**学校编码：10384 分类号 密级**

**学号：22320211151410 UDC**



**硕 士 学 位 论 文**

**基于深度学习自监督策略的水下工程脉冲声信号降噪**

**Research on Layered Routing Protocol for Underwater Acoustic Sensor Networks Based on Neural Networks**

**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**

**答辩委员会主席：**

**评 阅 人：**

**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它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.不保密，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

随着对海洋领域探测与研究的不断深入，亟需发展水声传感网以解决水下勘探、军事监视、数据搜集和水下灾害预防等诸多领域的应用需求。针对低带宽、长时延、低能耗供应的水声传感网，如何提升其网络性能是目前的主要研究方向。设计高效、低功耗的水声传感网路由协议，提高水声节点间数据传输高效可靠性能，则是当前水声传感网发展面临的技术难题。因此，水声传感网路由协议的研究具有重要意义。

水声传感网通过多个节点之间的多跳传输，可有效地提高水下远距离传输的可靠性。然而，水声节点的能量有限，端到端传输延迟大，网络负载不均，使用寿命短。针对上述问题，基于神经网络算法，从单条路由效率（时延小、中断率低）最优、网络整体能耗负载均衡、传输安全性提高三个方面，提出适用于动态分层结构下有内部攻击安全隐患的海洋环境水声传感网路由协议解决方案。论文主要内容包括：

1、查阅文献，整理国内外水声通信与组网、水声通信路由协议研究现状，比较了早期经典水声路由和新型水声路由协议的特点，阐述水声信道和水声传感网的特性，分析说明水声路由协议对水声通信的重要作用，阐明动态分层策略和人工智能算法对水声传感网路由的重要意义。

2、描述总结了水声信道的特性、分析了水声传感网特性和其与陆地无线传感网的区别、介绍了水声传感网所面临的效率和安全挑战，说明了利用神经网络和动态分层策略改进现有水声路由协议的可行性和有效性。

3、利用反向传播神经网络（BP-NN）算法，结合分层策略和自主水下航行器（AUV）手段，提出了水声传感网动态分层路由算法。利用AUV的机动性和BP-NN的分类识别功能更新节点的工作模式，以此实现路由的动态分层，使算法能够在保证网络效率的同时，有效缓解水声传感网的“能量空洞”和“通信空区”问题，延长网络寿命。

4、考虑水声传感器节点受到恶意攻击而出现网络内部安全问题，利用图神经网络（GNN）对非欧式结构的强表达能力和异常检测能力，引入节点信任预测机制，提出了一种水声分层安全路由方案。在所提方案中，充分利用水声传感网图结构特性，赋予图结构对应的物理意义，高效地检测水声传感网种的恶意节点，旨在避免恶意节点被路由选中，提高网络安全性和可靠性，同时进一步降低节点能耗、提高数据包传递率，延长网络寿命。

**关键词：**水声传感网；水下分层路由；水下安全路由；信任模型；神经网络

**Abstract**

With the continuous deepening of exploration and research in the marine field, there is an urgent need to develop underwater acoustic sensing networks (UASNs) to meet the application requirements in various fields such as underwater exploration, military surveillance, data collection, and underwater disaster prevention. Given the challenges posed by low bandwidth, long latency, and low energy consumption supply in UASNs, enhancing the network performance has become a primary research focus. Designing efficient and low-power routing protocols for UASNs to improve the efficiency and reliability of data transmission performance between underwater acoustic nodes is currently a technical challenge facing the development of UASNs. Thus, investigating routing schemes in UASNs holds considerable importance.

UASNs effectively improves the reliability of underwater long-distance transmission through multi-hop transmission among multiple nodes. However, the energy of underwater acoustic nodes is limited, the end-to-end transmission delay is significant, the network load is uneven, and the service life is short. In response to the above issues, based on neural network algorithms, this paper proposes a routing protocol solution for UASNs in a dynamic layering structure with internal security vulnerabilities in a marine environment, focusing on three aspects: optimizing the efficiency of individual routes (low latency and low interruption rate), balancing the overall energy consumption and load in the network, and enhancing transmission security. The main contents of the paper include:

1. Reviewing literature to summarize the current research status of underwater acoustic communication and networking, as well as routing protocols for UASNs at home and abroad, comparing the characteristics of classical and new underwater acoustic routing protocols, elaborating on the characteristics of underwater acoustic channels and UASNs, analyzing and explaining the important role of underwater acoustic routing protocols in underwater acoustic communication, and demonstrating the importance of dynamic layering strategies and artificial intelligence algorithms in UASN routing.

2. This article describes and summarizes the characteristics of underwater acoustic channels, analyzes the characteristics of UASNs and their differences from terrestrial wireless sensor networks, introduces the efficiency and security challenges faced by UASNs, and demonstrates the feasibility and effectiveness of using neural networks and dynamic layering strategies to improve existing underwater acoustic routing protocols.

3. A dynamic layered routing algorithm for UASNs is proposed using backpropagation neural network (BP-NN) algorithm, combined with layering strategy and autonomous underwater vehicles (AUVs). By utilizing the maneuverability of AUVs and the classification and recognition function of BP-NN, the working mode of nodes is updated to achieve dynamic layering of routing, enabling the algorithm to effectively alleviate the "energy holes" and "communication gaps" problems of UASNs while ensuring network efficiency, and extend network lifespan.

4. Considering the internal security issues caused by malicious attacks on underwater sensor nodes, a acoustic layering secure routing scheme is proposed by utilizing the strong expressive and anomaly detection capabilities of graph neural networks (GNNs) on non-Euclidean structures, introducing a node trust prediction mechanism. In the proposed scheme, the characteristics of UASNs’ graph structures are fully utilized, and the corresponding physical meanings are assigned to graph structures to efficiently detect malicious nodes in UASNs. The aim is to avoid malicious nodes being selected by routing, improve network security and reliability, and further reduce node energy consumption, improve packet delivery rate (PDR), and extend network lifespan.

**Key words:** Underwater Acoustic Sensor Networks (UASNs); Underwater Layered Routing; Underwater Secure Routing; Trust Model; Neural Network

**目录**

[摘 要 I](#_Toc193732900)

[第一章 绪论 1](#_Toc193732901)

[1.1 研究背景及意义 1](#_Toc193732902)

[1.2 国内外研究现状 3](#_Toc193732903)

[1.2.1 人为水下噪声研究现状 3](#_Toc193732904)

[1.2.2 人为水下噪声降噪技术 4](#_Toc193732905)

[1.3 论文主要内容 6](#_Toc193732906)

[第二章 人为水下声学信号特性分析 9](#_Toc193732907)

[2.1 海洋背景噪声的来源与特性 9](#_Toc193732908)

[2.1.1 背景噪声的分类 9](#_Toc193732909)

[2.1.2 海洋背景噪声的频谱特性分析 10](#_Toc193732910)

[2.2 人为水下声信号评估参数 11](#_Toc193732911)

[2.2.1 声压及其度量参数 11](#_Toc193732912)

[2.2.2 均方根声压级 12](#_Toc193732913)

[2.2.3 等效连续声级、声暴露级和累计声暴露级 13](#_Toc193732914)

[2.3 常见人为水下声信号类型概述 14](#_Toc193732915)

[2.3.1 桥梁噪声特性 14](#_Toc193732916)

[2.3.2 船舶噪声特性 15](#_Toc193732917)

[2.3.3 打桩噪声特性 17](#_Toc193732918)

[2.3.4 各噪声类型的比较与对比 19](#_Toc193732919)

[2.4 本章小结 20](#_Toc193732920)

[2.5 本章小结 22](#_Toc193732921)

[第三章 监测案例与人为水下声信号统计特性分析 23](#_Toc193732922)

[3.1 监测方案介绍 23](#_Toc193732923)

[3.1.1 监测数据采集系统 23](#_Toc193732924)

[3.1.2 船载式监测方案 24](#_Toc193732925)

[3.2 桥梁水下声信号监测 25](#_Toc193732926)

[3.2.1 监测环境 25](#_Toc193732927)

[3.2.2 监测布置 28](#_Toc193732928)

[3.2.3 时域分析结果及讨论 31](#_Toc193732929)

[3.2.4 频域分析结果及能量分布 33](#_Toc193732930)

[3.3 噪声特性讨论与自监督降噪启示 38](#_Toc193732931)

[第四章 基于深度学习的自监督降噪方案设计 40](#_Toc193732932)

[4.1 方案设计思路 40](#_Toc193732933)

[4.1.1 方法背景 40](#_Toc193732934)

[4.1.2 研究思路 **错误!未定义书签。**](#_Toc193732935)

[4.1.3 系统流程与总体框架 46](#_Toc193732936)

[4.2 模型架构设计 47](#_Toc193732937)

[4.2.1 编码器-解码器结构 48](#_Toc193732938)

[4.2.2 频谱动态聚焦单元（SDFU） 49](#_Toc193732939)

[4.2.3 复数注意力模块（CAM） 52](#_Toc193732940)

[4.2.4 编码端-解码端特征连接（CSFU） 55](#_Toc193732941)

[4.2.5 多尺度跳跃连接(CMFFU) 60](#_Toc193732942)

[4.3 自监督学习策略与算法细节 62](#_Toc193732943)

[4.3.1 伪“干净”目标构造方法 62](#_Toc193732944)

[4.3.2 损失函数设计与优化策略 63](#_Toc193732945)

[4.3.3 训练流程与超参数设置 **错误!未定义书签。**](#_Toc193732946)

[4.4 本章小结 65](#_Toc193732947)

[第五章 降噪实验与结果分析 66](#_Toc193732948)

[5.1 数据集与实验设置 66](#_Toc193732949)

[5.1.1 数据集构建与来源 67](#_Toc193732950)

[5.1.2 实验环境与预处理流程 72](#_Toc193732951)

[5.1.3 评价指标与实验方案 73](#_Toc193732952)

[5.2 降噪效果展示 74](#_Toc193732953)

[5.2.1 人造数据降噪效果 75](#_Toc193732954)

[5.2.2 实际数据降噪效果 75](#_Toc193732955)

[5.3结果讨论与对比分析 76](#_Toc193732956)

[5.3.1 传统方法对比 76](#_Toc193732957)

[5.3.2 模型各模块消融实验 76](#_Toc193732958)

[第六章 总结与展望 77](#_Toc193732959)

[6.1 论文的主要工作与创新点 77](#_Toc193732960)

[6.2 未来工作展望 77](#_Toc193732961)

[参考文献 78](#_Toc193732962)

[攻读硕士学位期间获得的成果 83](#_Toc193732963)

[致谢 85](#_Toc193732964)

[学位论文答辩委员会名单 88](#_Toc193732965)

**Catalog**

[摘 要 I](#_Toc193732966)

[第一章 绪论 1](#_Toc193732967)

[1.1 研究背景及意义 1](#_Toc193732968)

[1.2 国内外研究现状 3](#_Toc193732969)

[1.2.1 人为水下噪声研究现状 3](#_Toc193732970)

[1.2.2 人为水下噪声降噪技术 4](#_Toc193732971)

[1.3 论文主要内容 6](#_Toc193732972)

[第二章 人为水下声学信号特性分析 9](#_Toc193732973)

[2.1 海洋背景噪声的来源与特性 9](#_Toc193732974)

[2.1.1 背景噪声的分类 9](#_Toc193732975)

[2.1.2 海洋背景噪声的频谱特性分析 10](#_Toc193732976)

[2.2 人为水下声信号评估参数 11](#_Toc193732977)

[2.2.1 声压及其度量参数 11](#_Toc193732978)

[2.2.2 均方根声压级 12](#_Toc193732979)

[2.2.3 等效连续声级、声暴露级和累计声暴露级 13](#_Toc193732980)

[2.3 常见人为水下声信号类型概述 14](#_Toc193732981)

[2.3.1 桥梁噪声特性 14](#_Toc193732982)

[2.3.2 船舶噪声特性 15](#_Toc193732983)

[2.3.3 打桩噪声特性 17](#_Toc193732984)

[2.3.4 各噪声类型的比较与对比 19](#_Toc193732985)

[2.4 本章小结 20](#_Toc193732986)

[2.5 本章小结 22](#_Toc193732987)

[第三章 监测案例与人为水下声信号统计特性分析 23](#_Toc193732988)

[3.1 监测方案介绍 23](#_Toc193732989)

[3.1.1 监测数据采集系统 23](#_Toc193732990)

[3.1.2 船载式监测方案 24](#_Toc193732991)

[3.2 桥梁水下声信号监测 25](#_Toc193732992)

[3.2.1 监测环境 25](#_Toc193732993)

[3.2.2 监测布置 28](#_Toc193732994)

[3.2.3 时域分析结果及讨论 31](#_Toc193732995)

[3.2.4 频域分析结果及能量分布 33](#_Toc193732996)

[3.3 噪声特性讨论与自监督降噪启示 38](#_Toc193732997)

[第四章 基于深度学习的自监督降噪方案设计 40](#_Toc193732998)

[4.1 方案设计思路 40](#_Toc193732999)

[4.1.1 方法背景 40](#_Toc193733000)

[4.1.2 研究思路 **错误!未定义书签。**](#_Toc193733001)

[4.1.3 系统流程与总体框架 46](#_Toc193733002)

[4.2 模型架构设计 47](#_Toc193733003)

[4.2.1 编码器-解码器结构 48](#_Toc193733004)

[4.2.2 频谱动态聚焦单元（SDFU） 49](#_Toc193733005)

[4.2.3 复数注意力模块（CAM） 52](#_Toc193733006)

[4.2.4 编码端-解码端特征连接（CSFU） 55](#_Toc193733007)

[4.2.5 多尺度跳跃连接(CMFFU) 60](#_Toc193733008)

[4.3 自监督学习策略与算法细节 62](#_Toc193733009)

[4.3.1 伪“干净”目标构造方法 62](#_Toc193733010)

[4.3.2 损失函数设计与优化策略 63](#_Toc193733011)

[4.3.3 训练流程与超参数设置 **错误!未定义书签。**](#_Toc193733012)

[4.4 本章小结 65](#_Toc193733013)

[第五章 降噪实验与结果分析 66](#_Toc193733014)

[5.1 数据集与实验设置 66](#_Toc193733015)

[5.1.1 数据集构建与来源 67](#_Toc193733016)

[5.1.2 实验环境与预处理流程 72](#_Toc193733017)

[5.1.3 评价指标与实验方案 73](#_Toc193733018)

[5.2 降噪效果展示 74](#_Toc193733019)

[5.2.1 人造数据降噪效果 75](#_Toc193733020)

[5.2.2 实际数据降噪效果 75](#_Toc193733021)

[5.3结果讨论与对比分析 76](#_Toc193733022)

[5.3.1 传统方法对比 76](#_Toc193733023)

[5.3.2 模型各模块消融实验 76](#_Toc193733024)

[第六章 总结与展望 77](#_Toc193733025)

[6.1 论文的主要工作与创新点 77](#_Toc193733026)

[6.2 未来工作展望 77](#_Toc193733027)

[参考文献 78](#_Toc193733028)

[攻读硕士学位期间获得的成果 83](#_Toc193733029)

[致谢 85](#_Toc193733030)

[学位论文答辩委员会名单 88](#_Toc193733031)

[摘 要 I](#_Toc193733032)

[第一章 绪论 1](#_Toc193733033)

[1.1 研究背景及意义 1](#_Toc193733034)

[1.2 国内外研究现状 3](#_Toc193733035)

[1.2.1 人为水下噪声研究现状 3](#_Toc193733036)

[1.2.2 人为水下噪声降噪技术 4](#_Toc193733037)

[1.3 论文主要内容 6](#_Toc193733038)

[第二章 人为水下声学信号特性分析 9](#_Toc193733039)

[2.1 海洋背景噪声的来源与特性 9](#_Toc193733040)

[2.1.1 背景噪声的分类 9](#_Toc193733041)

[2.1.2 海洋背景噪声的频谱特性分析 10](#_Toc193733042)

[2.2 人为水下声信号评估参数 11](#_Toc193733043)

[2.2.1 声压及其度量参数 11](#_Toc193733044)

[2.2.2 均方根声压级 12](#_Toc193733045)

[2.2.3 等效连续声级、声暴露级和累计声暴露级 13](#_Toc193733046)

[2.3 常见人为水下声信号类型概述 14](#_Toc193733047)

[2.3.1 桥梁噪声特性 14](#_Toc193733048)

[2.3.2 船舶噪声特性 15](#_Toc193733049)

