

◇ 研究报告 ◇

# 基于 DEMON 谱和 LSTM 网络的水下运动目标 噪声基频检测\*

卢佳敏<sup>1,2,3</sup> 宋三明<sup>1,2†</sup> 景 严<sup>1,2</sup> 张 瑶<sup>1,2</sup> 谷 浪<sup>1,2</sup>  
鲁 帆<sup>4</sup> 胡志强<sup>1,2</sup> 李 硕<sup>1,2</sup>

(1 中国科学院沈阳自动化研究所 机器人学国家重点实验室 沈阳 110016)

(2 中国科学院机器人与智能制造创新研究院 沈阳 110169)

(3 中国科学院大学 北京 100049)

(4 中国科学院声学研究所 北京 100190)

**摘要:** 传统的船舶辐射噪声基频检测方法不仅依赖大量的先验知识,而且对背景噪声非常敏感。为了提高目标识别的稳定性和精确性,该文提出了一种基于深度神经网络的基频检测算法。首先从多通道水听器信号中提取 DEMON 谱,然后将二维谱特征矩阵输入由卷积神经网络和长短时记忆网络构成的级联网络,最后通过稠密层输出实现对基频的估计。仿真和外场实验结果初步表明:深度网络能够实现无先验知识和不同信噪比条件下的基频检测,具有良好的泛化性能。长短时记忆网络能够高效地从时序 DEMON 谱中提取统计特征,提高基频估计精度。输入信号的时间长短会影响网络的检测精度,更长时间的信号能够获得更好的检测结果。

**关键词:** 基频;深度网络;长短时记忆网络;卷积神经网络;水听器阵列;水下目标噪声

**中图分类号:** TB566 **文献标识码:** A **文章编号:** 1000-310X(2021)05-0745-09

**DOI:** 10.11684/j.issn.1000-310X.2021.05.013

## Fundamental frequency detection of underwater target noises using DEMON spectrum and LSTM network

LU Jiamin<sup>1,2,3</sup> SONG Sanming<sup>1,2</sup> JING Yan<sup>1,2</sup> ZHANG Yao<sup>1,2</sup> GU Lang<sup>1,2</sup>  
LU Fan<sup>4</sup> HU Zhiqiang<sup>1,2</sup> LI Shuo<sup>1,2</sup>

(1 State Key Laboratory of Robotics, Shenyang Institute of Automation, Chinese Academy of Sciences, Shenyang 110016, China)

(2 Institutes for Robotics and Intelligent Manufacturing, Chinese Academy of Sciences, Shenyang 110169, China)

(3 University of Chinese Academy of Sciences, Beijing 100049, China)

(4 Institute of Acoustics, Chinese Academy of Sciences, Beijing 100190, China)

**Abstract:** The traditional fundamental frequency(F0) detection methods not only rely on prior knowledge, but also are very sensitive to ambient noises. In this paper, a fundamental frequency detection algorithm based on deep neural network is proposed to improve the stability and accuracy of target recognition. The DEMON

2021-03-12 收稿; 2021-04-07 定稿

\*国家自然科学基金项目(61973297), 中国科学院先导专项子课题(XDC03060105), 中国科学院青年创新促进会课题(2020209), 机器人学国家重点实验室课题(2017-Z010)

作者简介: 卢佳敏(1996-), 女, 福建福州人, 硕士研究生, 研究方向: 水下噪声信号处理和水下机器人自主环境感知。

†通信作者 E-mail: songsanming@sia.cn

spectrum matrix, which is composed of spectral vector extracted from each single hydrophone signal, is directly fed into the cascaded network made up of convolutional neural networks (CNN) and long short-term memory (LSTM) networks. Then, with the one-hot vector from the final dense layer, the fundamental frequency is estimated. The following conclusions can be drawn from computer simulation and field experiments: The deep learning-based method works well when no prior knowledge is assumed or signal to noise ratio varies, having good generalization performance. LSTM network can effectively extract the statistical characteristics from the DEMON spectrum sequence and improve the accuracy of the F0 estimation. The detection precision depends on the input signal length, and a better detection result could be obtained when a longer signal is available.

**Keywords:** Fundamental frequency; Deep learning; Long short-term memory; Convolutional neural networks; Hydrophone array; Underwater target noises

## 0 引言

水中机动目标在航行时会辐射出具有特殊频谱分布的噪声,因此可以通过噪声分析来估计航行器的吨位、桨叶数等物理特征和轴频、航速等运动状态<sup>[1-2]</sup>,进而实现合作或对抗场景下的态势感知。轴频是螺旋桨转速的最直接反映<sup>[3]</sup>,对水下目标探测具有重要的意义,本文将重点研究如何从水听器阵列接收到的噪声中估计轴频信息。

轴频也称为基频,在数值上等于目标螺旋桨主轴的转动频率<sup>[4]</sup>,常用 DEMON 谱分析法提取调制谱特征,以获取基频及其谐波在内的低频线谱,进而检测目标的基频、叶频或桨叶数目。文献[5-6]评估了不同解调分量得到的 DEMON 谱的鉴别能力。文献[7-11]研究了 DEMON 谱净化与增强,提高了线谱的信噪比与基频估计的精度。其中,文献[7]基于线谱之间的倍频关系,提出基于最大公约数的基频提取方法,文献[12-13]对此进行了拓展并提出相应的改进方法。这些算法需要人为设置参数进行特征提取和基频估计,对信噪比较敏感。实际接收到的目标噪声信号,受到不同海洋波导传输过程的乘性作用和加性干扰,其信噪比变化大<sup>[14-15]</sup>。

