

## 基于深度学习的水声被动目标识别研究综述

张奇<sup>①②③</sup> 笄良龙\*<sup>①②</sup> 王超<sup>①②</sup> 张延厚<sup>①②</sup> 嵯江浩<sup>①②</sup>

<sup>①</sup>(海军潜艇学院 青岛 266199)

<sup>②</sup>(崂山实验室 青岛 266237)

<sup>③</sup>(青岛协同创新研究院 青岛 266071)

**摘要:** 被动声呐通过接收目标自身发出的辐射噪声信号进行目标探测。水声目标识别通过分析水声信号来判别目标个体,是水声工程领域的重点研究方向。深度学习作为近年来各领域的研究热点,其在水声目标识别领域中的应用引起了学者的广泛关注。该文以水声目标识别的步骤框架为切入,介绍了典型深度网络模型;总结了深度学习在水声目标识别领域中的两大内涵:围绕时频谱、梅尔倒谱系数等特征调研了近几年深度学习作为分类器的关键问题以及研究进展,围绕数据增强、数据降噪等信号处理手段调研了近几年深度学习作为信号处理工具的关键问题以及研究进展;并从数据驱动、特征驱动、模型驱动3个方面对该领域的发展趋势进行展望,旨在推动水声目标识别领域的发展。

**关键词:** 水声目标识别;深度学习;信号处理;特征提取

中图分类号: TN911.7; O427.9

文献标识码: A

文章编号: 1009-5896(2023)11-4190-13

DOI: 10.11999/JEIT221301

## An Overview on Underwater Acoustic Passive Target Recognition Based on Deep Learning

ZHANG Qi<sup>①②③</sup> DA Lianglong<sup>①②</sup> WANG Chao<sup>①②</sup> ZHANG Yanhou<sup>①②</sup>  
ZHUO Jianghao<sup>①②</sup>

<sup>①</sup>(Naval Submarine Academy, Qingdao 266199, China)

<sup>②</sup>(Laoshan Laboratory, Qingdao 266237, China)

<sup>③</sup>(Qingdao Institute of Collaborative Innovation, Qingdao 266071, China)

**Abstract:** Passive sonar detects targets by receiving radiated noise signals emitted from the targets. Underwater acoustic target recognition is an important research area in the underwater acoustic engineering field to identify individual targets by analyzing underwater acoustic signals. As a research hotspot in various fields in recent years, deep learning has attracted considerable attention from scholars for its application to the underwater acoustic target recognition field. Based on the step framework of underwater acoustic target recognition, two typical deep network models are introduced. Herein, two major implications of deep learning in the underwater acoustic target recognition field are summarized. The key issues and research progress in recent years are investigated for deep learning as a classifier based on features such as spectrograms and mel-frequency cepstrum coefficient and for deep learning as a signal processing tool based on signal processing methods such as data enhancement and data denoising. The development trend of this field is explored from three aspects, namely, data-driven, feature-driven, and model-driven, to promote the development of underwater acoustic target recognition.

**Key words:** Underwater acoustic target recognition; Deep learning; Signal processing; Feature extraction

收稿日期: 2022-10-14; 改回日期: 2023-03-31; 网络出版: 2023-04-06

\*通信作者: 笄良龙 liangld@126.com

基金项目: 国家重点研发计划(2021YFC3100900), 崂山实验室科技创新项目(LSKJ202201100), 青岛协同创新研究院创新计划(LYY-2022-05)

Foundation Items: The National Key R&D Program of China (2021YFC3100900), The Financially Supported by Laoshan Laboratory (LSKJ202201100), The Innovation Plan of Institute of Collaborative Innovation (LYY-2022-05)

# 1 引言

声呐设备利用水下声波确定目标的存在，通过信号处理完成对水声目标的探测、定位、跟踪与识别等任务<sup>[1]</sup>。近年来随着计算机硬件的更新迭代与深度学习理论不断创新<sup>[2,3]</sup>，深度学习在计算机视觉<sup>[4]</sup>、自然语言处理<sup>[5]</sup>、数据增强<sup>[6]</sup>、数据降噪<sup>[7]</sup>等领域取得了令人满意的效果，这为水声工程领域中的目标定位、目标检测、目标识别、噪声抑制等方向提供了新的解决思路。深度学习作为一种数据驱动方法，在水声目标识别领域中的基本思路是将水声信号与目标类别映射成一种高度非线性的拟合关系。在工程技术领域，水下无人平台与现代水声对抗装备正蓬勃发展，人们对多平台协同探测目标技术与智能化战争中OODA环加速的需求明显增加。同时，船舶降噪工艺水平的提高使得目标辐射噪声降低，水声数据受到“污染”，这对分类器性能的提升以及对数据增强、数据降噪、特征提取等信号处理算法提出了新的挑战。因此，水声目标智能识别技术成为必然发展趋势。所谓“智能”，即基于人工智能算法，减少人在回路中的作用。水声目标识别的本质是模式识别在水声工程领域中的应用，其流程如图1所示。水声目标识别系统包含数据获取、数据预处理、特征提取、分类器设计、类别决策等步骤。

目前国内外学者针对深度学习在水声目标识别中研究的主要内容集中在基于舰船辐射噪声、海洋环境噪声、海洋生物回声等。据公开论文显示，国内的研究机构集中在部分高校与科研院所，主要包括哈尔滨工程大学<sup>[8-13]</sup>，西北工业大学<sup>[14-18]</sup>，西北理工大学<sup>[19-21]</sup>，海军潜艇学院<sup>[22-25]</sup>，海军工程大学<sup>[26,27]</sup>，东南大学<sup>[28,29]</sup>，中国科学院声学所<sup>[30,31]</sup>等。国外的研究机构有巴西海军研究所<sup>[32,33]</sup>，伊朗科技大学等高校<sup>[34-37]</sup>，巴基斯坦国立科技大学<sup>[38]</sup>，墨西哥国立大学<sup>[39]</sup>，美国华盛顿大学的应用物理实验室(UW-APL)<sup>[40]</sup>、海洋学院，美国伍兹霍尔海洋研究所

(WHOI)<sup>[41]</sup>，美国加利福尼亚大学(UCSD)克利普斯海洋研究所<sup>[42]</sup>。国内外研究机构大多处于理论研究阶段，实际工程应用较少。

本文结合目前深度学习中具有代表性的模型框架，以深度学习分别作为分类器和信号处理工具两个应用为例，对近几年深度学习在水声目标识别中的应用及研究情况进行了具体介绍，并提供了后续研究中的解决思路。本文框架如图2所示。

## 2 典型深度学习模型

2006年，多伦多大学Hinton等人<sup>[2]</sup>提出了针对梯度消失问题的解决方案：“无监督预训练+有监督微调”，首次提出了无监督学习和有监督学习相结合的训练网络思想，“深度学习”术语被正式提出。本节将介绍两种在水声目标识别中常用的典型深度神经网络模型。

### 2.1 卷积神经网络

卷积神经网络(Convolutional Neural Network, CNN)是一类特殊前馈神经网络。自2012年Alex Krizhevsky等人<sup>[3]</sup>提出的AlexNet在ImageNet上取得了最佳分类效果后，CNN被广泛地应用于计算机视觉领域的各项任务。目前在目标识别领域，CNN的识别正确率和泛化性能均优于深度学习的其他参数量相当的标准网络模型<sup>[43]</sup>。CNN的网络模型如图3所示。CNN包含输入层、卷积层、池化层、全连接层和输出层。网络通过卷积操作获取不同卷积层的特征图(feature map)，通过反向传播算法训练卷积核与偏置。特征图的计算表达式如式(1)

