常出现窄时间、宽频带的高能量带状结构，SDFU能够沿频率轴拉伸卷积核以覆盖该宽频带区域，确保脉冲能量被完整捕获；同时沿时间轴细化采样，刻画其短暂出现的瞬态特征。通过这种动态聚焦，SDFU有效避免了下采样过程中目标脉冲频谱细节的丢失。实验结果表明，引入SDFU后，编码器各层提取的特征中脉冲信号的显著性得到增强，为后续CAM模块的深入挖掘和区分提供了包含更多关键信息的输入。

### 4.2.3 复数注意力模块（CAM）

复数注意力模块（Complex Attention Module, CAM）是AMFNet网络中另一项重要创新设计，旨在提升网络对复数频谱特征的提取和表达能力。传统注意力机制大多仅针对幅度谱进行加权，忽略了相位信息；而在水下脉冲信号提取任务中，信号与噪声在幅度和相位上往往具有不同特征，仅关注幅度可能无法充分区分两者。为此，CAM在复数域同时对频谱实部（幅度分量）和虚部（相位分量）建模，通过特定结构设计实现幅度-相位信息的交互融合，提高了对复数谱信息的利用效率。如图4-5所示，CAM模块的整体架构可分为两个核心单元：复数特征交互单元和跨空间学习单元，共同作用以实现对时频谱图局部细节与全局关联的自适应建模。

在复数特征交互单元中，CAM首先针对输入的复数特征张量结构进行重新排布和并行处理。具体来说，对于输入的复数特征张量 （最后一个维度分别代表复数特征的实部和虚部），CAM首先在最后一个维度将其拆分为实部和虚部两部分，随后将两部分沿通道维度重新连接，形成多个子特征组，以强化不同复数分量之间的交互：

|  |  |  |
| --- | --- | --- |
|  |  | (4-9) |

|  |
| --- |
| 图4-5 CAM组成结构示意图 |

为了最大限度地加强实部和虚部之间的交互，CAM设计了三个并行子分支结构：其中两个分支以1×1的共享卷积为主，用于跨通道的信息融合；第三个分支以3×3卷积为主，用于提取局部空间细节特征。

在前两个共享的1×1卷积分支中，CAM借鉴坐标注意力（Coordinate Attention, CA）机制[82]的思想，分别针对实部特征和虚部特征在时频谱图横向（时间轴）和纵向（频率轴）上的分布进行编码。具体过程是：通过在空间维度上分别进行水平方向和垂直方向的全局平均池化（Global Average Pooling, GAP），获取实部特征在各频率带的整体能量分布，以及虚部特征在各时间帧的整体相位分布。这相当于提取了实、虚部特征在两个正交方向上的全局描述。一方面，横向GAP捕获不同时间段上能量的变化趋势；另一方面，纵向GAP刻画不同频率分量上的总体激活水平。接着，将两个方向的全局特征向量分别通过共享的1×1卷积层进行投影和压缩，然后沿通道维度连接起来，形成综合的通道注意力向量。这一操作在降低特征维度的同时，实现了实部与虚部跨通道信息的融合交互，显著增强了两者之间的关联性。最终得到的通道注意力权重经过适当广播后，对原始特征图的各通道进行加权调制，自适应突出对目标信号有判别力的频带和特征通道。

与上述注重全局信息的两个分支并行，3×3卷积分支则用于捕捉复数特征中实部与虚部的局部空间细节。该分支对输入的实、虚部特征在局部邻域内进行卷积操作，获取时频谱图局部邻域的细微模式。例如，它可以关注到某一时刻相邻频点上共同出现的异常高幅度，或某一频带在相邻时刻的相位变化趋势。卷积输出经过 Sigmoid 非线性激活生成一张空间注意力权重图，用于细粒度地调节原始特征图对应位置的值——相当于根据局部邻域模式对输入特征做像素级的增强或抑制。通过1×1和3×3卷积分支的互补，CAM在同一模块中同时建模了复数频谱特征的局部相关 （短程依赖）与全局相关 （长程依赖）。实践证明，这种多尺度相关性的融合策略在处理复杂多变的频谱结构时具有明显优势[83]。

在复数特征交互单元处理后，CAM利用跨空间学习单元进一步融合来自不同尺度的注意力信息，以产生最终的注意力映射。具体做法是：首先对前述得到的多尺度特征应用二维全局平均池化（GAP）操作，在空间维度上对特征图进行全局聚合：

|  |  |  |
| --- | --- | --- |
|  |  | (4-10) |

这一步获得的向量 表示特征图在整个时频平面的平均激活水平，捕获了长距离的空间依赖关系。然后，对该全局特征向量利用 Softmax 函数进行归一化缩放，确保其元素和为1且与原特征尺度一致。这相当于从全局视角提炼出一个空间注意力权重系数 。接着，将这一系数与之前得到的局部空间注意力权重 通过矩阵点乘的方式融合，生成最终的复数空间注意力图。该注意力图综合了短程局部细节和长程全局依赖两种空间相关信息，对时频谱图上各位置的显著性进行评价。最后，利用 Sigmoid 激活将注意力图归一化至0-1范围，并作用于输入特征图，以精确突出目标频谱特征区域、有效抑制背景区域。通过以上一系列操作，CAM模块能够自适应调节时频谱图中不同位置和通道的响应，使与目标脉冲信号相关的谱区得到强化，而噪声占优的谱区被淡化。

综上，CAM模块通过实部-虚部信息交互和多尺度空间注意力融合，大幅提升了网络对频谱细节特征的敏感度和表达能力。在复杂海洋噪声背景下，CAM可以更稳定地捕捉那些区分脉冲信号与背景噪声的关键模式——例如短时突发的能量峰值、异常的相位变动等——从而提高模型对目标信号的辨别能力和鲁棒性。值得强调的是，在自监督训练情景下，没有干净目标信号直接指导模型关注哪些特征，CAM提供了一种有效机制引导模型自主关注关键的幅度和相位特征，这对于准确提取脉冲型海洋工程水下噪声信号至关重要。

### 4.2.4 语义重校准单元（SRU）

在AMFNet中，编码器与解码器的衔接部分（即网络底部的瓶颈层）是信号由高维压缩表示过渡到重构输出的关键环节，也是全局特征信息汇聚与分流的桥梁。为充分融合编码端提取的高级语义特征并为解码端提供最优信息表征，本文在编码器输出与解码器输入的交汇处设计了语义重校准单元（Semantic Recalibration Unit, SRU）。SRU的作用在于重新编码并加权调整编码端的顶层特征，使之既保留原始的高层语义信息，又经过优化更适配解码端逐层细化重建的需求。换言之，SRU对编码器末层提取的特征进行“语义校准”，突出对后续重构有用的成分，削弱冗余或有害的信息，从而提高最终重建结果的质量。

如图4-6所示，为实现上述目标，SRU融合了全局特征建模和局部细节增强两方面策略，对应引入了两个子模块：复数全局多层感知器（CGMLP）和复数局部注意力（CLA）。其中，CGMLP模块侧重建模长程依赖和全局语义信息，通过通道维度的深度变换来整合频谱的宏观结构；CLA模块侧重捕捉局部细节和短程依赖，通过一种字典注意力机制来发掘频谱局部的重要模式。两个模块分别处理特征的不同方面，最后在SRU中协同工作，达到对编码顶层特征全方位“校准”的效果。

|  |
| --- |
| 图4-6 SRU组成结构示意图 |

SRU首先对来自编码器的顶层复数特征 进行预处理，以保证后续模块可以更高效地加以利用。具体而言， 的实部 与虚部 分别通过一个5×5卷积层（步长为1）、BN和ReLU激活函数，然后沿通道维度连接：

|  |  |  |
| --- | --- | --- |
|  |  | (4-11) |

其中函数 的具体定义如下：

|  |  |  |
| --- | --- | --- |
|  |  | (4-12) |

其中， 代表5×5卷积运算， 为批量归一化层， 为ReLU激活函数。该预处理稳定了特征分布，减小了不同通道之间的统计差异，为后续信息交互奠定基础。随后，将标准化后的特征 并行输入CGMLP和CLA两个子模块，分别进行全局特征融合和局部特征增强。

（1）CGMLP：复数全局多层感知器。CGMLP由多个子层叠加而成，每个子层各司其职，共同对输入特征的全局语义进行提炼和强化。首先，深度卷积残差模块采用1×1深度卷积（Depthwise Convolution）来处理 ，在保持各通道独立的前提下提取每个频谱通道的深层模式。由于使用深度卷积，此步骤几乎不引入跨通道计算，在提高效率的同时保留了原始特征结构。紧接着，引入通道缩放操作（Channel Scaling, CS）对经过深度卷积的特征进行线性变换，调整各通道的幅值范围，以增强特征的泛化能力：

|  |  |  |
| --- | --- | --- |
|  |  | (4-13) |

其中， 表示组归一化， 为1×1的深度卷积， 表示通道缩放操作。随后，在特征交互模块中，使用1×1卷积将实部特征与虚部特征在通道维度重新融合：

|  |  |  |
| --- | --- | --- |
|  |  | (4-14) |

如此一来，原本独立的实、虚部通道产生交互，从中学习复数域全局信息的关联模式。幅度信息和相位信息在全局尺度上得以整合，使网络可以更好地感知那些分散在不同通道的全局噪声模式或信号模式的整体趋势。接下来，经过一个通道MLP残差模块（即在通道维度上的多层感知机变换），进一步挖掘通道与通道之间的高阶非线性关系：

|  |  |  |
| --- | --- | --- |
|  |  | (4-15) |

其中， 代表通道维度的多层感知机（Multi-Layer Perceptron, MLP）[84]，通过跨通道的映射增强了复数特征之间的交互，能够捕获更复杂的全局依赖关系，提高模型对不同特征组合的适应性[85]。

此外（如图4-7所示），为抑制过拟合并强化全局特征的稳健性，我们对通道MLP输出的实部与虚部又施加了一次通道平均池化：

|  |  |  |
| --- | --- | --- |
|  |  | (4-16) |

其中， 为通道数，该操作相当于对每个通道的响应求平均，提取全局背景的平均能量水平，使网络对整体噪声强度的估计更加准确稳定。经过以上层层处理，CGMLP模块有效建模了频谱全局的语义特征和长程依赖：例如，它能学习背景噪声在各频带的大致能量分布，提取出目标信号在整体上的微弱稳态特征。

|  |
| --- |
| 图4-7 通道平均池化 |

（2）CLA：复数局部注意力模块。CLA旨在对 中残留的细节信息进行强化，特别关注那些对重构有重要影响的局部模式。CLA引入了一种内置字典机制（Dictionary Mechanism）来实现这一目标。首先，将 通过1×1和3×3卷积层压缩通道维度，接着利用局部注意力计算相应特征的权重。特征信息 计算如下：

(4-17)

其中， 为局部编码向量， 表示缩放因子， 和 分别对应第 个输入向量的实部与虚部， 和 分别为第 个编码向量的实部与虚部。获得所有局部编码权重 后，CLA通过一个融合单元将它们整合，以突出那些对应关键模式的分量：

|  |  |  |
| --- | --- | --- |
|  |  | (4-18) |

其中， 由BN、ReLU、均值池化和全连接层构成。这一融合相当于统计所有局部模式出现的情况，找出最显著的几个模式类型并增强其权重。最后，CLA将得到的关注权重作用回原始顶层特征：具体地，将各局部注意力信号按照对应的位置映射回 ，并在实部和虚部通道间加入适当的交互，然后通过 Sigmoid 归一化后与原特征逐元素相乘，再残差式地叠加回原特征：

(4-19)

|  |  |  |
| --- | --- | --- |
|  |  | (4-20) |

其中， 为 Sigmoid 函数， 为通道维度点乘， 为通道维度加法。这一系列操作使与字典模式匹配的局部细节特征被显著强化，而无关细节被淡化。在水下脉冲信号场景中，这意味着CLA能够突出诸如脉冲峰值、瞬态尖锐变化等细节，为解码器提供更丰富准确的局部信息。