[2.3.3 打桩噪声特性 17](#_Toc193733050)

[2.3.4 各噪声类型的比较与对比 19](#_Toc193733051)

[2.4 本章小结 20](#_Toc193733052)

[2.5 本章小结 22](#_Toc193733053)

[第三章 监测案例与人为水下声信号统计特性分析 23](#_Toc193733054)

[3.1 监测方案介绍 23](#_Toc193733055)

[3.1.1 监测数据采集系统 23](#_Toc193733056)

[3.1.2 船载式监测方案 24](#_Toc193733057)

[3.2 桥梁水下声信号监测 25](#_Toc193733058)

[3.2.1 监测环境 25](#_Toc193733059)

[3.2.2 监测布置 28](#_Toc193733060)

[3.2.3 时域分析结果及讨论 31](#_Toc193733061)

[3.2.4 频域分析结果及能量分布 33](#_Toc193733062)

[3.3 噪声特性讨论与自监督降噪启示 38](#_Toc193733063)

[第四章 基于深度学习的自监督降噪方案设计 40](#_Toc193733064)

[4.1 方案设计思路 40](#_Toc193733065)

[4.1.1 方法背景 40](#_Toc193733066)

[4.1.2 研究思路 **错误!未定义书签。**](#_Toc193733067)

[4.1.3 系统流程与总体框架 46](#_Toc193733068)

[4.2 模型架构设计 47](#_Toc193733069)

[4.2.1 编码器-解码器结构 48](#_Toc193733070)

[4.2.2 频谱动态聚焦单元（SDFU） 49](#_Toc193733071)

[4.2.3 复数注意力模块（CAM） 52](#_Toc193733072)

[4.2.4 编码端-解码端特征连接（CSFU） 55](#_Toc193733073)

[4.2.5 多尺度跳跃连接(CMFFU) 60](#_Toc193733074)

[4.3 自监督学习策略与算法细节 62](#_Toc193733075)

[4.3.1 伪“干净”目标构造方法 62](#_Toc193733076)

[4.3.2 损失函数设计与优化策略 63](#_Toc193733077)

[4.3.3 训练流程与超参数设置 **错误!未定义书签。**](#_Toc193733078)

[4.4 本章小结 65](#_Toc193733079)

[第五章 降噪实验与结果分析 66](#_Toc193733080)

[5.1 数据集与实验设置 66](#_Toc193733081)

[5.1.1 数据集构建与来源 67](#_Toc193733082)

[5.1.2 实验环境与预处理流程 72](#_Toc193733083)

[5.1.3 评价指标与实验方案 73](#_Toc193733084)

[5.2 降噪效果展示 74](#_Toc193733085)

[5.2.1 人造数据降噪效果 75](#_Toc193733086)

[5.2.2 实际数据降噪效果 75](#_Toc193733087)

[5.3结果讨论与对比分析 76](#_Toc193733088)

[5.3.1 传统方法对比 76](#_Toc193733089)

[5.3.2 模型各模块消融实验 76](#_Toc193733090)

[第六章 总结与展望 77](#_Toc193733091)

[6.1 论文的主要工作与创新点 77](#_Toc193733092)

[6.2 未来工作展望 77](#_Toc193733093)

[参考文献 78](#_Toc193733094)

[攻读硕士学位期间获得的成果 83](#_Toc193733095)

[致谢 85](#_Toc193733096)

[学位论文答辩委员会名单 88](#_Toc193733097)

第一章 引言

1.1 研究背景及意义

近年来，随着近海基础设施建设的快速推进[1]，水下打桩、桥梁施工、隧道掘进等海工作业活动日趋频繁，其所产生的结构性脉冲声信号成为海洋声学监测、施工安全评估及生态保护研究中的重要关注对象[3]。在工程实践中，准确捕捉这些水下人为工程脉冲声信号，不仅有助于实现作业状态的智能识别与设备状态监测，还在海洋环境影响评估、声能传播建模和水下感知等应用领域中具有广泛价值[6]。

然而，实际采集过程面临着极大的挑战。由于水声传播环境本身具有强非平稳、多径混响与噪声源复杂等特点[7]，目标信号往往深受背景干扰的影响，尤其是在远距离接收条件下，结构性特征被严重削弱，信号的可识别性和可利用性显著下降。特别是海底回波、边界反射及多径路径造成的尾波混叠，不仅导致信号后部出现结构紊乱和能量冗余，还可能干扰打桩事件识别与声暴露评估，成为工程声信号处理中的一大难点[8][9]。

|  |
| --- |
| 图1-1 人为水下工程声事件场景图 |

传统方法如维纳滤波、小波阈值等，在平稳噪声背景下具备一定效果，但在处理这类短时突发、结构突显的脉冲声时，普遍存在结构信息压缩、频谱主带扭曲及尾波误保留等问题，难以满足工程任务对信号结构完整性与物理一致性的要求[11]。同时，水下工程场景中常常缺乏纯净参考信号，进一步限制了有监督深度模型的训练与泛化。

近年来，深度学习方法在语音降噪和语音增强任务中展现出强大的特征建模能力与去噪潜力。自监督学习框架的引入为无标签数据场景下的建模提供了可行路径[12]，但目前已有方法多聚焦于空气语音领域，对于脉冲类水下人为声源的建模能力、结构保持性与工程适应性仍研究不足，尤其在复杂背景干扰下的降噪稳定性与泛化能力尚有待突破。

针对上述问题，本文聚焦于水下人为工程脉冲声信号降噪这一关键任务，结合自监督建模思路，构建融合能量结构感知与频谱保持机制的深度神经网络模型，提出面向工程场景的结构性降噪框架。研究从两个方面出发：一方面，针对工程背景下难以获取纯净标签的现状，设计基于实际采集信号与累计能量段截取策略的训练数据构建流程；另一方面，从模型结构入手，引入频谱感知模块与复数注意力机制，提升对信号主能量段的还原能力和非结构噪声的抑制能力。

在实验层面，本文通过多个任务场景设置，包括累计能量段降噪、原始信号降噪与非训练类型脉冲声降噪等，综合考察模型在结构保持性、频谱一致性与时域精度等维度的性能表现。同时，通过与传统方法及主流深度模型的对比，以及模块消融实验的验证，有效揭示了模型各组成模块在降噪能力中的协同作用。研究结果表明，所提方法在保持信号结构特征的基础上显著提升了信噪比与可识别性，为复杂背景下的水下工程声信号降噪与海洋背景噪声的抑制提供了切实有效的技术方案。

综上，本文针对当前水下人为工程脉冲声信号降噪中的实际痛点与研究空白，提出了一种面向结构恢复与工程实用性的自监督深度学习模型，并通过系统实验验证了其在结构识别、尾波抑制与频谱保真的综合性能。研究成果对于推动水声信号处理的智能化转型、提升工程声源感知的准确性和鲁棒性以及海洋环境噪声监测具有重要的理论价值与实际意义。

1.2 国内外研究现状

水下声信号作为海洋工程监测与环境感知中的关键信息载体，其处理与降噪技术长期以来是水声信号研究的重要方向。与传统水声通信、语音增强等连续语音类信号不同，水下打桩、桥梁撞击等人为工程活动产生的结构性脉冲声信号具有突发性强、能量集中、主频明确、持续时间短等特点，这类信号在受到复杂背景干扰（如风浪声、远距离船舶噪声、生物声与设备运转声）影响时，更易出现信号淹没、结构模糊与尾波拖延等问题，严重影响其在事件识别、能量评估及传播建模等任务中的应用效果。因此，针对水下工程脉冲声的高保真降噪与结构保持问题，成为近年来国内外研究者重点关注的技术难题之一。

在早期研究中，降噪处理多依赖于传统信号处理方法，如频域滤波、谱减法、维纳滤波与小波变换等。这类方法在处理平稳背景噪声或连续语音信号方面具备一定实用性，但在面对非平稳、多频复杂背景及结构性目标信号时，往往存在目标信号被误削、频谱主结构丢失或伪波形引入等问题。例如，维纳滤波方法依赖对噪声统计特征的准确估计，在复杂水下环境中极易失效；小波变换在处理瞬态信号方面具有一定优势，但面临对噪声与目标频带边界划分不清的问题，难以保证结构性冲击声的能量聚焦效果[23]。国内相关研究中也尝试将经验模态分解（EMD）[15]、最小均方（LMS）等方法用于水声降噪，但对多径拖尾与非线性干扰的抑制效果仍不理想。

近年来，深度学习方法在语音增强与水下声信号处理领域取得显著进展，为非平稳噪声环境中的建模与降噪提供了全新路径。典型模型包括基于卷积神经网络（CNN）的降噪自动编码器、U-Net 结构、循环神经网络（RNN）与长短时记忆网络（LSTM）组合架构等。这些模型可通过端到端学习复杂特征映射关系，有效提升信噪比（SNR）与语音质量（PESQ）等指标，逐步成为水下语音降噪的新趋势。其中，U-Net 凭借其编码-解码结构及多尺度融合能力，被广泛应用于水下语音恢复与水下目标检测中，取得了良好效果。但这类方法在处理结构性冲击类信号时，往往侧重于频谱整体还原，忽略了脉冲主能量段的局部结构表达，在尾波干扰复杂的场景下易出现过度平滑或目标模糊现象，影响事件识别的准确性与稳定性。

为弥补有监督方法对标签依赖强的问题，近年来自监督学习（Self-Supervised Learning, SSL）逐步应用于语音与水声信号处理场景。自监督方法通过构造预训练任务，如信号掩蔽预测、时间重构与频谱填补，实现对大量未标注数据的特征提取与表示学习。例如，Facebook 推出的 wav2vec 系列模型在语音建模方面取得突破性进展；在水下领域，部分研究也开始探索基于噪声合成、对比学习或伪标签构建的自监督降噪框架。然而，目前多数方法仍面向连续语音信号或水下通信声调，对打桩声等工程结构冲击信号的结构保持能力关注不足，缺乏对非平稳背景下能量段识别与尾波压制的专门建模。

在国内，已有部分研究开始关注水下结构声事件识别、水下振动声源分离与打桩噪声监测等应用任务。例如，中科院声学所、清华大学、哈尔滨工程大学等单位针对水下人造声源的识别与定位开展了基于图谱学习、稀疏建模与阵列处理的工作，但多以特征提取为主，缺乏面向降噪本身的系统性研究。少量针对水下工程脉冲信号的降噪工作中，也多基于手工特征设计或实验室模拟，缺少结合实测数据、复杂背景与结构保持的系统框架。

综上所述，当前水下人为工程脉冲声降噪面临以下突出问题：（1）现有传统方法难以处理结构信号中的非平稳尾波与突发扰动；（2）已有深度模型在工程结构信号上结构保持能力不足，泛化性能有限；（3）针对缺乏纯净标签的数据环境，尚缺少系统化的自监督建模策略。为解决上述问题，亟需构建面向实际采集场景的水下结构声降噪模型，结合能量段识别机制、频域建模结构与自监督学习策略，提升对水下工程类脉冲信号的还原质量与工程适用性。

1.3 本文主要工作与技术路线

针对水下人为工程脉冲声信号在复杂背景干扰下结构信息易被掩盖、能量集中段难以准确恢复的问题，本文聚焦于提升此类信号的时频结构保持能力与工程适应性，设计并实现了一种基于自监督学习的水下降噪方法。研究从声信号特性出发，构建覆盖理论分析、模型设计、系统实现与实验验证的完整技术路线（如图1-2所示），具体工作可归纳为以下几个方面：

第一章阐明了本研究的背景与意义。随着海洋开发活动的频繁开展，海底打桩、桥梁施工等产生的大量短时强冲击声对环境与工程监测提出挑战。水下脉冲声信号的精准获取成为支撑生态评估与结构健康识别的重要环节。在此基础上，梳理国内外相关研究进展，明确目前方法在非平稳背景噪声干扰下存在结构还原不足、工程泛化性弱等问题，进而提出以结构保持为核心的水下降噪方案。

第二章围绕水下人为脉冲声信号的时频特征展开系统性分析。首先从声源机制与传播路径角度，分析海洋背景噪声的来源、频谱分布与变化规律。随后以实际采集的打桩信号为代表，通过均方根声压、等效连续声级与声暴露级等指标，探讨打桩类脉冲信号在能量集中性、瞬态时域结构等方面的特征，为模型设计提供理论依据。

第三章结合工程应用需求，介绍了完整的水下声信号采集系统与监测方案，包括采集方案、采集设备参数与水下声信号采集获取过程，并对实测信号进行结构性分析。通过对比不同采样场景下的脉冲事件与背景扰动，揭示水下环境中多径效应、反射干扰对声信号结构提取造成的影响，进一步明确降噪任务的目标特征与难点。

第四章提出了一种面向结构保持的自监督水下降噪方案。该方案在不依赖纯净监督样本的前提下，设计了伪标签生成机制与多层级损失函数构建策略，引入频谱动态聚焦单元（SDFU）、复数注意模块（CAM）与多尺度结构融合机制（CMS-SCU），以提升模型对非平稳背景下主能量段的识别能力与结构恢复精度。模型整体具备良好的可迁移性与泛化性，适用于多类冲击声信号的处理场景。

第五章通过三类实验对所提方法进行验证：首先构建了基于声暴露级理论的“累计能量段”打桩信号训练集，评估模型在结构约束条件下的恢复能力；随后将模型直接应用于未截断的原始打桩信号与桥梁振动类非训练样本，测试其对非理想输入的适应性与稳健性。实验采用多组噪声叠加条件与信噪比设置，结合波形图、频谱图、帧级SNR曲线等可视化方式进行对比，并以SNR、PSNR、RMSE三类定量指标为核心评估手段，全面衡量模型降噪性能。进一步引入传统方法（维纳滤波、小波变换）与经典深度学习网络进行对比分析，并开展模块消融实验，验证所提结构单元的有效性与整体架构的合理性。

第六章对全文工作进行了总结，展望其在以下方向的进一步拓展：一是面向多声源干扰环境的复合信号分离问题；二是在实际工程中开发可部署的实时降噪系统，实现“边采集边分析”的智能化平台，推动其在海洋监测、近岸施工评估及生态保护等任务中的落地应用。

|  |
| --- |
|  |
| 图1-2 技术路线图 |

第二章 水下声学特性分析

2.1 海洋背景噪声的来源与特性

2.1.1 背景噪声的分类

海洋背景噪声广泛存在于各种水下探测和信号处理的应用场景中，其复杂的来源和多样化的特性给水下目标信号的识别与提取带来了巨大挑战。从噪声的产生机制来看，海洋背景噪声主要可分为自然背景噪声与人为背景噪声两大类[25]。

自然背景噪声是由海洋环境自身的物理现象和生物活动引起的。其中最常见的是风浪噪声，它源于风力作用于海洋表面所产生的波浪运动，并伴随大量气泡的形成与破裂。随着风速增加，这类噪声的能量显著提升，表现出连续而宽带的特征[26]。此外，海洋生物活动产生的生物噪声也是重要的自然噪声之一，鲸类、鱼类、海豚及甲壳类动物的发声活动在不同频段上表现活跃，这些声音不仅是生态系统内部交流的重要途径，也构成了复杂的海洋声学环境[27]。同时，自然背景噪声还包括由海洋湍流、水体流动以及海洋地质活动（如地震等）所引起的背景声学干扰，进一步增加了海洋噪声环境的复杂性[25]。

与自然背景噪声相对应，人为背景噪声则主要来自人类在海洋中的各种活动。这类噪声的典型代表包括船舶航行产生的噪声、海洋工程施工引起的噪声（如打桩、桥梁施工）、声纳探测以及其他工业作业产生的声信号。其中，船舶航行所产生的噪声最具代表性，其主要来源为船体机械振动、发动机工作以及螺旋桨空化现象。该类噪声具有明显的周期性和结构化特征，对海洋生态环境以及水下声学探测系统产生持续而深远的影响[28]。同时，海洋工程活动如桥梁施工、打桩等也逐渐成为海洋人为噪声的重要来源，这类工程施工通常产生幅值高且呈现明显脉冲性的信号，对近海区域的水下环境形成强烈的局部性干扰[29]。

从整体来看，自然噪声与人为噪声的区别主要体现在其产生的机制、空间分布特性及随环境因素变化的规律性上。自然噪声普遍表现为广泛而均匀的空间分布，并且受气象条件、水文特征、季节变化等环境因素的影响较大；而人为噪声则表现为局部性更强、与特定人类活动区域紧密相关的特征，通常具有更明确的结构性和时空规律性[30]。