深度学习具有较强的泛化能力,近年来逐渐被研究人员引入水下目标识别领域。文献[16]通过提取 DEMON 谱中各谐波之间的结构特征关系建立模板库,根据模板库对深度神经网络(Deep neural networks, DNN)进行训练,并识别螺旋桨的叶片数目。文献[17]将梅尔频率倒谱系数(Mel-frequency cepstrum coefficient, MFCC)和频谱包络特征输入卷积神经网络(Convolutional neural networks, CNN)和深度置信网络(Deep belief networks, DBN),对船舶目标进行分类。文献[18]提出一个类似方案,将线性预测倒谱系数(Linear pre-

dictive cepstral coefficient, LPCC)和 MFCC 串联后输入 DBN 网络。文献[19]将梅尔频率特征分别输入支持向量机(Support vector machine, SVM)和 VGGish 网络中,对生物、潜艇和船只3种目标噪声进行分类。文献[20]使用堆叠降噪自编码器(Stacked denoising auto encoder, SDAE)对原始信号数据进行降维和逐层特征提取,然后使用 SVM 和 BP 网络进行分类。文献[21]开展了基于长短时记忆(Long short-term memory, LSTM)网络的声学目标分类,发现时域、频域和 MFCC 三种特征具有相当的性能。随后,他们对 DEMON 谱和 MFCC 等特征进行融合,提高了 LSTM 网络的分类准确率<sup>[22]</sup>。

大多数现有水下目标分类算法首先通过波束形成将水听器阵列信号约减为一维信号,然后提取频域特征输入深度网络进行训练。本文开展基于深度学习的基频估计,但是重点考虑如下两个问题:

(1) 在特征提取时,绝大多数研究都采用了 MFCC 特征。MFCC 反映能量在不同频率区间的分布,本质上是对噪声信号频谱包络的一种描述,关注的是连续谱而非调制谱特征。

(2) 在波束形成时,会丢失很多细节信息,且会引入新的阵形估计误差。

对于问题(1),由于调制谱是由基频及其谐波决定的,而一般采用 DEMON 谱来提取调制谱,所以本文在估计基频时候选择 DEMON 谱作为网络的输入。需要注意的是,在目标航速很低时噪声的调制效应很弱,因此调制谱分析更适合分析非低速航行下的目标噪声。

对于问题(2),一种更加合理的选择是将水听器阵列采集的多通道信号直接用于特征提取。然而,多通道数据不仅会导致特征数急速增长,而且还会引入随机噪声的谱特征。本文拟采用 CNN 网

络对多通道特征进行二次特征提取,实现特征降维和净化。最后,将多个时间窗对应的二次特征输入LSTM网络实现基频估计。

### 1 算法流程

本文提出的基频检测算法流程如图1所示。首先,对每个通道的噪声信号作加窗分帧处理,在每个时间窗内提取DEMON谱特征;然后,利用梳状滤波器对DEMON谱特征中包含的倍频关系进行增强;最后,将同一时间窗内所有通道特征组成二维矩阵经卷积层处理后输入LSTM网络。

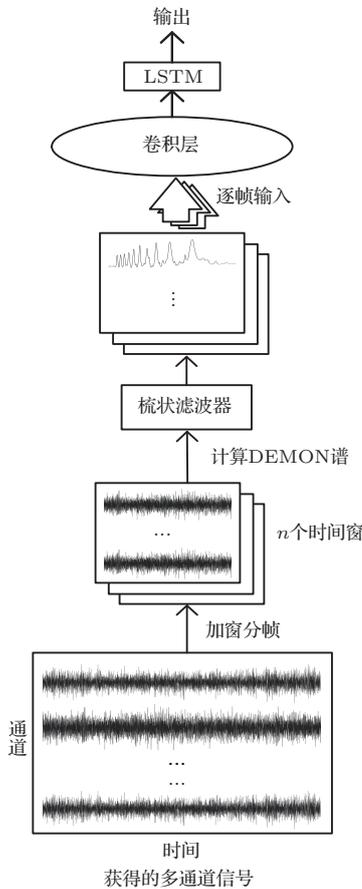


图1 基频检测算法流程图

Fig. 1 Flowchart of the proposed F0 detection algorithm

本文采用分类网络来估计基频。把基频取值范围分成几个频段,每个频段当作一个类别,将基频估计问题转化为分类问题。需要说明的是,也可以采用回归网络进行基频估计。但是,在很多研究领域中,回归网络或者分类网络在性能上并不存在本质的区别<sup>[23-24]</sup>。下面将详细介绍算法各个步骤的具