通过CGMLP和CLA的协同作用，SRU实现了对编码器顶层特征的全局-局部联合重校准：CGMLP从宏观上把握频谱整体结构和背景噪声水平，CLA从微观上雕琢关键脉冲细节特征，两者相互配合。经过SRU重校准后的特征既包含全局语义上下文（例如噪声分布、信号的大致形态），又保留局部判别细节（例如峰值位置、短时结构），为解码端的逐层信号重建奠定了坚实基础。在后续解码过程中，这些校准后的特征将指导每一层解码器更有针对性地恢复目标信号——既保证重构信号保留脉冲的瞬时特征，又有效压制背景噪声，最终实现高质量的去噪输出。

### 4.2.5 复数多尺度跳跃连接单元（CMS-SCU）

自监督学习场景下缺乏干净参考，模型必须充分挖掘混合信号中蕴含的各种线索来区分信号与噪声。其中，频谱的多尺度信息利用至关重要：水下环境中的工程脉冲信号通常伴随缓慢变化的低频背景噪声，同时叠加短暂尖锐的高频脉冲。单一尺度的特征往往无法同时描述这两类截然不同的频谱成分。如果模型仅依赖于浅层（高分辨率）特征进行重建，可能缺乏对全局低频结构的把握，导致背景噪声残留；反之，若仅依赖深层（低分辨率）特征，则细微的高频脉冲纹理容易被忽略。为此，AMFNet引入复数多尺度跳跃连接单元（Complex Multi-Scale Skip Connection Unit, CMS-SCU），用于在解码阶段有效融合编码端不同尺度的特征表示，确保模型同时关注局部细节和全局背景。

|  |
| --- |
| 图4-8 CMS-SCU组成结构示意图 |

如图4-8所示，CMS-SCU的核心思想是在U-Net[28]的跳跃连接基础上，引入多尺度特征融合策略：每一层解码器在利用对应编码器输出进行特征补充时，不再仅使用单一层的特征，而是将其相邻的多个尺度的编码器特征一并融合进来。具体而言，对于解码端第  层，需要融合编码端第 层的输出特征以及其相邻的上、下尺度（例如第 和第 层）的特征。为了使这些尺度不同的特征能够直接融合，CMS-SCU首先对它们进行空间尺度对齐：对尺度较小（分辨率低）的特征进行上采样、对尺度较大（分辨率高）的特征进行下采样，最终将三者调整到相同尺寸。然后，将对齐后的特征在通道维度拼接，并通过一个1×1卷积进行初步融合。1×1卷积的作用一方面在于压缩通道维度，减弱直接拼接带来的冗余信息；另一方面是混合不同来源特征，有助于提炼多尺度的综合表示。融合后的特征通过通道注意力机制（Channel Attention, CA）为融合特征的每一通道分配权重。这样，来自不同尺度的特征信息将通过注意力得到自适应的衡量：与目标信号关联度高的特征通道被赋予较大权重予以保留，而冗余或噪声相关的通道则被削弱。动态加权融合确保了网络在解码每层时，都能优先利用有益于脉冲信号重建的那部分编码器特征，同时抑制干扰信息。

因此，在自监督降噪任务中，CMS-SCU使得多尺度特征在解码过程中得以充分交互。从频谱角度看，低频部分的稳态噪声结构和高频部分的瞬态信号细节都不会在单一尺度的处理过程中被淹没或遗失，而是通过多尺度特征融合学习浅层与深层的特征，增强了模型对频谱全局结构的学习能力。

## 4.3 损失函数设计

### 4.3.1 时频损失

在自监督学习框架下，损失函数的设计对于模型的训练效果至关重要。针对水下声信号的特点，本文构建的损失函数综合了时域和频域的误差项，以确保模型能够有效学习水下噪声环境中的目标信号特征。

设模型输出的时域信号为 ，其对应的目标信号为 ，两者的STFT分别记为 和 。在每一帧中，时域信号长度为 ，频域采样点数为 ，总帧数为 。综合考虑时域信号的重建精度、频域信号的频谱一致性以及模型的正则化约束，定义如下损失函数：

|  |  |  |
| --- | --- | --- |
|  |  | (4-21) |

其中， 主要用于保证时域信号的整体波形重建精度， 约束模型对频谱特征的精准重构，而 则用于在训练过程中引入额外的正则化约束， 和 为各项损失的平衡系数。

首先，在时域损失部分，采用MSE衡量模型输出与目标信号之间的差异：

|  |  |  |
| --- | --- | --- |
|  |  | (4-22) |

该项损失确保模型重建的时域波形与目标信号尽可能贴近，从整体上控制去噪信号的波形失真度。

其次，在频域损失部分，为了确保模型在不同频率段的重建精度，引入了一种加权误差策略。在计算STFT得到 和 后，定义频域损失为两者差异的加权MSE：

|  |  |  |
| --- | --- | --- |
|  |  | (4-23) |

### 4.3.2 正则化损失

此外，为进一步提升模型的泛化能力，并确保自监督学习过程中网络能更稳定地提取有用信息，我们在损失函数中加入了正则化约束项 ，其定义如下：

(4-24)

其中， 表示训练的降噪网络， 和 为Nbr2Nbr策略生成的两个子采样信号对。该正则项约束模型输出的自一致性，即要求网络在不同子采样条件下的降噪结果保持稳定，并减少过拟合风险。加入这一正则化约束能够提升模型对水下脉冲信号的敏感度，同时降低子采样带来的信息缺失影响，使模型更关注信号的稳定成分而非背景噪声的随机波动。

4.4 本章小结

本章围绕海洋工程水下噪声信号的结构特征与降噪提取需求，设计并实现了一种自监督深度降噪网络AMFNet，系统构建了自监督训练方案、主干网络架构及关键功能模块，并制定了相关损失函数，在无干净数据条件下实现网络模型的训练。

首先，在网络训练方案方面，结合海洋场景中缺乏干净参考信号的现实问题，提出基于单通道背景噪声的子采样自监督策略，利用同源噪声片段间的随机扰动差异构建伪训练目标，有效规避了对高质量标注样本的依赖。

其次，在网络架构层面，AMFNet以U型结构为骨架，集成了多个具有工程语义的关键模块：频谱动态聚焦单元（SDFU）通过自适应信号频谱能量密度差异，调整卷积核形状，有效捕捉主能量区域；复数注意力模块（CAM）强化了幅度与相位之间的交互建模，提升了对信号频谱结构信息的恢复能力；语义重校准单元（SRU）对多尺度通道特征进行结构对齐，增强了跨尺度表达一致性；复数多尺度跳跃连接单元（CMS-SCU）弥合了深浅层间高低频特征的断层，提升了频谱结构重构能力。这些模块的设计均以海洋工程水下噪声信号的非平稳性、频谱局部性与复数域结构为理论支撑，构成了具结构感知的网络架构体系。

最后，损失函数的设计融合了频域幅度误差、能量差异和正则约束项，在自监督训练框架下充分引导网络学习信号主结构与稳定频谱特征，实现对背景噪声的有效抑制与目标信号频谱结构的高保真提取。

综上所述，本章完成了AMFNet网络的自监督训练方案、主干网络架构、关键功能模块与损失函数的全流程设计，并结合海洋工程水下噪声信号的结构规律，构建了实践可行性的自监督降噪方案，为后续实验验证与性能对比奠定了坚实基础。

# 第五章 降噪实验与结果分析

## 5.1 实验设置

为系统验证所提出方法在复杂水下环境中的海洋工程水下噪声信号提取与降噪性能，本研究设计了不同能量水平的系列实验。具体而言，针对典型强脉冲信号——打桩施工过程中产生的打桩声，基于近距离采集的高信噪比实测数据，通过叠加实测海洋背景噪声，控制混叠后信噪比水平（−10  dB、-5 dB、0  dB、10  dB），以模拟远距离传播引起的信号劣化场景，考察方法在信号强度衰减后的脉冲特征恢复与背景噪声抑制能力。同时，针对典型弱脉冲信号——跨海桥梁钢结构振动声，直接使用原始实测数据作为输入，因其本身能量较弱、初始信噪比较低，自然形成低信噪比环境，进一步验证方法在弱脉冲信号提取、尾波干扰抑制及噪声适应性方面的鲁棒性。

上述两类脉冲信号的设置，覆盖了从强能量脉冲到弱能量脉冲的典型应用场景，系统评估了所提方法在不同能量结构、不同信噪比条件及不同脉冲特性下的降噪增益、脉冲结构保持能力与跨场景泛化适应性。

### 5.1.1 实验设计内容

基于第二章对海洋工程水下噪声实测特征的分析，本研究在实验方案设计中聚焦于两类典型的水下脉冲噪声信号：一是打桩施工过程中释放的强脉冲型噪声，二是跨海桥梁钢制结构在车辆荷载作用下辐射的弱脉冲型噪声。两类信号在能量水平、频谱分布及时域结构方面存在明显差异。打桩脉冲噪声通常具有极高的瞬时能量，声压峰值大，频谱成分丰富，时域上表现为短脉冲快速衰减的特性；而桥梁振动脉冲信号能量较低，声压峰值小，频谱分布相对集中，时域脉冲形态更为平缓，衰减过程也更为缓慢。尽管特性不同，这两类脉冲噪声均以主脉冲成分为核心，并伴随明显的尾波振铃，即主脉冲结束后出现的逐渐减弱的余振。尾波不仅在时域上延长了信号持续时间，还在频域引入了额外低频能量，成为脉冲提取与分析过程中主要的干扰源，对降噪方法提出了兼顾结构保持与干扰抑制的双重要求。

针对低信噪比环境下脉冲信号的提取与降噪问题，本研究分别以打桩声与桥梁振动声为对象，设计了系统性的实验方案。首先，通过基于近距离实测打桩声叠加背景噪声，模拟远距离传播导致信噪比劣化的观测场景，以评估模型对强脉冲信号结构恢复能力的提升。其次，以初始信噪比较低的桥梁振动声作为测试对象，检验方法在弱脉冲信号条件下对脉冲特征提取与背景噪声抑制的鲁棒性，从而覆盖典型强脉冲与弱脉冲信号的实际应用需求。具体设计详见表5-1。

首先开展了累计能量段打桩声降噪实验，以应对主脉冲能量集中而尾波能量分散所带来的提取挑战。该实验从原始打桩脉冲信号中截取累计能量达到90%的主脉冲段，叠加实测海洋背景噪声，并设置-10 dB、-5 dB、0 dB、5 dB、10 dB五档初始信噪比，系统模拟远距离传播后信号质量劣化的情形。训练基于该数据完成，旨在强化模型在低信噪比条件下对主能量结构的恢复能力与尾波干扰的抑制效果，为后续复杂环境下的应用奠定基础。

其次，为在真实海洋环境条件下检验所提方法的整体降噪效果，设置原始打桩声信号降噪实验。该实验直接使用未截取处理的完整打桩脉冲波形叠加背景噪声，信噪比设置与累计能量段实验保持一致。通过在同档信噪比的累计能量段数据上完成训练后，迁移至完整输入波形进行测试，实现训练与测试一一对应。该设计不仅延续了基于近场实测数据模拟远距信噪比劣化的思路，而且进一步考察了模型在面对复杂尾波与自然噪声混合背景时的结构还原与能量重构性能。

最后，为考察方法在不同脉冲特性下的泛化适应性，设计了非训练类型信号降噪实验。选取在累计能量段打桩声0 dB混叠数据上训练得到的模型，直接应用于桥梁振动声叠加背景噪声的测试集，检验其在未见过弱脉冲分布情况下的结构恢复能力。考虑到桥梁振动声本身能量较低，且叠加背景噪声后信噪比进一步恶化，该实验自然构成了严苛的低信噪比测试环境，为评估所提方法在弱脉冲场景下的脉冲辨识与噪声抑制鲁棒性提供了重要依据。

综上，通过累计能量段实验、完整打桩声实验与非训练类型信号实验三种递进式设计，本研究系统考察了所提自监督降噪方法在不同信号能量结构、不同信噪比水平及不同脉冲特性条件下的综合性能，确保实验设置能够高度贴合海洋工程水下噪声提取与降噪应用的实际需求。