$$h_i = f(h_{i-1} \otimes w_i + b_i) \quad (1)$$

$h_i$ 代表第*i*层特征图， $w_i$ 代表第*i*层卷积核， $\otimes$ 符号代表卷积运算， $b_i$ 是偏置向量， $f$ 是激活函数。

CNN适用于处理图像数据，因此常用来对水声信号的LOFAR谱、Mel功率谱、小波谱等2维特征图进行分类。

### 2.2 循环神经网络

循环神经网络(Recurrent Neural Networks,

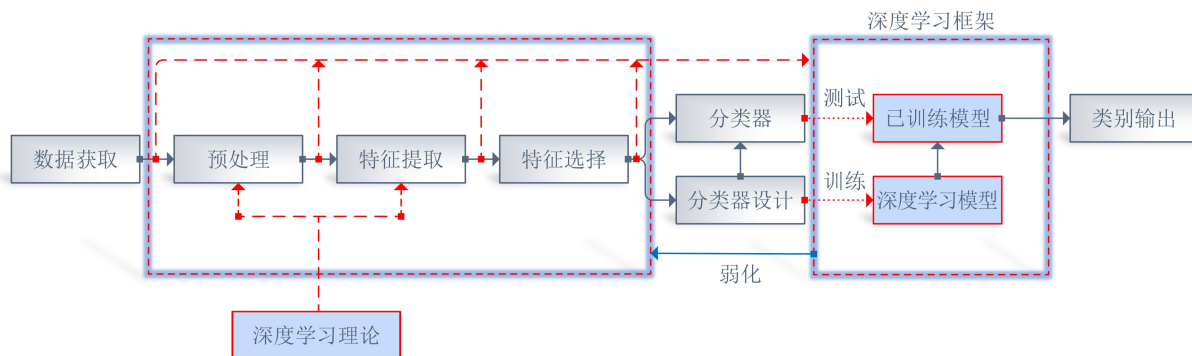


图 1 水声目标识别流程框架

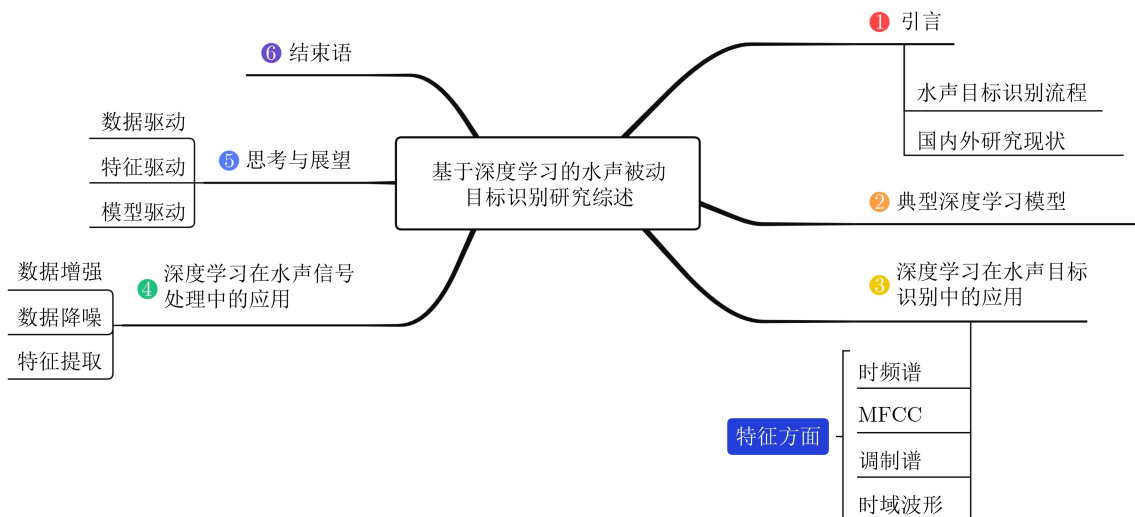


图2 本文框架

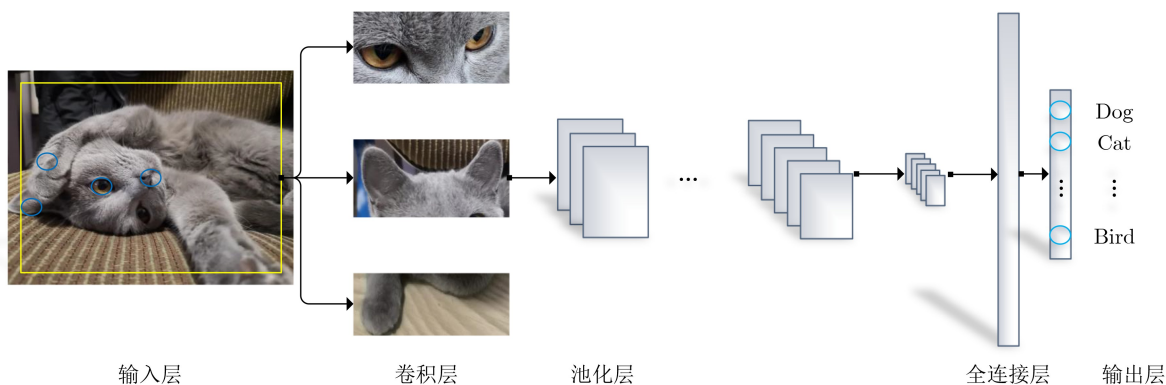


图3 CNN结构示意图

RNN)相对于传统神经网络增加了记忆功能,在时序问题和自然语言处理等领域取得了巨大的成功。目前RNN中常用的变式是长短时记忆网络(Long Short-Term Memory, LSTM)<sup>[44]</sup>,相对于常规RNN在一定程度上解决了RNN的长时依赖问题。LSTM由3个门来控制,分别是输入门、遗忘门和输出门,其内部结构如图4所示。

对于 $t$ 时刻, LSTM的输入有:前一个时刻的细胞状态 $C_{t-1}$ ,前一个时刻的网络输出 $h_{t-1}$ 和当前时刻的输入向量 $x_t$ ;输出有:当前时刻细胞状态 $C_t$ ,当前时刻网络输出 $h_t$ 。

首先计算遗忘门:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (2)$$

其中,  $[\ ]$ 代表连接 $h_{t-1}$ 和 $x_t$ 两个向量,  $\sigma$ 是sigmoid激活函数,上式表示遗忘门的衰减系数。

然后计算输入门:

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (3)$$

$$\tilde{C}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \quad (4)$$

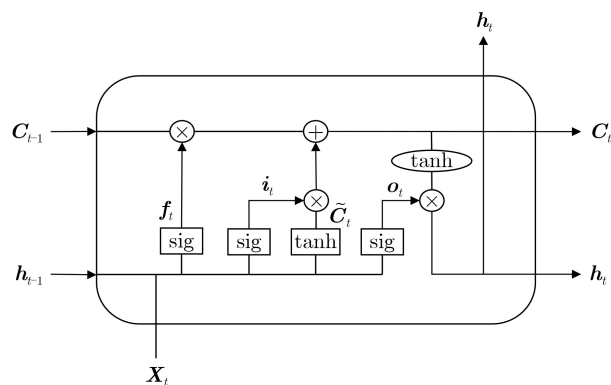


图4 LSTM内部结构

其中,  $i_t$ 是 $t$ 时刻输入门的衰减系数,  $\tilde{C}_t$ 是输入门的输入,  $\tanh$ 是激活函数。

最后计算输出门:

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \quad (5)$$

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (6)$$

$$h_t = o_t * \tanh(C_t) \quad (7)$$

其中,  $o_t$ 是 $t$ 时刻输出门的衰减系数,  $h_t$ 是输出门

的输出。式(2)、式(3)、式(4)、式(6)中 $W_f$ ,  $W_i$ ,  $W_c$ ,  $W_o$ 和 $b_f$ ,  $b_i$ ,  $b_c$ ,  $b_o$ 分别代表网络训练过程中的权重矩阵和偏置向量。

需要指出的是, 遗忘门、输入门、输出门是一种网络结构, 并不单指某一函数或衰减系数。另外, 递归神经网络(Recursive Neural Network)作为深度学习算法中的另一种网络, 不建议简称为RNN<sup>[10]</sup>, 以免混淆, 在深度学习领域中, 不作特殊说明的情况下RNN一般默认为循环神经网络<sup>[45]</sup>。LSTM常用于处理序列数据, 因此常用来对水声信号的MFCC、DEMON谱、小波系数、包络系数等1维特征序列进行分类。

### 3 深度学习在水声目标识别中的应用

通常而言, 水声信号的声压通道为20 Hz~20 kHz的1维时域数据。不同于一般的模式识别, 如基于深度学习的图像识别可直接将图像数据作为网络的输入, 对于时间长、采样率高的水声信号若在不做分帧处理的条件下直接将原始数据作为输入可能会使网络负担过重。另外, 原始数据中往往包含着噪声与干扰, 更多的是海洋噪声与干扰信息, 对目标的表达能力较弱, 而变换域特征能够更好地表达目标, “拉开”不同类别的类间距。因此, 由于水声信号的特异性, 目前大部分学者对于水声目标识别的研究仍是以先提取特征, 然后将人工提取的特征作为深度学习的输入进行分类为主。本节以水声信号的常用特征为切入, 调研了目前常用水声特征作为网络输入的研究进展。同时评价了不同特征在参与网络训练时的优劣, 并总结了特征融合这一概念在水声目标智能识别中的应用。

#### 3.1 基于时频谱特征的水声目标智能识别

短时傅里叶变换(Short Time Fourier Transform, STFT)在雷达、通信、水声领域应用广泛, 具有多种优点。STFT扩展了特征维度, 将原来对整段信号求频谱的方式变为了加窗求傅里叶变换, 增加了时间维度。同时, STFT也是无源声呐信号分类过程中最基本、最流行的信号处理方法<sup>[35]</sup>。

文献[14]提出了一种水声目标辐射噪声音色感知的深度卷积神经网络模型ASTEM\_DCNN。该模型由频率感知模块、音色感知模块和决策模块组成。频率感知模块通过1维CNN提取不同频率的线谱分量, 音色感知模块通过卷积层合成不同频率分量的信号并对信号作时频变换提取信号时频谱, 决策模块是2维CNN用于分类。通过原始时域数据、时频谱两种网络输入和不同网络模型的对比, 文章所提方法的识别准确率达78.2%, 比其他深度学习模型高出2.1%~13.5%。值得一提的是, 该模型是

一种端到端的网络模型, 即输入是原始时域数据, 时频特征是在网络结构中自动提取的。文献[14]从某种程度上说明了原始数据包含着样本最多的信息, 深度学习具有分类原始数据的能力。文献[46]提取了水声信号的MFCC、LOFAR谱图特征作为网络输入, 比较了CNN, LSTM和机器学习中的SVM 3种方法在不同信噪比下的识别结果, 表明CNN的识别效果最优, 达到了95.22%的识别率, 3种分类器的AUC值分别为0.991 4, 0.989 2, 0.953 6, 对于舰船辐射噪声仿真信号, 在-10 dB条件下CNN和LSTM可达到近80%的识别率。文献[38]利用在印度洋海域采集的4种不同类型的船舶样本数据进行分类, 使用了“LOFAR谱图输入+CNN网络分类”的方法, 最终准确率达99.4%。