体实现方法。

#### 1.1 DEMON 谱

DEMON 谱的计算流程如图2所示。



图2 DEMON 谱计算流程图

Fig. 2 The DEMON spectrum extraction procedure

调制信号对不同频带噪声的信号调制强度是不同的<sup>[7]</sup>。可以将信号的频谱划分成多个频段,然后将不同频段的线谱进行融合来提高线谱质量。具体包括如下几个步骤:(1)在高频端选取若干个子带,对原始噪声信号进行时域带通滤波;(1)对每个子带进行平方解调,得到时域包络;(2)对包络做快速傅里叶变换(Fast Fourier transform, FFT)得到频谱,并利用双向 $\alpha$ 滤波器去除趋势项得到各个子带的线谱;(3)根据线谱根数将各子带的线谱进行加权融合,得到最终的线谱特征。考虑到水中运动目标的实际工况,后面实验中将基频的搜索范围限制在5~10 Hz之间。而且,由于DEMON谱的线谱阶数不会太高<sup>[7]</sup>,本文以10阶为限,仅保留100 Hz以下的DEMON谱。

#### 1.2 梳状滤波器

受到不稳定调制的影响,线谱之间的倍频关系可能会减弱或消失。本文采用文献[25]提出的梳状滤波器对倍频关系进行增强。梳状滤波器由许多按等频率间隔排列的通带和阻带组成,能够很好地保留信号中的倍频关系。

对于理想的DEMON谱,其对数频域上基频及倍频的分布可以表示为

$$Y(q) = \sum_{k=1}^K b_k \delta(q - \log k - \log f_0) + N(q), \quad (1)$$

其中,  $b$  为线谱幅度,  $k$  表示第  $k$  个谐波,  $f_0$  为基频,  $N(q)$  为加性噪声。

理想状态下,各次谐波线谱之间遵守严格的倍数关系。当梳状滤波器  $h(q)$  的第一个通带移动至某个线谱频率处时,其余各通带会落在该频率对应的各次谐波上,对各个谐波进行累积。在第一个通带位于基频  $f_0$  处时达到峰值。相应的理想滤波器  $h(q)$  可以用式(2)描述:

$$h(q) = \sum_{k=1}^K \delta(q - \log k). \quad (2)$$

但在实际中,基频及谐频之间的倍频关系并不严格,如谐频位置存在偏差或存在一定的谱峰宽度。因此,DEMON谱不能直接用式(1)描述,相应的也不能用式(2)的滤波器进行滤波。一种考虑了偏差和谱峰宽度的滤波器为

$$h(q) = [g(q)]_+, \quad (3)$$

其中,  $g(q)$  为

$$g(q) = \begin{cases} \frac{1}{\gamma - \cos(2\pi\epsilon^q)} - \beta, & \log 0.5 < q < \log(K + 0.5), \\ 0, & \text{其他,} \end{cases} \quad (4)$$

参数  $K$  是谐波个数,  $\gamma$  代表谱峰宽度。在后面的实验中  $K$  取 10,  $\gamma$  取 1.8, 取  $\beta$ , 使

$$\int \left( \frac{1}{\gamma - \cos(2\pi\epsilon^q)} - \beta \right) dq = 0.$$

利用梳状滤波器进行滤波时,首先将特征信号映射至对数域,  $q = \log f$ , 然后将 DEMON 谱与滤波器做卷积得到增强后的线谱特征:

$$Y'(q) = Y(q) \otimes g(-q). \quad (5)$$

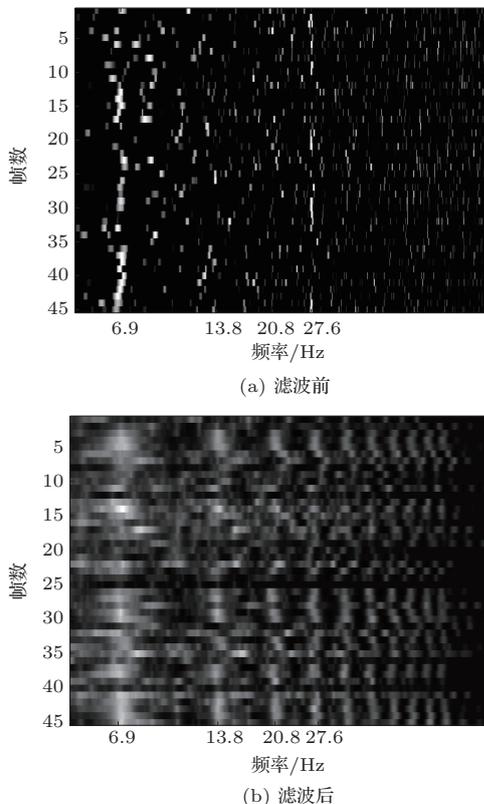


图3 梳状滤波器对 DEMON 谱的增强效果

Fig. 3 Enhancement to DEMON spectrum by the comb filtering

梳状滤波前后的 DEMON 谱如图 3 所示,滤波后谐波结构(线谱间的倍频关系)明显增强。

### 1.3 深度分类网络

提取 DEMON 谱特征后,可以通过训练分类网络来估计基频。作者在文献[26]中提出过一种基于 CNN 网络的检测方法,利用小波变换对 DEMON 谱特征去噪,然后将净化谱特征输入 CNN 网络来估计基频。但是,去噪可能会造成谱特征信息的丢失,而且单纯采用 CNN 网络需要对较长的时间序列进行融合,影响实时性。本文去掉了去噪步骤,同时在 CNN 网络上增加了 LSTM 网络,以期利用 LSTM 网络的时序建模能力去捕获低信噪比条件下基频及其倍频的统计特性,提高基频估计的稳定性和实时性。