表5-1 实验环节设置

|  |  |  |
| --- | --- | --- |
| 实验环节名称 | 数据构成 | 数据用途 |
| 累计能量段打桩声降噪实验 | 90 % 能量段打桩声信号  + 海洋背景噪声 | 训练/验证/测试 |
| 原始打桩声信号降噪实验 | 原始打桩声信号  + 海洋背景噪声 | 测试 |
| 非训练类型声信号降噪实验 | 桥梁振动声信号  + 海洋背景噪声 | 测试 |

### 5.1.2 数据集构建与来源

本研究依据“目标信号采集—背景噪声采集—子集设计—信号混叠”四个步骤构建实验数据集，旨在系统评估所提出自监督降噪模型在复杂海洋噪声环境下的适应性与鲁棒性。针对低信噪比环境下脉冲信号提取与降噪的问题，分别以打桩声与桥梁振动声为对象，涵盖从强脉冲到弱脉冲信号的典型应用场景。

首先，目标工程声信号包括海上打桩声与跨海桥梁振动声两类，两者均采用B&K 8105型水听器进行采集，采样率为65536 Hz，精度为32 bit。具体布放方案与现场参数已在第2.3节详细阐述。为满足后续网络训练中对片段化输入的需求，采集的打桩声依据单次脉冲事件切片，平均时长约0.8 s，最终得到4887段脉冲型打桩声样本；桥梁振动声按照3 s为单位切片，得到3544段弱脉冲型振动声样本，二者共同构成降噪建模与泛化测试的基础数据资源。

其次，为保证背景噪声样本的多样性与代表性，在相邻海域非施工时段布设同型号水听器，进行长周期静态采集，涵盖春夏秋冬不同季节及昼夜时段。总采集时长约30小时，采样现场示意如图5-1所示。采集完成后，将原始背景噪声数据按5秒为单位切片，构建了海洋背景噪声数据库 。其中样本总数 ，为后续数据集构建提供充足的海洋背景噪声样本。

|  |  |  |
| --- | --- | --- |
|  |  |  |
| 230320杏林2  (a) | 230418莆田  (b) | 269607fd6070b8023e50c44bc00b452  (c) |
| 230926厦门2  (d) | 240319厦门2  (e) | 图片2  (f) |
| 240902厦门2  (g) | 241022厦门  (h) | 241203厦门  (i) |
| 图 5-1海洋背景噪声采集场景 | | |

基于上述信号资源，结合强弱脉冲信号的实验设计要求，构建了三类结构化数据集，以支撑不同测试需求：

首先，构建累计能量段打桩声数据集。为突出主脉冲成分并减少尾波干扰，该子集从原始打桩声 ，中提取归一化累计能量5 % –95 % 区间，方法如下：设每采样点能量为 ，其归一化累计能量定义为：

|  |  |  |
| --- | --- | --- |
|  |  | (5-1) |

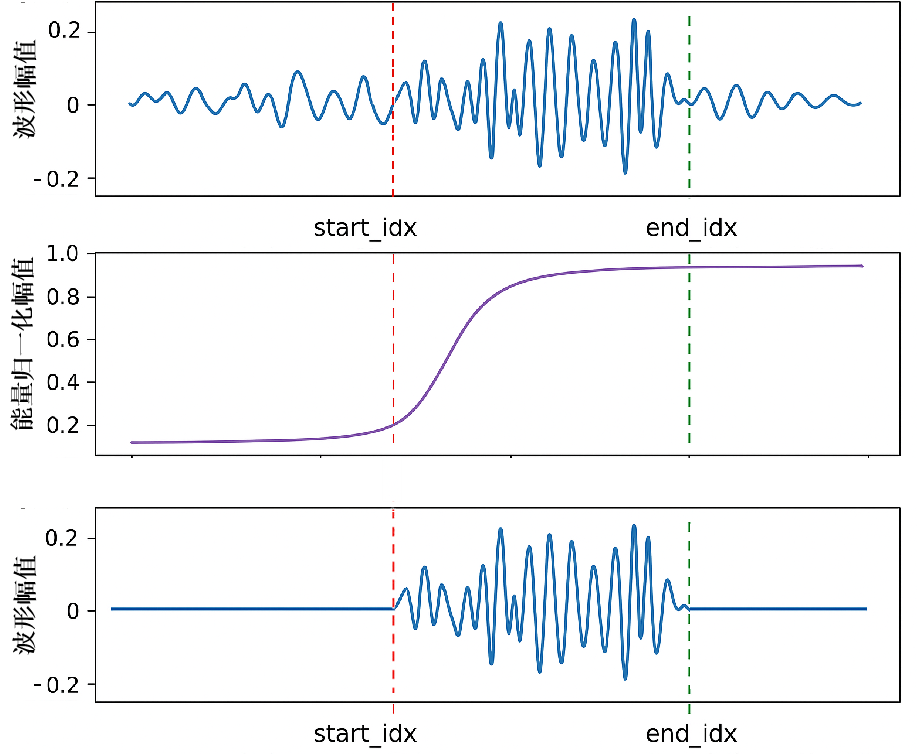
设起始点 和终止点 满足：

|  |  |  |
| --- | --- | --- |
|  |  | (5-2) |

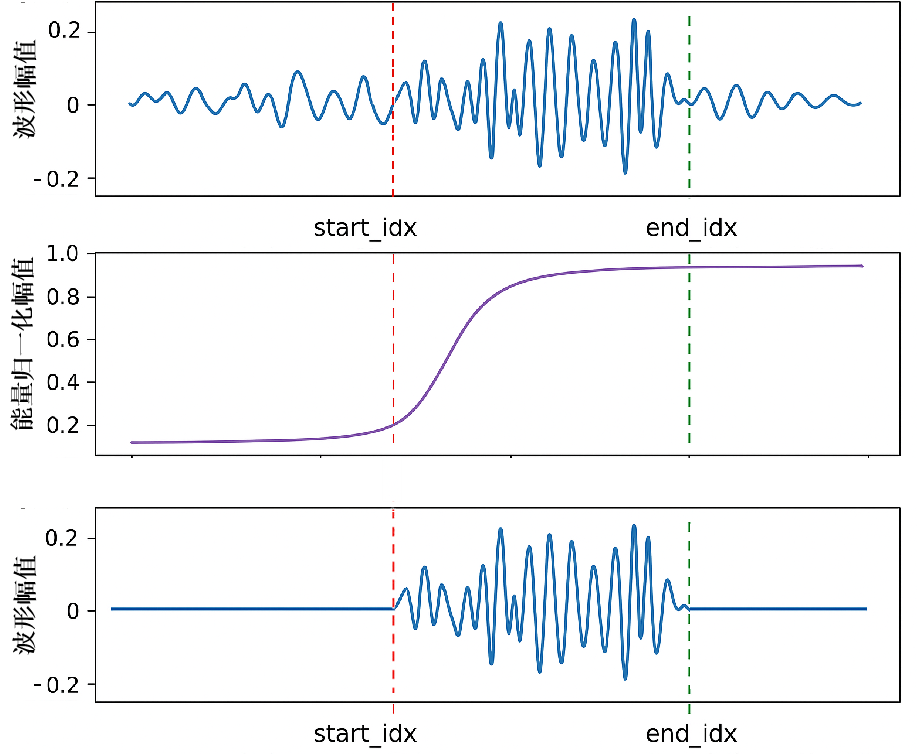
则构建主能量波形段 ：

|  |  |  |
| --- | --- | --- |
|  |  | (5-3) |

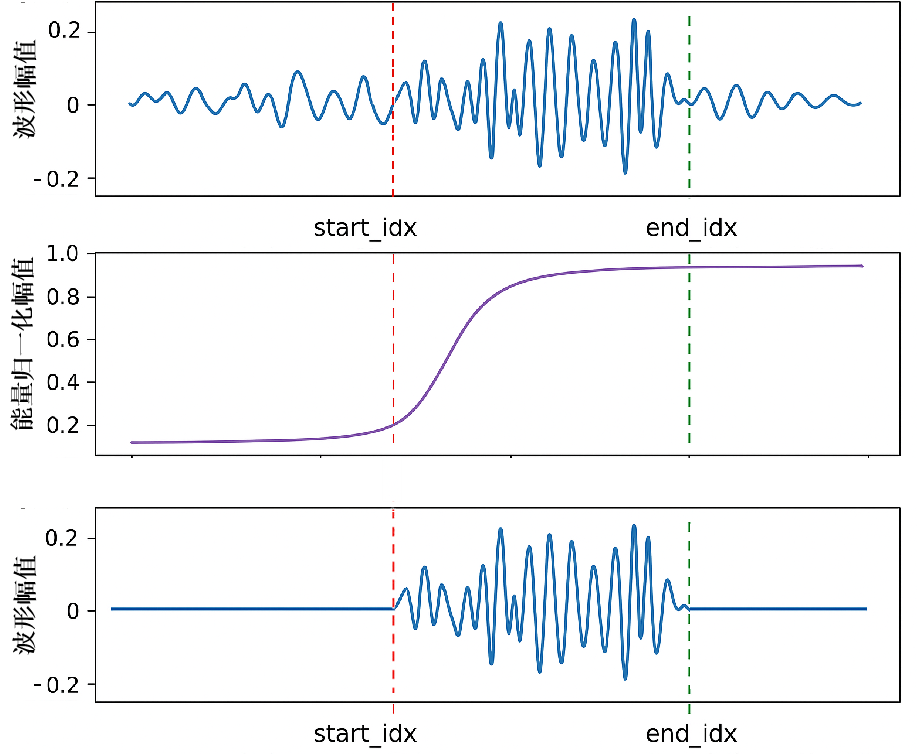
如图5‑2所示，上图为原始波形，中图为累计能量曲线，下图为提取后的有效主脉冲段。最终得到4887段主脉冲信号，按8:1:1比例划分为训练集、验证集与测试集，是模型训练的数据来源。此部分数据通过叠加背景噪声，控制信噪比（−10 dB、−5 dB、0 dB、5 dB、10 dB），模拟远距离传播引起的信噪比劣化环境，重点评估模型在低信噪比下对强脉冲结构的恢复能力。



(a)



(b)



(c)

图 5-2：打桩信号有效能量段提取过程：（a）原始信号；（b）累计能量曲线；（c）截取后的有效信号段

其次，构建原始打桩声数据集。该子集保留完整波形及尾波振铃，采用与累计能量段打桩声相同的采样段数（4887段），仅用于外部测试。信号同样叠加背景噪声并设置信噪比分档，与训练集一致，确保训练与测试条件对应，重点考察模型在复杂尾波干扰与背景噪声叠加条件下的整体降噪效果与工程适应性。

最后，构建桥梁振动声数据集。直接采用3544段原始振动声信号样本，不再进行能量截取，也不与背景噪声混叠。由于桥梁振动声本身能量水平较低、初始信噪比较差，天然构成了弱脉冲信号的低信噪比测试环境。该数据集用于检验模型在未见弱脉冲类型下的迁移能力与泛化鲁棒性。

在信号混叠过程中，仅对前两类打桩声数据集（累计能量段与原始波形）进行噪声叠加。具体步骤为：首先利用PyDub[86] 工具对背景噪声片段进行截断或循环拼接，使其时长与目标信号一致；随后根据设定信噪比 dB 计算比例系数：

|  |  |  |
| --- | --- | --- |
|  |  | (5-4) |

并生成混合信号：

|  |  |  |
| --- | --- | --- |
|  |  | (5-5) |

其中，每段背景噪声 从海洋背景噪声库 中随机抽取。对于累计能量段打桩声与原始打桩声中来源于同一脉冲事件的样本，混叠时保持使用相同的背景噪声片段和信噪比设置，以保证对比评估的一致性。

至此，本研究构建了涵盖多种信噪比工况与脉冲特性差异的实验数据集，为后续模型训练、性能验证与跨场景适应性评估奠定了坚实的数据基础。各数据集构建方式与用途汇总见表5-2。

表5-2数据集构建与属性说明

|  |  |  |  |
| --- | --- | --- | --- |
| 数据集名称 | 构建方式 | 样本总数 | 数据用途 |
| 累计能量段打桩声数据集 | 混叠背景噪声 | 4887 | 训练/验证/测试 |
| 原始打桩声信号数据集 | 混叠背景噪声 | 4887 | 测试 |
| 桥梁振动声信号数据集 | 原始信号 | 3544 | 测试 |