单一特征所包含的信息有限, 多特征融合往往包含着目标更多的信息, 相比于单一特征更有利于对数据的表达。信息融合的3个层级包括: 数据级、特征级和决策级, 特征级融合是信息融合的第2层级。通常而言, 频谱图是STFT的幅度谱, 舍弃了相位信息, 文献[22]利用了水声数据的STFT幅度谱、相位谱和双谱3种特征进行融合, 包含了更加丰富的信息, 且首次将双谱特征用于基于深度学习的水声目标识别。双谱特征具有抑制高斯白噪声的特点, 可以有效避免加性高斯白噪声。该文献利用3个CNN网络进行集成, 在Softmax层使用了蛙跳算法进行决策级融合, 相较于单一特征, 在4类实测水声数据集取得了更高的识别率, 达90.66%。值得注意的是, 在该文献中STFT相位谱的识别率为69.19%, 超过25%, 说明STFT相位谱包含了目标的可分信息。舰船辐射噪声在频谱上的信息分布并不均匀, 大部分信息都集中在低频段, 而目前大部分所提取的频谱图特征其坐标轴刻度都是线性的, 基于此, 文献[26]提出了基于双对数谱的识别方法, 即将LOFAR谱图的频率轴从线性刻度映射到对数刻度, 提取对数LOFAR谱作为CNN的输入, 相比于线性谱识别率提高了2.87%。文献[33]使用了CNN中的MobileNetV2网络进行分类, 将网络的输入从 $224 \times 224 \times 3$ 的大小改成了 $513 \times 513 \times 1$ , 以适应输入的LOFAR谱图, 与线性核SVM、高斯核SVM进行了对比, 在测试集上的识别率达到99.4%, 线性核SVM与高斯核SVM的识别率分别为61.6%, 98.6%, 由此说明了深度学习分类器对比一般机器学习分类器的优越性。文献[35]将STFT做了改进, 提出了短时分数阶傅里叶变换(Short Time Fractional Fourier Transform, STFrFT)的特征提取方法, 利用径向基神经网络(Radial Basis



Function, RBF)对比了分数参数 $\alpha$ 在不同取值下的识别率,发现 $\alpha = 0.96$ 是最优分数阶。

### 3.2 基于梅尔倒谱系数特征的水声目标智能识别

MFCC反映了人对语音的感知特性,是在Mel标度频率提取出来的倒谱系数。MFCC更符合人耳的听觉特性,因此广泛应用于语音识别领域,在水声目标识别领域同样流行。

由于MFCC特征是一组向量,因此“MFCC+LSTM”的水声目标识别方法较为常见。文献[23]将实测水声数据分为了水面和水下两类,提取了频谱、时域波形、MFCC3种特征作为LSTM网络的输入,结果发现MFCC作为特征向量时识别率最高,水面与水下的识别率分别为84%和74%。文献[24]是基于文献[23]方法的进一步研究,在MFCC单一特征的基础上,提出了将目标噪声时域包络、DEMON谱和MFCC3种特征进行融合的方式进行识别,首先利用LSTM网络对3种特征进行训练得到分类模型并保存参数,其次将测试数据的3种特征分别用3种训练好的模型进行预测,得到3个目标所属类别的概率,最后通过D-S证据理论进行决策级融合判别目标类型。文章利用了信息融合中特征级融合与决策级融合相结合的方式对水声目标进行分类,较单一特征的识别方法有更高的识别率和更低的虚警率。然而,时域包络和DEMON谱这两种特征分别是同一种信息在不同域上的表达,对于能够自动提取目标信息的深度学习而言含有冗余信息,可能会加深网络负担,识别效果提升不高,但文章中信息融合与深度学习相结合的思想值得借鉴。文献[30]采用了将实测水声数据的频谱、MFCC、功率谱3种特征进行融合,与文献[24]融合方式不同的是,该文献将网络训练后的不同抽象特征通过连接层进行融合后,再通过多层神经网络对连接特征进行学习分类,即只在特征级融合。同样地,文献[30]中的频谱和功率谱对于深度学习存在着信息冗余。

将不同网络进行组合的网络模型也常被用作分类器<sup>[47-50]</sup>,往往较单一网络具有更高的识别效果。文献[47]将1D\_CNN(1维卷积神经网络)与LSTM级联成联合网络,提取MFCC、Mel频率、色谱图、光谱对比度、调性网格5种水声信号的音频特征作为网络输入的联合特征向量,首先将联合特征输入1D\_CNN,然后将1维卷积后的数据输入LSTM网络进行识别,这里前端的1D\_CNN起到了特征提取和降维的作用,1D\_CNN、LSTM和联合网络在测试集的识别率分别为84.18%, 76.10%, 92.14%,验证了联合网络具有更高的识别性能。文献[49]将LSTM和CNN作为并行网络,将“1维时域+

LSTM”和“2维频谱+CNN”的输出结果联合Softmax进行识别。“1维时域+LSTM”、“2维频谱+CNN”和联合网络的识别结果分别为87.2%, 85.4%, 96.4%。需要注意的是,虽然上述两篇文章都是CNN和LSTM两种网络的组合,但组合方式完全不同,可以将文献[47]的组合方式理解为CNN和LSTM“串联”,文献[49]的组合方式理解为“并联”。文献[50]采用了类似文献[49]的“并联”网络结构,将水声信号的MFCC特征分别作为CNN和LSTM网络的输入,将两种网络所提取的特征进行组合分类,对比单一网络提高了识别率。

此外,有学者<sup>[34,36,37]</sup>将智能优化算法用于训练网络以提升分类性能。例如,文献[36,37]使用了黑猩猩优化算法(ChOA)来训练神经网络。ChOA通过模拟攻击者、驱赶者、拦截者和追逐者4类黑猩猩协同狩猎行为来达到求解问题的目的。文献[36]提出了基于ChOA训练多层感知机(MLP)的方法,即MLP-ChOA。利用ChOA训练一个具有单隐层的MLP神经网络。所提出的训练方法存在两个关键问题:第一,MLP神经网络的权重参数在ChOA中的表示,即将每只黑猩猩被编码为一个1维向量,该1维向量由MLP的权重和偏置向量组成;其次用均方差来表示损失函数。该文献提取了被动声呐信号的MFCC特征作为网络输入,对比了离子运动算法(IIMO)、灰狼优化算法(GWO)、粒子群优化-引力搜索混合算法(PSO-GSA)3种方法,比较了分类准确率和损失函数值,结果显示所提出的MLP-ChOA算法具有更好的分类效果,识别率达到97.95%。文献[37]在文献[36]的基础上进行了改进,提出了基于模糊黑猩猩算法(Fuzzy-ChOA)和MLP的海洋生物分类模型FChOA-MLPNN, FChOA-MLPNN利用模糊逻辑来调整ChOA的控制参数,该文献提取海洋生物声音的Mel频率作为输入特征,比较算法包括ChOA、冠状病毒优化算法<sup>[51]</sup>(CHIO,灵感来源于应对2019冠状病毒大流行的群体免疫概念)、哈里斯鹰优化算法(HHO)、黑寡妇优化算法(BWO)、卡尔曼滤波算法,比较基准包括收敛速度、损失函数、分类率、ROC,结果表明所设计FChOA-MLPNN分类模型的各种性能均优于其他智能优化算法。文献[34]还提出了自适应最佳GSA优化算法训练MLP,对被动声呐信号的MFCC特征进行分类,分类准确率达95.20%。

### 3.3 基于调制谱特征的水声目标智能识别

DEMON谱包含了船舶的轴频、叶频以及桨叶数特征。桨叶数作为稳健特征,是水声目标识别的重要判别依据。DEMON谱的提取方法如图5所

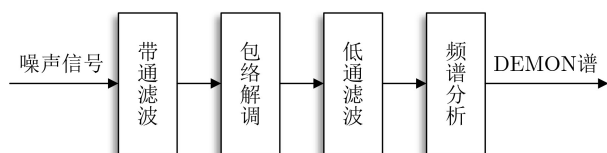


图5 DEMON谱提取流程

示。结合目前对DEMON谱的研究现状, 本文总结了两大类基于DEMON谱的桨叶数判别方法: 一是基于专家系统的桨叶数提取方法, 二是基于轴频谐波线谱的桨叶数提取方法。基于专家系统的桨叶数提取方法直接通过观察DEMON谱信息人工判别桨叶数, 如通过构建谐波簇理论模型建立桨叶数识别专家系统<sup>[52]</sup>; 通过DEMON谱相位耦合特性提取双谱切片谱来识别桨叶数<sup>[53]</sup>; 通过自适应滤波等理论算法来增强DEMON谱谐波线谱<sup>[54]</sup>等, 该方法完全依赖于专家经验, 受船舶工况、海洋环境等因素影响较大, 只能描述某些典型情况, 稳健性较低。基于轴频谐波线谱的桨叶数提取方法通过建立轴频谐波线谱等特征模板库, 利用模式识别方法对桨叶数进行识别, 结合了专家经验与机器识别, 稳健性较高。目前常用的方法有模板匹配、SVM、深度学习等。

文献<sup>[55]</sup>构建了一个DEMON谱的6级10阶模板库, 即将每阶轴频谐波的幅度分为6级, 共提取前10阶谐波。该模板是一个封闭集, 即使如此, 状态总数共有 $6^{10} \approx 6 \times 10^7$ 个。利用神经网络对已知样本进行训练, 较模板匹配法得到了更高的识别率。文献<sup>[11]</sup>提取了300组海试数据信号的DEMON谱、1.5维谱和MFCC特征, 利用CNN和SVM两种分类器作对比, 6种识别方法中, MFCC+CNN的识别率最高, 为96.0%, 并且CNN在不同特征下的分类性能都优于SVM。文献<sup>[28]</sup>利用CNN和决策树两种分类器分别对DEMON谱和连续谱进行识别, 结果发现决策树的识别准确率远低于CNN。与文献<sup>[55]</sup>不同的是, 文献<sup>[11,28]</sup>直接将DEMON谱的1维向量作为网络的输入, 并没有提取轴频谐波簇建立模板。