LSTM 是循环神经网络 (Recurrent neural network, RNN) 的一种改进结构,具有 RNN 处理时间序列信号能力的同时解决了 RNN 存在的长时依赖问题,被广泛用于处理各种语音任务。LSTM 基本单元结构如图 4 所示。其中,  $x_t$  为各时间步输入 LSTM 单元的特征,  $h_t$  是各时间步 LSTM 单元的输出, sig 为 sigmoid 函数。

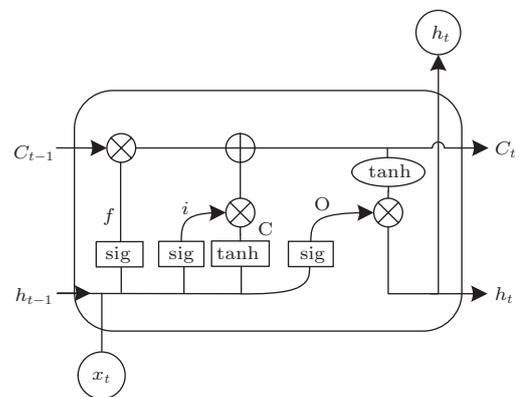


图4 LSTM 基本单元

Fig. 4 Cell of LSTM

网络的整体结构如图 5 所示,包含一个拥有 8 个  $3 \times 3$  卷积核的卷积层、一个卷积核尺寸为  $2 \times 2$  的池化层、一个含有 256 个神经元的 LSTM 层以及一个大小为 128 的隐藏层和大小为 25 的 softmax 分类层。注意,输出是 one-hot 向量,当基频的真实值落在相应频率范围内时,该类值为 1,其余为 0。损失函数采用交叉熵函数,网络使用 mini-batch 进行训练。其中卷积层对输入的多通道 DEMON 谱特征进

行降维,LSTM层用于提取输入信号的时序特征,随后输入分类网络进行分类。

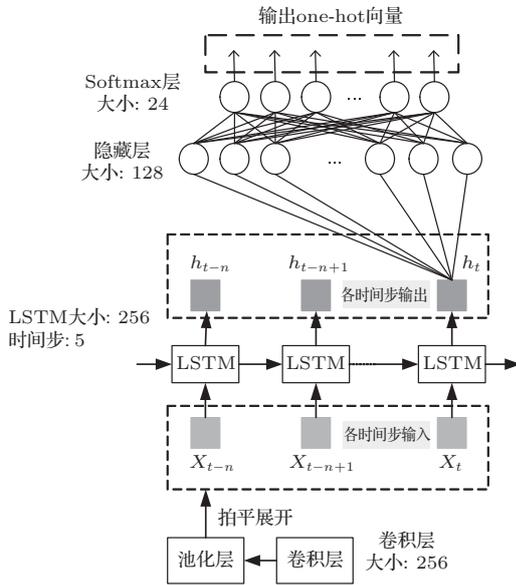


图5 网络结构示意图

Fig. 5 The proposed deep-learning network structure

## 2 数据仿真

训练深度网络需要大量标定好的数据。但是,真实舰船噪声数据往往比较稀缺,而且实测数据标定是一项繁琐且容易出错的工作。因此本文首先使用仿真数据对网络进行训练,在湖试中利用实测舰船噪声数据对预训练好的网络进行微调后,再进行基频估计。本节介绍舰船噪声信号的仿真方法。

对于仿真舰船噪声信号来说,周期性局部平稳过程<sup>[14]</sup>和准周期性随机声脉冲序列模型<sup>[15]</sup>是描述其时域信号较为客观接近实际的两种数理模型。准周期性随机声脉冲序列模型可以模拟以不同噪声源为主导的不同工况下的目标辐射噪声信号,其脉冲形状和脉冲出现的周期性程度可取不同值,使其具有不同的宽带连续谱形状和不同线谱幅度、调制深度,具有良好的普适性,因此本文根据文献<sup>[15]</sup>对目标噪声信号进行仿真。

### 2.1 舰船噪声准周期性随机声脉冲序列模型

舰船噪声的频谱可以看作由连续谱和线谱叠加组成,噪声的时域波形可以看作由准周期性随机脉冲序列构成,其中第 $n$ 个声脉冲的出现时刻(脉冲前沿)为

$$t_n^S = nT + \Delta_n, \quad (6)$$

其中, $T$ 为声脉冲序列之间的平均间隔, $\Delta_n$ 为脉冲序列在周期点附近的随机摆幅,服从均值为零、方差为 $\delta\Delta$ 的正态分布。

准周期性随机声脉冲序列指出噪声辐射线谱相对于连续谱的幅值仅取决于脉冲的随机摆幅,脉冲序列的准周期性越差,则线谱幅值越低。对于不同工况的船舶,通过选择周期性随机序列脉冲的各个参数可以获得不同频率和幅度的线谱,以及谱峰频率不同、连续谱形状和衰减规律不同的功率谱<sup>[15]</sup>。