准确的海洋背景噪声分类不仅为信号处理提供了理论支撑，也为相关噪声治理与生态保护研究奠定了基础。通过有效识别并区分各类噪声的频谱与时域特性，可以显著提高水下声学系统对目标信号的检测与提取能力[31]。例如，近年来的深度学习模型已被广泛应用于区分不同来源的背景噪声，包括生物声、船舶声及工程噪声等，证明了通过频谱特征识别背景噪声的可行性与有效性[32]。因此，系统地理解和区分不同类别的海洋背景噪声，对于改进自监督降噪算法与提升水下信号处理的整体性能具有重要的理论和实践意义。

2.1.2 海洋背景噪声的频谱特性分析

海洋背景噪声具有显著的频谱差异特性，不同频段内的噪声来源存在明显差异，深入理解这些频谱特性对于水下声信号传播规律的掌握以及降噪策略的有效设计具有重要意义。通常根据频率的不同，海洋背景噪声可划分为低频（<100 Hz）、中频（100 Hz至1 kHz）和高频（>1 kHz）三个典型频段，各频段内噪声来源的主导因素有所不同。

低频背景噪声主要由远距离传播的船舶噪声和海洋水体运动引起，其中螺旋桨空化现象被广泛认为是低频背景噪声的主要来源，典型的噪声主频约为50 Hz[33]。这一频段的噪声具有传播距离远、能量衰减缓慢的特点，因此在远离人类活动区域的海洋环境中依然能够检测到明显的船舶活动噪声。Yang等（2024）通过对Zengmu盆地的长期实测数据分析发现，在低于200 Hz的频率范围内，船舶噪声始终占据主导地位，而风浪噪声在此频段内的能量则明显较低[34]。

中频背景噪声（100 Hz至1 kHz）以风浪噪声和部分近距离的人为活动噪声为主。在浅海区域，风力作用于海面形成的波浪和气泡运动，所产生的噪声逐渐在中频频段上占据主导地位。研究表明，随着频率的提高，船舶噪声的主导性逐渐降低，而风浪噪声开始占据更重要的位置，特别是在超过400 Hz之后，风浪噪声几乎成为这一频段噪声能量的主要贡献源[34]。此外，在近岸工程施工区域，如桥梁施工、打桩活动引发的脉冲式工程噪声，也常集中分布在中频频段。这类人为噪声在时间上具有明显的脉冲特征，能量密度高且频谱范围较宽，尤其在工程活动集中的区域，中频噪声的强度和干扰性明显增强。

高频背景噪声（>1 kHz）则主要来源于降雨、生物活动和水体湍流，其中降雨噪声在高频段表现出明显的频谱集中性和瞬态波动性。Meng等（2017）在浅海环境中对10至35 kHz的高频背景噪声进行了长时间观测，结果显示环境状态显著影响了高频噪声的波动特征，且环境因素变化（如降雨强度、风速变化等）会引起该频段噪声的快速波动[35]。类似的研究还发现，在具有不同植被覆盖的浅水环境中，如南葡萄牙的Ria Formosa泻湖区域，不同种类海草覆盖区域的背景噪声频谱特性表现出明显的一致性，尽管植物种类各异，但整体高频噪声的频率变化趋势却相对稳定，这提示高频背景噪声在不同浅水生态系统中的普适性特征[36]。

在深海区域，Iqbal等（2019）采用Monte Carlo方法对印度洋深海区域的背景噪声频谱特性进行了研究，结果发现深海区域的宽带噪声频谱更接近于高斯分布，而窄带背景噪声则表现出明显的非对称性，这表明窄带噪声频谱需要更复杂的统计分布模型进行描述[37]。针对高背景噪声场的建模问题，Zhang等（2018）提出了基于随机信号生成与声传播理论的噪声场建模方法，可有效模拟大范围海洋区域的复杂干扰噪声场[38]。

2.2 常见人为水下声信号类型概述

2.2.1 桥梁振动声信号来源与特性

桥梁在运营过程中，主要噪声源包括车辆行驶引起的结构振动和机械噪声。车辆通过桥面时，轮胎与路面的接触、发动机运转以及车辆自身的振动都会引发桥梁结构的振动。这些振动通过桥梁的支撑结构传递至水中，形成水下噪声。此外，如图2-4桥梁上的机械设备，如伸缩缝、支座等在运行过程中也可能产生机械噪声，这些噪声同样会通过结构传导至水体中[42]。



**图2-4厦门某海域桥梁观测图**

桥梁运营期间产生的水下噪声频谱特性与噪声源的类型和桥梁结构密切相关。车辆引起的振动通常包含低频成分，主要集中在20 Hz至200 Hz之间，这是由于车辆重量和桥梁结构的固有频率所致[42]。机械设备运行产生的噪声则可能包含中高频成分，频率范围在200 Hz以上。总体而言，桥梁运营期间的水下噪声频谱呈现出低频为主、高频衰减的特点，具体可分为时间与空间两个方面的特性。

桥梁水下噪声的时间变化主要受交通流量和车辆类型的影响。在高峰时段，车辆密度增加，桥梁结构承受的动态载荷增大，导致水下噪声水平升高。相反，在夜间或交通低谷期，噪声水平相对降低。此外，重型车辆如卡车、大巴等通过桥梁时，产生的振动幅度更大，导致瞬时噪声水平的显著提升[43]。因此，桥梁水下噪声具有明显的时间波动性，与交通状况密切相关。

桥梁运营期间的水下噪声在空间上的分布受多种因素影响，包括桥梁结构形式、水深、底质类型以及水流速度等。一般而言，噪声源强度在桥梁下方和附近水域最高，随着距离的增加，噪声强度逐渐衰减。水下噪声的传播还受到水体声速剖面和海床反射特性的影响，可能出现声影区或声聚焦区，导致噪声分布的不均匀性[44]。此外，水流的存在可能改变噪声的传播路径，使得下游方向的噪声水平高于上游方向。

桥梁在运营期间产生的水下噪声是多种因素共同作用的结果。车辆引起的结构振动和机械设备运行是主要的噪声源，其频谱特性以低频为主，时间变化与交通流量密切相关，空间分布则受到环境条件的影响。深入研究这些特性，有助于制定有效的噪声控制措施，减轻桥梁运营对水生生态环境的影响。

2.2.2 打桩声信号来源与特性

如图2-5，在水下打桩施工过程中，冲击噪声是最显著的人为水下噪声来源之一，其主要由打桩设备施加在桩基上的高强度瞬时冲击力所引发。这一冲击力通过桩体向下传导，并激发了周围水体中的声波辐射，形成了具有高能量、宽频带的水下声场[53]。打桩噪声的瞬时特性表现为极高的声压级和快速的上升时间，每次打击会在极短的时间内释放出巨大的声能。这类噪声的声压级通常可达200 dB re 1 μPa以上，特别是在打桩初期，声压峰值显著高于后续打击的声强[56]。



**图2-5厦门某海域打桩现场施工观测图**

打桩过程中产生的水下噪声在时频特征上表现出高度的复杂性，通常包含宽频带和高强度的瞬时脉冲特征。打桩产生的每次冲击信号会在水下形成具有特定频率成分的脉冲噪声，这些频率成分主要集中在低频至中频范围内，通常覆盖从数十赫兹到几千赫兹的频率区间[55]。其中，低频段的能量分布最为显著，主要集中在100 Hz至1 kHz之间，这与打桩设备的能量传递特性和水体对声能的吸收规律密切相关[57]。在高频区间，尽管声能衰减较快，但仍然会形成一定的能量峰值，特别是在打桩材料和海底地质结构对声波的散射作用下，高频噪声信号可能在局部区域得到增强[58]。

从时间域的角度来看，打桩噪声具有明显的非平稳性和脉冲性特征，每次打击会产生一个持续时间较短但能量极高的声脉冲。这些脉冲信号具有快速上升的前缘和较长的尾部衰减特性，形成所谓的“尖峰-尾随”时域特征[60]。打桩频率和节奏的不同也会影响噪声在时间上的分布特性，例如在高频率打桩作业中，声波脉冲会出现重叠现象，导致局部区域的背景噪声水平显著升高[59]。

在传播特性方面，打桩噪声的传播路径和能量衰减受到多种环境因素的影响，包括水深、海底地质结构、水体声速剖面以及水流的存在等。研究表明，打桩噪声在传播过程中表现出明显的方向性特征，其能量主要沿桩体轴向向外扩散，形成以打桩点为中心的同心圆形传播模式[54]。在浅水区，由于海底反射和水体多径传播效应的存在，打桩噪声会形成复杂的干涉现象，导致特定区域内的噪声强度增强或削弱[55]。此外，声波在不同传播距离上的衰减速率也存在显著差异，通常在近场区域（距离打桩点数百米以内），声压级随距离增加呈现快速衰减趋势，而在远场区域（数千米之外），由于多次反射和散射作用，噪声衰减速度趋于平缓[53]。

打桩噪声对水生生态系统的影响也不容忽视，尤其是在频率范围与海洋生物感知频带重叠的情况下，打桩噪声会干扰海洋生物的行为模式、通讯系统甚至导致生理损伤[56]。例如，研究表明，印太驼背豚会在打桩作业期间远离施工区域，表现出明显的回避行为[57]。因此，在进行打桩作业时，必须充分评估噪声传播对周围环境和生态系统的潜在影响，并采取相应的缓解措施[59]。

综上所述，打桩过程中产生的冲击噪声具有瞬时强度高、频率覆盖广、时域特性明显的特点。在传播过程中，打桩噪声会受到水体环境因素的多重影响，形成复杂的传播模式。深入研究打桩噪声的时频特性和空间传播规律，对于制定有效的噪声控制策略、保护水生生态环境具有重要的理论和实践意义。

2.3 水下工程声信号评估参数

水下工程声信号的声学特性通常由多个声学参数进行量化描述，包括声压级（Sound Pressure Level, ）、峰值声压级（Peak Sound Pressure Level,）、均方根声压级（Root Mean Square Sound Pressure Level）以及声暴露级（Sound Exposure Level, SEL）等。其中，人为水下声信号通常表现为高源强的宽频脉冲信号，适用于描述的参数分别为峰值声压级、均方根声压级和声暴露级[40]。

2.3.1 声压及其度量参数

声波是由于介质振动产生的，当声波在介质中传播时，会引起介质内部的压力变化，这种压力变化被定义为 声压（Pressure, ）。声压是描述声信号强弱的重要物理量，随时间变化，通常以帕斯卡（Pa）或微帕（μPa）为单位。在海洋环境中，例如海洋冲击打桩这类人为水下声信号的测量通常采用水中声压的幅值表示。其中，声压幅值的最大绝对值称为峰值声压（），该值反映声压波动的极大程度，数值通常较大，可达到数个数量级。

为了便于量化和比较声信号的强度，常采用 进行表征。声压级通过对数标准化方式表示声压大小，其单位为分贝（dB），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。将峰值声压 入该公式，即可得到 峰值声压级（Peak Sound Pressure Level, ），用于描述声信号的瞬时最大声压。

2.3.2 均方根声压级

均方根声压级（Root-mean-square Sound Pressure Level, ）用于表征声信号的均方根声压（Root-mean-square Sound Pressure, RMS），即声压随时间变化的均方根值的对数表达形式。其单位为dB，数学定义如下：

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

其中， 为均方根声压，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）。均方根声压级 的单位为 dB。

在脉冲声信号分析中，均方根声压级有时也被称为有效值声压级（Effective Sound Pressure Level），但其计算方式可能因声源类型的不同而有所区别。对于脉冲噪声和脉冲噪声类信号，计算 有效值声压级 时通常会考虑信号能量的主要贡献部分，而排除两端能量较低的部分。例如，常见的计算方法是取 信号总能量的90% 进行计算，即剔除波形两端能量较小的前5 % 与后5 % 部分。这一方法能够更准确地反映信号的有效声压水平，从而提高数据分析的稳定性和可比性[41]。

在人为水下声信号监测与分析中，均方根声压级是衡量声信号能量分布 及其时均强度的重要指标，能够为噪声影响评估、环境监测以及工程应用提供关键参数。

2.3.3 声暴露级

声暴露级（Sound Exposure Level, ）是用于衡量单一或离散的水下声事件的声能量积累，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）。单位同样为dB re 1 μPa²s。

声暴露级主要用于瞬态或间歇性脉冲声信号，如海上冲击打桩等水下声源的评估。对于脉冲信号，常采用90% 能量持续时间法进行计算，即仅考虑波形中累积 90% 能量的有效时间段（），以更准确地反映声能的真实积累情况。

2.4 本章小结

本章围绕水下声学信号的基本特性展开了系统性研究，分别从海洋背景噪声的来源与频谱特性、典型人为工程信号的来源与结构特征，以及水下声学信号评估参数三个方面进行了深入探讨与分析。

首先，系统阐述了海洋背景噪声的分类与特性。根据其产生机制与空间分布特征，背景噪声主要分为自然背景噪声与人为背景噪声两大类。自然背景噪声以风浪、生物活动和环境湍流为主要来源，频谱宽、空间分布较为均匀且受环境因素影响明显；人为背景噪声则以船舶航行、海洋工程施工与声纳活动为代表，具备明显的局部性、结构性及明确的时空分布特征。频谱特性分析进一步揭示，不同频段的背景噪声具有明显的主导来源差异，低频段以船舶噪声为主，中频段主要为风浪及工程施工噪声，高频段则以降雨、生物活动和湍流噪声为主导。

其次，本章对桥梁振动声与打桩声两类典型人为水下工程信号的产生机制及传播特性进行了详细阐述。桥梁振动声源于车辆通过桥梁时产生的结构振动与机械噪声，频谱主要集中于低频段，时间分布随交通流量变化，空间分布受水深及环境因素明显影响。打桩声表现出强脉冲特性，能量高度集中，频谱范围宽广，传播过程中受到多径效应、海底反射及环境条件影响，形成复杂的传播模式，对水生生态产生显著干扰。

最后，本章阐述了水下工程声信号的重要评估参数，包括峰值声压级 、均方根声压级 与声暴露级 。这些参数从不同角度量化了水下声信号的特性，为后续声信号分析与工程应用提供了必要的理论基础和方法工具。

在上述分析基础之上，水下工程声信号的实际采集、特征识别与详细分析成为进一步研究的关键内容。下一章将围绕实际工程场景，开展水下工程脉冲声信号的现场数据采集与深度分析，以明确信号的具体结构特征、传播规律及降噪需求，为后续降噪方案与网络模型设计奠定基础。

第三章 水下工程声信号采集与分析

3.1 采集方案介绍

3.1.1 数据采集系统

本研究的数据采集系统主要由水听器、数据采集卡及笔记本电脑等设备组成，用于高精度获取和分析人为水下声信号。系统的核心组件包括水听器用于声信号接收，数据采集卡用于信号转换，笔记本电脑用于处理相关数据。

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。其频率响应范围为0.1 Hz 至160 kHz，涵盖了绝大多数水下人为声源的频段，且最大工作声压可达263 dB（9.8×10⁶ Pa），各项技术参数均满足采集工作需求，具体规格详见表 3-1。

表3-1 B&K 8105型水听器主要参数

|  |  |
| --- | --- |
| 项目 | 参数 |
| 水听器型号 | 8105 |
| 标称电压灵敏度 | 56 μV/Pa |
| 电容 | 7250 pF |
| 频率响应 | 0.1 Hz - 160 kHz: +3.5/-10.0 dB |
| 垂直方向性 | 0° - 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。

3.1.2 采集方案

如图3-1（a）所示，为了多维度分析水下工程声信号的传播特性，采集主要依托调查船作为平台（或海上平台），在水体中垂直布设多层水听器，并通过高精度数据采集系统记录水下声信号。

|  |  |
| --- | --- |
| (a) | (b) |
| 图3-1 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 | |

在数据采集过程中，调查船或调查平台搭载数据采集设备，并负责水听器的布放和回收。为了避免调查船自身噪声对数据的干扰，需关闭机械设备并抛锚，以减少船舶动力系统产生的噪声，使采集环境尽可能接近无干扰状态。水听器布设采用多层分布方式，即分别在不同水深处布设表层水听器、中层水听器和底层水听器，以同步获取不同深度的水下声信号。所有水听器均通过电缆连接至调查船上的数据采集系统，实现实时声信号传输。此外，为保证水听器在水中的垂直稳定性，在水听器阵列的底端挂设铅锤重物，以抵抗水流扰动，确保水听器始终保持稳定的姿态，避免因晃动导致数据失真，见图3-1（b）。