文献<sup>[56]</sup>首先利用梳状滤波器(comb filter)增强DEMON谱, 然后用CNN进行目标分类。梳状滤波器<sup>[57]</sup>被广泛应用于语音识别领域的基音检测中, DEMON谱中的轴频检测也属于基音检测的一种。结果发现基于CNN和梳状滤波器的DEMON谱识别率为92%。同样地, 作者于文献<sup>[58]</sup>中对文献<sup>[56]</sup>作了改进, 在梳状滤波器增强DEMON谱的基础上, 使用的分类网络为CNN+LSTM结构, 先利用CNN的卷积操作提取特征, 再用LSTM进行分类,

即前文提到的联合网络中“串联”的方法。实验通过与改进的最大公约数<sup>[59]</sup>和序列匹配法<sup>[60]</sup>作对比, 发现CNN+LSTM网络的识别率最高, 海试数据和湖试数据的准确率分别为90.6%和98.3%。此外, 还有前文<sup>[24]</sup>提到的将DEMON谱作为融合特征中的一种特征训练网络。

### 3.4 基于时域波形的水声目标智能识别

时域波形通常也是原始时域数据, 即水下目标辐射噪声的原始信号。原始数据包含着目标最原始、最多的信息, 在深度学习算法流行后, 也有学者尝试直接将原始数据作为深度网络的输入, 如前文提到的文献<sup>[14,23]</sup>。直接利用时域波形进行目标识别的优势在于实现了端对端的识别方式, 并且减少了人工特征提取这一环节, 实现了算法轻量化的目的。

文献<sup>[8]</sup>提出了深度可分离卷积神经网络模型自动提取目标特征, 将1维原始数据、小波特征、MFCC、Mel频率、HHT、非线性听觉特征、频谱、倒谱特征与所提模型自动提取的特征进行了比较, 结果发现文章所提出的模型识别率最高, 达90.1%, 高于其他特征与分类器相组合的识别率。文献<sup>[61]</sup>则完全使用了水声信号的时域数据进行目标识别, 并在网络端进行了改进。该文献使用了1维CNN, 引入卷积核为1的卷积层, 针对CNN在全连接层会丢失特征位置信息的问题, 使用了全局平均池化层(Global Average Pooling, GAP)代替全连接层, GAP<sup>[62]</sup>是目前深度学习领域较为流行的手段, 即直接将来自上一层的特征图取全局平均后映射到下一层, 该方法既保留了上一层特征的位置信息, 又大大减少了网络的参数。通过与高阶统计特征(HOS)、传统CNN对比, 总识别率达到了91.7%, 超过了HOS和传统CNN的85.0%和69.8%。文献<sup>[18]</sup>提出了多尺度稀疏简单循环单元(SRU)对3类实测水声目标辐射噪声原始数据进行分类, 对比CNN具有更高的识别率。

### 3.5 基于其他特征的水声目标智能识别

除上述提到的在水声目标智能识别中常用的特征之外, 文献<sup>[16]</sup>将小波分析与CNN相结合, 提出了WAVEDEC\_CNN的水声目标识别方法。特征方面, 该文献直接将分解后的小波系数作为网络输入; 网络方面, 在卷积层和池化层之间添加了目前深度学习领域中流行的批量归一化层(Batch Normalization, BN)<sup>[63]</sup>。与MFCC+CNN、无预处理的小波系数结合传统CNN、经过预处理后的小波系数结合CNN、经验模态分解(EMD)结合CNN 4种方法进行了对比, 识别率分别提高了15.38%, 4.41%,



3.23%, 12.81%。文献[31]提取了湖试数据的FBANK特征,使用时延神经网络(Time Delay Neural Network, TDNN)进行分类,对比SVM分类器,取得了更高的分类准确率。

除此之外,还有基于频谱[32]、六分之一倍频程谱[39]、水声干涉条纹图像[19]等特征的水声目标识别方法,识别流程都是基于“特征+深度学习模型”的方式,文章在此不再赘述。

## 4 深度学习在水声信号处理中的应用

深度学习作为一种强大的分类器,其识别性能和前端输入数据的质量息息相关。基于目前水声数据稀缺、信噪比低、特征提取会造成信息损失的现状,本节总结了深度学习在水声数据增强、降噪、特征提取中的应用。深度学习作为信号处理工具,其最终目的仍是提升识别性能。

### 4.1 深度学习在水声数据增强中的应用

数据增强(Data Augmentation)具有扩充训练数据集、提高模型泛化能力的作用。生成对抗网络(GAN)[6]是2014年Ian Goodfellow等人提出的一种生成模型,核心思想是从训练样本中学习所对应的概率分布,以期望通过概率分布函数获取更多的“生成”样本来实现数据扩充。GAN具有生成不存在于真实世界的的能力,可以从一定程度上解决水声数据稀缺的问题,实现数据增强,满足深度学习对于大数据的需求。

文献[12]从时域和频域特征出发,用GAN实现数据增强。基于时域的数据增强中,该文献在生成器的结构上引入了时序卷积(TCN)、转置卷积和Embedding结构,实验中利用生成模型对12类船舶目标的每种类型生成500个样本,在原始数据集中的每一类随机抽取300个样本组成新的训练集,对比了TCN, Conv1d, LSTM3种模型在测试集的识别率,分别为87.23%, 80.23%, 83.13%。基于频域的数据增强中,频域特征选择LOFAR谱,采用深度卷积生成对抗网络(DCGAN)和条件生成对抗网络(CGAN)相结合的网络结构,DCGAN和GAN原理相同,在生成器和判别器上采用了卷积模块,用网络所生成的样本对识别网络进行微调。使用数据增强后,DCGAN+GAN模型在测试集上的识别率影响不大,而ACGAN和SA两种模型在测试集上识别率提升了约5%,证明了利用GAN作数据增强的有效性。文献[9]改进了条件DCGAN模型,将网络的全连接层用K-邻近算法(K-Nearest Neighbor, KNN)、随机森林(Random Forest, RF)、SVM等集成分类器代替,生成伪DEMON谱数据来缓解数据不足与样本不均衡的问题,用查准率、查全率、

F1分数作为评价指标,通过与原始样本数据集、用传统的SMOTE算法进行数据增强后的数据集的分类结果进行对比,证明了所提出的改进DCGAN模型在数据量较少时能够扩充数据集、有效提高分类器的识别率。文献[17]通过GAN对实测水声数据进行增强,并将GAN模型本身当作分类器,该文献基于3类实测水声数据,每一类水声样本时长在150 s左右,属于小样本数据集,在小样本数据集下对比了GAN、DBN、自编码网络3种深度学习分类器的识别率,其性能均优于MFCC+Softmax方法,3种方法的识别率分别为96.31%, 93.54%, 90.72%,并通过在测试集添加噪声的方式验证了该模型具有更强的鲁棒性能。

### 4.2 深度学习在水声数据降噪中的应用

目前,基于深度学习的水声数据降噪研究较少,且都是基于CNN对信号STFT特征进行重构的方式展开研究。文献[64]针对海洋环境存在大量非高斯噪声的特性研究了基于CNN的水声信号降噪重构方法,该方法首先提取传感器接收带噪信号的STFT,然后通过去噪CNN模型学习带噪信号STFT和残差谱之间的映射关系在时频谱上去除环境噪声,最后将去噪信号的STFT通过短时傅里叶逆变换(ISTFT)得到纯净信号。该文献对比了传统的LMS滤波去除噪声的方法,通过仿真数据与实测数据的对比,发现深度学习方法在数据降噪任务中表现出了良好的泛化能力。文献[29]应用了与文献[64]同样的降噪方法,采用去噪CNN模型对水声信号LOFAR谱图进行降噪,从而达到线谱增强的目的,该文献与传统的维纳滤波算法作对比,使用去噪后的信噪比作为评价指标,所采用的深度学习去噪方法能够保持在18 dB以上的增益,对比维纳滤波算法具有较大的优势。

上述两篇文献研究的深度学习降噪方法都来自Park等人[7]于2016年提出的基于全卷积神经网络的语音数据降噪方法,该方法属于信号映射域的降噪(相对于时域端对端降噪),即通过深度网络学习信号变换域中的特征进行降噪,文献[7,29,64]都是针对信号的STFT特征进行研究的。然而,该方法需要纯净信号与噪声干扰信号的先验知识,目前海洋环境复杂,不同时空下的海洋噪声分布难以预测,只能通过仿真或在纯净信号的基础上添加噪声的方式进行研究。

### 4.3 深度学习在水声数据特征提取中的应用

前文中所提到的特征都是物理意义明确或是不同域中可解释的特征,而人工特征提取都会不可避免地造成信息损失。深度学习具有自动提取数据特

征的能力,在网络训练过程中的每一层都是样本数据的特征,该特征具有“不可解释性”,因此也可称为抽象特征,在CNN中称为特征图,在自编码等全连接神经网络中称为特征向量,网络的层次越深所提取的特征层级也越深。基于深度学习所提取的抽象特征信息损失较小,因此在特征提取阶段也经常使用深度学习算法。