以声脉冲形状为指数衰减型为例给出具体实现过程。第 $n$ 个声脉冲的形状为<sup>[15]</sup>

$$\zeta_n^t = \begin{cases} A_n u\left(\frac{t-t_n^S}{\tau_n}\right), & t_n^S \leq t \leq t_n^E, \\ 0, & \text{其他,} \end{cases} \quad (7)$$

其中, $t_n^S$ 为脉冲起始时刻, $t_n^E$ 为脉冲结束时刻;指数衰减函数为

$$u(t) = e^{-\gamma t}, \quad (8)$$

其中, $\gamma$ 为衰减系数, $\tau$ 是脉冲持续时间。脉冲幅度分布遵循式(9):

$$A_n \sim \frac{x}{\mu^2} e^{-\frac{x}{\mu^2}}. \quad (9)$$

每个脉冲的起始时间在桨叶周期附近服从高斯分布:

$$t_n^S \sim N(nT, \Delta), \quad t_n^E = (n+1)T. \quad (10)$$

图6给出一段仿真信号的频谱。

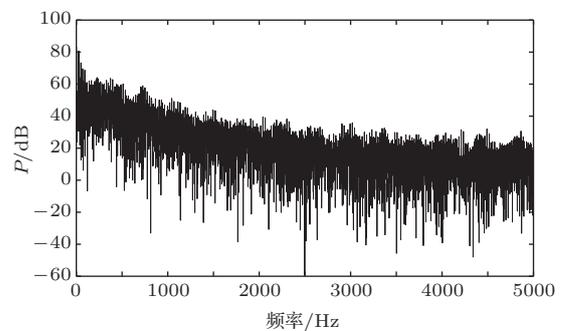


图6 一段基频为5 Hz的噪声仿真信号的频谱

Fig. 6 The frequency spectrum of a synthesized ship-radiated noise signal with the shaft frequency being 5 Hz

## 2.2 阵列信号仿真

海洋环境噪声和水声信道传输过程会降低接收点处接收信号的信噪比,本节介绍如何仿真受海洋环境影响的接收点信号。

本文主要面向接收点与目标之间距离为中近距离的情形,侧重考虑随距离变化的声能平均传播损失、海底与海面反射引起的相干多途干扰以及海洋环境噪声加性干扰。采用基于射线声学的bellhop模型对传播损失和多途干扰进行仿真,声源深度10 m,接收6元阵,接收深度10 m,为垂直线阵,接收阵水平距离2000 m,海水密度1.0 g/cm<sup>3</sup>,海底声速1680 m/s,海底密度1.8 g/cm<sup>3</sup>,吸收系数0.6 dB/λ。声速剖面分别选用等声速梯度剖面、负声速梯度剖面、含有温跃层的浅海声速剖面、深海Munk声速剖面等,水深分别选用50 m、500 m、5000 m,来模拟不同环境下的相干多途传播信道。声源使用2.1节中合成的仿真信号。海洋环境噪声使用不同信噪比高斯白噪声模拟。

合成的仿真噪声数据与真实工况下的噪声并不完全一致,但可以使用仿真数据完成网络训练后,再利用真实数据对网络进行微调来弥补二者之间的差别。

## 3 实验结果及分析

本文设计了4个实验来测试本文提出的基频检测算法。

(1) 分别采用本文提出的深度学习算法(数据驱动)和传统的模型驱动方法对仿真噪声信号进行基频检测,以测试数据驱动方法在抗噪性和适应性上的提升。

(2) 本文在文献[26]的基础上,去掉了去噪步骤,增加了LSTM网络。为此,本文设计了一组对比实验来比较二者在准确性和实时性上的性能差异。

(3) 为了测试LSTM网络在提取时序谱特征上的可行性,分别采用CNN和CNN+LSTM网络开展时序基频检测。实验中保持两者卷积层部分的结构相同。

(4) 分别采用本文提出的深度学习算法(数据驱动)和传统的模型驱动方法对外场实测噪声数据进行基频检测,以测试本文方法在实测数据上的检测性能。

在仿真数据准备阶段,总共生成了长度大约

为12 h的数据,采样频率为10 kHz,数据的基频范围在5~10 Hz之间,按0.2 Hz的间隔划分为25个类别。估计的基频值定义为预测类别的中点。例如:划分到5~5.2 Hz类别的基频预测值为5.1 Hz。仿真的噪声数据中桨叶数的取值为3~7叶。此外,本文将80%的样本作为训练数据集,剩余20%的数据作为测试样本集。评价指标包括: $P$ ,定义为误差小于0.1 Hz的预测值比例,计算公式为 $P = (N - N_e)/N \times 100\%$ ,其中 $N_e$ 为预测误差超过0.1 Hz的数据个数, $N$ 为总数据量;平均误差 $\bar{E} = \frac{1}{N} \sum_{i=1}^N |f_{ig} - f_{ip}|$ ,其中 $f_{ig}$ 表示第 $i$ 段数据的真实基频值, $f_{ip}$ 表示第 $i$ 段数据的预测基频值。

### 3.1 数据驱动方法与模型驱动方法

传统检测方法存在对先验知识的依赖,尤其DEMON谱提取过程和倍频统计步骤需要人工设定很多参数。本文方法简化预处理过程,并利用神经网络建模的方式降低检测算法对先验知识的依赖。为了比较两类方法的性能,本文选取文献[12]和文献[13]中提出的两种基于DEMON谱的基频提取算法作为对比。这两种方法计算基频时,首先利用波束形成将阵列信号约减为一维信号,再进行计算。