### 5.1.3 实验环境与预处理流程

本实验在高性能计算服务器上进行，以确保模型训练和推理的计算效率。服务器配备NVIDIA RTX 4090 GPU，采用Linux Ubuntu 20.04操作系统，并运行Python 3.8作为开发环境。深度学习框架选用 PyTorch 1.10.0，并结合CUDA 11.3进行并行加速计算，以优化大规模数据处理性能。在模型训练完成后，最佳模型的参数将用于测试集推理，以评估其在不同数据集上的降噪性能。

数据预处理阶段，为确保所有音频样本具有一致的输入尺寸，所有音频数据统一截取或零填充为65280个采样点，再将整体信号幅值按固定比例因子统一缩小，以规范输入尺度、提升训练稳定性。

随后，缩小的信号再经STFT转换为复数频谱，便于网络学习频谱特征。该转换过程利用PyTorch中的torch.stft函数实现，具体变换公式如下：

|  |  |  |
| --- | --- | --- |
|  |  | (5-6) |

其中， 表示在时间帧 和频率索引 处的复数频谱， 为输入音频信号， 为汉宁窗（Hanning Window），窗口长度设置为 1022，步长设置为256，以保证时间分辨率和频率分辨率的平衡。STFT转换得到的复数频谱被拆分为实部和虚部，并分别作为网络输入的两个通道，以适应AMFNet复数卷积网络的输入格式。

在模型训练采用Adam优化器，初始学习率设置为0.001。若验证集损失在一个epoch内无显著下降，则将学习率按衰减因子0.1递减；连续5个epoch无改进时触发早停以防止过拟合。批大小设为16，在保证 GPU 显存利用率的同时兼顾模型的泛化性能。每个epoch结束后计算验证损失并更新最优权重，直至满足早停条件或达到预设训练轮次。

### 5.1.4 评价指标

为全面评估所提出自监督降噪方案在降噪任务中的表现，本文采用了多种主流且具有代表性的信号质量评估指标。考虑到水下声信号的非平稳特性及打桩声的瞬态冲击特征，本节选用的评价指标不仅关注整体信号增强效果，还兼顾信号保真度与重建误差的量化，具体包括信噪比（SNR）、峰值信噪比（PSNR）和均方根误差（RMSE）。

（1）信噪比（Signal-to-Noise Ratio, SNR）

SNR是衡量处理后信号中有用成分相对于噪声成分的比例，定义如下：

|  |  |  |
| --- | --- | --- |
|  |  | (5-7) |

其中， 为原始干净信号， 为去噪后信号。SNR越高，说明信号中的噪声成分越少，去噪效果越好。

（2）峰值信噪比（Peak Signal-to-Noise Ratio, PSNR）

衡量最大信号值与重建误差之间的比例，尤其敏感于局部异常和信号峰值的还原质量，计算公式如下：

|  |  |  |
| --- | --- | --- |
|  |  | (5-8) |

其中， 表示信号可能的最大幅值， 为均方误差。对于具有明显脉冲或峰值特征的打桩信号， 能有效评估模型对关键信息的保留能力。

（3）均方根误差（Root Mean Square Error, RMSE）

是衡量去噪信号与真实信号之间整体偏差的重要指标，能有效反映模型重建误差的绝对水平。其计算公式如下：

|  |  |  |
| --- | --- | --- |
|  |  | (5-9) |

值越小，表示模型的重建误差越低，性能越优。

综上，实验通过上述三项指标从多角度对模型性能进行量化评估，以保证对自监督去噪模型在复杂海洋噪声背景下的有效性做出全面而可靠的判断。所有波形信号在预处理阶段，统一通过固定比例因子进行幅值缩放。该操作不会影响SNR和PSNR的计算结果，但会对RMSE产生数量级上的缩小效应。

## 5.2 实验结果

### 5.2.1 累计能量段打桩声降噪实验

为系统评估本研究提出的方法在低信噪比条件下提取强脉冲信号的性能，本节基于第5.1.2节中构建的累计能量段打桩声数据集开展降噪实验。目标信号为通过SEL分析提取的90%累计能量波形段，噪声源为实际采集的海洋背景噪声。通过在近距离实测打桩声信号上叠加不同强度的背景噪声，模拟形成−10 dB、−5 dB、0 dB、5 dB和10 dB五种初始信噪比条件，覆盖了从极端恶劣到较佳的海洋背景噪声环境，旨在复现打桩声在远距离传播过程中信噪比劣化后的提取挑战。

采用本研究提出的方法对各初始信噪比条件下的测试信号进行降噪处理，并通过时域波形、频域频谱以及帧级信噪比曲线的可视化对比，系统分析降噪效果。设置多档混叠信噪比进行测试，旨在全面考察所提方法在不同噪声水平下的鲁棒性与有效性，为后续工程应用提供可靠的实验依据。

|  |  |
| --- | --- |
| (a) | (b) |
| (c) | (d) |
| (e) | (f) |
| 图 5-3 SNR = -10 dB 条件下累计能量段打桩声信号的波形图与时频谱图: （a）目标信号波形图；（b）目标信号时频谱图；（c）混叠信号波形图；（d）混叠信号时频谱图；（e）降噪信号波形图；（f）降噪信号时频谱图 | |

首先，通过时域波形对比分析降噪前后的信号变化，结果如图5-3所示。如图 5-3（c）所示，在极端噪声条件下（SNR = -10 dB），原始打桩声的波形几乎被背景噪声完全湮没，冲击脉冲淹没于剧烈的随机波动中，难以辨识明显特征。如图 5-3（e）所示，经降噪处理后，波形变得平稳有序，背景噪声显著削弱，恢复了打桩冲击信号清晰的脉冲结构。在频谱域分析中，对比图 5-3（d）和图5-3（f），混叠信号频谱呈现出宽带随机噪声特性，主频成分被背景掩盖；降噪后，频谱能量重新集中在打桩冲击的主要频率区间，主频峰值显著突出，频谱结构得到有效恢复。

|  |  |
| --- | --- |
| (a) | (b) |
| (c) | (d) |
| (e) | (f) |
| 图 5-4 SNR = 0 dB 条件下累计能量段打桩声信号的波形图与时频谱图: （a）目标信号波形图；（b）目标信号时频谱图；（c）混叠信号波形图；（d）混叠信号时频谱图；（e）降噪信号波形图；（f）降噪信号时频谱图 | |

类似地，在信噪比0 dB的条件下，图5-4展示了SNR = 0 dB 混叠条件下的降噪结果。如图 5-4（c）所示，原始波形虽依稀可见打桩冲击的轮廓，但仍叠加有明显的噪声起伏；如图 5-4（d）所示降噪后，背景噪声干扰被显著消除，频谱脉冲结构的起止边界变得更加分明。由图 5-3 和图 5-4 可以看出，即使在极低信噪比情况下模型仍有出色的降噪效果。对比分析图 5-4（d）和图5-4（f）的频域特征以进一步评估降噪对信号频谱结构的影响，原始频谱可观察到打桩脉冲结构些许中低频能量部分，但仍有较高的背景噪声能量分布于各频段；降噪后，海洋背景噪声得到大幅削减，信号的主能量频带更加凸显，整个频谱结构变得干净。由频域对比结果可见，本方法可在不同信噪比条件下有效滤除宽频带噪声干扰，显著提升打桩声信号频谱的清晰度和主要成分的辨识度：即使在极低信噪比下依然能够恢复出信号频谱的主体结构特征，在中高信噪比下则进一步清除残余噪声使频谱更加纯净。

为了进一步从时间局部角度量化降噪效果，图 5-5 绘制了两种典型混叠条件（SNR = -10 dB与0 dB）下的帧级信噪比（Frame-Level SNR）随时间变化曲线。蓝色折线表示混叠信号，绿色折线表示降噪信号，红色阴影区域标注打桩主脉冲发生时段（约0.45至0.55秒）。

|  |
| --- |
|  |
| (a)    (b)  图 5-5不同混叠条件下帧级SNR可视化对比图（红色区域为打桩段）：（a）SNR = -10 dB；（a）SNR = 0 dB |

在主脉冲前后仅含背景噪声的静默区段（t < 0.45 s 与t > 0.55 s），混叠信号的帧级 SNR 长期维持在 -80 dB（-10 dB 混叠场景）至 -70 dB（0 dB 混叠场景）之间。经过降噪处理后，该区段SNR整体抬升至-50 dB ~ -40 dB，平均增益可达20 至35 dB；与此同时，曲线波动幅度显著收敛，反映出算法对低能量背景噪声具有一致且持续的抑制能力。在红色阴影标注的冲击区段，混叠信号的帧级 SNR 最高仅约5 dB（-10 dB 混叠场景）与 10 dB（0 dB 混叠场景），仍受尾波噪声拖尾影响。降噪后，该段SNR被进一步提升至20 dB，增益约10 dB，表明本方法能够在强化主脉冲能量的同时有效抑制低能量背景噪声。

表 5-3 进一步列出了不同初始信噪比条件下降噪后的定量性能指标。在最极端的 -10 dB 场景下，输出SNR由 -10.00 dB 显著提升13.07 dB至3.07 dB；对应PSNR增益 13.48 dB，RMSE 则由0.0173 降至 0.0037，误差削减约79 %。当输入SNR提升至 -5 dB 与 0 dB 时，模型仍分别取得 11.53 dB 与 10.25 dB 的SNR提升，同时将RMSE进一步压缩73 % 与 78 %；与之匹配的PSNR增益亦维持在 11.50 dB 与 12.52 dB。即便在原始信噪比较高的5 dB 与 10 dB条件下，算法仍能挖掘出8.48 dB 与 6.78 dB的额外SNR余量，并将RMSE进一步降低73 % 与 65 %，充分说明模型对残留噪声的深度抑制能力。

表5-3 模型降噪性能指标对比（累计能量段信号）

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 混叠信噪比(dB) | 类型 | SNR (dB) | ΔSNR (dB) | PSNR (dB) | ΔPSNR (dB) | RMSE | ΔRMSE |
| -10 | 混叠信号 | -10.00 | — | 15.22 | — | 0.0173 | — |
| **降噪信号** | **3.07** | **+13.07** | **28.70** | **+13.48** | **0.0037** | **-0.0136** |
| -5 | 混叠信号 | -5.00 | — | 20.21 | — | 0.0097 | — |
| **降噪信号** | **6.53** | **+11.53** | **31.71** | **+11.50** | **0.0026** | **-0.0071** |
| 0 | 混叠信号 | 0.00 | — | 23.90 | — | 0.0064 | — |
| **降噪信号** | **10.25** | **+10.25** | **36.42** | **+12.52** | **0.0014** | **-0.0050** |
| 5 | 混叠信号 | 5.00 | — | 28.65 | — | 0.0037 | — |
| **降噪信号** | **13.48** | **+8.48** | **39.65** | **+11.00** | **0.0010** | **-0.0027** |
| 10 | 混叠信号 | 10.00 | — | 33.12 | — | 0.0020 | — |
| **降噪信号** | **16.78** | **+6.78** | **42.65** | **+9.53** | **0.0007** | **-0.0013** |

综上所述，本节实验结果充分验证了所提出方法在近距离实测打桩声叠加背景噪声模拟的远距低信噪比场景中，具有显著的脉冲结构恢复与背景噪声抑制能力。无论是在-10 dB等低信噪比的严苛噪声条件下，还是在10 dB等较高信噪比的背景下，均能够有效提升信号可辨识度，恢复冲击信号主体特征，体现了方法在复杂水下噪声环境中实际应用的良好潜力与研究价值。

### 5.2.2 原始打桩声信号降噪实验

为进一步系统检验所提出方法在原始海洋工程水下噪声号输入下的适应性，本节基于第5.1.2节中构建的原始打桩声数据集开展降噪实验。目标信号直接采用近距离实测的完整打桩脉冲波形（含主脉冲与尾波结构），噪声源为实际采集的海洋背景噪声。通过叠加不同强度的背景噪声，构建信噪比为 -10  dB、-5  dB、0  dB、5  dB、10  dB的五组混叠测试集，以模拟打桩声在远距离传播后信噪比劣化的实际观测条件。