文献[25]在利用栈式自编码(SA)对水声目标进行识别的过程中将SA逐层所提取的特征向量进行了展现,该文献设计了5层AE,每一层的神经元数量递减,在最后一层用Softmax分类器进行分类。文献[15]提出了一种门控循环单元和卷积自编码(GRU-CAE)协同深度学习网络模型,GRU-CAE结合了CAE和GRU的优点,CAE可以提取船舶辐射噪声频谱图的空间信息,GRU可以提取船舶辐射噪声频谱图的时间信息。该文献首先构建GRU-CAE模型,利用该模型提取深度协同特征;其次构建特征模板;最后利用模板匹配分类器进行目标分类。与LSTM网络相比,GRU对时间结构更敏感,结构更简单,训练速度也更快,实验结果证明GRU-CAE在实测水声数据集上有更好的识别性能。文献[65]采用了特征融合策略提取了舰船辐射噪声经不同子带中小波分解的时间特征、统计特征、频谱特征、倒谱特征、希尔伯特谱特征、小波特征、特定特征(归一化线谱特征、线谱密度、频带功率、LPC系数)和DBN所提取的抽象特征,采用2维特征融合策略构成特征矩阵,该文献将DBN提取的特征与其他可解释性特征一起融合,采用KNN作为分类器,实验结果表明所提出的特征融合方案可以显著提高船舶辐射噪声的识别性能。由于域的变化会造成信息的损失,文献[13]使用了1维CNN对原始时域数据进行特征提取,并将所提取的特征用ELM进行分类,在实际民船水声数据集上对比了不同特征与不同分类器的分类效果,特征包括HOS、小波特征、MFCC、HHT、Mel频率、非线性听觉感知特征、倒谱特征、CNN对原始数据自动提取的特征,分类器包括SVM, DNN, CNN和ELM,结果发现CNN提取特征+ELM识别率最高,达90.9%。

## 5 思考与展望

从识别步骤而言,基于深度学习的水声目标识别所面临的挑战包括3个方面,分别是水声数据、特征提取和网络模型构建。以大数据为主来提升识别性能的方式为数据驱动;以多特征融合为主来提升识别性能的方式为特征驱动;以网络模型为主来提升识别性能的方式为模型驱动。水声目标识别的

未来发展趋势将是围绕数据驱动、特征驱动、模型驱动这3个方面来展开的。

### 5.1 数据驱动

数据是海洋的生命线,由于水下环境复杂多变、水声数据获取难度较大、密级较高、各国研究机构不公开等,水声目标辐射噪声样本稀缺,目前大部分水声目标识别的研究都是基于各自研究机构所获得的数据进行展开的。而深度学习作为一种数据驱动算法需要大量数据,数据样本稀缺与不均衡是目前水声目标识别的一个关键问题。

目前大多学者都采用对水声信号分帧的方式进行识别,将每一帧作为一个训练或者测试样本。由于采样率的不同,帧长的选取会有所差异,其原则应是样本中尽量包含水声数据的全部信息,如轴频为4 Hz的舰船噪声信号,帧长大于0.25 s时才会包含目标的螺旋桨信息。帧长过短则会丢失部分水声目标信息,帧长过长则会使网络负担加重。基于目前的水声数据现状,小样本、类别不均衡的水声目标识别技术将是该领域的未来发展趋势,其中,GAN仍是效果较好且值得借鉴的数据增强方法。

### 5.2 特征驱动

前文提到对水声数据进行人工特征提取后再进行训练主要有以下两点考虑:一是原始数据维数较高使得网络负担较重;二是原始数据中往往掺杂着各种背景噪声与干扰使得网络学习到大量与水声目标无关的噪声数据,造成Garbage in, garbage out。因此针对不同类型分布的数据进行预处理与特征提取是必要的。在特征提取的过程中,应尽量提取类内间距小、类间间距大,包含目标足够信息的特征,且特征之间应包含较少的信息冗余。值得一提的是,在特征提取过程中会使研究人员对水声信号的理解逐渐深刻。

目前,水声目标智能识别处于起步阶段,工程应用较少,人在回路中的作用仍然明显,专家经验较为重要。图6为本团队于2022年5月份在中国东海某海域利用水下无人平台采集到的一段水声样本数据。结合听音判情与频谱图分析,除了海洋噪声外,该段音频中还包含了:(1)某远程低频探测声呐;(2)商船磨轴声音;(3)主动声呐。在同一时间段内出现了3种类型的信号,且各种信号频段与带宽各不相同。以上3种不同类型的水声信号都是通过专家经验分析得出的,但基于目前的深度学习算法进行识别仍较为困难。

因此,人在回路的人机交互模型仍是目前水声目标识别的主流,通过人工提取特征仍是目前水声目标识别的关键。单一特征所包含的目标信息有



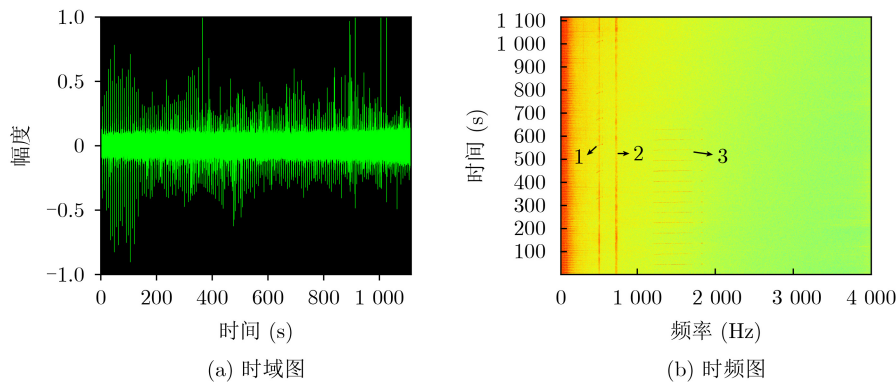


图6 中国东海海域某实测水声样本

限,目前水声信号的特征融合大多选取2~3个特征,基于多特征融合的特征驱动方法或将在工程实践中成为水声目标识别领域的主流。例如,通过仿照水声数据库构建水声特征库,提取水声样本的音频特征(包括时频谱、MFCC、DEMON谱等)、声矢量特征、能量特征、运动特征等多特征数据构建水声样本特征库,将所有或部分稳健特征通过深度学习进行特征融合,这对提升目标识别率、加深学者对水声数据样本的理解都是值得探索和研究的。

### 5.3 模型驱动

深度网络设计对识别效果起着至关重要的作用,随着计算机算力的提升与网络模型的不断扩大,深度学习算法的学习能力越来越强,由此逐渐弱化了人工特征提取的步骤。近年来,用图来表示数据的图神经网络(GNN)<sup>[66]</sup>、引入注意力机制的Transformer<sup>[4,5,67]</sup>等新兴网络模型在深度学习领域逐渐成为主流。谷歌于2021年推出了Switch Transformer<sup>[67]</sup>架构,其参数量达到 $1.6 \times 10^{12}$ 次,由此带动了深度网络步入大模型时代。由于水声样本量和计算资源的限制,目前还尚未有研究人员将GNN或Transformer应用于水声目标识别。

大模型网络需考虑算力成本、时间成本以及高能耗等问题,因此目前在硬件条件受限的条件下,轻量化网络的设计与构建符合实际工程应用与智能化战争中对OODA环加速的需求。随着未来硬件(显卡、CPU等)的发展与算力成本的下降,在水声数据集充足的条件下,基于模型驱动的水声目标识别方法或将逐步取代以特征提取、小规模网络训练等为代表的算法驱动。

## 6 结束语

本文主要将深度学习在水声目标识别中的应用进行了总结,包括深度学习作为分类器和信号处理工具这两大内涵。与传统方法相比,深度学习取得了较好的识别效果。其次从数据驱动、特征驱动、

模型驱动3个方面进行了展望,提出了下一步的研究方向,以供学者在后续研究中作参考借鉴。另外,深度学习作为黑盒模型的不可解释性也为水声目标识别带来了阻碍,网络所提取的各层特征的可解释性探索仍是目前以至将来水声目标识别所面临的挑战。

### 参考文献

- [1] 乌立克 R J, 洪申, 译. 工程水声原理[M]. 北京: 国防工业出版社, 1972: 1-2.  
URICK R J, HONG Shen. translation. Principles of Engineering Underwater Sound[M]. Beijing: National Defense Industry Press, 1972: 1-2.
- [2] HINTON G E and SALAKHUTDINOV R R. Reducing the dimensionality of data with neural networks[J]. *Science*, 2006, 313(5786): 504-507. doi: [10.1126/science.1127647](https://doi.org/10.1126/science.1127647).
- [3] KRIZHEVSKY A, SUTSKEVER I, and HINTON G E. Imagenet classification with deep convolutional neural networks[C]. Proceedings of the 25th International Conference on Neural Information Processing Systems, Lake Tahoe, USA, 2012.
- [4] LIU Ze, LIN Yutong, CAO Yue, *et al.* Swin transformer: Hierarchical vision transformer using shifted windows[C]. 2021 IEEE/CVF International Conference on Computer Vision (ICCV), Montreal, Canada, 2021: 9992-10002. doi: [10.1109/ICCV48922.2021.00986](https://doi.org/10.1109/ICCV48922.2021.00986).
- [5] VASWANI A, SHAZEER N, PARMAR N, *et al.* Attention is all you need[C]. Proceedings of the 31st International Conference on Neural Information Processing Systems, Long Beach, USA 2017.
- [6] GOODFELLOW I J, POUGET-ABADIE J, MIRZA M, *et al.* Generative adversarial nets[C]. Proceedings of the 27th International Conference on Neural Information Processing Systems, Montreal, Canada, 2014.
- [7] PARK S R and LEE J. A fully convolutional neural network for speech enhancement[C]. Interspeech 2017, 18th