监测过程中，水听器阵列采集的水下声信号直接传输至调查船上的数据采集系统，并通过高精度数据采集卡进行数字化处理，并详细记录水深、海况等海洋环境参数。

3.2 打桩工程水下声信号采集

3.2.1 采集环境与布置

如图3-2所示，信号采集自厦门某海域的工程打桩现场。施工过程中采用直径为2.1 m 至3.0 m 的钢管复合桩，当钢管桩遇到较硬的海底地层无法继续压沉时，需借助液压冲击锤进行打桩作业。该过程中产生的冲击声是水下打桩声信号的主要来源。

|  |
| --- |
| 图3-2 厦门某海域打桩现场图 |

如图3-3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动态移动，并抛锚固定采集点位，实现对不同离桩中心距离处的声信号监测。

|  |
| --- |
| 图3-3 打桩工程水下声信号采集布置 |

3.2.2 信号特性分析与讨论

为全面揭示水下打桩声信号的传播特性，本节从时域波形结构、时频能量分布特征及频带能量分布规律三方面，对固定点（距打桩锤约50 m）采集到的声信号进行深入分析。

图3-4所示为该信号的时域波形，呈现出典型的瞬态脉冲特征：每个脉冲在极短时间内迅速达到峰值，随后快速衰减。主脉冲之后普遍伴有明显的尾波振荡现象，表明声波在传播过程中受到水面、海底等界面及周围障碍物的多次反射，形成多径传播效应，从而显著延长了信号的衰减时间。

|  |
| --- |
|  |
| 图3-4 打桩平台测得打桩声信号时域图 |

图3-5给出了该打桩声信号的频谱图，打桩声信号的能量在脉冲出现时呈现宽频带分布特征，在脉冲初始阶段存在明显的频率扩展现象并向高频延伸，随后在低频范围仍保持较长时间的余波振荡，印证了时域中观测到的多径效应。

|  |
| --- |
|  |
| 图3-5 打桩平台测得打桩声信号频谱图 |

为进一步分析打桩声信号的能量分布特性，在同一采集点位记录了非打桩时段的声信号，作为海洋背景噪声进行对比。图3-6（a）展示了打桩声信号与背景噪声的功率谱密度（PSD）对比结果。可以看出，打桩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）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，体现出其明显的低频主导特性。随着频率的升高，打桩声与背景噪声之间的差异逐渐减小，说明其高频分量衰减较为明显。

|  |  |
| --- | --- |
|  | |
| (a) | (b) |
| 图3-6 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 | |

为进一步量化不同频带内的能量分布，图3-6（b）展示了打桩声与背景噪声的1/3倍频程频带声压级（SPL）对比结果。可以看出，打桩声信号的能量主要集中在4 kHz 以下频段，且在各频带内均显著高于背景噪声。由于海洋背景噪声受水体运动影响，其能量主要分布在中低频范围，因此在低频段，两者差异并不明显。相比之下，在约400 Hz 频带附近，打桩声与背景噪声的差值达到峰值，约为40 dB；随着频率的进一步升高，二者的差异逐渐趋于稳定，维持在约25 dB。

综上所述，水下打桩声信号呈现显著的瞬态脉冲及多径传播特征，尾波振荡明显，能量分布集中在中低频段，尤其在中频段（100 Hz 至1000 Hz）表现出与背景噪声的明显差异。这些特征的深入分析为后续水下工程声信号的降噪方法设计和信号特征识别研究奠定了重要的理论基础。

3.3 跨海桥梁振动水下声信号采集

3.3.1 采集环境与布置

本研究选取集美大桥作为桥梁振动水下声信号采集的目标。如图3-7所示，桥梁全长约8.4千米，其为多桥墩多跨结构，桥面设置有钢制伸缩结构。

|  |
| --- |
| 图3-7 集美大桥 |

如图3-8（a）所示，使用走船式布置进行水下声信号采集，在水下1 m、4 m和接近河床处（约7~8 m）的深度布放水听器进行同步采集。如图3-8（b）所示布设了A、B、C三个测点，分别距最近的桥墩32 m、34m与32m，测点B位于桥梁钢制结构下方。

|  |  |
| --- | --- |
| (a) | (b) |
| 图3-8采集布置与采集点位示意图：(a) 走船式采集布置；(b) 集美大桥桥墩及监测点A、B、C位置 | |

3.3.2 信号特性分析与讨论

各测点在不同水深下的桥梁交通水下噪声峰值声压级、均方根声压级和声暴露级如表3-2所示。监测点A的三类声学指标平均值分别为：峰值声压级155.2 dB，均方根声压级120.4 dB，声暴露级125.5 dB；监测点B对应的平均值分别为158.5 dB、127.5 dB 和133.0 dB；监测点C则分别为157.7 dB、122.9 dB 和126.3 dB。

表3-2 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** |

从对比结果可见，监测点B在三个声学指标上均为最高，监测点A则为最低，其中峰值声压级最大差值为3.3 dB，均方根声压级差值为7.1 dB，声暴露级差值为7.5 dB。此外，监测点A和B的声压级最大值均出现在4 m水深处，而监测点C的最大声压级出现在7 m水深处。

为全面揭示水下桥梁振动声信号的传播特性，本节从时域波形结构、时频能量分布特征及频带能量分布规律三方面，对测点B（位于桥梁钢制结构下方，距桥墩约34 m）在4 m 水深采集到的声信号进行深入分析。

如图3-9所示，桥梁振动产生的水下声信号在时域上表现出显著的非平稳特性。从中可以观察到，背景噪声水平相对稳定。水下声压会出现尤其明显的短时脉冲峰值：在约58 s处捕捉到数次显著的脉冲，红框突出显示了1 s 内声压幅值的变化情况，其声压瞬时峰值达到约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）。车辆轮胎冲击伸缩缝垂直面的瞬态激励，是桥梁机械结构振动的主要来源[61]。特别是大型车辆通过时，引起桥墩与地基振动，向水下辐射声波，其中桥墩振动对水下声信号总声级的贡献最大[62]。

|  |
| --- |
| 图3-9 桥梁振动声信号时域图 |

如图3-10所示，该时段声信号的频谱图进一步展示了其频域特征。图中红框部分突出标示了在约58 s 处捕捉到的多次显著脉冲信号所引起的频谱变化。可观察到，这些脉冲信号在频域上主要集中于中低频段（1 kHz 以下）呈现出典型的短脉冲信号柱状频谱特征。整体来看，背景噪声主要分布于中低频段，表现较为平稳，但期间叠加了间歇性出现的高能量短脉冲成分。结合表3-2中的数据可见，位于桥梁钢结构下方的测点B，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含有强烈的瞬态脉冲特征。

|  |
| --- |
|  |
| 图3-10 桥梁振动声信号频谱图 |

为进一步揭示桥梁振动声信号的组成成分，在远离桥梁的海域（距离集美大桥约680 m，离岸最短距离约784 m，且采集过程中无其他船舶经过）采集了水下声信号，作为该海域的环境背景噪声用于对比分析。图3-11展示了水深4 m处桥梁振动声信号与背景噪声的PSD曲线及三分之一倍频程声压级对比结果。

|  |  |
| --- | --- |
|  | |
| (a) | (b) |
| 图3-11 桥梁振动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3-11（a）所示，桥梁振动声信号的能量主要集中在中低频段（1 kHz以下）。在约100 Hz 频率范围内，PSD曲线出现显著峰值，峰值声压级超过110 dB。随着频率升高，信号能量迅速衰减：频率高于100 Hz 时谱级降至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 以上频段已接近背景噪声水平（低于80 dB）。

为进一步量化不同频带内的能量分布，图3-11（b）展示了桥梁振动声信号与背景噪声的1/3倍频程频带声压级（SPL）对比结果。可以明显看出，在1 kHz以下的中低频段，二者声压级存在显著差异，尤其在约500 Hz处达到最大差值，约为25 dB。这一频谱特征表明，桥梁振动声信号的能量主要集中于中低频范围，且随着频率进一步升高，其声压级逐渐趋近于海洋背景噪声水平。

桥梁振动声信号的这一中低频特性与其传播机制密切相关。该类声源主要源于车辆荷载作用引发的桥梁结构振动，通过桥面板等钢制结构单元与水体发生耦合辐射，从而形成低频主导的水下声场。尤其对于钢结构桥梁，由于其刚性较强、传导性良好，当车辆快速通过时，桥体结构中易激发强烈的低频结构波，并在与水体耦合过程中转化为水下声波，形成持续或间歇性的声辐射。

3.4 降噪分析与自监督策略引入

基于第3.2节与第3.3节对打桩声与桥梁振动声两类水下工程脉冲信号的分析结果可知，两者在结构特征与频谱分布方面具有明显共性。首先，这类水下工程脉冲信号均表现出短时高幅度的瞬态冲击特征，在时域上具有突发性、非周期性以及主能量集中的特征；在频域上则主要集中于中低频段，频率能量分布多集中于100 Hz至1 kHz之间。

然而，水下复杂传播环境引起的多径效应、边界反射和尾波冗余等现象，使得信号结构易受非结构性干扰，增加了降噪任务的难度。传统降噪方法（如维纳滤波与小波变换）在理论上适用于平稳或局部瞬态噪声的场景，但在水下环境的多径干扰与频谱交叠条件下，表现出明显的局限性。具体而言，维纳滤波在频谱估计准确性不足时易导致有效信号的削弱，小波变换则在复杂尾波干扰下难以实现稳定的主结构恢复。

另一方面，近年来深度学习技术在语音增强与音频分离领域已展现出显著的性能提升，但监督学习方法通常依赖大量高质量、标注明确的纯净信号作为训练标签。然而由于真实海洋环境中难以获取绝对纯净的目标信号，传统监督学习方法对高质量标注数据的依赖成为主要制约因素。尤其对于突发性、结构性显著的水下脉冲事件，采集难度大、样本分布不均等问题进一步加剧了数据瓶颈。

鉴于上述传统方法和监督学习方法存在的问题，引入无需纯净标签的自监督降噪策略成为一种具有明显优势的路径。自监督方法能够从数据本身的结构特征出发构建训练目标，有望规避数据稀缺与标注困难带来的瓶颈，并更贴合本文所分析的水下结构性脉冲信号特征。下一章将在此基础上进一步提出基于信号能量分布特性的结构定义，构建具体的自监督训练策略。

3.5 本章小结

本章围绕水下工程脉冲声信号的实测数据采集与频谱特性分析展开，聚焦于打桩作业与桥梁振动两类典型工况，构建了面向不同工程场景的结构声数据集，并对其时域波形、频谱分布及传播特性进行了系统分析。研究表明，水下工程脉冲声信号普遍具有突发性强、持续时间短、主能量集中的结构特征，频谱能量主要集中于100 Hz 至 1 kHz 的中低频段；主脉冲结束后常伴随明显的尾波拖延和非结构性干扰，增加了信号结构的复杂性。

分析结果进一步揭示了当前水下工程脉冲声信号降噪任务所面临的关键挑战，即如何在保留主结构能量的同时，有效抑制冗余尾波等干扰成分。传统降噪方法如维纳滤波、小波变换在处理此类结构性复杂信号时效果有限，难以兼顾信号还原与噪声抑制。而监督学习方法则受到水下纯净数据难以获取的限制，面临显著的数据瓶颈问题。

为克服上述困难，本章进一步分析和探讨了自监督降噪方法的可行性。自监督方法不依赖于人工标注的纯净信号，而是通过数据自身的结构规律来建立训练目标，可望有效突破数据稀缺问题，同时又符合本文所揭示的水下工程信号的结构特性与降噪需求。因此，引入自监督策略为下一章设计结构保持型深度降噪方法奠定了必要的理论与实践基础。

综上，第三章完成了水下工程脉冲声信号从实测采集到特性分析的基础性工作，明确了信号的典型结构特征与降噪难点，并引入了自监督降噪策略的可行性与必要性。下一章将在此基础上，系统展开自监督训练策略与网络模型的设计与构建，进一步推进适用于复杂背景下水下结构性脉冲信号的降噪方案研究。

第四章 自监督降噪方案设计

本章的核心目标是提出一种基于自监督深度学习的背景噪声抑制方法，以提升水下脉冲信号（如打桩、桥梁振动信号）的提取能力。具体内容包括方法的总体思路、模型架构、伪“干净”目标构造策略、训练优化方案等。

4.1 方案设计思路

经过第三章对人为水下噪声的时域、频域特征分析（见第三章相关内容），可以发现人为水下声信号具有明显的低频噪声主导以及脉冲噪声突出的特性。其中，海洋背景噪声在低频段能量强盛（例如风浪、湍流等）[64]，而人为水下噪声（例如水下打桩、桥梁伸缩缝结构振动等）等事件产生的脉冲信号在时域上表现为高幅值的瞬态，在频域上具有宽频带分布但缺乏连续性。这些特征给传统降噪方法带来挑战：一方面，持续的低频噪声容易掩蔽目标信号；另一方面，短暂脉冲信号频谱的连续性缺失使得平滑滤波等方法难以同时保留瞬态细节并去除背景噪声干扰。

此外，水下环境复杂，获取有效的纯净样本训练数据难度与成本极高，导致传统监督学习模型易受噪声干扰且泛化性能下降。为解决上述难题，本章设计了一种基于深度学习的自监督降噪方案，利用仅含噪声的水下声学数据进行模型训练，在无需纯净样本的情况下实现对目标信号（如桥梁振动噪声脉冲成分）的提取与降噪。

4.1.1 方法背景

水下声信号降噪的传统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纯净（Noise-to-Clean, N2C）策略[66]。在这种策略中，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纯净信号的映射关系：

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

其中， 为神经网络对含噪音频输入 的估计输出， 为对应地纯净参考信号。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[65]。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。

针对上述问题，典型的自监督学习策略Noise2Noise (N2N) 通过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[67]。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：

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

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 (4-2) 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 (4-2) 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：

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

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[69]。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两个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) 策略[70] ，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[71]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 ：

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

通过引入偏差项的差值 ，其中 是噪声项，并假设 ，可以得到：

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

考虑到理想的降噪网络 是用纯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：

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

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。

4.1.2 自监督学习策略的局限性

自监督策略引入子采样器虽然解决了N2N方法的数据配对局限性，但带来了新的挑战，即子采样导致信号频谱结构连续性丢失。具体来说，子采样会破坏频谱图中结构与纹理的自然流动，使信号在时间和频率两个维度上的信息被割裂，影响降噪后音频信号的整体完整性与真实度。

短时傅里叶变换（STFT）是一种广泛应用于音频信号处理中的时频分析方法，它通过窗函数将长信号划分成多个短帧，对每个短帧分别进行傅里叶变换，从而提供信号在时间和频率两个维度上的精细表示。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：

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

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：

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

图4-1（a）展示了原始音频信号的频谱图，其中红色方框突出显示了沿特定时间轴的音频柱状频谱，捕捉了20个时间帧的音频频谱信息。而在图4-1（b）中，红色方框显示了对应的柱状音频频谱信息被压缩到仅10个时间帧内。原本在频谱上体现出连续结构或瞬时脉冲的能量特征，在新的时间分辨率下被压缩至更少的时间帧中，导致频谱中细节信息和结构连续性遭到严重破坏。这种信息的损失尤其影响对具有短暂脉冲特征的信号的特征提取效果。例如，第三章分析的桥梁振动引起的瞬态脉冲信号，其频谱特征能量被压缩后会呈现明显的模糊与失真，不利于网络准确学习完整的目标信号特征，也降低了后续频谱重建的质量。

|  |
| --- |
| （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）DMFNet网络处理频谱卷积核采样点分布情况 |

在STFT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。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：

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

由公式（4-9）可知，当FFT点数 保持不变而采样率 减半时，频率分辨率 和奈奎斯特频率同时减半。具体而言，频谱图的频率轴会被压缩，导致频率范围缩窄。此时，相邻频率点之间的间距增大，使原本分布在多个频率点上的谱成分集中至更少的频率点上，造成不同频率成分之间的区分度降低，影响频谱的解析能力。此外，超过子采样后奈奎斯特频率一半的频率成分可能因混叠效应而折叠至低频部分，进而影响频谱图的完整性与准确性。