表1给出了两种不同信噪比(Signal to noise ratio, SNR)背景噪声强度下对仿真数据进行基频检测的结果,两种模型驱动算法的检测性能相当,而本文方法在误差小于0.1 Hz的预测值比例 $P$ 上有较大提升,且平均误差更小。相对于传统方法从DEMON谱提取人工特征然后计算基频而言,基于深度网络的方法提取DEMON谱的高层特征再进行基频估计,能够降低噪声的影响,从而改善检测结果。而传统方法提取的浅层特征对噪声线谱的抗干扰能力相对较弱,因此对于SNR较低、线谱质量较差的DEMON谱,本文方法能够取得较常规方法更好的检测结果。但是,当DEMON谱质量较好,干扰线谱很少的时候,常规方法也能获得比较理想的检测结果。

### 3.2 本文算法与文献[26]的检测结果对比

文献[26]提出了一种基于CNN网络的基频检测方法。而本文优化了网络结构,去掉了DEMON谱去噪和多帧融合步骤,使得特征提取部分更加简单。两种方法的检测结果对比如表2所示。

表1 本文方法与传统方法的对比结果

Table 1 Comparison between classic methods and the proposed method

	$P/\%$		平均误差/Hz	
	SNR = 3 dB	SNR = -2 dB	SNR = 3 dB	SNR = -2 dB
CNN+LSTM	84.4	73.9	0.23	0.41
改进的最大公约数 <sup>[12]</sup>	74.3	62.6	0.62	1.04
序列匹配法 <sup>[13]</sup>	76.7	64.5	0.58	0.65

表2 与文献[26]进行对比

Table 2 Comparison with the method in Ref. [26]

	每个输出对应的输入 信号时长/s	$P/\%$ (SNR = 3 dB)
	10	77.4
	3	73.8
CNN+LSTM	6	84.4
	10	86.5

从表2可以看出,当信号输入时长上由20 s下降至6 s时,本文的检测算法几乎达到同样的准确率;而输入时长增加至10 s时,误差小于0.1 Hz的预测值比例 $P$ 会略高于CNN网络。因此,尽管简化了预处理程序,但是LSTM网络却能够以较短的信号输入时长获得大致相同的 $P$ 值;从不同长时的检测结果能够看出,本文的方法在时效性上有了较大的提升。对于本文提出的算法,对比时长为3 s和6 s的实验结果,可以发现时长较长的信号能够提供更多稳定的检测结果。由于计算代价和信号长度成正比,这也提示我们可以通过调整输入信号的时长来满足不同任务对准确率和实时性的不同需求。

### 3.3 CNN和CNN+LSTM的检测结果对比

为了验证LSTM网络从时序DEMON谱中提取统计特征的能力,分别采用CNN和CNN+LSTM网络进行基频检测。表3列出了两种网络结构在不同的输入信号长度上的实验结果。CNN的输入为一帧语音信号计算的DEMON谱,添加LSTM层的网络每输入5个时间帧输出一个检测结果,其中帧间重叠为50%。

当帧长为1 s时,CNN网络的 $P$ 值在53%左右,添加LSTM结构后增加至74%;当帧长增加至2 s

时,前者的 $P$ 值上升至75%左右,而后者提高至84%左右。由实验结果能够看出,随着帧长的增加,输入网络的特征信息将增多,基频的检测精度上升;而在相同的帧长情况下,添加LSTM层后精度有较大的提升,表明LSTM网络能够较大幅度提升基频检测的准确率。

表3 CNN与CNN+LSTM的实验结果对比

Table 3 Comparison between CNN and CNN + LSTM

	每帧信号时长/s	$P/\%$ (SNR = 3 dB)
	2	74.9
CNN+LSTM	1	73.8
	2	84.4

### 3.4 外场实验数据分析

分别利用湖试数据和海试数据对所提算法进行测试。其中,湖试数据于千岛湖外场实验中获得,目标船只吨位为几十吨级,接收距离约1 km,共采集216段数据;海试数据于三亚外场实验中获得,为百吨级渔船噪声,接收距离约2 km,共采集162段数据。两次实验的接收阵均为24阵元的水平线阵,阵元距离1.875 m,位于水下15 m,采样频率为20 kHz。数据共包含5种不同转速,基频范围为5~10 Hz。对使用湖试数据进行基频估计前,先利用实测噪声数据对神经网络的权值进行微调学习。为了扩充数据集,本文采用加噪、平移、缩放等方式对实测数据进行数据集扩充。使用20%的数据作为训练集,其余80%数据作为测试集。

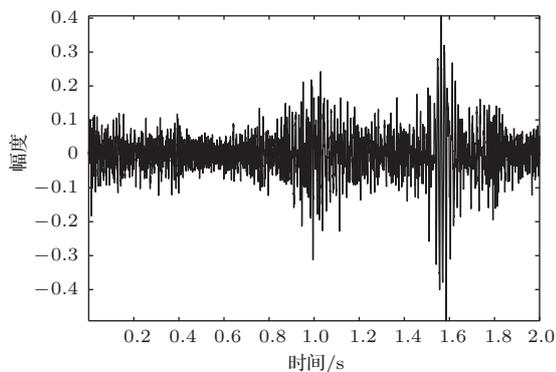
将无目标情况下采集到的背景噪声放大后叠加到原始信号上,叠加后信号的信噪比大约为0 dB,用于验证加噪情况下算法的性能。