- Annual Conference of the International Speech Communication Association, Stockholm, Sweden, 2017. doi: [10.21437/Interspeech.2017-1465](https://doi.org/10.21437/Interspeech.2017-1465).
- [8] HU Gang, WANG Kejun, and LIU Liangliang. Underwater acoustic target recognition based on depthwise separable convolution neural networks[J]. *Sensors*, 2021, 21(4): 1429. doi: [10.3390/s21041429](https://doi.org/10.3390/s21041429).
- [9] 李理, 李向欣, 殷敬伟. 基于生成对抗网络的舰船辐射噪声声分类方法研究[J]. *电子与信息学报*, 2022, 44(6): 1974–1983. doi: [10.11999/JEIT211077](https://doi.org/10.11999/JEIT211077).
- LI Li, LI Xiangxin, and YIN Jingwei. Research on classification algorithm of ship radiated noise data based on generative adversarial network[J]. *Journal of Electronics & Information Technology*, 2022, 44(6): 1974–1983. doi: [10.11999/JEIT211077](https://doi.org/10.11999/JEIT211077).
- [10] 卢安安. 基于深度学习方法的水下声音目标识别研究[D]. [硕士学位论文], 哈尔滨工程大学, 2017.
- LU Anan. Underwater acoustic classification based on deep learning[D]. [Master dissertation], Harbin Engineering University, 2017.
- [11] 郝宇昕. 基于深度神经网络的舰船辐射噪声特征分类技术[D]. [硕士学位论文], 哈尔滨工程大学, 2019.
- HAO Yuxin. Ship radiated noise classification method based on deep neural network[D]. [Master dissertation], Harbin Engineering University, 2019.
- [12] 姜岩松. 基于生成对抗网络的水声信号识别与分离研究[D]. [硕士学位论文], 哈尔滨工程大学, 2021. doi: [10.27060/d.cnki.ghbcu.2021.000182](https://doi.org/10.27060/d.cnki.ghbcu.2021.000182).
- JIANG Yansong. Research on underwater acoustic signal recognition and separation based on generative adversarial network[D]. [Master dissertation], Harbin Engineering University, 2021. doi: [10.27060/d.cnki.ghbcu.2021.000182](https://doi.org/10.27060/d.cnki.ghbcu.2021.000182).
- [13] 胡钢. 基于深度学习的水下目标识别和运动行为分析技术研究[D]. [博士学位论文], 哈尔滨工程大学, 2021. doi: [10.27060/d.cnki.ghbcu.2021.001397](https://doi.org/10.27060/d.cnki.ghbcu.2021.001397).
- HU Gang. Research on underwater target recognition and motion behavior analysis technology based on deep learning[D]. [Ph. D. dissertation], Harbin Engineering University, 2021. doi: [10.27060/d.cnki.ghbcu.2021.001397](https://doi.org/10.27060/d.cnki.ghbcu.2021.001397).
- [14] LI Junhao and YANG Honghui. The underwater acoustic target timbre perception and recognition based on the auditory inspired deep convolutional neural network[J]. *Applied Acoustics*, 2021, 182: 108210. doi: [10.1016/j.apacoust.2021.108210](https://doi.org/10.1016/j.apacoust.2021.108210).
- [15] YANG Honghui, ZHENG Kaifeng, and LI Junhao. Open set recognition of underwater acoustic targets based on GRU-CAE collaborative deep learning network[J]. *Applied Acoustics*, 2022, 193: 108774. doi: [10.1016/j.apacoust.2022.108774](https://doi.org/10.1016/j.apacoust.2022.108774).
- [16] 黄擎, 曾向阳. 小波分解和改进卷积神经网络相融合的水声目标识别方法[J]. *哈尔滨工程大学学报*, 2022, 43(2): 159–165. doi: [10.11990/jheu.202011040](https://doi.org/10.11990/jheu.202011040).
- HUANG Qing and ZENG Xiangyang. An underwater acoustic target recognition method combining wavelet decomposition and an improved convolutional neural network[J]. *Journal of Harbin Engineering University*, 2022, 43(2): 159–165. doi: [10.11990/jheu.202011040](https://doi.org/10.11990/jheu.202011040).
- [17] 薛灵芝, 曾向阳, 杨爽. 基于生成对抗网络的水声目标识别算法[J]. *兵工学报*, 2021, 42(11): 2444–2452. doi: [10.3969/j.issn.1000-1093.2021.11.018](https://doi.org/10.3969/j.issn.1000-1093.2021.11.018).
- XUE Lingzhi, ZENG Xiangyang, and YANG Shuang. Underwater acoustic target recognition algorithm based on generative adversarial networks[J]. *Acta Armamentarii*, 2021, 42(11): 2444–2452. doi: [10.3969/j.issn.1000-1093.2021.11.018](https://doi.org/10.3969/j.issn.1000-1093.2021.11.018).
- [18] 杨爽, 曾向阳. 基于多尺度稀疏简单循环单元模型的水声目标识别方法[J]. *哈尔滨工程大学学报*, 2022, 43(7): 958–964. doi: [10.11990/jheu.202105048](https://doi.org/10.11990/jheu.202105048).
- YANG Shuang and ZENG Xiangyang. Underwater acoustic target recognition method based on the multi-scale sparse simple recurrent unit model[J]. *Journal of Harbin Engineering University*, 2022, 43(7): 958–964. doi: [10.11990/jheu.202105048](https://doi.org/10.11990/jheu.202105048).
- [19] ZHOU Xingyue, YANG Kunde, and DUAN Rui. Deep learning based on striation images for underwater and surface target classification[J]. *IEEE Signal Processing Letters*, 2019, 26(9): 1378–1382. doi: [10.1109/LSP.2019.2919102](https://doi.org/10.1109/LSP.2019.2919102).
- [20] IRFAN M, ZHENG Jiangbin, 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. doi: [10.1016/j.eswa.2021.115270](https://doi.org/10.1016/j.eswa.2021.115270).
- [21] IRFAN M, ZHENG Jiangbin, IQBAL M, *et al.* Brain inspired lifelong learning model based on neural based learning classifier system for underwater data classification[J]. *Expert Systems with Applications*, 2021, 186: 115798. doi: [10.1016/j.eswa.2021.115798](https://doi.org/10.1016/j.eswa.2021.115798).
- [22] ZHANG Qi, DA Lianglong, ZHANG Yanhou, *et al.* Integrated neural networks based on feature fusion for underwater target recognition[J]. *Applied Acoustics*, 2021, 182: 108261. doi: [10.1016/j.apacoust.2021.108261](https://doi.org/10.1016/j.apacoust.2021.108261).
- [23] 张少康, 王超, 田德艳, 等. 长短时记忆网络水下目标噪声智能识别方法[J]. *舰船科学技术*, 2019, 41(12): 181–185. doi: [10.3404/j.issn.1672-7649.2019.12.035](https://doi.org/10.3404/j.issn.1672-7649.2019.12.035).
- ZHANG Shaokang, WANG Chao, TIAN Deyan, *et al.* Intelligent recognition of underwater target noise based on long short-term memory networks[J]. *Ship Science and*