这种子采样的方式带来了频谱结构连续性受损的缺点。第三章分析表明，桥梁振动引起的水下声信号在频谱上表现为低频连续噪声伴随瞬时宽带脉冲噪声。子采样处理使频谱的局部纹理和结构被严重破坏，表现为时间分辨率和频率分辨率的降低，影响了网络模型对瞬时宽带脉冲等关键结构信息的有效捕捉。这种频谱连续性缺失不仅削弱了网络的特征提取能力，也降低了去噪后信号的质量与真实感。

在自监督策略受限的情况下，本研究提出了一种新型的自监督动态多重聚焦音频降噪网络（Dynamic Multi-Focusing Network，DMFNet），以解决自监督策略子采样后频谱图信息丢失的问题。在图4-1（a）、（b）和（c）中，黄色点代表卷积核的中心，红色点代表卷积核的采样点。如图4-1（c）中白色方框所示，传统卷积核在水平与垂直方向上仍然对“压缩”后的频谱位置进行固定的采样，无法有效捕捉子采样后频谱的变化特征。而图4-1（c）中红色方框展示了我们网络的效果：通过自适应调整卷积核采样点位置，网络可以更聚焦于目标信号的频谱特征。具体来说，在水平方向上根据目标信号频谱结构自适应调整，从而准确捕捉沿时间维度分布的频谱柱状结构；在垂直方向上显著增加采样点密度，弥补子采样后频谱压缩造成的信息缺失。此外，DMFNet在逐层设置了特征提取模块（详见4.2节），主动弥补子采样带来的信息损失，使模型能够有效识别并提取水下人为声信号中瞬态脉冲的频谱特征，同时抑制海洋环境中的连续背景噪声，显著提升整体降噪性能。

4.1.3 方案概述

本研究提出的自监督深度学习降噪方案整体流程如图4-2所示。在训练阶段，DMFNet仅使用含噪音频数据进行网络训练。具体来说，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 |

在网络模型训练阶段，我们强制网络学习两个子采样输入之间的共同信号结构，自动忽略随机扰动带来的差异部分，从而实现网络对脉冲信号结构性特征的捕捉与对背景随机噪声的有效抑制。当模型通过持续的损失最小化逐渐收敛时，模型即学习到了水下人为噪声的稳定内在特征，如桥梁振动噪声所对应的脉冲频谱结构。充分训练后的模型在测试阶段仅需将原始含噪频谱输入，即可完成高质量的降噪重构输出。

最后，网络输出的频谱经由ISTFT重构回时域，形成最终降噪后的信号输出。整个自监督降噪方案全程无需外部纯净数据辅助，完全基于含噪数据自身驱动，实现模型的有效训练与降噪重构。。

4.2 网络模型架构

为有效从复杂海洋背景噪声中提取人为水下声信号（如桥梁振动噪声、打桩声等脉冲信号），本研究提出了一种自监督降噪模型架构——DMFNet。该模型采用编码器-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，有针对性地对水下脉冲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DMFNet网络架构 |

如图4-3所示，DMFNet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，DMFNet能够在无纯净标签监督的情况下，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DMFNet整体架构以及各关键模块的设计原理与优势。

4.2.1 编码器-解码器结构

DMFNet总体采用对称的编码器-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一个创新模块——频谱动态聚焦单元（SDFU）。SDFU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）。这一设计对于桥梁振动、打桩等脉冲信号尤为重要，因为此类人为水下声信号的频谱通常呈现不连续的高能量瞬时脉冲（例如在时频谱上表现为短暂出现的纵向高能量柱状结构）。借助SDFU对卷积感受野的动态调整，当频谱中出现这类瞬时突出的能量特征时，网络能够更有效地捕捉细节脉冲信息，避免其在后续下采样过程中被遗漏。

在SDFU预处理之后，每级编码器模块由复数注意力模块（CAM）和下采样操作组成。CAM模块旨在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（详细原理见4.2.3节）。与只关注幅度的传统注意力机制不同，CAM对频谱的实部（幅度）和虚部（相位）特征同时赋予关注，以充分利用复数谱图中的信息。这使得网络在特征提取时能够兼顾频谱宏观能量分布和微观相位细节，更有效地区分目标信号与背景噪声。例如，当桥梁振动产生瞬时振动脉冲时，频谱上会同时出现幅度的突增和相位的剧烈变化；CAM可以敏锐捕捉并强调这些复数域特征，从而增强网络对目标脉冲的辨识能力，抑制环境噪声干扰。经过CAM处理后的特征再通过下采样模块进行空间尺寸压缩（例如使用步长为2的卷积或池化），以输入下一尺度的编码器。

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：

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

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[72]，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简单拼接，本模型在每层解码前引入复数多尺度跳跃连接单元（CMS-SCU）来替代直接跳跃连接。CMS-SCU通过对编码端不同尺度的特征进行对齐融合（详细见4.2.5节），实现编码特征与当前解码特征的深度整合。融合后，解码器模块内部还会再次使用CAM模块，以进一步细化频谱重建过程中的特征交互与优化。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。例如，我们使用两个3×3、步长为2的实值转置卷积核分别作用于复数特征的实部和虚部，实现空间上采样的同时重新引入高频细节；并配合CBN和CReLU对重建特征进行归一化和非线性变换，提升信号重建质量。

综上，DMFNet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，实现多尺度下特征的自适应融合。在复杂水下噪声的自监督降噪任务中，该架构有效克服了子采样引起的频谱不连续性问题，在无需纯净参考信号的前提下突出并恢复出稳定的目标脉冲特征，显著提升了模型的降噪性能和鲁棒性。

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DMFNet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[73]，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。

|  |
| --- |
| 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：

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

受到可变形卷积[73]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：

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

同样的，沿着垂直轴（频率轴）方向，卷积核 位置定义变为：

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

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：

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

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：

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

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）[74]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：

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

其中符号 表示通道维度的拼接操作，函数 的具体定义如下：

|  |  |  |
| --- | --- | --- |
|  |  | (4-17) |

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DMFNet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，共同作用以实现对频谱图局部细节与全局关联的自适应建模。

|  |
| --- |
| 图4-5 CAM组成结构示意图 |

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）（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：

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

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[75] 的思想，分别针对实部特征和虚部特征在频谱图横向（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，第三个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[76]。

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：

|  |  |  |
| --- | --- | --- |
|  |  | (4-19) |

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，没有纯净标签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，这对于准确提取水下脉冲型人为声信号至关重要。

4.2.4 语义重校准单元（SRU）

在DMFNet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：

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

其中函数 的具体定义如下：

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

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：

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

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：

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

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：

|  |  |  |
| --- | --- | --- |
|  |  | (4-24) |

其中， 代表通道维度的多层感知机（Multi-Layer Perceptron, MLP）[77]，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[78]。

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：

|  |  |  |
| --- | --- | --- |
|  |  | (4-25) |

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，它能识别出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。设第 个编码子相关的所有特征信息 计算如下：

(4-26)

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：

|  |  |  |
| --- | --- | --- |
|  |  | (4-27) |

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：

(4-28)

|  |  |  |
| --- | --- | --- |
|  |  | (4-29) |

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纯净参考，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：水下环境中的人为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，本文提出了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，从而提升降噪性能。

|  |
| --- |
| 图4-8 CMS-SCU组成结构示意图 |

如图4-8所示，CMS-SCU的核心思想是在传统跳跃连接基础上，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$l$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，有助于提炼跨尺度的综合表示。融合后的特征再送入一个通道注意力机制（Channel Attention, CA）进行加权筛选。通道注意力为融合特征的每一通道分配一个权重（范围0到1），表示该通道对当前重建任务的重要程度。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（例如代表脉冲结构的通道），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，实现更纯净的重建。

借助CMS-SCU，多尺度特征在解码过程中得以充分交互和互补。特别地，在存在低频背景噪声与高频瞬态脉冲共存的情况下，CMS-SCU允许网络灵活调整感受野大小：对于描述缓变背景的全局模式，网络可以依赖融合进来的粗尺度特征进行刻画；而对于定位瞬时脉冲峰值这样的细节，网络则可以借助细尺度特征进行精确重构。这一融合策略保证了不同频率成分的信息在重建时被全面保留。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，而是通过跨尺度的信息流得到统一考虑。这不仅增强了模型对频谱全局结构的理解力（确保背景噪声的整体形态被正确建模），也提高了其对局部细节的分辨力（确保脉冲信号的尖峰细节被精确重现）。因此，在自监督降噪任务中，CMS-SCU提供了一种稳定而高效的特征表示方式：即使缺乏明确的监督信号，网络也能凭借多尺度融合获得对目标信号更完整的表征，使最终输出的去噪频谱同时具备正确的整体噪声水平和细节信号特征，实现对水下目标脉冲信号更出色的提取与还原。

4.3 自监督学习策略与算法细节

4.3.1 伪“干净”目标构造方法

基于DMFNet的网络架构，本节聚焦模型的自监督训练策略，包括训练数据的构造方法及优化目标设定。由于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。我们借鉴了Nbr2Nbr策略[70] ，在单通道水下噪声数据上应用随机子采样策略，以生成用于自监督训练的数据对。

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：

|  |  |  |
| --- | --- | --- |
|  |  | (4-30) |

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：

|  |  |  |
| --- | --- | --- |
|  |  | (4-31) |

同时，设每组的起始索引序列为：

|  |  |  |
| --- | --- | --- |
|  |  | (4-32) |

则每一组的子采样信号可分别表示为：

|  |  |  |
| --- | --- | --- |
|  |  | (4-33) |
|  |  | (4-34) |

最后，将每组结果重塑为 的矩阵，得到两个子采样信号：

|  |  |  |
| --- | --- | --- |
|  |  | (4-35) |

这种子采样方法确保在每一组中，每个采样点均被利用，并通过随机交叉分配，使得两次子采样在局部时域上的细节信息具有互补性。此外，由于子采样对整个时间段均匀覆盖，背景噪声的整体能量得以保持一致，从而避免了随机丢弃可能带来的能量偏差问题。该方法采用矢量化操作实现，可高效地在GPU上运行，提升模型训练效率。

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4.1.1小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（即非随机的人为水下脉冲信号）。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。

4.3.2 损失函数设计

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，并在频域部分引入了针对不同频带的加权策略，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：

|  |  |  |
| --- | --- | --- |
|  |  | (4-36) |

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：

|  |  |  |
| --- | --- | --- |
|  |  | (4-37) |

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：

|  |  |  |
| --- | --- | --- |
|  |  | (4-38) |

其中， 为频率加权函数，用于在不同频率段施加不同的关注度。考虑到人为水下声信号的主要能量通常集中在中高频区域，而低频段往往包含较强的背景噪声，我们在加权策略上对高频区域施以更大权重，以确保模型更精确地重构这些关键频带的信号特征；同时对低频段施以较小权重，以降低背景噪声对损失的影响，避免模型过度拟合低频噪声成分。加权策略可定义如下：

|  |  |  |
| --- | --- | --- |
|  |  | (4-39) |

其中， 和 分别表示高频段和低频段的加权系数， 为频率阈值。通过该策略，模型在优化过程中会倾向于关注高频脉冲信号，对低频噪声的建模保持适度约束，从而更好提取脉冲信号特征。

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：

(4-40)

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

本章提出了一种面向复杂海洋背景噪声抑制的自监督深度学习降噪方法DMFNet，以有效提取水下脉冲信号（如打桩噪声、桥梁振动噪声）。首先，我们引入Nbr2Nbr子采样策略，在无纯净参考的条件下构造伪“干净”目标信号；针对子采样策略引入的频谱信息缺失问题，设计了DMFNet网络架构。DMFNet采用U型编码器-解码器结构，融合频谱动态聚焦单元（SDFU）、复数注意力模块（CAM）、语义重校准单元（SRU）和复数多尺度跳跃连接单元（CMS-SCU）等创新模块，有效捕获瞬态脉冲信号的频谱特征并抑制连续背景噪声。随后，详细介绍了自监督训练的数据构造方法与损失函数设计，确保模型在无监督条件下仍能稳定收敛并具备良好泛化性能。综上，本章提出的自监督降噪方案在无需真实干净信号的情况下，实现了对水下人为脉冲信号的有效提取和背景噪声的抑制。下一章将通过实验验证该方法的降噪性能，并与传统方法进行对比分析。

第五章 降噪实验与结果分析

5.1 实验设置

5.1.1 实验内容

为全面评估所提出自监督深度学习模型在水下人为脉冲声信号降噪任务中的有效性，特别是在复杂背景干扰与信号多样性条件下的鲁棒性与泛化能力，本研究设计了三个互为补充的实验环节。其中两个实验环节以水下打桩声信号为目标，海洋背景噪声为干扰源，用于验证模型在典型目标信号条件下的降噪性能；另一个实验环节则选取未参与训练的脉冲类声信号作为目标，用于评估模型在非训练类型信号上的适应能力与泛化表现。

实验环节一为累计能量段打桩声降噪实验，旨在构建训练基础并进行初步性能测试。实验选取距离实际打桩作业点约50米处采集的打桩声信号，作为“近似纯净”的目标信号。相关分析表明，该距离范围内信号信噪比较高，时频结构稳定，具备良好的工程代表性。考虑到水下声传播过程中的多径效应，信号的主要能量集中于初始时段，而尾部常由反射、绕射等路径引起，结构混乱、信噪比低，干扰较大。通过第三章的分析，针对水下工程监测中对信号能量时域分布的精确捕捉需求，本研究引入声暴露级（Sound Exposure Level, SEL）的能量积累特性，对信号进行积分分析，提取累计能量达到90%的有效波形段。该处理策略有助于去除尾部多径扰动，提高训练信号的代表性与稳定性。随后将该波形段与实际采集的海洋背景噪声进行叠加，构建训练集与基础测试集。

实验环节二为原始打桩声信号降噪实验。在前一实验的基础上，构建由原始未截断的打桩声信号与相同海洋背景噪声混叠而成的测试数据集，模拟系统直接接收未经预处理的原始含噪信号的工程场景。该实验用于验证模型在面对结构完整、但冗余干扰较多的原始信号时，是否仍具备良好的降噪能力。同时，该设置也用于考察模型是否能够在无需人工裁剪的前提下，自动识别并突出信号中的主能量波形段，并有效抑制尾部的多径干扰。该能力对于水下监测系统中前端降噪与目标提取模块的实用性验证具有重要意义。

实验环节三为非训练类型脉冲信号降噪实验，面向模型在非训练信号类型上的泛化能力评估。实验选取水下桥梁振动声信号作为目标，该类信号在时域结构和频谱特性上均与打桩声存在明显差异，未在模型训练阶段出现。将其与相同的背景噪声叠加构成测试数据集，可用于检验模型在面对“未见类型”信号时的适应能力与泛化表现，评估其在多样化水下声源处理任务中的应用潜力。为此，实验选取距离实际打桩作业点约50米处采集的打桩声信号作为“近似纯净信号”。由第三章的分析可知，该距离范围内的目标信号声压级显著高于背景噪声，信噪比较大，且能够有效规避机械干扰与声波传播畸变影响，具备稳定的时频结构和工程可行性，广泛被水声研究与监测实践所认可。

综上，三个实验环节从控制变量的训练验证，到原始复杂信号的工程测试，再到任务迁移的泛化评估，构建了一个完整的降噪性能测试体系，为后续系统部署和实际应用提供了系统性支撑与理论依据。

表5-1 不同实验环节的数据构建方式与属性说明

|  |  |  |
| --- | --- | --- |
| 实验环节名称 | 数据构建 | 数据属性 |
| 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 | 训练集 + 测试集 |
| 原始打桩声信号降噪实验 | 原始打桩声信号  + 海洋背景噪声 | 测试集 |
| 非训练类型脉冲信号降噪实验 | 桥梁振动声信号  + 海洋背景噪声 | 测试集 |

5.1.2 数据集构建与来源

为全面评估所提出的自监督水下降噪模型在不同类型噪声背景与信号结构条件下的适应能力与降噪效果，实验基于实际海上打桩作业与跨海桥梁营运期采集的水下声信号（如图5-1所示），结合多组自然海洋背景噪声数据，构建了三类实验数据集，分别对应于累计能量段打桩声降噪实验、原始打桩声信号降噪实验以及非训练类型脉冲信号降噪实验。

|  |  |
| --- | --- |
|  |  |
| f36909a7667eb41215e13e02c3dfe1a  (a) | 7feccc0d0e91b90d81b7dcd3abfbfec  (b) |
| (c) | 55f2e1911dd4ef7ebe19775141ceb28  (d) |
| 图5-1水下目标信号采集场景图：(a) 打桩船；(b) 打桩平台；(c) 跨海桥梁；  (d) 数据采集场景 | |