表4给出了两组外场实验数据的基频检测结果。其中原始数据中目标距离较近,信号的信噪比相对较高。因此,本文方法及常规基频检测方法均获得了较好的检测结果,但本文方法的平均误差更小。对于加噪后的信号,本文方法较常规方法在精度上有较大的提升,并且平均误差在可接受的范围;常规方法的精度较低且常出现倍频半频错误的缘故导致平均误差很大。

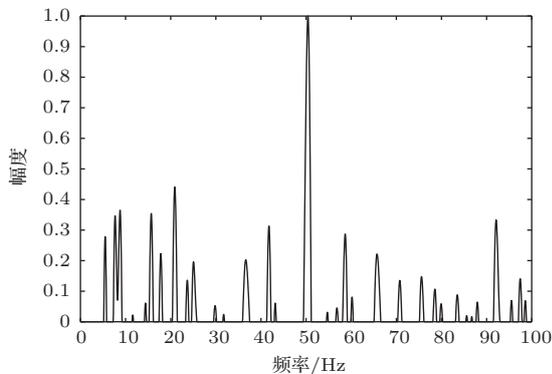
表4 本文方法与传统方法在实测数据上的检测性能对比

Table 4 Comparison between classic methods and the proposed method on the field dataset

	P/%			平均误差/Hz		
	海试	湖试	加噪信号	海试	湖试	加噪信号
CNN+LSTM	90.6	98.3	83.4	0.12	0.10	0.31
改进的最大公约数 <sup>[12]</sup>	81.3	85.5	62.3	0.31	0.25	1.53
序列匹配法 <sup>[13]</sup>	83.8	87.1	68.8	0.26	0.19	0.82



(a) 时域波形



(b) DEMON谱

图7 某段基频为8.4 Hz的渔船噪声数据

Fig. 7 A segment of sea trial data with shaft frequency being 8.4 Hz

图7给出了一段实测噪声的时域波形及其DEMON谱,从DEMON谱中能够看到较多的干扰线谱。表5给出了该段数据分别使用3种方法得到的检测结果。可以看到,常规方法由于受到干扰线谱的影响,估计出的基频值误差较大。

表5 图7的基频检测结果

Table 5 Fundamental frequency detection results of the signal segment in Fig.7

	基频/Hz
CNN+LSTM	8.5
改进的最大公约数法	8
序列匹配法	8.8
真实值	8.4

## 4 结论

本文提出了一种基于CNN和LSTM网络的水下目标辐射噪声基频检测方法。在网络输入上,与传统方法利用波束形成融合多通道信息不同,本文提取每个通道的DEMON谱并以二维矩阵的方式直接送入网络中,充分利用各个通道的信息;在特征提取部分,去掉了先前工作中的去噪和多帧融合步骤,在简化预处理步骤的同时提高算法的实时性,也摆脱了传统模型驱动算法在预处理过程中对先验知识的依赖。最后,将经过CNN+LSTM提取的特征输入全连接层实现基频估计。仿真实验和实验数据表明,本文方法能够较好适应模拟的不同环境和不同信噪比的目标噪声,具有良好泛化性能,可望实际推广;在一定信噪比下,能够从时序DEMON谱特征中估计出较为准确的基频值,可供进一步研究和应用参考。

需要注意的是,本文主要关注较高航速下目标噪声的基频检测。当航行速度很低时,无法从DEMON谱中提取到基频的调制信息,所以本文算法不适用于低速状态的航行器的基频检测。且限于实测数据规模的限制,本文使用仿真数据进行训练。仿真数据并不能完好地模拟实测数据,因此对实测数据的检测结果可能有一定的影响。后续工作中随着实测数据数量的增加,将尽可能利用实测数据对网络进行训练,进一步提高网络的性能。

## 参考文献

- [1] 程玉胜, 邱家兴, 刘振. 水声被动目标识别技术挑战与展望[J]. 应用声学, 2019, 38(3): 653-659.