- Technology*, 2019, 41(12): 181–185. doi: [10.3404/j.issn.1672-7649.2019.12.035](https://doi.org/10.3404/j.issn.1672-7649.2019.12.035).
- [24] 张少康, 王超, 孙芹东. 基于多类别特征融合的水声目标噪声识别分类技术[J]. 西北工业大学学报, 2020, 38(2): 366–376. doi: [10.1051/jnwpu/20203820366](https://doi.org/10.1051/jnwpu/20203820366).
- ZHANG Shaokang, WANG Chao, and SUN Qindong. Underwater target noise recognition and classification technology based on multi-classes feature fusion[J]. *Journal of Northwestern Polytechnical University*, 2020, 38(2): 366–376. doi: [10.1051/jnwpu/20203820366](https://doi.org/10.1051/jnwpu/20203820366).
- [25] 程玉胜, 邱家兴, 刘振, 等. 水声被动目标识别技术挑战与展望[J]. 应用声学, 2019, 38(4): 653–659. doi: [10.11684/j.issn.1000-310X.2019.04.023](https://doi.org/10.11684/j.issn.1000-310X.2019.04.023).
- CHENG Yusheng, QIU Jiaying, LIU Zhen, *et al.* Challenges and prospects of underwater acoustic passive target recognition technology[J]. *Journal of Applied Acoustics*, 2019, 38(4): 653–659. doi: [10.11684/j.issn.1000-310X.2019.04.023](https://doi.org/10.11684/j.issn.1000-310X.2019.04.023).
- [26] 徐源超, 蔡志明, 孔晓鹏. 基于双对数谱和卷积网络的船舶辐射噪声分类[J]. 电子与信息学报, 2022, 44(6): 1947–1955. doi: [10.11999/JEIT211407](https://doi.org/10.11999/JEIT211407).
- XU Yuanchao, CAI Zhiming, and KONG Xiaopeng. Classification of ship radiated noise based on bi-logarithmic scale spectrum and convolutional network[J]. *Journal of Electronics & Information Technology*, 2022, 44(6): 1947–1955. doi: [10.11999/JEIT211407](https://doi.org/10.11999/JEIT211407).
- [27] 徐源超, 蔡志明. 水声目标分类算法性能评估[J]. 哈尔滨工程大学学报, 2020, 41(10): 1559–1565. doi: [10.11990/jheu.202007114](https://doi.org/10.11990/jheu.202007114).
- XU Yuanchao and CAI Zhiming. Performance evaluation on the algorithm of underwater acoustic target classification[J]. *Journal of Harbin Engineering University*, 2020, 41(10): 1559–1565. doi: [10.11990/jheu.202007114](https://doi.org/10.11990/jheu.202007114).
- [28] 徐萍. 水声目标辐射噪声特征提取与识别技术研究[D]. [硕士学位论文], 东南大学, 2019. doi: [10.27014/d.cnki.gdnau.2019.002039](https://doi.org/10.27014/d.cnki.gdnau.2019.002039).
- XU Ping. Research on feature extraction and recognition technology of underwater acoustic target radiated noise[D]. [Master dissertation], Southeast University, 2019. doi: [10.27014/d.cnki.gdnau.2019.002039](https://doi.org/10.27014/d.cnki.gdnau.2019.002039).
- [29] 张昊. 水声目标辐射噪声信号增强与特征辨识技术研究[D]. [硕士学位论文], 东南大学, 2021. doi: [10.27014/d.cnki.gdnau.2021.003088](https://doi.org/10.27014/d.cnki.gdnau.2021.003088).
- ZHANG Hao. Research on enhancement and feature recognition technology of radiated noise signal of underwater acoustic target[D]. [Master dissertation], Southeast University, 2021. doi: [10.27014/d.cnki.gdnau.2021.003088](https://doi.org/10.27014/d.cnki.gdnau.2021.003088).
- [30] 倪俊帅, 赵梅, 胡长青. 基于深度学习的舰船辐射噪声多特征融合分类[J]. 声学技术, 2020, 39(3): 366–371. doi: [10.16300/j.cnki.1000-3630.2020.03.019](https://doi.org/10.16300/j.cnki.1000-3630.2020.03.019).
- NI Junshuai, ZHAO Mei, and HU Changqing. Multi-feature fusion classification of ship radiated noise based on deep learning[J]. *Technical Acoustics*, 2020, 39(3): 366–371. doi: [10.16300/j.cnki.1000-3630.2020.03.019](https://doi.org/10.16300/j.cnki.1000-3630.2020.03.019).
- [31] 李琛, 黄兆琼, 徐及, 等. 使用深度学习的多通道水下目标识别[J]. 声学学报, 2020, 45(4): 506–514. doi: [10.15949/j.cnki.0371-0025.2020.04.007](https://doi.org/10.15949/j.cnki.0371-0025.2020.04.007).
- LI Chen, HUANG Zhaoqiong, XU Ji, *et al.* Multi-channel underwater target recognition using deep learning[J]. *Acta Acustica*, 2020, 45(4): 506–514. doi: [10.15949/j.cnki.0371-0025.2020.04.007](https://doi.org/10.15949/j.cnki.0371-0025.2020.04.007).
- [32] FILHO W S, DE SEIXAS J M, and DE MOURA N N. Preprocessing passive sonar signals for neural classification[J]. *IET Radar, Sonar & Navigation*, 2011, 5(6): 605–612. doi: [10.1049/iet-rsn.2010.0157](https://doi.org/10.1049/iet-rsn.2010.0157).
- [33] DE B A, BARROS R E, and EBECKEN N F F. Development of a ship classification method based on Convolutional neural network and Cyclostationarity Analysis[J]. *Mechanical Systems and Signal Processing*, 2022, 170: 108778. doi: [10.1016/j.ymsp.2021.108778](https://doi.org/10.1016/j.ymsp.2021.108778).
- [34] MOSAVI M R, KHISHE M, NASERI M J, *et al.* Multi-layer perceptron neural network utilizing adaptive best-mass gravitational search algorithm to classify sonar dataset[J]. *Archives of Acoustics*, 2019, 44(1): 137–151. doi: [10.24425/aoa.2019.126360](https://doi.org/10.24425/aoa.2019.126360).
- [35] JAHROMI M S, BAGHERI V, ROSTAMI H, *et al.* Feature extraction in fractional fourier domain for classification of passive sonar signals[J]. *Journal of Signal Processing Systems*, 2019, 91(5): 511–520. doi: [10.1007/s11265-018-1347-x](https://doi.org/10.1007/s11265-018-1347-x).
- [36] KHISHE M and MOSAVI M R. Classification of underwater acoustical dataset using neural network trained by Chimp Optimization Algorithm[J]. *Applied Acoustics*, 2020, 157: 107005. doi: [10.1016/j.apacoust.2019.107005](https://doi.org/10.1016/j.apacoust.2019.107005).
- [37] SAFFARI A, KHISHE M, and ZAHIRI S H. Fuzzy-ChOA: An improved chimp optimization algorithm for marine mammal classification using artificial neural network[J]. *Analog Integrated Circuits and Signal Processing*, 2022, 111(3): 403–417. doi: [10.1007/s10470-022-02014-1](https://doi.org/10.1007/s10470-022-02014-1).
- [38] BAQAR M and ZAIDI S S H. Performance evaluation of linear and multi-linear subspace learning techniques for object classification based on underwater acoustics[C]. 2017 14th International Bhurban Conference on Applied Sciences and Technology, Islamabad, Pakistan, 2017: 675–683. doi: [10.1109/IBCAST.2017.7868124](https://doi.org/10.1109/IBCAST.2017.7868124).
- [39] GONZÁLEZ-HERNÁNDEZ F R, SÁNCHEZ-FERNÁNDEZ L P, SUÁREZ-GUERRA S, *et al.* Marine mammal sound