其中，打桩声信号均采集自福建厦门周边海域的海上打桩施工现场。水听器布设于距离打桩锤约 50 米处，以尽可能获得信噪比较高、结构清晰的打桩声信号，并有效避免机械干扰和传输畸变。水听器部署位置包括打桩平台与打桩船，相关采集时间及海况信息如表 5-1 所示，采集所用的相关设备已在第 3.1 节中详细介绍。

表5-1 打桩信号采集时间及环境参数

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 序号 | 采集日期 | 水深 | 潮汐状态 | 海况 |
| 1 | 2024年5月10日 | 23.5 m | 涨潮 | 1级 |
| 2 | 2024年5月31日 | 23.2 m | 涨潮 | 1级 |
| 3 | 2024年9月2日 | 13.7 m | 涨潮 | 1级 |
| 4 | 2024年10月22日 | 14.4 m | 涨潮 | 1级 |
| 5 | 2024年12月3日 | 4.8 m | 涨潮 | 1级 |

累计能量段打桩声降噪实验旨在从原始信号中提取包含主能量的有效波形段。具体做法如下：设打桩声信号为 ，每个采样点能量为 ，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) |

该处理过程如图5-2所示，最终保留主能量波形段，置零两端冗余部分。

|  |
| --- |
| 图 5-2：打桩信号有效能量段提取过程 上：原始信号；中：累计能量曲线；下：截取后的有效信号段 |

“原始打桩声信号降噪实验”保留完整打桩波形结构，同样与海洋背景噪声叠加，模拟工程系统中原始采集到的人为水下声信号输入情形，用于评估模型的原始信号适应能力与自动主段提取能力。

“非训练类型脉冲信号降噪实验”则使用实际采集的桥梁振动声信号作为目标信号，信号来源为某跨海桥梁基础结构，记录该结构在潮流、交通荷载等作用下产生的水下振动声。该类信号与打桩声在时频特征上存在明显差异，未在训练集中出现，用于检验模型在未见信号上的迁移适应能力。相关采集时间及环境参数见表 5-2。

表5-2 水下桥梁振动声信号采集时间及环境参数

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 序号 | 采集日期 | 水深 | 潮汐状态 | 海况 |
| 1 | 2022年5月19日 | 9.2 m | 涨潮 | 1级 |
| 2 | 2022年10月26日 | 11.9 m | 涨潮 | 1级 |

背景噪声数据采集于多个非施工时段海域，涵盖风浪、远距离船舶、低频生物声等多种自然噪声类型，具有代表性和环境多样性。部分采集场景如图 5-3 所示，具体采集信息见表 5-3。

|  |  |  |
| --- | --- | --- |
|  |  |  |
| 230320杏林2  (a) | 230418莆田  (b) | 230529漳州  (c) |
| 269607fd6070b8023e50c44bc00b452  (d) | 230926厦门2  (e) | 240319厦门2  (f) |
| 图片2  (g) | 240902厦门2  (h) | 241022厦门  (i) |
| 241203厦门  (j) | 250121福清  (k) | 250318厦门  (l) |
| 图 5-3 海洋背景噪声采集场景图（部分） | | |

表5-3 海洋背景噪声采集时间及环境参数

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 序号 | 采集日期 | 水深 | 潮汐状态 | 海况 | 场景图 |
| 1 | 2022年10月26日 | 4.8 m | 涨潮 | 1级 | — |
| 2 | 2023年3月20日 | 17.6 m | 涨潮 | 1级 | 图5-3(a) |
| 3 | 2023年4月18日 | 30.2 m | 涨潮 | 1级 | 图5-3(b) |
| 4 | 2023年5月29日 | 6.0 m | 涨潮 | 1级 | 图5-3(c) |
| 5 | 2023年9月14日 | 22.9 m | 涨潮 | 1级 | 图5-3(d) |
| 6 | 2023年9月26日 | 7.8 m | 涨潮 | 1级 | 图5-3(e) |
| 7 | 2024年3月19日 | 7.1 m | 涨潮 | 1级 | 图5-3(f) |
| 8 | 2024年5月10日 | 23.6 m | 涨潮 | 1级 | 图5-3(g) |
| 9 | 2024年5月31日 | 23.2 m | 涨潮 | 1级 | — |
| 10 | 2024年9月2日 | 13.7 m | 涨潮 | 1级 | 图5-3(h) |
| 11 | 2024年9月26日 | 5.8 m | 涨潮 | 1级 | — |
| 12 | 2024年10月22日 | 9.1 m | 涨潮 | 1级 | 图5-3(i) |
| 13 | 2024年12月3日 | 4.8 m | 涨潮 | 1级 | 图5-3(j) |
| 14 | 2024年12月24日 | 12.6 m | 涨潮 | 1级 | — |
| 15 | 2025年1月21日 | 16.3 m | 涨潮 | 1级 | 图5-3(k) |
| 16 | 2025年3月18日 | 7.2 m | 涨潮 | 1级 | 图5-3(l) |

综上，三类实验数据集在保持目标信号高质量的前提下，通过能量聚焦、结构保留与类型泛化的方式，系统模拟了多种水下应用场景，为后续模型训练与降噪性能评估提供了坚实的数据支撑。

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，我们将所有音频数据统一截取或填充为65280个采样点，并采用STFT将时域波形转换为频域表示，以便网络更有效地学习水下噪声信号的频谱特征。该转换过程利用PyTorch中的torch.stft函数实现，具体变换公式如下：

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

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DMFNet复数卷积网络的输入格式。

在模型训练过程中，我们采用Adam优化器进行参数更新，以确保梯度计算的稳定性并提升收敛速度。初始学习率设置为0.001，并采用指数衰减策略：每个epoch结束后，若验证集损失无显著下降，则学习率按衰减因子0.1进行调整。这一策略保证了训练的稳定性并防止过拟合。此外，为避免梯度爆炸，本研究在训练中引入了梯度裁剪（Gradient Clipping）策略，将梯度范数的最大阈值设为5。当梯度范数超过该阈值时，将其归一化，以保持训练过程稳定。

训练批次大小（Batch Size）设置为16，以平衡计算效率与模型泛化能力。在训练过程中，每个epoch对完整训练集进行一次遍历，并计算验证集上的损失以监控模型性能。在训练过程中，如连续5个epoch验证损失无显著下降，则触发早停机制（Early Stopping），以避免模型在训练数据上过拟合，并确保模型在泛化性能最佳时终止训练。

5.1.4 评价指标

为全面评估所提出的自监督深度学习模型在水下打桩声去噪任务中的表现，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：

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

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：

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

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：

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

值越小，表示模型的重建误差越低，性能越优。

综上，实验通过上述四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。

5.2 实验结果

5.2.1 累计能量段打桩声降噪实验

本实验基于第5.1.2节中构建的累计能量段打桩声数据集，目标信号为通过SEL分析所提取的90%主能量波形段，噪声源为实际采集的自然海洋背景噪声。该实验旨在评估所提出模型在结构清晰、能量集中的典型脉冲信号条件下的降噪性能，尤其关注其对主能量区域的还原能力与干扰成分的抑制能力。

为构建实验数据集，实验将目标信号与背景噪声分别按照信噪比0 dB、5 dB 和10 dB 使用PyDub工具[79]进行三组混叠。为保证混叠精度，背景噪声音频被自动裁剪或循环拼接，以覆盖目标信号的完整时长，从而模拟不同干扰强度下的水下信号接收场景。所有混叠样本随后输入模型进行预测，输出结果用于分析模型在时域结构保持、频谱特征一致性以及多指标量化性能上的表现。

图5-4展示了信噪比为0 dB 条件下，目标信号、混叠信号与模型降噪输出的波形对比结果。如图5-4（a）所示，原始目标信号在0.5 s 附近具有典型的脉冲形态，结构紧凑、能量集中。如图5-4（b）所示，混叠信号中的海洋背景噪声显著干扰了脉冲结构识别，尤其是前后段被大量非结构噪声覆盖，信号起止模糊，波形细节几乎被完全掩盖。如图5-4（c）所示，模型输出结果呈现出显著改善：脉冲信号结构清晰恢复，两端的高频噪声与随机扰动被有效抑制，整体信号在时间轴上重新集中在目标段，体现出强大的波形重构能力与非结构成分抑制能力。

|  |
| --- |
| (a) |
| (b) |
| (c) |
| 图 5-4：SNR = 0 dB 条件下累计能量段打桩声信号波形图: (a) 目标信号；  (b) 混叠海洋背景噪声信号；(c) 降噪信号 |

进一步的频谱图展示见图5-5。如图5-5（a）所示，目标信号具有明显的脉冲特征，能量集中度高、带宽窄。如图5-5（b）所示，混叠信号频谱则被中低频随机噪声全面污染，几乎无法辨识出有效脉冲事件，主频段已与噪声混叠在一起。如图5-5（c）所示，模型输出展现出高度还原能力：主频段能量再次集中，原始脉冲特征基本重建，同时有效压制了非目标频段的能量泄露，表现出良好的频谱聚焦能力与中低频噪声抑制效果。

|  |
| --- |
| (a) |
| (b) |
| (c) |
| 图 5-5：SNR = 0 dB 条件下的频谱图: (a) 目标信号；  (b) 混叠海洋背景噪声信号；(c) 降噪信号 |

从工程声信号监测角度分析，水下打桩声信号属于典型的短时脉冲类信号，在时域上呈现出明确的突发能量集中结构。在实际应用中，声波在海水中传播过程中易受到多径效应、边界反射及海底回波等复杂干扰的影响，使得信号后半部分常出现幅度减弱但结构紊乱的“尾波”成分。此类成分虽具高能冗余，但对事件判别与声源分析价值有限，反而可能干扰声暴露评估与信号识别模型的稳定性。

基于SEL理论，声事件的有效波形段通常定义为累计能量达到90% 的时间区间，核心在于真实反映事件主体能量释放的过程，剔除环境响应引起的非结构成分。因此，该实验在数据构建上即采用累计能量截取方法提取目标信号，模型输出则应尽可能与此结构保持一致。

从图5-4的时域结果来看，模型显著提高了信号起止边界的清晰度，恢复了主段结构，同时在两端噪声主导区域实现了尾波成分的有效压制，这一特性对基于事件检测或能量累计的应用尤为关键；而从图5-5的频谱表现来看，模型输出不仅保留了目标信号在中低频段的能量主峰，还抑制了混叠信号中由背景噪声引入的高频干扰，增强了信号的频域可分性与结构一致性。时域与频域双重维度的降噪效果共同验证了模型对水下工程声信号的强适配能力，为打桩事件的自动识别、暴露级计算与传播路径分析等下游任务提供了坚实基础。

此外，为更细致地评估降噪模型在不同时刻的表现，本文采用帧级信噪比（frame-level SNR）可视化方式，对目标信号与背景噪声在时序上的分离效果进行深入分析。将音频信号以30毫秒的帧长进行分帧，并设置75%的帧重叠率，可在时间轴上获得连续的帧级信号段。随后分别计算原始含噪信号与模型降噪输出在每一帧上的SNR值，并绘制成曲线图，从而实现对模型时域性能的可视化追踪。

图5-6所示为打桩声信号在0–1秒范围内的帧级SNR对比结果，其中蓝色曲线表示原始含噪信号的帧级SNR，绿色曲线表示模型降噪后的帧级SNR，红色阴影区域则标注了打桩信号帧段（即5.1.2节处理帧段），时间约0.5 s 附近。

从图像整体趋势可以观察到，在非打桩帧段中，原始含噪信号的帧级SNR长期维持在较低水平，约为 -60 dB，而模型输出信号的帧级SNR稳定提升至约 -40 dB，说明模型能够在背景区域有效降低噪声干扰。在打桩帧段开始时，模型输出信号的帧级SNR出现明显提升，表明模型不仅成功保留了冲击类打桩信号的能量特征，而且显著抑制了其混杂背景。随着打桩段结束，帧级SNR值快速回落至稳定水平，未观察到明显的残留伪影，显示模型对瞬态事件具有较好的响应能力和时间分辨性。

|  |
| --- |
| 图 5-6帧级SNR可视化对比图（蓝线：噪声信号；绿线：降噪信号；红色区域为打桩段） |

为系统评估模型性能，本文在三种信噪比（0 dB / 5 dB / 10 dB）混叠条件下，分别计算噪声信号与降噪输出在SNR、PSNR与RMSE三项指标下的表现。各项指标的定义与计算方法遵循语音增强与音频重建领域的通用规范。结果汇总见表5-4。

表5-4 模型降噪性能指标对比（累计能量段信号）

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 混叠信噪比(dB) | 信号类型 | SNR (dB) | PSNR (dB) | RMSE |
| 0 | 噪声信号 | -1.20 | 23.90 | 0.0064 |
|  | 降噪信号 | **10.25** | **36.42** | **0.0014** |
| 5 | 噪声信号 | 3.75 | 28.65 | 0.0037 |
|  | 降噪信号 | **13.48** | **39.65** | **0.0010** |
| 10 | 噪声信号 | 8.45 | 33.12 | 0.0020 |
|  | 降噪信号 | **16.78** | **42.65** | **0.0007** |

5.2.2 原始打桩声信号降噪实验

为进一步验证模型在非预处理条件下的降噪能力，本节采用未经累计能量截断处理的原始打桩声信号作为目标信号，与实际采集的海洋背景噪声进行混叠，构建新的测试数据集，并开展降噪实验。该实验旨在评估模型在面对更复杂、未规整的原始信号输入时，是否仍具备对关键能量区域的自动提取能力，并能有效抑制由于水声多径传播导致的尾部混响干扰，为水下声信号采集监测等工程应用提供更具通用性的降噪策略支持。

相比累计能量段信号，原始打桩声信号包含完整的打桩事件响应过程，前期为能量突发阶段，中后段则常出现因海底或水面反射引发的尾波与混响。该类尾部信号虽具一定能量，但其结构往往紊乱、分布离散，且难以用于事件强度评估或源定位分析，属于典型的多径冗余信号。因此，降噪模型若能自动识别主能量段并弱化尾波结构，将在实用性与鲁棒性上表现更佳。

图5-7展示了信噪比为0 dB 条件下，原始打桩声信号的混叠波形与模型降噪结果。从图中可见，原始信号前段呈现出强烈脉冲结构，而尾段则被背景噪声与多径混响共同覆盖，导致时域轮廓模糊、边界不清。模型输出信号则实现了前段结构的清晰恢复，并在中后段显著抑制了多余尾波成分，有效聚焦于主事件区域，体现出对打桩事件的自动识别与结构提取能力。

|  |
| --- |
| (a) |
| (b) |
| (c) |
| 图 5-7：SNR = 0 dB 条件下原始打桩声信号的波形图: (a) 目标信号；  (b) 混叠海洋背景噪声信号；(c) 降噪信号 |

图5-8展示了该信号在频谱域的可视化结果。从图5-8（a）可见，原始打桩信号在约0.5s处的主频段能量集中，但中后段受多径传播效应影响，在1–3 kHz 区间频谱表现出明显的模糊扩展特征，尾部混响能量扩散明显，反映出典型的声波在水中遇边界反射或海底折返造成的混响效应，但结构离散，对频谱贡献有限。如图5-8（c）所示，模型输出信号频谱结构更为集中，不仅保留了主要冲击事件频率成分，还实现了对尾部非结构能量的有效压制。特别是在中高频区域，模型显著削弱了多径混响造成的尾段杂散频率分布，使频谱恢复紧凑，轮廓清晰，进一步提升了后续能量评估、事件识别与频率特征提取的可靠性。

|  |
| --- |
| (a) |
| (b) |
| (c) |
| 图 5-8：SNR = 0 dB 条件下条件下原始打桩声信号的频谱图: (a) 目标信号；(b) 混叠海洋背景噪声信号；(c) 降噪信号 |

进一步地，图5-9展示了原始信号条件下的帧级信噪比可视化结果。图中，蓝色曲线表示原始含噪信号帧级SNR，绿色曲线为降噪后信号，红色阴影区域则标注了打桩信号帧段（即5.1.2节处理帧段），时间约0.5 s 附近。可观察到：在背景噪声段，模型能稳定将帧级SNR提升约20 dB；在打桩主频段附近，模型帧级SNR峰值提升至 +30 dB 左右，而原始信号峰值仍不足 15 dB。在尾段，降噪信号帧级SNR快速回落至稳定水平，尾部混响未形成伪影拖尾。以上表明模型在保留信号主结构特征的同时，显著降低了背景噪声与多径混响效应的干扰。