与前述累计能量段打桩声实验相比，原始信号在保留完整能量结构的同时，包含复杂的尾波混响和多径效应，显著增加了信号提取与降噪的挑战。实验旨在评估模型在更高难度场景下的性能，重点考察其对打桩声主要冲击脉冲的恢复能力及对多径尾波干扰的抑制能力，从而全面验证方法的实用性与鲁棒性。

图 5-6 展示了在SNR = -10 dB 条件下的时域波形与频谱对比。其中，图5-6（a）为原始目标信号时域波形，呈现清晰的冲击脉冲结构，而中后段尾波因多径反射而较为复杂。如图 5-6（c）为混合海洋噪声后的含噪信号时域波形，主脉冲已被强烈噪声干扰所淹没，尾部轮廓变得模糊不清。如图 5-6（e）所示，模型输出的降噪信号准确恢复了前部的主脉冲形态：前段能量突发被清晰还原，主能量段突显，中后段尾波成分显著减弱，时域轮廓更加干净。可以观察到模型能够在极低信噪比下提取出信号的主要冲击事件，并有效抑制尾波部分的多余振荡。图 5-6（b）展示了目标信号的频谱，可见主冲击频率段能量集中，而尾波引起的高频能量呈现扩散特征。图 5-6（d）中含噪信号的频谱由于低信噪比影响，能量分布更加模糊，杂散频率成分增多。相应地，见图 5-6（f），降噪后频谱的能量重心重新集中：保持了主脉冲频率带宽，抑制了尾波所带来的高频扩散。尤其在中高频段，模型显著削弱了非结构化噪声成分，使频谱轮廓更加紧凑、清晰。由此可见，模型在 -10 dB 条件下成功提高了信号主脉冲的时频特征强度，同时有效降低了尾波导致的干扰。

|  |  |
| --- | --- |
| (a) | (b) |
| (c) | (d) |
| (e) | (f) |
| 图 5-6 SNR = -10 dB 条件下原始打桩声信号的波形图与时频谱图: （a）目标信号波形图；（b）目标信号时频谱图；（c）混叠信号波形图；（d）混叠信号时频谱图；（e）降噪信号波形图；（f）降噪信号时频谱图 | |

图 5-7 展示了SNR = 0 dB条件下的时域和频谱对比。图 5-7（a）为原始目标信号时域波形，其前部冲击脉冲依然突显，但尾波能量略大。图 5-7（c）为对应的含噪信号时域波形，相较于 -10 dB条件下的结果，主脉冲在噪声中更加可辨，虽然仍存在一定程度的噪声覆盖。如图 5-7（c）所示，模型输出在该条件下继续准确恢复了主脉冲结构，并进一步抑制了尾部残余噪声，使前后段过渡更为平滑干净。与 -10 dB情形相比，由于输入信噪比提高，降噪后信号的主脉冲与目标信号之间的吻合度显著增强，误差降低。在频谱方面，图 5-7（b）中目标信号的频谱依然主要集中于主冲击事件对应频段。图 5-7（d）所示含噪信号频谱保留了主要频率成分，但可见一些噪声引起的次要杂散分布。如图 5-7（f）所示，模型输出的频谱与目标信号频谱高度一致：主要频率带宽得到很好保持，高频非结构化能量成分进一步减少，频谱轮廓更加紧凑。总体来看，在0 dB条件下模型的降噪效果与 -10 dB情况下相似，但由于原始信噪比较高，输出信号更接近目标，噪声影响更小，信噪比提升幅度相对略有减小。

|  |  |
| --- | --- |
| (a) | (b) |
| (c) | (d) |
| (e) | (f) |
| 图 5-7 SNR = 0 dB 条件下原始打桩声信号的波形图与时频谱图: （a）目标信号波形图；（b）目标信号时频谱图；（c）混叠信号波形图；（d）混叠信号时频谱图；（e）降噪信号波形图；（f）降噪信号时频谱图 | |

图 5-8 对比了不同信噪比条件下的帧级SNR变化情况，图 5-8（a）对应 -10 dB，图 5-8（b）对应0 dB。曲线中蓝色为原始含噪信号的帧级SNR，绿色为降噪后信号，红色阴影表示打桩主脉冲发生的帧段。无论何种输入信噪比，模型均能显著提高帧级SNR。在背景噪声段，模型将帧级SNR稳定提升约20 dB；在打桩主脉冲段附近，提升尤为明显： -10 dB条件下原始信号帧级SNR峰值不足0 dB，经降噪后提高到约 +15 dB；0 dB条件下则从约15 dB提升到 +30 dB左右。值得注意的是，在尾波段，降噪后帧级SNR迅速回落到稳定水平，但没有产生额外的伪影拖尾。这些结果表明，模型对打桩信号的关键帧段（红色区域）赋予了更高的能量权重，同时保持了尾波段的相对平滑。

|  |
| --- |
|  |
| (a) |
| (b) |
| 图 5-8不同混叠条件下帧级SNR可视化对比图（红色区域为打桩段）：（a）SNR = -10 dB；（b）SNR = 0 dB |

表 5-4 列出了五种输入信噪比条件下的定量对比指标，与表 5-3 的累计能量段相比，此处信号完整保留了尾波与弱幅度拖尾，因此评测难度更高，但算法仍在 -10 dB 至 10 dB 的五种混叠条件下取得了显著而稳定的性能收益，具体分析如下：在最不利的 -10 dB 场景中，输出SNR由 -10.00 dB 抬升12.74 dB 至 2.74 dB；同时，PSNR 增益 13.07 dB，RMSE从 0.0180 降至 0.0040，误差削减 78 %。当输入SNR提升至 -5 dB 与 0 dB 时，算法分别贡献 11.49 dB 和 9.12 dB 的额外 SNR，并将RMSE 进一步压缩 74 % 与 68 %，PSNR 增益仍保持在约11 dB。对于 5 dB 与 10 dB 的中高信噪比条件，模型仍能提升7.84 dB 与6.32 dB的 SNR，RMSE进一步下降 71 % 与 64 %，证明其对背景噪声的抑制能力。

表5-4 模型降噪性能指标对比（原始打桩声信号）

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 混叠信噪比(dB) | 类型 | SNR (dB) | ΔSNR (dB) | PSNR (dB) | ΔPSNR (dB) | RMSE | ΔRMSE |
| -10 | 混叠信号 | -10.00 | — | 14.85 | — | 0.0180 | — |
| **降噪信号** | **2.74** | **+12.74** | **27.92** | **+13.07** | **0.0040** | **-0.0140** |
| -5 | 混叠信号 | -5.00 | — | 19.85 | — | 0.0101 | — |
| **降噪信号** | **6.49** | **+11.49** | **31.67** | **+11.82** | **0.0026** | **-0.075** |
| 0 | 混叠信号 | 0.00 | — | 24.10 | — | 0.0060 | — |
| **降噪信号** | **9.12** | **+9.12** | **34.28** | **+10.18** | **0.0019** | **-0.0041** |
| 5 | 混叠信号 | 5.00 | — | 27.56 | — | 0.0041 | — |
| **降噪信号** | **12.84** | **+7.84** | **38.17** | **+10.61** | **0.0012** | **-0.0029** |
| 10 | 混叠信号 | 10.00 | — | 31.95 | — | 0.0025 | — |
| **降噪信号** | **16.32** | **+6.32** | **41.68** | **+9.73** | **0.0009** | **-0.0016** |

综上，本节实验结果表明，即使在极低信噪比与复杂尾波混响条件下，本研究提出的方法仍能准确提取打桩声主冲击脉冲，显著抑制多径尾波和背景噪声干扰。验证了方法在更高复杂度真实环境下的稳健性与普适性，进一步支撑其在实际海洋工程声环境中的应用潜力。

### 5.2.3 非训练类型声信号降噪实验

为进一步检验所提出方法在非训练场景中的泛化能力与类型迁移适应性，本节选取实际海洋工程中另一类典型水下噪声源——水下桥梁振动声信号作为测试对象，输入已训练完成的模型进行推理，并分析其降噪效果。该类信号具有突发性、瞬时性及多频成分特征，且其时域形态、频谱结构与背景环境均未在模型训练阶段出现，因此可用于系统评估方法在未知类型脉冲事件下的适应表现与泛化性能。

考虑到实验样本为实际采集的自然混合信号，缺乏对应的干净参考信号，故本节无法采用SNR、PSNR、RMSE等传统定量指标进行性能评估。为确保分析的科学性与工程实用性，本研究采用可视化分析的方法，综合评估模型在结构保持、背景干扰抑制与尾波削弱等方面的效果，并结合桥梁结构声传播特点与脉冲事件识别需求，重点考察模型在未训练场景下能否有效突出主能量段并压制背景扰动与多径尾波干扰。

从图 5-9（a）所示的原始信号波形可以看出，在1 s 的时间范围内出现了一系列明显的脉冲振动，伴随逐渐衰减的尾部震荡结构，但脉冲振动主波形被海洋背景噪声淹没，使得脉冲结构轮廓紊乱。相比之下，图 5-9（c）所示的降噪后波形保留了脉冲振动的主结构，同时显著抑制了尾部的多余成分，尾波变得更加平缓清晰；整个时域波形的背景噪声干扰也明显降低，体现出较高的信噪比。

|  |  |
| --- | --- |
| (a) | (b) |
| (c) | (d) |
| 图 5-9：水下桥梁振动声信号降噪结果: （a）原始信号波形图；（b）原始信号时频谱图；（c）降噪信号波形图；（d）降噪信号时频谱图 | |

频域对比结果进一步验证了上述观察。如图5-9（b）所示，原始信号频谱显示在脉冲发生时段存在能量集中的主频带，但后续存在大量分散的低能量杂散频率成分，反映了背景噪声与结构混响的叠加效应。图 5-9（d）展示了降噪后频谱情况，主频能量得以有效保留并突出显示，杂散频率成分显著削弱，频谱能量集中度显著提升，频谱轮廓变得更加紧凑清晰。上述对比结果表明，模型在未见过的水下桥梁振动声信号上仍能准确提取其主要能量结构并有效抑制噪声。

从工程声学监测的角度来看，有效提取此类脉冲信号中的主能量段对于声场计算、传播路径分析、异常检测及海洋生态影响评估具有重要意义。尤其在未经过该类型信号训练的情况下，仍能实现脉冲主结构的恢复与背景干扰的显著抑制，进一步验证了本方法在实际复杂工程场景中具备良好的通用性与鲁棒性，为其在未来海洋工程声学监测与环境评估等领域的应用提供了有力支撑。

## 5.3 模型有效性与降噪方法对比分析

本章将对提出的模型进行进一步的有效性验证，通过设计不同实验场景，分别针对纯海洋背景噪声信号与典型海洋工程水下噪声信号（桥梁振动信号）进行降噪效果分析，并与传统降噪方法的表现进行对比，深入评估本文模型在工程实际应用场景中的优势。此外，通过消融实验验证模型内部模块的实际贡献，凸显所提出方法在架构设计上的科学性与有效性。

### 5.3.1 模型有效性对比分析

为了进一步验证本文所提出自监督深度学习模型在实际复杂环境下的有效性，本文设计了两组对比实验。在纯海洋背景噪声环境下进行第一组实验，以评估模型的噪声抑制性能；在典型海洋工程水下噪声（桥梁振动声）环境下进行第二组实验，该信号包含突发冲击成分且未参与模型训练，用以考察模型对非训练信号的泛化适应性以及对脉冲特征的恢复能力。两组实验均将本文模型与传统降噪方法（维纳滤波、小波变换）的结果进行对比，通过分析时域波形和频谱（见图5-10至图5-13），全面评估各方法的降噪效果和信号保真度。

在纯海洋背景噪声实验中，各方法降噪后的时域波形对比如图5-10所示。如图 5-10（a）所示，原始噪声信号呈现显著的随机波动和突发尖峰干扰，整体上呈典型的非平稳随机噪声特性。如图 5-10（b）所示，维纳滤波仅对背景噪声的随机波动实现了轻微抑制，信号中仍存在明显的随机扰动成分，未能显著提升信号整体的平稳性。如图 5-10（c）所示，小波变换同样未能有效减弱随机波动，波形依然较为紊乱，降噪效果有限。如图 5-10（d）所示，相比之下，本文模型降噪后的波形变得非常平稳，随机噪声被大幅度压制，波动幅度显著减小，展现出极佳的噪声抑制效果和工程适用价值。

|  |  |
| --- | --- |
| (a) | (b) |
| (c) | (d) |
| 图5-10海洋背景噪声降噪时域波形图对比：（a）原始信号；（b）维纳滤波降噪；（c）小波变换降噪；（d）本文模型降噪 | |