- classification based on a parallel recognition model and octave analysis[J]. *Applied Acoustics*, 2017, 119: 17–28. doi: [10.1016/j.apacoust.2016.11.016](https://doi.org/10.1016/j.apacoust.2016.11.016).
- [40] ZHONG Ming, TORTEROTOT M, BRANCH T A, *et al.* Detecting, classifying, and counting blue whale calls with Siamese neural networks[J]. *The Journal of the Acoustical Society of America*, 2021, 149(5): 3086–3094. doi: [10.1121/10.0004828](https://doi.org/10.1121/10.0004828).
- [41] COLE A M. Automated open circuit scuba diver detection with low cost passive sonar and machine learning[D]. [Master dissertation], Massachusetts Institute of Technology, 2019. doi: [10.1575/1912/24459](https://doi.org/10.1575/1912/24459).
- [42] BIANCO M J, GERSTOFT P, TRAER J, *et al.* Machine learning in acoustics: Theory and applications[J]. *The Journal of the Acoustical Society of America*, 2019, 146(5): 3590–3628. doi: [10.1121/1.5133944](https://doi.org/10.1121/1.5133944).
- [43] 周飞燕, 金林鹏, 董军. 卷积神经网络研究综述[J]. *计算机学报*, 2017, 40(6): 1229–1251. doi: [10.11897/SP.J.1016.2017.01229](https://doi.org/10.11897/SP.J.1016.2017.01229).
- ZHOU Feiyan, JIN Linpeng, and DONG Jun. Review of convolutional neural network[J]. *Chinese Journal of Computers*, 2017, 40(6): 1229–1251. doi: [10.11897/SP.J.1016.2017.01229](https://doi.org/10.11897/SP.J.1016.2017.01229).
- [44] GRAVES A. Long short-term memory[M]. GRAVES A. Supervised Sequence Labelling with Recurrent Neural Networks. Berlin: Springer, 2012: 37–45.
- [45] GOODFELLOW I, BENGIO Y, and COURVILLE A. Deep Learning[M]. Cambridge: MIT Press, 2016: 243.
- [46] 张健. 基于深度学习的水下目标识别的研究[D]. [硕士论文], 电子科技大学, 2020.
- ZHANG Jian. Research on underwater target recognition based on deep learning[D]. [Master dissertation], University of Electronic Science and Technology of China, 2020.
- [47] 任晨曦, 王黎明, 韩星程, 等. 基于联合神经网络的水声目标识别方法[J]. *舰船科学技术*, 2022, 44(1): 136–141. doi: [10.3404/j.issn.1672-7649.2022.01.026](https://doi.org/10.3404/j.issn.1672-7649.2022.01.026).
- REN Chenxi, WANG Liming, HAN Xingcheng, *et al.* Underwater acoustic target recognition method based on joint neural network[J]. *Ship Science and Technology*, 2022, 44(1): 136–141. doi: [10.3404/j.issn.1672-7649.2022.01.026](https://doi.org/10.3404/j.issn.1672-7649.2022.01.026).
- [48] HAN Xingcheng, REN Chenxi, WANG Liming, *et al.* Underwater acoustic target recognition method based on a joint neural network[J]. *PloS One*, 2022, 17(4): e0266425. doi: [10.1371/journal.pone.0266425](https://doi.org/10.1371/journal.pone.0266425).
- [49] 曾赛, 杜选民. 水下目标多模态深度学习分类识别研究[J]. *应用声学*, 2019, 38(4): 589–595. doi: [10.11684/j.issn.1000-310X.2019.04.016](https://doi.org/10.11684/j.issn.1000-310X.2019.04.016).
- ZENG Sai and DU Xuanmin. Multimodal underwater target recognition method based on deep learning[J]. *Journal of Applied Acoustics*, 2019, 38(4): 589–595. doi: [10.11684/j.issn.1000-310X.2019.04.016](https://doi.org/10.11684/j.issn.1000-310X.2019.04.016).
- [50] 杨路飞, 章新华, 吴秉坤, 等. 基于MFCC特征的被动水声目标深度学习分类方法[J]. *舰船科学技术*, 2020, 42(10): 129–133. doi: [10.3404/j.issn.1672-7649.2020.10.025](https://doi.org/10.3404/j.issn.1672-7649.2020.10.025).
- YANG Lufei, ZHANG Xinhua, WU Bingkun, *et al.* Research on the classification method of passive acoustic target depth learning based on MFCC[J]. *Ship Science and Technology*, 2020, 42(10): 129–133. doi: [10.3404/j.issn.1672-7649.2020.10.025](https://doi.org/10.3404/j.issn.1672-7649.2020.10.025).
- [51] AL-BETAR M A, ALYASSERI Z A A, AWADALLAH M A, *et al.* Coronavirus herd immunity optimizer (CHIO)[J]. *Neural Computing and Applications*, 2021, 33(10): 5011–5042. doi: [10.21203/rs.3.rs-27214/v1](https://doi.org/10.21203/rs.3.rs-27214/v1).
- [52] 史广智, 胡均川. 舰船噪声调制谐波族结构特性理论分析[J]. *声学学报*, 2007, 32(1): 19–25. doi: [10.3321/j.issn:0371-0025.2007.01.004](https://doi.org/10.3321/j.issn:0371-0025.2007.01.004).
- SHI Guangzhi and HU Junchuan. Theoretical analysis of the structure law of ship radiated-noise demodulation spectrum harmonic clan feature[J]. *Acta Acustica*, 2007, 32(1): 19–25. doi: [10.3321/j.issn:0371-0025.2007.01.004](https://doi.org/10.3321/j.issn:0371-0025.2007.01.004).
- [53] 程玉胜, 张宝华, 高鑫, 等. 船舶辐射噪声解调谱相位耦合特性与应用[J]. *声学学报*, 2012, 37(1): 25–29. doi: [10.15949/j.cnki.0371-0025.2012.01.004](https://doi.org/10.15949/j.cnki.0371-0025.2012.01.004).
- CHENG Yusheng, ZHANG Baohua, GAO Xin, *et al.* Phase-coupling characteristics of ship radiated-noise demodulation spectrum and application[J]. *Acta Acustica*, 2012, 37(1): 25–29. doi: [10.15949/j.cnki.0371-0025.2012.01.004](https://doi.org/10.15949/j.cnki.0371-0025.2012.01.004).
- [54] 张奇, 张延厚, 贾书阳, 等. 一种脉冲分布噪声下DEMON谱提取方法[C]. 中国声学学会水声学会2021-2022年学术会议论文集, 青岛, 2021: 261–263. doi: [10.26914/c.cnkihy.2022.062057](https://doi.org/10.26914/c.cnkihy.2022.062057).
- ZHANG Qi, ZHANG Yanhou, JIA Shuyang, *et al.* A DEMON spectrum extraction method under pulse distribution noise[C]. Hydroacoustics Branch of Chinese Acoustical Society, Qingdao, China, 2021: 261–263. doi: [10.26914/c.cnkihy.2022.062057](https://doi.org/10.26914/c.cnkihy.2022.062057).
- [55] 刘振, 邱家兴, 程玉胜. 深度神经网络在螺旋桨叶片数识别中的应用[J]. *声学技术*, 2019, 38(4): 459–463. doi: [10.16300/j.cnki.1000-3630.2019.04.017](https://doi.org/10.16300/j.cnki.1000-3630.2019.04.017).
- LIU Zhen, QIU Jiaying, and CHENG Yusheng. Application of deep neural network in blade-number recognition of ship propeller[J]. *Technical Acoustics*, 2019, 38(4): 459–463. doi: [10.16300/j.cnki.1000-3630.2019.04.017](https://doi.org/10.16300/j.cnki.1000-3630.2019.04.017).
- [56] LU Jiamin, SONG Sanming, HU Zhiqiang, *et al.* Fundamental frequency detection of underwater acoustic target using DEMON spectrum and CNN network[C]. 2020 3rd International Conference on Unmanned Systems, Harbin, China, 2020: 778–784. doi: [10.1109/ICUS50048](https://doi.org/10.1109/ICUS50048).

- 2020.9274949.
- [57] GONZALEZ S and BROOKES M. A pitch estimation filter robust to high levels of noise (PEFAC)[C]. 19th European Signal Processing Conference, Barcelona, Spain, 2011: 451–455.
- [58] 卢佳敏, 宋三明, 景严, 等. 基于DEMON谱和LSTM网络的水下运动目标噪声基频检测[J]. 应用声学, 2021, 40(5): 745–753. doi: [10.11684/j.issn.1000-310X.2021.05.013](https://doi.org/10.11684/j.issn.1000-310X.2021.05.013).  
LU Jiamin, SONG Sanming, JING Yan, *et al.* Fundamental frequency detection of underwater target noises using DEMON spectrum and LSTM network[J]. *Journal of Applied Acoustics*, 2021, 40(5): 745–753. doi: [10.11684/j.issn.1000-310X.2021.05.013](https://doi.org/10.11684/j.issn.1000-310X.2021.05.013).
- [59] 白敬贤, 高天德, 夏润鹏. 基于DEMON谱信息提取算法的目标识别方法研究[J]. 声学技术, 2017, 36(1): 88–92. doi: [10.16300/j.cnki.1000-3630.2017.01.016](https://doi.org/10.16300/j.cnki.1000-3630.2017.01.016).  
BAI Jingxian, GAO Tiande, and XIA Rungeng. Target recognition based on the information extraction algorithm of DEMON spectrum[J]. *Technical Acoustics*, 2017, 36(1): 88–92. doi: [10.16300/j.cnki.1000-3630.2017.01.016](https://doi.org/10.16300/j.cnki.1000-3630.2017.01.016).
- [60] 杨日杰, 郑晓庆, 韩建辉, 等. 基于序列匹配的螺旋桨轴频自动提取方法[J]. 振动与冲击, 2018, 37(16): 57–61. doi: [10.13465/j.cnki.jvs.2018.16.009](https://doi.org/10.13465/j.cnki.jvs.2018.16.009).  
YANG Rijie, ZHENG Xiaoqing, HAN Jianhui, *et al.* An automatic extraction method of propeller shaft frequency based on sequence matching[J]. *Journal of Vibration and Shock*, 2018, 37(16): 57–61. doi: [10.13465/j.cnki.jvs.2018.16.009](https://doi.org/10.13465/j.cnki.jvs.2018.16.009).
- [61] 王小宇, 李凡, 曹琳, 等. 改进的卷积神经网络实现端到端的水下目标自动识别[J]. 信号处理, 2020, 36(6): 958–965. doi: [10.16798/j.issn.1003-0530.2020.06.018](https://doi.org/10.16798/j.issn.1003-0530.2020.06.018).  
WANG Xiaoyu, LI Fan, CAO Lin, *et al.* End to end underwater targets recognition using the modified convolutional neural network[J]. *Journal of Signal Processing*, 2020, 36(6): 958–965. doi: [10.16798/j.issn.1003-0530.2020.06.018](https://doi.org/10.16798/j.issn.1003-0530.2020.06.018).
- [62] LIN Min, CHEN Qiang, and YAN Shuicheng. Network in network[EB/OL]. <https://arxiv.org/abs/1312.4400#>, 2013.
- [63] IOFFE S and SZEGEDY C. Batch normalization: Accelerating deep network training by reducing internal covariate shift[C]. Proceedings of the 32nd International Conference on International Conference on Machine Learning, Lille, France, 2015: 448–456.
- [64] 李悦, 马晓川, 王磊, 等. 非高斯环境下的深度学习脉冲信号去噪与重构[J]. 应用声学, 2021, 40(1): 131–141. doi: [10.11684/j.issn.1000-310X.2021.01.016](https://doi.org/10.11684/j.issn.1000-310X.2021.01.016).  
LI Yue, MA Xiaochuan, WANG Lei, *et al.* Using deep learning to de-noise and reconstruct pulse signals in non-Gaussian environment[J]. *Journal of Applied Acoustics*, 2021, 40(1): 131–141. doi: [10.11684/j.issn.1000-310X.2021.01.016](https://doi.org/10.11684/j.issn.1000-310X.2021.01.016).
- [65] KE Xiaoquan, YUAN Fei, and CHENG En. Integrated optimization of underwater acoustic ship-radiated noise recognition based on two-dimensional feature fusion[J]. *Applied Acoustics*, 2020, 159: 107057. doi: [10.1016/j.apacoust.2019.107057](https://doi.org/10.1016/j.apacoust.2019.107057).
- [66] WU Zonghan, PAN Shirui, CHEN Fengwen, *et al.* A comprehensive survey on graph neural networks[J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2021, 32(1): 4–24. doi: [10.1109/TNNLS.2020.2978386](https://doi.org/10.1109/TNNLS.2020.2978386).
- [67] FEDUS W, ZOPH B, and SHAZEER N. Switch transformers: Scaling to trillion parameter models with simple and efficient sparsity[J]. *The Journal of Machine Learning Research*, 2023, 23(1): 120.
- 张 奇: 男, 博士生, 研究方向为水声目标识别、深度学习。  
笪良龙: 男, 博士, 教授, 研究方向为海洋环境效应。  
王 超: 男, 博士, 副研究员, 研究方向为水声信号处理。  
张延厚: 男, 博士生, 研究方向为信息融合。  
嵇江浩: 男, 博士, 助理研究员, 研究方向为水声信号处理。

责任编辑: 于 青