图 5-9帧级SNR可视化对比图（蓝线：噪声信号；绿线：降噪信号；红色区域为打桩段）

表5-5汇总了在三种信噪比下，原始信号条件下模型降噪前后的定量指标（SNR、PSNR、SDR、RMSE）对比情况。对比第5.2.1节中表5-4可见，累计能量段信号在0dB 混叠条件下降噪后的SNR达10.25 dB，略高于原始信号条件下的9.12 dB，表明模型在结构规整输入下能获得更佳性能。PSNR呈现相同趋势，主要归因于原始信号包含完整尾部波形，存在更多由水声多径传播引起的混响干扰，结构更复杂，降噪难度提升。但在信噪比为5 dB 和10 dB 情况下，模型在两类信号上的性能差距显著缩小，表现出较强的一致性和适应性，说明所提出模型具备良好的泛化能力，能够在结构理想与实际复杂场景中均实现有效降噪。

表5-5 模型降噪性能指标对比（累计能量段信号）

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 混叠信噪比(dB) | 信号类型 | SNR (dB) | PSNR (dB) | RMSE |
| 0 | 噪声信号 | -1.20 | 24.10 | 0.0060 |
|  | 降噪信号 | **9.12** | **34.28** | **0.0019** |
| 5 | 噪声信号 | 3.16 | 27.56 | 0.0041 |
|  | 降噪信号 | **12.84** | **38.17** | **0.0012** |
| 10 | 噪声信号 | 8.03 | 31.95 | 0.0025 |
|  | 降噪信号 | **16.32** | **41.68** | **0.0009** |

5.2.3 非训练脉冲信号降噪实验

为进一步检验所提出模型在非训练场景中的泛化能力与类型迁移适应性，本节选取实际工程中另一类典型的人为冲击类声源——水下桥梁振动声信号，作为测试样本，输入模型进行推理并分析降噪效果。该类信号具有突发性、瞬时性与多频成分特征，但其时域形态、频谱结构及背景环境均未参与模型训练，故可用于评估模型在“未知类型脉冲事件”场景下的适应表现。

由于该实验样本为实际采集的噪声混合信号，缺乏对应的纯净标签信号，因此无法采用SNR、PSNR、RMSE等传统定量指标进行性能评估。为了确保分析的科学性与工程有效性，实验采用可视化分析法对模型的结构保持能力与干扰抑制效果进行综合评估，结合桥梁结构声传播特点与脉冲事件识别需求，重点考察模型是否能突出恢复主能量段，并有效剔除背景扰动与结构尾波。

图5-10展示了某一桥梁振动信号片段的时域波形降噪效果对比。从图中可见，原始信号中存在强背景底噪干扰，脉冲主段被大量非结构能量覆盖，信号边界模糊。降噪结果中，模型显著提升了信号的对比度，脉冲主段结构更加清晰，两侧底噪得到明显削弱，具备良好的冲击事件聚焦能力。

|  |
| --- |
| (a) |
| (b) |
| 图 5-10：水下桥梁振动声信号降噪前后波形对比图: (a) 原始采集含噪信号；  (b) 降噪信号 |

图5-11为对应的频谱分析结果，显示降噪前信号频谱能量在全频段扩散，低频部分尤其受到环境噪声和混响的显著污染；而模型输出的频谱结构表现为能量向中低频主段回归、杂散频率成分明显压缩，整体频谱轮廓更为集中，反映出模型对结构主频段的识别与保持能力。

|  |
| --- |
| (a) |
| (b) |
| 图 5-11：水下桥梁振动声信号降噪频谱对比图: (a) 原始采集含噪信号；  (b) 降噪信号 |

从工程声学监测角度出发，桥梁振动信号属于结构瞬态冲击声，常用于评估振动响应强度、结构健康状态与激励传播路径。有效提取此类信号中的主能量段对于事件识别、冲击定位及异常检测具有关键意义。模型在未训练该类型信号的条件下，仍能准确识别其脉冲核心段并压制背景干扰，说明所提出架构具备良好的结构通用性与时频自适应能力。尤其在模型未显式学习该类事件的语义特征、传播路径或信号形式的情况下，仍可实现工程有效的信号增强，进一步验证了模型的类型迁移能力与弱监督鲁棒性。

5.3 模型有效性与降噪方法对比分析

本章将对提出的自监督水下降噪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水下工程信号（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 海洋背景噪声降噪效果对比实验

为验证本文提出的自监督深度学习模型在实际复杂海洋环境中的降噪性能，本节选取典型的实际海洋背景噪声信号，分别应用传统降噪方法（维纳滤波、小波变换）与本文所提出模型进行降噪实验，通过波形和频谱的详细对比分析，深入评估各方法的降噪能力与实际工程适用性。

图5-12 展示了不同方法降噪后的时域波形对比。原始的海洋背景噪声波形（图5-12(a)）呈现明显的随机波动与尖峰扰动，整体表现为典型的非平稳随机噪声特性。维纳滤波方法（图5-12(b)）对背景噪声的随机波动仅有轻微的抑制效果，随机扰动成分仍明显存在，未能有效改善信号的整体稳定性。小波变换方法（图5-12(c)）表现出与维纳滤波类似的问题，随机波动未能显著减轻，信号整体仍然表现为较为杂乱的状态，降噪效果有限，无法有效提高后续水声信号识别分析的可靠性。相比之下，本文提出的自监督模型（图5-12(d)）的降噪效果极为显著，波形变得非常平稳，随机噪声成分被大幅压制，波动幅值明显减小，展现出极佳的噪声抑制能力和工程适用价值。

|  |  |
| --- | --- |
| (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 | |

进一步通过频谱分析（图5-13）验证各方法的降噪性能差异。从原始噪声频谱（图5-13(a)）可以明显观察到能量广泛分布于整个频带内，随机扰动强烈且频谱结构复杂。维纳滤波降噪后的频谱（图5-13(b)）尽管对部分高频噪声稍有抑制，但频谱整体杂乱程度并未显著改善，随机扰动特征依然突出，显示了维纳滤波在非平稳随机背景中的局限性。小波变换降噪后的频谱（图5-13(c)）也表现出类似的特征，随机噪声仍较明显，频谱整体未出现明显的结构改善或扰动压制迹象，难以有效改善信号品质。而本文所提出模型的降噪频谱结果（图5-13(d)）表现出极为显著的改善，频谱能量高度集中在低频区域，随机扰动和高频噪声成分几乎被完全抑制，频谱清晰、稳定且干净，显著提升了频谱结构的辨识度与可靠性，展现出该模型在非平稳随机噪声降噪方面的卓越性能。

|  |  |
| --- | --- |
| (a) | (b) |
| (c) | (d) |
| 图5-13海洋背景噪声降噪频谱图对比：(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 | |

综合以上详细的时域与频谱域分析可得，传统方法（维纳滤波、小波变换）在面对非平稳随机海洋背景噪声时，存在明显的适用性问题，表现为随机噪声压制效果弱、频谱结构无法有效改善，降噪性能受限。与之对比，本文提出的自监督深度学习模型则充分体现了其结构稳定性与频谱高效聚焦优势，不仅有效抑制了随机扰动，还极大提升了频谱结构清晰度与可靠性，具备明显的工程应用价值与优势。

5.3.2 海洋非训练脉冲信号降噪效果对比实验

为进一步验证本文提出的自监督深度学习模型在非训练信号上的结构适应能力与工程实用性，本节选取实际采集的水下桥梁振动声信号作为输入样本，分别采用维纳滤波、小波变换和本文模型进行降噪处理。该类信号在实际采集中常伴随海洋背景噪声与结构传播干扰，信号特征表现为非周期突发冲击与背景拖尾并存，结构不稳定、干扰复杂，具备显著的非平稳性与脉冲性特征。

图5-14展示了各方法处理下的时域波形图。原始信号（图5-14(a)）包含多个幅度不一的冲击波动，并夹杂有持续性的背景噪声干扰，信号结构边界模糊。维纳滤波（图5-14(b)）压制了部分高频噪声成分，但对于突发冲击结构仍存在模糊趋势，主能量段辨识度不高；小波变换结果（图5-14(c)）虽然保留了部分结构边缘，但对背景扰动抑制有限，且在多个位置引入高幅值伪信号，削弱了原始结构清晰度。相比之下，本文模型处理结果（图5-14(d)）有效突出了多个冲击事件的结构波动，信号主体成分突出，起止界限明确，同时对背景波动实现了更强的压制，波形清晰紧凑。

|  |  |
| --- | --- |
| (a) | (b) |
| (c) | (d) |
| 图5-14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 | |

频谱分析结果如图5-15所示，进一步验证了各方法的频域降噪能力。原始信号（图5-15(a)）呈现多个短时冲击频率带叠加现象，但同时存在大量高频杂散分量，频谱能量分布松散。维纳滤波（图5-15(b)）削弱了部分高频能量，但未能聚焦真实频率结构，频谱依旧混杂；小波变换（图5-15(c)）频谱结构略有收敛，但信号主体频带仍被背景所淹没，难以提供有效事件支撑。本文模型频谱结果（图5-15(d)）则呈现出良好的频谱聚焦性，冲击成分在频域中被完整还原，背景频带能量下降显著，主频段结构突出，特别在低频段表现出更强的能量集中度，说明模型在频谱层面具有更优异的分离与还原能力。

|  |  |
| --- | --- |
| (a) | (b) |
| (c) | (d) |
| 图5-15海洋桥梁振动声信号降噪频谱图对比：(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 | |

从图5-14 和图5-15的联合观察中可进一步发现，在时域约0.85 s 处存在明显的非平稳背景噪声扰动，频谱图中该处信号不具备明确主频结构，属于典型的海洋背景噪声波动。在原始波形图中，该波动以突增扰动的形式出现，对信号整体判断构成干扰。传统方法如维纳滤波与小波变换均无法有效区分该背景扰动与真实冲击信号，导致残余能量保留甚至误识为主信号成分。而本文模型在该区域展现出较强的抑制能力，成功压制该无效扰动，使主能量段的结构边界更加清晰，频谱中也未出现对应杂散频带，进一步体现出模型在时频域上的协同降噪优势。

综上所述，本文模型在面对非训练脉冲信号时，仍能准确保留其主结构波动并有效压制背景扰动，表现出良好的时域结构保持能力与频域聚焦能力。传统方法在该类非平稳信号处理上存在明显的残留与结构模糊问题，难以满足工程信号处理中的实际需求。本文方法不仅具备高适应性与强鲁棒性，同时展现出在复杂水下背景中的广泛适用潜力。

5.3.3 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对信噪质量的提升能力。

在模块消融部分，保持模型整体结构与训练配置一致，依次去除SDFU、CAM、CMS-SCU，构建三种对照变体，并与完整模型在相同数据集上进行性能评估。表5-5汇总了各模型在SNR、PSNR和RMSE三项指标下的表现情况。结果显示，完整模型在三项指标上均表现最优，SNR达10.25 dB，PSNR为 36.42 dB，RMSE降至 0.0014，相较于原始噪声信号（SNR = -1.20 dB，PSNR = 23.90 dB，RMSE = 0.0064）实现了大幅提升。去除残差连接后模型的结构对齐能力减弱，SNR 降为8.31 dB，通道注意力模块缺失则影响主能量聚焦能力，频域模块剥离则显著提升RMSE至0.0025。实验验证了多模块协同在复杂结构信号恢复中的关键作用。

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）

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 模型编号 | 模型结构 | SNR (dB) | PSNR (dB) | RMSE |
| 原始信号 | 噪声信号 | -1.20 | 23.90 | 0.0064 |
| 模型A | **完整模型** | **10.25** | **36.42** | **0.0014** |
| 模型B | 去除SDFU | 8.31 | 33.66 | 0.0021 |
| 模型C | 去除CAM | 7.54 | 32.83 | 0.0023 |
| 模型D | 去除CMS-SCU | 6.74 | 31.85 | 0.0027 |

为进一步论证模型的工程实用性，如表5-6所示，本节引入近年自监督领域音频降噪模型、维纳滤波与小波变换，构建完整性能对比体系，结果如表5-6所示。维纳滤波在平稳背景中具有一定抑噪能力，但面对非结构冲击成分时出现明显频谱模糊现象，SNR仅为3.20 dB；小波变换虽在某些脉冲信号上保留边缘，但抑制能力不稳定，RMSE为0.0061。U-Net具备良好的整体建模能力，但对非稳态尾波扰动及信号边界定位仍显不足。相比之下，本文模型在三项指标上均实现性能最优，结构还原与能量重构能力显著优于对照方法，充分说明模型结构设计的针对性与有效性。

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）

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 方法类型 | 模型方法 | SNR (dB) | PSNR (dB) | RMSE |
| — | 原始含噪信号 | -1.20 | 23.90 | 0.0064 |
| 传统方法 | 维纳滤波 | 1.17 | 26.35 | 0.0048 |
| 小波变换 | 0.02 | 25.2 | 0.0055 |
| 自监督策略模型 | U-Net | 5.47 | 30.65 | 0.0030 |
| NerNT | 7.18 | 32.42 | 0.0025 |
| N2N | 7.82 | 33.07 | 0.0022 |
| ONT | 8.63 | 33.81 | 0.0020 |
| **本文模型** | **10.25** | **36.42** | **0.0014** |

5.4 本章小结

本章围绕所提出的自监督水下结构脉冲声降噪模型，系统设计并实施了一系列具有工程针对性与科学严谨性的实验，全面评估模型在多种复杂场景下的适应能力与降噪效果。具体工作涵盖数据集构建、信噪控制实验、模型对比与消融验证三个方面，具体内容如下：

首先，在数据集构建方面，基于实测的水下打桩声与海洋背景噪声信号，设计了三类数据集与对应的降噪实验任务：累计能量段打桩声降噪实验、原始打桩声信号降噪实验以及非训练类型脉冲信号降噪实验。其中，累计能量段实验聚焦于SEL理论下的主结构能量恢复，通过提取信号中累计达到90%能量的有效波形段，体现模型在结构性特征提取方面的适应性；原始信号实验则考察模型在未进行预处理情况下对完整信号的结构自适应能力及其抗尾波干扰性能；非训练类型实验引入水下桥梁振动声作为泛化验证样本，用于分析模型在跨类型冲击声处理中的拓展能力。

其次，在降噪性能评估方面，针对不同实验任务，从波形对齐、频谱保持及帧级SNR可视化等维度开展分析。结果显示，在累计能量段实验中，模型能够准确增强主能量段、有效抑制尾波扰动，输出波形具备良好可视解释性，频谱主带重建清晰，帧级SNR在打桩声段显著提升，体现出较强的瞬态识别能力。在原始信号实验中，模型表现出良好的结构起始定位能力、频谱抖动修复效果及尾部多径干扰抑制能力，展现出高度稳健性。在非训练样本实验中，尽管声源结构与训练数据不同，模型仍能有效识别脉冲峰值、压制背景非稳态噪声，验证了其良好的泛化能力与工程适应性。

此外，为进一步验证模型结构设计的有效性与先进性，本章基于累计能量段数据集开展了模块消融实验与多方法对比实验。消融实验结果表明，完整模型在SNR、PSNR和RMSE等指标上均优于去除任一模块的简化结构，凸显多模块协同设计的重要性。对比实验结果也显示，本文所提出模型在整体性能上显著优于传统方法（如维纳滤波、小波变换）以及现有自监督深度模型，体现出其在结构保持与噪声抑制方面的综合优势。

综上所述，本章通过多维度、系统性的实验验证，充分论证了所提模型在复杂海洋环境下对结构性脉冲声信号的识别与降噪能力，尤其在低信噪比条件下表现出更高的稳定性与工程应用价值。为后续在实际水下工程场景中的部署与应用奠定了坚实的技术基础。

第六章 总结与展望

6.1 论文的主要工作与创新点

本论文围绕水下人为工程脉冲声信号的降噪问题，针对传统方法在复杂背景干扰条件下结构保持能力不足、现有深度模型对监督样本依赖强以及模型工程适应性差等问题，提出了一种融合结构感知机制的自监督水下降噪方法，并围绕数据构建、模型设计与性能验证展开系统研究，构建了完整的工程任务闭环方案。本文的主要研究工作与创新点概括如下：