图5-11展示了频谱结果。如图 5-11（a）所示，原始噪声随机扰动的频谱能量遍布整个频带。图 5-11（b）展示了维纳滤波降噪后的频谱，随机扰动特征依旧突出，反映了维纳滤波在非平稳随机背景噪声下的局限性。图 5-11（c）展示了小波变换降噪后的频谱，未见明显的结构改善。图 5-11（d）展示了本文模型降噪结果：随机扰动和高频噪声基本消除，频谱呈现出稳定且干净的结构。

|  |  |
| --- | --- |
| (a) | (b) |
| (c) | (d) |
| 图5-11海洋背景噪声降噪时频谱图对比：（a）原始信号；（b）维纳滤波降噪；（c）小波变换降噪；（d）本文模型降噪 | |

接下来，为进一步验证本文模型对非训练信号的适应性与实用性，选取一段实际采集的水下桥梁振动声信号进行测试。该信号通常伴有海洋背景噪声和结构传播干扰，表现为非周期性的突发冲击与背景拖尾并存，结构不稳定、干扰复杂，具有明显的非平稳性和脉冲性特征。

|  |  |
| --- | --- |
| (a) | (b) |
| (c) | (d) |
| 图 5-12 海洋背景噪声降噪时域波形图对比：（a）原始信号；（b）维纳滤波降噪；（c）小波变换降噪；（d）本文模型降噪 | |

图 5-12 给出了各方法处理该信号后的时域波形对比。如图 5-12（a）所示，原始信号包含多个幅度不一的冲击波，并夹杂有持续性的背景噪声，导致信号结构边界模糊。见图 5-12（b），维纳滤波处理后，部分高频噪声得到抑制，但突发冲击结构仍显模糊，主能量段的辨识度不高。图 5-12（c）展示了小波降噪结果，在保留部分信号边缘的同时，对背景扰动的压制力度有限，且在多处引入了较高幅值的伪冲击信号，削弱了原始结构的清晰度。如图 5-12（d）所示相比之下，本文模型处理后的波形清晰突出地呈现了各次冲击事件的波动，信号主体成分明显，起止边界分明，同时背景波动被更强地抑制，整个波形显得紧凑而清晰。

|  |  |
| --- | --- |
| (a) | (b) |
| (c) | (d) |
| 图5-13海洋桥梁振动声信号降噪时频谱图对比：（a）原始信号；（b）维纳滤波降噪；（c）小波变换降噪；（d）本文模型降噪 | |

频谱分析结果如图5-13所示，进一步比较了各方法的频域降噪效果。如图5-13（a）所示，原始信号的频谱呈现出多个短暂冲击所对应的频带叠加现象，但同时存在大量高频杂散成分，频谱能量分布较为分散。图 5-13（b）展示了维纳滤波后的频谱，削弱了部分高频能量，但未能聚焦出真实的频率结构，频谱仍然混乱。如图5-13（c）所示，小波降噪后的频谱结构略有收敛，但信号主体频带仍被背景噪声淹没，难以提供清晰的事件特征。如图5-13（d）所示，相比之下，本文模型的频谱结果展现出良好的频谱聚焦特性：冲击信号成分在频域中被完整保留，背景噪声的能量显著降低，主频段结构清晰突出，尤其在低频段表现出更高的能量集中度，说明本模型在频域分离与信号还原方面具有更优越的能力。

通过联合观察图 5-13 和图 5-14，可以进一步发现时域约 0.85 s处存在一次明显的非平稳背景噪声扰动。该时刻的频谱缺乏清晰的主频结构，属于典型的海洋背景噪声波动。在原始波形中，这一扰动以突增形式出现，干扰了对真实信号的判断。传统方法（维纳滤波、小波变换）均无法有效区分该背景扰动与真实冲击信号，导致残余干扰被错误保留甚至误判为信号成分。而本文模型在该区域表现出很强的干扰抑制能力，成功压制了这一无效扰动，使主信号的结构边界更加清晰，且在频谱中未出现对应的杂散频带，进一步体现了模型在时频协同降噪方面的优势。

综上所述，通过以上两种典型背景下的对比实验可以看出，本文所提出的降噪模型在各项性能上均明显优于传统方法。面对非平稳的随机海洋背景噪声，本文模型能够有效抑制噪声并显著提高信号的时域平稳性和频谱清晰度；在处理未参与训练的突发脉冲信号时，模型仍能准确保留信号的主要冲击特征并消除背景干扰，表现出良好的结构保真能力和泛化适应性。相较之下，维纳滤波和小波变换等传统方法在上述复杂环境中易出现噪声残留和信号结构模糊等问题，难以满足实际工程应用需求。综合而言，本文自监督深度学习模型不仅具备卓越的噪声抑制和脉冲特征恢复能力，还具有高度的适应性与鲁棒性，展现出在复杂水下噪声环境中的广泛应用潜力和综合优势。

### 5.3.2 模块消融与多方法对比分析

为进一步验证本文模型结构设计的合理性与性能表现的先进性，本节在累计能量段打桩信号数据集（混叠信噪比为0 dB）上开展模块消融实验与多方法对比实验，并在比较中引入原始含噪信号的性能指标作为参考基准，以更直观展示各模型的降噪能力。

表5-5展示了本文模型在累计能量段打桩信号 (输入SNR = 0 dB) 条件下进行模块消融实验的性能指标对比。可以看出，相比原始噪声信号（SNR约为0 dB，PSNR仅23.90 dB，RMSE高达0.0064），完整模型AMFNet（方案A） 的降噪性能提升最为显著：输出信号的SNR提高到10.25 dB，PSNR升至36.42 dB，分别较原始噪声提升了约10.25 dB和12.52 dB；同时RMSE降至0.0014，较原始噪声减少0.0050，说明降噪后信号误差显著降低。相比之下，当依次移除模型中的关键模块时，降噪性能均出现不同程度的下降。其中，去除“SDFU”模块的方案B 输出SNR为8.31 dB（较方案A降低约2 dB），PSNR为33.66 dB（降低约2.8 dB），RMSE上升至0.0021；去除“CAM”模块的方案C 输出SNR降至7.54 dB，PSNR降为32.83 dB，RMSE增至0.0023；而去除“CMS-SCU”模块的方案D 的性能下降最为明显，输出SNR仅6.74 dB，PSNR为31.85 dB，RMSE升至0.0027。可见，删除任意模块都会导致降噪效果变差，其中以去除CMS-SCU模块的影响最大，SNR和PSNR降幅分别超过3.5 dB和4.5 dB，表明该模块对提升信噪比贡献突出；去除CAM模块次之，性能也有明显下降；去除SDFU模块的方案B虽性能略高于C和D，但相比完整模型仍有约2 dB的SNR降低和RMSE增大0.0007，同样体现出SDFU模块的重要作用。上述消融实验结果充分证明了模型各组成模块对降噪性能的积极贡献：各模块协同作用使得完整模型取得最佳效果，印证了本文所提模型结构设计的合理性和有效性。

表5-5 模型结构消融实验性能指标对比（累计能量段打桩信号，SNR = 0 dB）

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 模型 | 类型 | SNR (dB) | ΔSNR (dB) | PSNR (dB) | ΔPSNR (dB) | RMSE | ΔRMSE |
| — | 噪声信号 | 0.00 | — | 23.90 | — | 0.0064 | — |
| A | AMFNet | **10.25** | **+10.25** | **36.42** | **+12.52** | **0.0014** | **-0.0050** |
| B | 去除SDFU | 8.31 | +8.31 | 33.66 | +9.76 | 0.0021 | -0.0043 |
| C | 去除CAM | 7.54 | +7.54 | 32.83 | +8.93 | 0.0023 | -0.0041 |
| D | 去除CMS-SCU | 6.74 | +6.74 | 31.85 | +7.95 | 0.0027 | -0.0037 |

表5-6进一步列出了 本文模型与多种其他降噪方法 在相同信号条件下的性能对比。从中可以看出，传统降噪方法对该复杂噪声信号的改善非常有限：维纳滤波 输出SNR仅约2.37 dB，PSNR为26.35 dB，较原始噪声分别提高2.37 dB和2.45 dB，RMSE降至0.0038；小波变换方法SNR约提升1.22 dB，PSNR提高1.30 dB至25.20 dB，RMSE降为0.0048。相较而言，这两种传统方法的SNR/PSNR提升幅度都不足3 dB，降噪后信号仍残留较强噪声，说明它们难以有效应对非平稳随机噪声。相比之下，深度学习模型展现出显著更强的降噪能力：基于U-Net结构的降噪模型[28]输出SNR约为5.47 dB，PSNR提升至30.75 dB，已明显优于传统算法；另一种自监督模型NerNT[54]将SNR提高到7.18 dB，PSNR达32.42 dB；Noise2Noise方法（N2N）[55]则进一步取得约7.82 dB的SNR和33.07 dB的PSNR；此外，文献提出的ONT模型[56]表现更佳，输出SNR达到8.63 dB，PSNR为33.81 dB。可以发现，随着降噪模型的改进，深度学习方法的性能指标逐步提升，整体降噪效果远超传统方法。

表5-6 模型与其他方法降噪性能对比（累计能量段打桩信号，SNR = 0 dB）

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 类型 | 类型 | SNR (dB) | ΔSNR (dB) | PSNR (dB) | ΔPSNR (dB) | RMSE | ΔRMSE |
| — | 噪声信号 | — | — | 23.90 | — | 0.0064 | — |
| 传统方法 | 维纳滤波 | 2.37 | +2.37 | 26.35 | +2.45 | 0.0038 | -0.0026 |
| 小波变换 | 1.22 | +1.22 | 25.20 | +1.30 | 0.0048 | -0.0016 |
| 自  监  督  模  型 | U-Net[28] | 5.47 | +5.47 | 30.65 | +6.75 | 0.0030 | -0.0034 |
| NerNT[54] | 7.18 | +7.18 | 32.42 | +8.52 | 0.0025 | -0.0039 |
| N2N[55] | 7.82 | +8.31 | 33.07 | +9.76 | 0.0022 | -0.0042 |
| ONT[56] | 8.63 | +7.54 | 33.81 | +8.93 | 0.0020 | -0.0044 |
| **AMFNet** | **10.25** | **+10.25** | **36.42** | **+12.52** | **0.0014** | **-0.0050** |

在众多对比方法中，本文提出的AMFNet模型的性能指标突出地优于其它所有方法：其输出信号SNR达到10.25 dB，PSNR高达36.42 dB，即使对比先进的自监督降噪模型（如N2N或ONT），AMFNet仍实现了至少约1.5～2 dB的额外SNR提升和约2～3 dB的PSNR提升。这一系列对比结果表明，本文模型在噪声抑制和信号保真方面均取得了当前最优的表现：更高的SNR和PSNR意味着噪声被更彻底地消除、信号细节得到更完整的保留，而更低的RMSE则体现了降噪结果与原始信号的误差更小。

综上，通过模块消融实验和多方法对比实验验证本文所提模型的有效性和设计合理性。一方面，由于引入SDFU、CAM、CMS-SCU等模块的协同作用，才使得模型能够充分挖掘噪声信号特征并实现卓越的降噪性能。另一方面，与多种传统及深度学习降噪方法的对比显示，AMFNet模型在提高信噪比、峰值信噪比以及降低重构误差等方面均取得了最优成绩，在复杂水下噪声信号降噪任务中展现出显著优势。这些实验结果充分表明，本文提出的模型在降噪性能上具有极大提升，模块设计切实可行且贡献明确，不仅提升了水下工程噪声信号的降噪提取效果，也为相关领域提供了一种性能卓越的降噪提取新方案，具有重要的理论意义和工程应用价值。

## 5.4 本章小结

本章围绕低信噪比条件下海洋工程水下噪声的提取与降噪这一核心问题，设计并开展了一系列系统性的实验研究，旨在全面验证所提出的自适应多重聚焦降噪网络在复杂海洋声环境中的实际性能。

