

# 基于多类别特征融合的水声目标噪声识别分类技术

张少康<sup>1,2,3</sup>, 王超<sup>1,2</sup>, 孙芹东<sup>1,2</sup>

(1.海军潜艇学院, 山东 青岛 266000; 2.青岛海洋科学与技术试点国家实验室, 山东 青岛 266000;  
3.国防大学 联合作战学院, 河北 石家庄 050000)

**摘要:**目标噪声信号作为当前水声目标识别的主要信号源之一,由于目标信号来源单一,难以像多传感器探测不同角度表征目标特性,导致目标识别分类正确率低、虚警率高,严重制约水声探测系统功能的发挥。针对这一问题,采用长短时记忆网络,建立多层 LSTM 水声目标噪声特征提取模型,学习提取目标噪声时域包络、DEMON 线谱、梅尔倒谱系数等信息特征,构建多类别特征子集;在此基础上,建立了基于多类别特征子集的特征级融合识别分类模型和基于 D-S 证据理论的决策级融合识别分类模型;利用样本库数据对上述模型进行了测试,对比多类别特征融合判别与单一类别特征识别分类的差异,并使用港池相关试验数据对上述模型进行了测试验证。测试结果表明,提出的基于多类别特征融合的水声目标噪声智能识别分类方法判别效果更好,对水声目标噪声信号识别分类的正确率和虚警率等相关指标均优于单一类别特征判别方法。

**关键词:**水声目标识别;水声目标噪声;多类别特征融合;特征级融合;决策级融合  
**中图分类号:**TB566 **文献标志码:**A **文章编号:**1000-2758(2020)02-0366-11

目标噪声信号作为当前水声目标识别的主要信号源之一,所承载的目标特性信息有限,难以像多传感器探测不同角度表征目标特性,但经过对目标噪声信号进行分析处理,可获取不同类别的数据信息,通过综合利用这些数据信息进行融合判别,能够有效提高识别分类的正确率,并降低虚警。近年来,水声目标噪声非线性特征提取研究工作发展迅速,目前的研究主要采用希尔伯特-黄变换<sup>[1-3]</sup>、小波分析<sup>[4-5]</sup>、高阶谱分析<sup>[6-8]</sup>等现代信号处理方法提取特征,并取得了一定的研究成果,但相关算法过于复杂,由于在实际工程应用中信号处理的实时性要求,上述算法在短时间内难以得到实际应用。随着近几年深度学习相关理论的发展,应用深度学习模型解决水声目标自主识别问题已逐渐成为研究重点。如卷积神经网络<sup>[9-13]</sup>、深度置信网络<sup>[14-15]</sup>、堆栈式自编码<sup>[16]</sup>等模型均被用来解决水声目标特征提取和识别问题,但相关方面研究仍处于起步阶段。另外,不少文献在进行识别分类研究时,使用的样本数据比较单一,研究对象多为某一特定类别的噪声信号,训

练和测试数据多来自于同一水声目标,因此在对相关模型进行测试时均取得了相对较好的分类效果。但由于水声目标信号类型繁多且分布广泛,加之数据来源困难,导致类别分布不平衡。因此,其实际应用效果仍有待进行进一步验证。

通常,多信息源数据融合过程可概括为:利用计算机技术对按时序获得的若干传感器的观测信息在一定准则下加以自动分析、综合以完成所需的决策和估计任务而进行的信息处理过程<sup>[17]</sup>。因此通常而言,信息融合多针对于多源或多传感器情况。而本文研究主要关注单传感器多类别目标噪声信息的融合情况,虽然这与传统多信息源数据融合概念有本质区别,但其相关方法理论仍有借鉴意义。参照信息融合相关定义,本研究中多类别特征融合判别过程可概括为:利用计算机技术对按时序获取的目标噪声信号的若干类别数据特征在特征级层面和决策级层面加以融合,以完成目标属性判别任务。