（1）从实际采集的打桩与背景噪声数据出发，结合声压级、声暴露级等典型声学参数，系统分析了水下脉冲类信号的能量分布规律与非结构噪声的干扰特性。在此基础上，设计基于累计能量段提取的结构性参考信号构建方案，采用实际采集的海洋背景噪声与工程脉冲声信号进行混叠，取代传统高斯噪声模拟方式，构建贴近真实场景的自监督训练集，解决了无纯净标签条件下模型训练目标难以定义的问题。

（2）针对脉冲类信号的主能量段识别与尾波干扰压制需求，构建了一种融合频域建模与通道注意力机制的深度神经网络结构。模型引入频谱动态聚焦模块、复数域注意力机制与多尺度结构融合策略，实现对非平稳噪声背景下信号主结构的多维感知与显著增强。各子模块具备明确的功能分工与协同作用，显著提升了模型在复杂背景中的结构保持能力与误差控制能力。

（3）论文设计了多层级、多场景的实验验证体系，涵盖基于累计能量段的模型训练、原始未截断信号的降噪效果验证，以及非训练类型脉冲信号（如桥梁振动声）的泛化能力测试，构建了从结构恢复到工程应用的完整性能评估框架。同时通过与传统、神经网络等方法进行对比，并辅以模块消融实验，全面验证了模型的结构设计合理性与性能优势。

综上，本文在模型结构设计、训练目标构建与实验验证框架三方面形成系统创新，具体体现在以下三点：

（1）提出基于声暴露级理论的结构性参考信号构建机制，采用累计能量段提取目标信号有效结构，解决了无纯净标签条件下自监督训练目标缺失的问题。

（2）设计融合频谱建模与注意力机制的结构保持型降噪网络，面向工程脉冲信号特性，实现对主能量段的精确恢复与尾波干扰的有效压制。

（3）构建覆盖训练内与训练外样本的多场景验证体系，实现模型在结构恢复精度、泛化能力与工程适应性等方面的系统性评估。

这些创新为水下工程声信号的智能降噪处理提供了切实可行的技术路径，为复杂水声环境中的结构事件监测、能量评估与声源感知奠定了方法基础与工程支撑。

6.2 未来工作展望

本文提出了一种面向水下人为工程脉冲声信号的自监督结构保持型降噪方法，围绕信号结构感知、背景噪声抑制及无标签训练机制进行了系统设计与验证，在多个任务场景下取得了良好的实验效果。然而，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几个方面展开：

（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、设备轰鸣、环境扰动等多类背景噪声，这些声源之间存在频谱重叠、时序交错与非结构混合等现象，传统基于单一目标的降噪方法在此类环境中易产生误削与信号混淆。未来研究可结合语义建模与特征判别机制，增强模型在复杂声源场景中的目标聚焦能力与干扰识别能力。

（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。未来可探索模型压缩、子网络裁剪与推理优化策略，使模型具备实时运行能力，并可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检测、噪声屏蔽与结构感知闭环处理。

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，进一步提升降噪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. 张景全.我国海洋强国建设面临的机遇与挑战[J].人民论坛,2023,(20):26-29.
3. 谢宝剑,李庆雯.新质生产力驱动海洋经济高质量发展的逻辑与路径[J].东南学术,2024,(03):107-118+247.
4. 王宏.以建设海洋强国新作为推进中国式现代化[N].学习时报,2023-09-22(001).
5. Akyildiz I F, Pompili D, Melodia T. Underwater acoustic sensor networks: research challenges[J]. Ad hoc networks, 2005, 3(3): 257-279.
6. Lurton X. An introduction to underwater acoustics: principles and applications[M]. Springer Science & Business Media, 2002.
7. Hildebrand J A. Anthropogenic and natural sources of ambient noise in the ocean[J]. Marine Ecology Progress Series, 2009, 395: 5-20.
8. Urick R J. Sound propagation in the sea[J]. 1997.
9. Erbe C, Marley S A, Schoeman R P, et al. The effects of ship noise on marine mammals—a review[J]. Frontiers in Marine Science, 2019, 6: 606.
10. Popper A N, Hastings M C. The effects of anthropogenic sources of sound on fishes[J]. Journal of fish biology, 2009, 75(3): 455-489.
11. Popper A N, Hawkins A D. An overview of fish bioacoustics and the impacts of anthropogenic sounds on fishes[J]. Journal of fish biology, 2019, 94(5): 692-713.
12. Stojanovic M, Preisig J. Underwater acoustic communication channels: Propagation models and statistical characterization[J]. IEEE communications magazine, 2009, 47(1): 84-89.
13. Gao R, Liang M, Dong H, et al. Underwater Acoustic Signal Denoising Algorithms: A Survey of the State-of-the-art[J]. arXiv preprint arXiv:2407.13264, 2024.
14. Wenz G M. Acoustic ambient noise in the ocean: spectra and sources[J]. The journal of the acoustical society of America, 1962, 34(12): 1936-1956.
15. Hildebrand J A. Anthropogenic and natural sources of ambient noise in the ocean[J]. Marine Ecology Progress Series, 2009, 395: 5-20.
16. 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.
17. 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.
18. 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.
19. 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.
20. Wojciechowski S. Hydroacoustic Parametric Study of Pile Driving-Induced Anthropogenic Sound[D]. Virginia Tech, 2024.
21. Tsouvalas A. Underwater noise emission due to offshore pile installation: A review[J]. Energies, 2020, 13(12): 3037.
22. 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.
23. Lippert T, von Estorff O. On a hybrid model for the prediction of pile driving noise from offshore wind farms[J]. Acta Acustica United with Acustica, 2014, 100(2): 244-253.
24. Jensen F B, Kuperman W A, Porter M B, et al. Computational ocean acoustics[M]. New York, NY: Springer New York, 2011.
25. Hildebrand J A. Anthropogenic and natural sources of ambient noise in the ocean[J]. Marine Ecology Progress Series, 2009, 395: 5-20.
26. Veeriayan V B, Rajendran V. Underwater Ambient Noise[M]//Noise and Environment. IntechOpen, 2020: 153.
27. Cato D H. Ambient noise and its significance to aquatic life[J]. Bioacoustics, 2008, 17(1-3): 21-23.
28. 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.
29. Yori Fernandez A. Underwater assessment of anthropogenic noise sources using a field recording method[J]. Acta Acustica united with Acustica, 2018, 104(1): 13-24.
30. 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.
31. Nie L, Zhang Y, Wang H B. Classification of underwater soundscapes using raw hydroacoustic signals[J]. The Journal of the Acoustical Society of America, 2023, 154(4\_supplement): A304-A304.
32. Mishachandar B, Vairamuthu S. Diverse ocean noise classification using deep learning[J]. Applied Acoustics, 2021, 181: 108141.
33. Wittekind D, Schuster M. Propeller cavitation noise and background noise in the sea[J]. Ocean Engineering, 2016, 120: 116-121.
34. 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.
35. 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.
36. 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.
37. Iqbal J, Khan R, Ahmed F, et al. A Comparative Analysis of the Latest Statistical Characterization Trends of Narrowband & Broadband Frequency Spectrum of Deep Sea Ambient Noise Using Monte Carlo Simulation Method[C]//2019 16th International Bhurban Conference on Applied Sciences and Technology (IBCAST). IEEE, 2019: 900-907.
38. Zhang L, Meng C, Na J. Modeling of high background noise in large area ocean based on measured data[C]//2018 IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC). IEEE, 2018: 1-4.
39. 赵润晗.海上打桩噪声对中华白海豚whistles信号的掩蔽效应评估[D].厦门大学,2022.DOI:10.27424/d.cnki.gxmdu.2022.003062.
40. Chuang Z, Zhong-Chang S, Yu Z. Sound reception pathway of the Indo-Pacific humpback dolphin[J]. ACTA PHYSICA SINICA, 2020, 69(23).
41. Reeder D B, Joseph J E, Rago T A. Underwater sound generated by motor vehicle traffic in an underwater tunnel[J]. The Journal of the Acoustical Society of America, 2020, 148(3): EL215-EL220.
42. MacGillivray A, Warner G, Racca R, et al. Tappan Zee Bridge construction hydroacoustic noise modeling[J]. prepared by JASCO Applied Sciences for ABCOM, New York, final report P001116-001, version 1.0, 2011.
43. Haxel J, Zang X, Martinez J, et al. Underwater noise measurements around a tidal turbine in a busy port setting[J]. Journal of Marine Science and Engineering, 2022, 10(5): 632.
44. 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.
45. Nagarkar A V. Duct-borne Noise & Vibration onboard Maritime Vessels for Underwater Radiated Noise Management[D]. Savitribai Phule Pune University, 2022.
46. Bilgin Güney C, Kinaci O K. Ecological effects and mitigation strategies for underwater noise from shipping activities[J]. Ships and Offshore Structures, 2024: 1-13.
47. McIntyre D. Predicting cavitation-induced noise from marine propellers[D]. , 2021.
48. Miglianti L P. Modelling of the cavitating propeller noise by means of semi-empirical and data driven approaches[J]. 2020.
49. Ianniello S, Muscari R, Di Mascio A. Ship underwater noise assessment by the acoustic analogy. Part I: nonlinear analysis of a marine propeller in a uniform flow[J]. Journal of marine Science and technology, 2013, 18: 547-570.
50. Zhu C, Gaggero T, Makris N C, et al. Underwater sound characteristics of a ship with controllable pitch propeller[J]. Journal of Marine Science and Engineering, 2022, 10(3): 328.
51. Yang C, Kuang F, Wang T, et al. Friction-Induced Vibration and Noise in Water-Lubricated Stern Bearings: A Comprehensive Review of Mechanisms and Design[C]//International Conference on Offshore Mechanics and Arctic Engineering. American Society of Mechanical Engineers, 2024, 87837: V05BT06A053.
52. Traverso F, Gaggero T, Tani G, et al. Parametric analysis of ship noise spectra[J]. IEEE Journal of Oceanic Engineering, 2016, 42(2): 424-438.
53. McPherson C, Quijano J. FPSO Facility Anchor Piling Acoustic Modelling[J]. 2017.
54. Huo X, Zhang P, Feng Z. Study of underwater sound propagation and attenuation characteristics at the Yangjiang offshore wind farma[J]. Ecological Informatics, 2024, 84: 102919.
55. 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.
56. Wang Z, Wu Y, Duan G, et al. Assessing the underwater acoustics of the world's largest vibration hammer (OCTA-KONG) and its potential effects on the Indo-Pacific humpbacked dolphin (Sousa chinensis)[J]. PloS one, 2014, 9(10): e110590.
57. Leunissen E. Underwater noise from pile-driving and its impact on Hector's dolphins in Lyttelton Harbour, New Zealand[D]. University of Otago, 2018.
58. Zhou W, Xu X, Tu X, et al. Preliminary exploration for effects of sound stimulus on the movement behavior of Litopenaeus vannamei[C]//2016 IEEE/OES China Ocean Acoustics (COA). IEEE, 2016: 1-6.
59. Polyzos E, Gortsas T, Tsinopoulos S, et al. A Numerical Study on the Underwater Noise Mitigation in Wind Turbine Pilling[M]//Recent Developments of Soil Mechanics and Geotechnics in Theory and Practice. Cham: Springer Nature Switzerland, 2024: 121-140.
60. Amaral J L. Characterization of Impact Pile Driving Signals and Fin Whale Vocalizations at the Block Island Wind Farm Site[M]. University of Rhode Island, 2021.
61. Bohatkiewicz J, Jukowski M, Hałucha M, et al. Influence of the acoustic cover of the modular expansion joint on the acoustic climate in the bridge structure surroundings[J]. Materials, 2020, 13(12): 2842.
62. Kawade A N, Shinde V M, Shastri R K, et al. Analysis of ship noise from underwater ambient noise[C]//2016 Conference on Advances in Signal Processing (CASP). IEEE, 2016: 265-269.
63. 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.
64. 郭新毅,李凡,铁广朋,等.海洋环境噪声研究发展概述及应用前景[J].物理,2014,43(11):723-731.
65. Chen T, Kornblith S, Norouzi M, et al. A simple framework for contrastive learning of visual representations[C]//International conference on machine learning. PmLR, 2020: 1597-1607.
66. Lehtinen J, Munkberg J, Hasselgren J, et al. Noise2Noise: Learning image restoration without clean data[J]. arXiv preprint arXiv:1803.04189, 2018.
67. 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.
68. Alamdari N, Azarang A, Kehtarnavaz N. Improving deep speech denoising by noisy2noisy signal mapping[J]. Applied Acoustics, 2021, 172: 107631.
69. 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.
70. Wu J, Li Q, Yang G, et al. Self-supervised speech denoising using only noisy audio signals[J]. Speech Communication, 2023, 149: 63-73.
71. Trabelsi C, Bilaniuk O, Zhang Y, et al. Deep complex networks[J]. arXiv preprint arXiv:1705.09792, 2017.
72. Dai J, Qi H, Xiong Y, et al. Deformable convolutional networks[C]//Proceedings of the IEEE international conference on computer vision. 2017: 764-773.
73. 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.
74. 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.
75. 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.
76. 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.
77. 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.
78. Robert J, Webbie M. Pydub (2018)[J]. URL http://pydub. com.

攻读硕士学位期间获得的成果

1. **Xiuling Zhu**, Yougan Chen, Lei Wan, Weidi Huang, Ge Lu, Xiaomei Xu, Dynamic Layered Routing Protocols Based on BP-NN for Underwater Acoustic Sensor Networks, *Applied Acoustics*, vol. 211, Art.no. 109454, Aug. 2023. (SCI/EI收录)
2. Shenao Tu, **Xiuling Zhu**, Yougan Chen, Xiaomei Xu, A Q-Learning and Data Priority-Based Routing Protocol with Dynamic Computing Cluster Head for Underwater Acoustic Sensor Networks, *2022 IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC), Xi'an, China*, 2022, pp. 1-5, 2022. （EI收录）
3. **Xiuling Zhu**, Ge Lu, Yuan Luo, Yanhan Dong, Yougan Chen, Tao Yi, A Secure Routing Protocol Based on Graph Neural Networks for Underwater Acoustic Sensor Networks, *2023 IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC), Zhengzhou, China*, 2023, pp. 1-5, 2023. （EI收录）
4. Yihao Zhao, Ge Lu, Shilu Tan, **Xiuling Zhu**, Zhixian Song, Yougan Chen, A Clustering Guiding-Network Based Routing Protocol for Underwater Acoustic Sensor Networks, *2023 IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC), Zhengzhou, China*, 2023, pp. 1-5, 2023. （EI收录）
5. 陈友淦, **朱秀玲**, 罗圆, 杜坤芸, 高榜君, 吴剑明, 许肖梅. 一种AUV辅助的水声传感网动态分层路由方法, 国家发明专利, 申请号202211240126.7, 申请日期2022.10.11, 公告号CN 115589625 A, 公开日期2023.01.10.
6. 陈友淦, **朱秀玲**, 卢鸽, 江涵希, 涂申奥, 陶毅, 许肖梅. 一种BP神经网络的水声传感网动态分层路由方法, 国家发明专利, 申请号202211240085.1, 申请日期2022.10.11, 公告号CN 115604740 A, 公开日期2023.01.13.
7. 陈友淦, 罗圆, **朱秀玲**, 董妍函, 陈哲扬, 谭诗路, 许肖梅. 水下动态环境中考虑时延和 AUV 能耗的混合数据搜集方法, 国家发明专利, 申请号202310095023.4, 申请日期2023.02.10, 公告号CN 116074915 A, 公开日期2023.05.05.
8. 陈友淦, 涂申奥, 周娜娜, **朱秀玲**, 江涵希, 熊艺程, 高榜君, 许肖梅. 基于Q学习和数据分级的水声网络动态计算簇头路由方法, 国家发明专利, 申请号202210638987.4, 申请日期2022.06.07, 公告号CN 115002865 A, 公开日期2022.09.02.
9. 陈友淦, **朱秀玲**, 董妍函, 赵矣昊, 杜坤芸, 陶毅, 许肖梅. 一种基于图神经网络的水声传感器网络信任模型构建方法, 国家发明专利, 申请号202311456189.0, 申请日期2023.11.03, 公告号CN 117354806 A, 公开日期2024.01.05.

致谢

123

**学位论文答辩委员会名单**

|  |  |  |  |
| --- | --- | --- | --- |
| 主 席： | 许肖梅 | 教授 | 厦门大学海洋与地球学院 |
| 委 员： | 万磊 | 副教授 | 厦门大学信息学院 |
|  | 陶毅 | 助理教授 | 厦门大学海洋与地球学院 |