实验设置阶段，首先依据第二章对实际采集的打桩声与桥梁振动声信号特征的分析，构建了覆盖强脉冲与弱脉冲典型场景的三类实验数据集，包括累计能量段打桩声、原始完整打桩声和桥梁振动声信号。其中，累计能量段数据集截取了原始打桩声信号中累计能量达到90%的有效主脉冲段，通过叠加不同强度的海洋背景噪声模拟形成了从−10 dB到10 dB不等的初始信噪比场景，以复现远距离传播导致信号严重劣化的实际情形；原始完整数据集则保留了打桩信号整体的主脉冲与尾波结构；而桥梁振动声则天然构成了典型的低信噪比弱脉冲测试环境，为泛化性能验证提供了严苛条件。

在实验任务与性能评估方面，本章通过累计能量段、原始信号以及非训练类型信号的降噪实验，分别系统考察了所提模型在不同信号能量结构、不同信噪比水平及不同脉冲特性条件下的性能表现。具体而言，在累计能量段实验中，模型在极端恶劣条件（−10 dB）下的输出SNR相较输入提高了13.07 dB（从−10.00 dB提升至3.07 dB），RMSE指标显著降低约79%，信号从难以识别的状态被恢复到清晰可辨的结构；而在0 dB条件下，模型依然取得了10.25 dB的SNR提升，且有效重建了主脉冲频谱结构，背景噪声被显著抑制。在原始完整打桩声实验中，即便在未预处理的复杂尾波干扰条件下，−10 dB初始信噪比时的输出SNR也提高了12.74 dB至2.74 dB，RMSE同样降低了约78%，体现了模型对主脉冲结构恢复与尾波抑制的稳健能力。对于未参与训练的桥梁振动声信号，模型表现出良好的泛化适应性，准确恢复了原始信号中的突发脉冲结构并显著降低了背景与尾波干扰，进一步证明了方法在复杂实际场景中的可靠性与通用性。

此外，为深入验证所提模型结构的合理性与优越性，本章还分别开展了模块消融实验与多方法对比实验。消融实验表明，去除频谱动态聚焦单元（SDFU）、复数注意力模块（CAM）或复数多尺度特征融合单元（CMS-SCU）后，模型输出的SNR相较完整结构分别下降了约2 dB至4 dB不等，其中去除CMS-SCU模块影响最大，突显了各个模块在降噪性能中的关键作用。在与传统降噪方法（如维纳滤波、小波变换）及现有先进自监督模型（如U-Net、N2N、ONT）的对比实验中，所提出的AMFNet模型在0 dB初始信噪比条件下实现了最高的输出SNR（10.25 dB）与最低的RMSE（0.0014），且均优于其他自监督模型，充分证实了模型在信号结构保持与噪声抑制方面的显著优势。

综上所述，本章通过严格的实验设计和系统深入的性能分析，充分验证了本文所提出的AMFNet在低信噪比海洋工程场景中的有效性与先进性。实验结果表明，本方法能够在不同脉冲类型、不同初始信噪比条件下稳健实现水下噪声的有效抑制与目标信号结构的高保真恢复，体现出明显的应用潜力与工程价值，也为海洋工程实际声学监测提供了有力的技术支撑与参考依据。

# 第六章 总结与展望

## 6.1 论文的主要工作与创新点

本论文围绕低信噪比条件下海洋工程水下噪声信号的提取与降噪问题，针对传统降噪方法在复杂非稳态海洋背景干扰下结构保持与噪声抑制能力不足、现有深度学习方法严重依赖高质量监督样本且工程泛化性能有限等关键问题，提出了一种基于自监督深度学习的海洋工程水下噪声提取方法，并从数据集构建、网络模型设计到性能验证开展了系统深入的研究，构成了面向海洋工程实际需求的完整技术方案。具体而言，本论文的主要研究工作与创新点如下：

（1）基于实际采集的打桩声、桥梁振动声和海洋背景噪声，深入分析了海洋工程水下噪声的能量分布特征与背景噪声的非稳态扰动特性，明确了信号在低信噪比条件下的特征结构。并在此基础上提出了一种无需干净样本的自监督训练方法，通过结构化随机子采样与交叉掩码策略，在单通道水下噪声数据中生成伪配对信号，实现低信噪比场景下高效、可靠的自监督模型训练，为后续模型设计与性能验证提供了扎实的数据与方法支撑。

（2）针对低信噪比条件下海洋工程水下噪声主能量段提取与复杂尾波干扰抑制的难题，本文创新性地提出了自适应多重聚焦降噪网络（AMFNet）。该网络结合频谱动态聚焦单元（SDFU）、复数注意力机制（CAM）、语义重校准单元（SRU）以及复数多尺度特征融合单元（CMS-SCU），实现对海洋工程水下噪声在复杂非稳态的海洋背景噪声干扰下的精准特征提取与能量重构。

（3）设计了多层次、多场景的实验验证体系，包括累计能量段打桩声信号实验、原始未截断打桩声信号实验，以及非训练类型桥梁振动声信号的泛化验证实验。通过设置严格的信噪比分档条件和严苛的测试环境，实验体系构建了从模型结构恢复到实际工程性能评估的完整验证闭环。此外，还通过与维纳滤波、小波变换等传统方法及现有先进自监督降噪模型（如U-Net、Noise2Noise、ONT）进行全面对比，并辅以模块消融实验，深入论证了本文模型在结构保持、噪声抑制与泛化适应性上的突出优势。

综上所述，本论文在数据构建方法、网络模型设计与实验验证体系三个维度实现了系统创新，具体体现在以下三个方面：

（1）提出了一种无需干净样本的结构化自监督训练方法，通过结构化随机子采样与交叉掩码策略，在单通道水下噪声数据中生成具有能量守恒特性的伪配对信号，实现低信噪比场景下高效、可靠的自监督模型训练。

（2）创新性地设计了自适应多重聚焦降噪网络AMFNet，引入频谱动态聚焦、多尺度特征融合与复数域注意力机制，有效解决了自监督方法频谱结构细节丢失的难题，显著提升了模型在低信噪比条件下海洋工程水下噪声信号提取与降噪能力。

（3）搭建了真实海洋工程场景的混叠数据集，并配套了严谨的多方法对比、模块消融与多尺度指标评估实验体系，全方位验证了所提AMFNet在各类低信噪比条件下的鲁棒性、泛化能力与工程适用性。

上述研究成果不仅为复杂海洋环境中的海洋工程水下噪声信号提取与降噪问题提供了新的技术途径，也为进一步开展海洋工程领域的水下噪声监测、声源识别和传播建模研究奠定了坚实的技术基础与理论支撑。

## 6.2 未来工作展望

考虑到实际应用需求的复杂性与模型推广的可行性，当前研究仍存在若干值得进一步拓展的方向，未来的研究工作可在以下两个方面展开：

（1）多声源干扰下的复杂信号建模与分离仍是亟待解决的重要问题。实际海洋施工环境中常同时存在船舶运行、环境扰动等多类背景噪声，这些声源之间重叠交错，基于单一目标的方法在此类环境中易产生信号混淆。未来研究可结合语义建模与特征判别机制，增强在复杂声源场景中的精确目标提取能力。

（2）当前研究仍以离线处理为主，模型部署与工程系统的结合程度有限。随着智能传感技术与边缘计算设备的持续发展，面向实时降噪与系统集成的模型轻量化研究将成为关键方向。未来可探索嵌入于水下观测平台或声学监测终端，实现打桩声等事件的在线提取检测与噪声抑制等处理。

综上，未来工作将在多源干扰感知、工程系统集成等方面持续推进，进一步提升模型在复杂水声环境中的实用性与工程应用价值，为智能海洋监测、生态评估与近岸作业声控系统提供更加稳健的技术支撑。