综上所述,本文针对当前水声目标噪声识别分类正确率相对较低、虚警率高的问题,采用长短时记

忆网络,自动学习提取目标噪声时域包络、DEMON 谱、梅尔倒谱系数等信息的数据特征,形成多类别特征子集,建立了基于多类别特征子集的特征级融合识别分类模型和基于 D-S 证据理论的决策级融合识别分类模型,并使用样本库数据和港池试验数据对上述模型进行了测试验证,对比了多类别特征融合判别与单一类别特征识别分类的差异。

## 1 多类别特征融合识别分类模型

### 1.1 长短时记忆网络提取声学数据特征

#### 1.1.1 长短时记忆网络结构单元

本文主要基于长短时记忆网络(long short term memory, LSTM),研究构建目标噪声识别分类模型,并通过借助网络结构中的“参数共享”机制,大幅降低模型参数量,以达到实际应用的目的。长短时记忆网络是神经网络的变种模型,对于处理时间序列数据问题有较好的表现效果,其基本单元如图 1 所示<sup>[18]</sup>,由 1 个中心节点和 3 个门控单元组成。中心节点通常被称为记忆细胞,用以存储当前网络状态,3 个门控单元分别为输入门、输出门和遗忘门,用以控制记忆块内信息流动。在前向传播过程中,输入门用以控制输入到记忆细胞的信息流,输出门用以控制记忆细胞到网络其他结构单元的信息流;在反向传播过程中,输入门用以控制迭代误差流出记忆细胞,输出门用以控制迭代误差流入记忆细胞。而遗忘门则用以控制记忆细胞内部的循环状态,决定信息的取舍或遗忘。通过这种门控机制,LSTM 网络得以控制单元内信息流动,使其具备了保存长时间信息的能力,即“记忆”能力,并使其在训练过程中能够防止内部梯度受外部干扰,避免了梯度弥散和梯度爆炸问题。

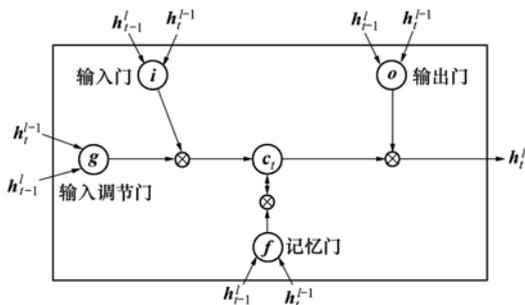


图 1 LSTM 基本结构单元

设单个 LSTM 记忆块的输入向量为  $x_t$ , 输出向

量为  $y_t$ , 前向传播公式可表述为<sup>[19]</sup>:

① 长期记忆单元  $C_t$  更新过程

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

$$i_t = \sigma(W_i \times [h_{t-1}, x_t] + b_i) \quad (2)$$

$$\tilde{C}_t = \tanh(W_c \times [h_{t-1}, x_t] + b_f) \quad (3)$$

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t \quad (4)$$

式中:  $f_t$  代表遗忘门,  $i_t$  代表输入门; 在每一个时刻, 遗忘门会控制上一时刻记忆的遗忘程度, 而输入门则控制新记忆  $\tilde{C}_t$  写入长期记忆的程度;  $f_t, i_t, \tilde{C}_t$  都是与上一时刻的短期记忆  $h_{t-1}$  和当前时刻输入  $x_t$  相关的函数, 其中,  $\sigma$  是 sigmoid 函数, 取值范围  $[0, 1]$ ,  $\tanh$  函数取值范围  $[-1, 1]$ ; 另外 (1) ~ (2) 式中  $W_f, W_i, W_c$  分别为遗忘门、输入门及  $C_t$  更新过程权重参数,  $b_f, b_i, b_c$  分别为这 3 个过程对应的偏置参数。