- Cheng Yusheng, Qiu Jiaying, Liu Zhen. Challenges and prospects of underwater acoustic passive target recognition technology[J]. *Journal of Applied Acoustics*, 2019, 38(3): 653–659.
- [2] Nielsen R O. Cramer-rao lower bounds for sonar broadband modulation parameters[J]. *IEEE Journal of Oceanic Engineering*, 1999, 24(2): 285–290.
- [3] Lourens J G, du Preez J A. Passive sonar ml estimator for ship propeller speed[J]. *IEEE Journal of Oceanic Engineering*, 1998, 23(3): 448–453.
- [4] Song L, Yang D, Zhang H. Experimental study of underwater propeller low-frequency noise field[C]. *IEEE/OES China Ocean Acoustics*, 2016.
- [5] 杨宏, 李亚安, 李国辉. 基于集合经验模态分解的舰船辐射噪声能量分析[J]. *振动与冲击*, 2015, 34(16): 55–59.
- Yang Hong, Li Ya'an, Li Guohui. Energy analysis of ship radiated noise based on ensemble empirical mode decomposition[J]. *Journal of Vibration and Shock*, 2015, 34(16): 55–59.
- [6] Pollara A, Sutin A, Salloum H. Improvement of the detection of envelope modulation on noise (DEMON) and its application to small boats[C]. *OCEANS 2016 MTS/IEEE Monterey*. IEEE, 2016.
- [7] 殷敬伟, 惠俊英, 姚直象, 等. 基于DEMON线谱的轴频提取方法研究[J]. *应用声学*, 2005, 24(5): 369–374.
- Yin Jingwei, Hui Junying, Yao Zhixiang, et al. Extraction of shaft frequency based on the DEMON line spectrum[J]. *Journal of Applied Acoustics*, 2005, 24(5): 369–374.
- [8] 童峰, 陆信人, 方世良. 水声目标识别中一种轴频提取方法[J]. *声学学报*, 2004, 29(4): 398–402.
- Tong Feng, Lu Jiren, Fang Shiliang. A method to estimate the propeller shaft rate in underwater acoustic vessel classification[J]. *Acta Acustica*, 2004, 29(4): 398–402.
- [9] 骆国强, 尚金涛, 杨柳. 一种时频综合的DEMON谱融合方法[C]. 2016'中国西部声学学术交流会议论文集, 2016.
- [10] 高鑫, 程玉胜. 舰船螺旋桨轴频估计中线谱要素提取算法[J]. *应用声学*, 2010, 29(5): 443–448.
- Gao Xin, Cheng Yusheng. A line spectrum element extraction algorithm in ship-propeller-shaft frequency estimations[J]. *Journal of Applied Acoustics*, 2010, 29(5): 443–448.
- [11] Hanson D, Antoni J, Brown G, et al. Cyclostationarity for passive underwater detection of propeller craft: a development of DEMON processing[J]. *Proceedings of Acoustics*, 2008: 1–6.
- [12] 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.
- Bai Jingxian, Gao Tiande, Xia Rumpeng. Target recognition based on the information extraction algorithm of DEMON spectrum[J]. *Technical Acoustics*, 2017, 36(1): 88–92.
- [13] 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.
- 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.
- [14] 吴国清, 王美刚, 陈耀明. 水声波导中包络线谱强度数值预报[J]. *声学学报*, 2012, 37(3): 432–439.
- Wu Guoqing, Wang Meigang, Chen Yaoming. Numerical prediction of envelope line spectrum intensity in underwater acoustic waveguide[J]. *Acta Acustica*, 2012, 37(3): 432–439.
- [15] 蒋国健, 林建恒, 孙军平, 等. 海洋信道随机性引起的船舶噪声线谱衰减[J]. *声学学报*, 2015, 40(1): 170–177.
- Jiang Guojian, Lin Jianheng, Sun Junping, et al. Attenuation of ship noise line-spectrum caused by the randomness of ocean channel[J]. *Acta Acustica*, 2015, 40(1): 170–177.
- [16] 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(3): 459–463.
- Liu Zhen, Qiu Jiaying, Cheng Yusheng. Application of deep neural network in blade-number recognition of ship propeller[J]. *Technical Acoustics*, 2019, 38(3): 459–463.
- [17] 朱可卿, 田杰, 黄海宁, 等. 基于深度学习的船舶辐射噪声识别研究[J]. *应用声学*, 2018, 37(1): 238–245.
- Zhu Keqing, Tian Jie, Huang Haining, et al. Ship-radiated noise recognition research based deep learning[J]. *Journal of Applied Acoustics*, 2018, 37(1): 238–245.
- [18] 王鹏. 基于神经网络的水中目标识别研究[D]. 哈尔滨: 哈尔滨工程大学, 2018.
- [19] Artusi E, Chaillanb F. Automatic recognition of underwater acoustic signature for naval applications[C]. *Maritime Situational Awareness Workshop*, 2019.
- [20] 陈越超, 徐晓男. 基于降噪自编码器的水中目标识别方法[J]. *声学电子工程*, 2018(1): 30–33.
- [21] 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(23): 181–185.
- 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(23): 181–185.
- [22] 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(1): 366–376.
- Zhang Shaokang, Wang Chao, Sun Qindong. Underwater target noise recognition and classification technology based on multi-classes feature fusion[J]. *Journal of Northwestern Polytechnical University*, 2020, 38(1): 366–376.
- [23] Tang Z, Kanu J, Manocha D, et al. Regression and classification for direction-of-arrival estimation with convolutional recurrent neural networks[C]. *Interspeech*, 2019.
- [24] van den Oord A, Kalchbrenner N, et al. Pixel recurrent neural networks[C]. *International Conference on Machine Learning*, 2016.
- [25] Gonzalez S, Brookes M. PEFAC—A pitch estimation algorithm robust to high levels of noise[J]. *IEEE ACM Transactions on Audio Speech and Language Processing*, 2014, 22(1): 518–530.
- [26] Lu J, Song S, Hu Z, et al. Fundamental frequency detection of underwater acoustic target using DEMON spectrum and CNN network[C]. *IEEE International Conference on Unmanned Systems*, 2020.