# 参考文献

1. 金永明.新时代海洋强国战略的演进历程与重要成就[J].国家治理,2024,(02):40-45.
2. 王春娟,辛庞晨雨,刘大海.中国海洋工程装备国产化进程及其高质量发展趋势[J].中国软科学,2024,(S1):379-387+413.
3. 张然.海上风电场水下打桩噪声研究[D].厦门大学,2019.
4. 章蔚.海上风电场水下噪声传播研究[D].上海海洋大学,2016.
5. 王晨卉,陶毅,许肖梅.海上风电场运营期水下噪声的仿真研究[C]//中国声学学会水声学分会,山东声学学会,中国造船工程学会船舶仪器仪表学术委员会.中国声学学会水声学分会2021～2022年学术会议论文集.厦门大学海洋与地球学院;水声通信与海洋信息技术教育部重点实验室;,2022:429-431.DOI:10.26914/c.cnkihy.2022.062110.
6. Lurton X. An introduction to underwater acoustics: principles and applications[M]. Springer Science & Business Media, 2002.
7. Gussen C M G, Diniz P S R, Campos M L R, et al. A survey of underwater wireless communication technologies[J]. J. Commun. Inf. Sys, 2016, 31(1): 242-255.
8. 王晨卉.海上风电场运营期水下噪声研究[D].厦门大学,2022.DOI:10.27424/d.cnki.gxmdu.2022.002902.
9. Ou H, Allen J S, Syrmos V L. Frame-based time-scale filters for underwater acoustic noise reduction[J]. IEEE Journal of oceanic engineering, 2011, 36(2): 285-297.
10. Raj K M, Murugan S S, Natarajan V, et al. Denoising algorithm using wavelet for underwater signal affected by wind driven ambient noise[C]//2011 International Conference on Recent Trends in Information Technology (ICRTIT). IEEE, 2011: 943-946.
11. Liu F, Li G, Yang H. Application of multi-algorithm mixed feature extraction model in underwater acoustic signal[J]. Ocean Engineering, 2024, 296: 116959.
12. Aslam M A, Zhang L, Liu X, et al. Underwater sound classification using learning based methods: A review[J]. Expert Systems with Applications, 2024, 255: 124498.
13. Zhou A, Li X, Zhang W, et al. A novel cross-attention fusion-based joint training framework for robust underwater acoustic signal recognition[J]. IEEE Transactions on Geoscience and Remote Sensing, 2023, 61: 1-16.
14. Jaiswal A, Babu A R, Zadeh M Z, et al. A survey on contrastive self-supervised learning[J]. Technologies, 2020, 9(1): 2.
15. Liu S, Mallol-Ragolta A, Parada-Cabaleiro E, et al. Audio self-supervised learning: A survey[J]. Patterns, 2022, 3(12).
16. Zhang W, Yang X, Leng C, et al. Modulation recognition of underwater acoustic signals using deep hybrid neural networks[J]. IEEE Transactions on Wireless Communications, 2022, 21(8): 5977-5988.
17. Gao R, Liang M, Dong H, et al. Underwater acoustic signal denoising algorithms: A survey of the state-of-the-art[J]. IEEE Transactions on Instrumentation and Measurement, 2025.
18. Wu Y, Ta X, Xiao R, et al. Survey of underwater robot positioning navigation[J]. Applied Ocean Research, 2019, 90: 101845.
19. Zhang W, Wu W, Teng Y, et al. An underwater docking system based on UUV and recovery mother ship: design and experiment[J]. Ocean Engineering, 2023, 281: 114767.
20. Xie Y, Ren J, Xu J. Unraveling complex data diversity in underwater acoustic target recognition through convolution-based mixture of experts[J]. Expert Systems with Applications, 2024, 249: 123431.
21. Jiang J, Wu Z, Huang M, et al. Detection of underwater acoustic target using beamforming and neural network in shallow water[J]. Applied Acoustics, 2022, 189: 108626.
22. Chen J, Benesty J, Huang Y, et al. New insights into the noise reduction Wiener filter[J]. IEEE Transactions on audio, speech, and language processing, 2006, 14(4): 1218-1234.
23. Hong J, Bae I, Seok J. Wiener filtering-based ambient noise reduction technique for improved acoustic target detection of directional frequency analysis and recording sonobuoy[J]. The Journal of the Acoustical Society of Korea, 2022, 41(2): 192-198.
24. Raj K M, Murugan S S, Natarajan V, et al. Denoising algorithm using wavelet for underwater signal affected by wind driven ambient noise[C]//2011 International Conference on Recent Trends in Information Technology (ICRTIT). IEEE, 2011: 943-946.
25. Gur B M, Niezrecki C. Autocorrelation based denoising of manatee vocalizations using the undecimated discrete wavelet transform[J]. The Journal of the Acoustical Society of America, 2007, 122(1): 188-199.
26. Ganapathi S, Kumar S M, Deivasigamani M. Noise reduction in underwater acoustic signals for tropical and subtropical coastal waters[C]//2016 IEEE/OES China Ocean Acoustics (COA). IEEE, 2016: 1-6.
27. Liu F, Li G, Yang H. Application of multi-algorithm mixed feature extraction model in underwater acoustic signal[J]. Ocean Engineering, 2024, 296: 116959.
28. Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation[C]//Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18. Springer international publishing, 2015: 234-241.
29. Xu Y, Du J, Dai L R, et al. A regression approach to speech enhancement based on deep neural networks[J]. IEEE/ACM transactions on audio, speech, and language processing, 2014, 23(1): 7-19.
30. Rethage D, Pons J, Serra X. A wavenet for speech denoising[C]//2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2018: 5069-5073.
31. Botinhao C V, Wang X, Takaki S, et al. Investigating RNN-based speech enhancement methods for noise-robust text-to-speech[C]//9th ISCA speech synthesis workshop. 2016: 159-165.
32. Weninger F, Erdogan H, Watanabe S, et al. Speech enhancement with LSTM recurrent neural networks and its application to noise-robust ASR[C]//Latent Variable Analysis and Signal Separation: 12th International Conference, LVA/ICA 2015, Liberec, Czech Republic, August 25-28, 2015, Proceedings 12. Springer International Publishing, 2015: 91-99.
33. Kim J, El-Khamy M, Lee J. T-gsa: Transformer with gaussian-weighted self-attention for speech enhancement[C]//ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2020: 6649-6653.
34. Wang K, He B, Zhu W P. TSTNN: Two-stage transformer based neural network for speech enhancement in the time domain[C]//ICASSP 2021-2021 IEEE international conference on acoustics, speech and signal processing (ICASSP). IEEE, 2021: 7098-7102.
35. Zhou A, Zhang W, Li X, et al. A novel noise-aware deep learning model for underwater acoustic denoising[J]. IEEE Transactions on Geoscience and Remote Sensing, 2023, 61: 1-13.
36. Zhou A, Zhang W, Xu G, et al. DBSA-net: Dual branch self-attention network for underwater acoustic signal denoising[J]. IEEE/ACM Transactions on Audio, Speech, and Language Processing, 2023, 31: 1851-1865.
37. Song Y, Liu F, Shen T. A novel noise reduction technique for underwater acoustic signals based on dual‐path recurrent neural network[J]. IET Communications, 2023, 17(2): 135-144.
38. Zhou A, Li X, Zhang W, et al. An attention-based multi-scale convolution network for intelligent underwater acoustic signal recognition[J]. Ocean Engineering, 2023, 287: 115784.
39. Santos-Domínguez D, Torres-Guijarro S, Cardenal-López A, et al. ShipsEar: An underwater vessel noise database[J]. Applied Acoustics, 2016, 113: 64-69.
40. Ashraf H, Shah B, Soomro A M, et al. Ambient-noise free generation of clean underwater ship engine audios from hydrophones using generative adversarial networks[J]. Computers and Electrical Engineering, 2022, 100: 107970.
41. Xu J, Xie Y, Wang W. Underwater acoustic target recognition based on smoothness-inducing regularization and spectrogram-based data augmentation[J]. Ocean Engineering, 2023, 281: 114926.
42. Yang S, Jin A, Zeng X, et al. Underwater acoustic target recognition based on sub-band concatenated Mel spectrogram and multidomain attention mechanism[J]. Engineering Applications of Artificial Intelligence, 2024, 133: 107983.
43. Chen L, Luo X, Zhou H. A ship-radiated noise classification method based on domain knowledge embedding and attention mechanism[J]. Engineering Applications of Artificial Intelligence, 2024, 127: 107320.
44. Irfan M, Jiangbin Z, Ali S, et al. DeepShip: An underwater acoustic benchmark dataset and a separable convolution based autoencoder for classification[J]. Expert Systems with Applications, 2021, 183: 115270.
45. Zhu P, Zhang Y, Huang Y, et al. Underwater acoustic target recognition based on spectrum component analysis of ship radiated noise[J]. Applied Acoustics, 2023, 211: 109552.
46. Sayigh L, Daher M A, Allen J, et al. The Watkins marine mammal sound database: an online, freely accessible resource[C]//Proceedings of Meetings on Acoustics. AIP Publishing, 2016, 27(1).
47. Du X, Hong F. QiandaoEar22: a high-quality noise dataset for identifying specific ship from multiple underwater acoustic targets using ship-radiated noise[J]. EURASIP Journal on Advances in Signal Processing, 2024, 2024(1): 96.
48. Wilkes D R, Gourlay T P, Gavrilov A N. Numerical modeling of radiated sound for impact pile driving in offshore environments[J]. IEEE Journal of Oceanic Engineering, 2016, 41(4): 1072-1078.
49. Tsouvalas A. Underwater noise emission due to offshore pile installation: A review[J]. Energies, 2020, 13(12): 3037.
50. Peng Y, Tsouvalas A, Stampoultzoglou T, et al. A fast computational model for near-and far-field noise prediction due to offshore pile driving[J]. The Journal of the Acoustical Society of America, 2021, 149(3): 1772-1790.
51. Caron M, Touvron H, Misra I, et al. Emerging properties in self-supervised vision transformers[C]//Proceedings of the IEEE/CVF international conference on computer vision. 2021: 9650-9660.
52. Liu S, Mallol-Ragolta A, Parada-Cabaleiro E, et al. Audio self-supervised learning: A survey[J]. Patterns, 2022, 3(12).
53. Gui J, Chen T, Zhang J, et al. A survey on self-supervised learning: Algorithms, applications, and future trends[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2024.
54. Fujimura T, Koizumi Y, Yatabe K, et al. Noisy-target training: A training strategy for DNN-based speech enhancement without clean speech[C]//2021 29th european signal processing conference (EUSIPCO). IEEE, 2021: 436-440.
55. Kashyap M M, Tambwekar A, Manohara K, et al. Speech denoising without clean training data: A noise2noise approach[J]. arXiv preprint arXiv:2104.03838, 2021.
56. Wu J, Li Q, Yang G, et al. Self-supervised speech denoising using only noisy audio signals[J]. Speech Communication, 2023, 149: 63-73.
57. Hildebrand J A. Anthropogenic and natural sources of ambient noise in the ocean[J]. Marine Ecology Progress Series, 2009, 395: 5-20.
58. Dahl P H, Miller J H, Cato D H, et al. Underwater ambient noise[J]. Acoustics Today, 2007, 3(1): 23-33.
59. Cato D H. Ambient noise and its significance to aquatic life[J]. Bioacoustics, 2008, 17(1-3): 21-23.
60. Berdnikova J, Klauson A, Mustonen M, et al. Underwater ship noise pattern detection and identification[J]. The Journal of the Acoustical Society of America, 2017, 142(4\_Supplement): 2686-2686.
61. Rodrigo F J, Ramis J, Carbajo J, et al. Underwater anthropogenic noise pollution assessment in shallow waters on the south-eastern coast of Spain[J]. Journal of Marine Science and Engineering, 2022, 10(9): 1311.
62. Wittekind D, Schuster M. Propeller cavitation noise and background noise in the sea[J]. Ocean Engineering, 2016, 120: 116-121.
63. Cui X, Yang H, Hu Q, et al. Noise Sources and Statistical Characteristics of the Marine Ambient Noise in the Zengmu Basin[C]//2024 OES China Ocean Acoustics (COA). IEEE, 2024: 1-5.
64. Chunxia M, Dan S, Feng C, et al. Statistical characteristic of spectrum for ambient noise at high frequencies in shallow water[C]//2017 IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC). IEEE, 2017: 1-4.
65. Felisberto P, Silva J P, Silva J, et al. Background noise in areas covered by marine plants in the Ria Formosa lagoon during the summer[C]//2018 OCEANS-MTS/IEEE Kobe Techno-Oceans (OTO). IEEE, 2018: 1-5.
66. Bailey H, Senior B, Simmons D, et al. Assessing underwater noise levels during pile-driving at an offshore windfarm and its potential effects on marine mammals[J]. Marine pollution bulletin, 2010, 60(6): 888-897.
67. Yang L, Xu X, Huang Z, et al. Recording and analyzing underwater noise during pile driving for bridge construction[J]. Acoustics Australia, 2015, 43: 159-167.
68. Dahl P H, de Jong C A F, Popper A N. The underwater sound field from impact pile driving and its potential effects on marine life[J]. Acoustics Today, 2015, 11(2): 18-25.
69. Reinhall P G, Dahl P H. Underwater Mach wave radiation from impact pile driving: Theory and observation[J]. The Journal of the Acoustical Society of America, 2011, 130(3): 1209-1216.
70. Dahl P H, Dall'Osto D R, Farrell D M. The underwater sound field from vibratory pile driving[J]. The Journal of the Acoustical Society of America, 2015, 137(6): 3544-3554.
71. 时文静,王志陶,方亮,等.打桩水下噪声对长江江豚影响初探[J].水生生物学报,2015,39(02):399-407.
72. 于晋源,许肖梅,张鑫海,等.水下打桩噪声场的有限元分析[C]//中国声学学会水声学分会.中国声学学会水声学分会2019年学术会议论文集.厦门大学海洋与地球学院;水声通信与海洋信息技术教育部重点实验室;,2019:406-408.
73. 陆炜.跨江桥梁水下振动噪声特性研究[D].东南大学,2022.DOI:10.27014/d.cnki.gdnau.2022.001807.
74. Song X, Yin L, Xiong W, et al. Underwater noise prediction and control of a cross-river subway tunnel: An experimental and numerical study[J]. International Journal of Environmental Science and Technology, 2024, 21(4): 4045-4062.
75. Song X, Zhang X, Xiong W, et al. Experimental and numerical study on underwater noise radiation from an underwater tunnel[J]. Environmental Pollution, 2020, 267: 115536.
76. Song X, Lu W, Xiong W, et al. Sound contribution of the low frequency underwater noise radiated from a suspension bridge[J]. Journal of Low Frequency Noise, Vibration and Active Control, 2024, 43(1): 3-19.
77. Lehtinen J, Munkberg J, Hasselgren J, et al. Noise2Noise: Learning image restoration without clean data[J]. arXiv preprint arXiv:1803.04189, 2018.
78. Huang T, Li S, Jia X, et al. Neighbor2neighbor: Self-supervised denoising from single noisy images[C]//Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2021: 14781-14790.
79. Trabelsi C, Bilaniuk O, Zhang Y, et al. Deep complex networks[J]. arXiv preprint arXiv:1705.09792, 2017.
80. Dai J, Qi H, Xiong Y, et al. Deformable convolutional networks[C]//Proceedings of the IEEE international conference on computer vision. 2017: 764-773.
81. Hu J, Shen L, Sun G. Squeeze-and-excitation networks[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2018: 7132-7141.
82. Zhang Y, Li K, Li K, et al. Image super-resolution using very deep residual channel attention networks[C]//Proceedings of the European conference on computer vision (ECCV). 2018: 286-301.
83. Wang X, Girshick R, Gupta A, et al. Non-local neural networks[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2018: 7794-7803.
84. Tolstikhin I O, Houlsby N, Kolesnikov A, et al. Mlp-mixer: An all-mlp architecture for vision[J]. Advances in neural information processing systems, 2021, 34: 24261-24272.
85. Nosrati L, Fazel M S, Ghavami M. Improving indoor localization using mobile UWB sensor and deep neural networks[J]. IEEE Access, 2022, 10: 20420-20431.
86. Robert J, Webbie M. Pydub (2018)[J]. URL http://pydub. com.

# 攻读硕士学位期间获得的成果



# 致谢

穹海潜心探幽微，碧涛深处觅芳菲。师恩如岳扶舟楫，益友同行共忘归。

**学位论文答辩委员会名单**