② 短期记忆单元  $h_t$  更新过程

$$o_t = \sigma(W_o \times [h_{t-1}, x_t] + b_o) \quad (5)$$

$$h_t = o_t \cdot \tanh(C_t) \quad (6)$$

式中:  $o_t$  表示输出门, 控制着短期记忆如何受长期记忆影响; (5) 式中  $W_o, b_o$  分别为输出门的权重和偏置参数。

#### 1.1.2 多层 LSTM 提取声学数据特征

模型主要通过多层 LSTM 网络自主学习声学数据特征, 图 2 给出了多层 LSTM 网络结构图。

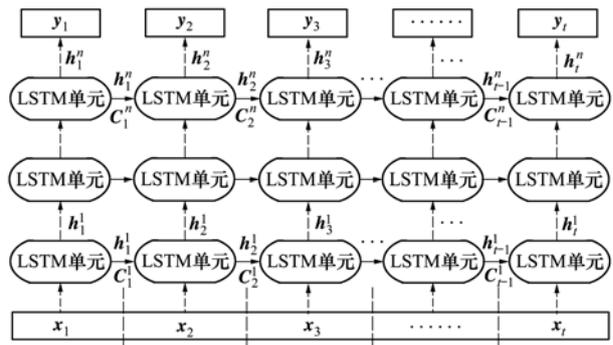


图 2 多层 LSTM 网络结构图

定义网络输入为  $X$ , 输出为  $Y$ , 网络层数为  $n$ , 每一层网络的时间步数为  $t$ 。按时间步数对输入序列  $X$  进行数据分割, 得到  $X = \{x_1, x_2, \dots, x_t\}$ , 依次输入第一层隐层网络, 根据公式 (1) ~ (6), 各层网络结构单元按时间步数  $t$  展开依次计算当前时刻单元短期记忆  $h_t^l$  及长期记忆  $C_t^l$  (变量中上标表示当前层数, 下标表示当前时间步数), 作为下一时刻结构单

元的输入,最终获取第一层网络输出  $Y^1 = \{h_1^1, h_2^1, \dots, h_i^1\}$ , 作为下一层 LSTM 隐层输入时间序列。设网络隐层数量为  $n$ , 由第一层至第  $n$  层网络逐层计算, 最终获取多层 LSTM 输出序列  $Y = \{y_1, y_2, \dots, y_i\} = \{h_1^n, h_2^n, \dots, h_i^n\}$ , 即为输入序列  $X$  的特征向量。

结合公式(1)~(6), 多层 LSTM 前向传播表达式为<sup>[20-21]</sup>

$$\begin{pmatrix} i_t^l \\ f_t^l \\ o_t^l \\ \tilde{C}_t^l \end{pmatrix} = \begin{pmatrix} \sigma \\ \sigma \\ \sigma \\ \sigma \end{pmatrix} T_{2n, 4n} \begin{pmatrix} h_{t-1}^{l-1} \\ h_{t-1}^{l-1} \end{pmatrix} \quad (7)$$

$$C_t^l = f_t^l \cdot C_{t-1}^l + i_t^l \cdot \tilde{C}_t^l \quad (8)$$

$$h_t^l = o_t^l \cdot \tanh(C_t^l) \quad (9)$$

式中:  $i_t^l, f_t^l, o_t^l, \tilde{C}_t^l, C_t^l, h_{t-1}^{l-1}, h_t^l$  分别为第  $l$  层单元节点的输入门、遗忘门、输出门、新记忆、长期记忆及短期记忆状态变量;  $h_{t-1}^{l-1}$  为上一层网络输出;  $T_{2n, 4n}$  为第  $l$  层网络权重参数和偏置向量集合。

将水声目标噪声信号经数据预处理过程, 分别获取时域包络、DEMON 谱及 MFCC 数据集, 使用上述 3 类训练数据集分别有监督预训练时域包络-LSTM、DEMON-LSTM 及 MFCC-LSTM 并保存模型参数, 使用保存的模型参数按照公式(7)~(9)计算多层 LSTM 输出, 形成多类别特征子集。其中, DEMON 数据集由 1 000 Hz 以内归一化 DEMON 谱数据构成; 在获取 MFCC 数据集时, 首先需对噪声信号样本作分帧处理, 获取各帧 36 阶 MFCC 特征向量  $C$ , 包括 MFCC 参数、一阶差分梅尔频率倒谱系数、二阶差分梅尔频率倒谱系数, 按各帧时间先后顺序, 生成各样本 MFCC 特征数据  $M$  作为 MFCC-LSTM 网络输入向量, 即  $M = \{c_1, c_2, \dots, c_n\}$ , 式中  $n$  为单个样本帧数。

### 1.2 基于多类别特征融合的目标噪声识别分类模型

按照融合过程中数据的抽象层次, 可将信息融合划分为数据级、特征级、决策级 3 个层级, 其中, 数据级融合为最低层级, 决策级为最高层级。理论上融合层级越低, 目标特性信息就越全面, 判别效果越好, 但同时数据量、模型规模等均相对较大, 模型对数据轻微变化也会相对比较敏感。本文主要从特征级融合判别和决策级融合判别 2 个方面分别构建识别分类模型, 进而对比 2 种判别模型的识别分类效果。

#### 1.2.1 基于多类别特征子集特征级融合的分类模型

多类别特征子集特征级融合以最大化分类效果为目的, 将反映噪声信号不同表征面的多个特征子集进行数学关联、分析、整合, 以便形成统一的融合识别分类框架。图 3 给出了基于多类别特征子集特征级融合的噪声信号识别分类框架。

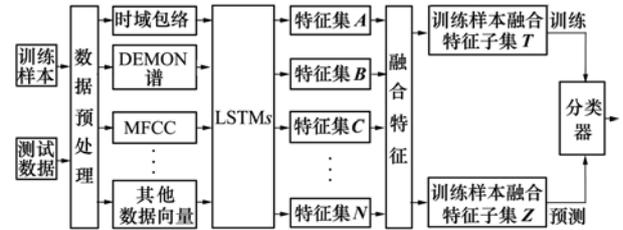


图 3 多类别特征子集特征级融合识别分类框架

由图 3 可知, 多类别特征子集特征级融合识别分类的具体内容为: 采用特征级信息直接融合策略  $f(\cdot) = \phi\{\beta_a(\cdot), \beta_b(\cdot), \dots, \beta_n(\cdot)\}$  获取多类别融合特征。式中  $f(\cdot)$  表示特征级融合框架, 是多类别子集特征融合的主要研究对象;  $\phi$  表示融合运算符, 是特征融合框架的核心, 本文主要采用了基于“形式”组合的串联特征融合运算符, 将多种类别形式的特征子集通过串联形式进行直接融合;  $\beta$  表示子集预处理算子, 为特征融合前准备工作, 本文中  $\beta$  主要通过数据预处理手段获取目标噪声归一化数据向量(如时域包络数据、DEMON 谱数据、MFCC 系数等)的过程。

本文基于多类别特征子集特征级融合对目标噪声进行识别分类的具体步骤为: ①在模型训练阶段, 经数据预处理, 获取训练样本同源多类别数据向量, 如时域包络数据、DEMON 谱数据、MFCC 数据以及其他类别的数据向量; ②以上述同源多类别数据作为网络输入, 训练多层 LSTM, 分别获取各类别数据条件下的特征子集; ③通过  $\phi$  融合算子, 以串联形式直接融合多类别特征子集, 获取训练样本融合特征子集  $T$ ; ④使用  $T$  训练分类器用以进行目标识别分类; ⑤在预测分类阶段, 重复上述过程, 获取测试样本融合特征子集  $Z$ , 使用训练好的分类器判别目标类别。

#### 1.2.2 基于 D-S 证据理论决策级融合的分类模型

设  $\Theta = \{\theta_1, \theta_2, \dots, \theta_n\}$  为当前所研究问题的识别框架, 其中包含  $n$  个完备且互斥的独立命题  $\theta_i$ , 根

据 D-S 融合理论,在当前识别框架下以  $M$  表示  $\Theta$  内所有元素及元素间并集构成的集合,其对应了所有目标类型的判别情况。在对水上、水下 2 类目标进行识别分类时,  $\Theta = \{\theta_1, \theta_2\}$ , 其中  $\theta_1$  表示“水上类别”,  $\theta_2$  表示“水下类别”。并且有  $M = \{\emptyset, \theta_1, \theta_2, \theta_1 \cup \theta_2\}$ , 其中,  $\theta_1 \cup \theta_2$  表示既可能为水下又可能为水上的情况。根据 D-S 理论,当前识别框架下 mass 函数可表述为

$$m(\emptyset) = 0 \text{ 且 } \sum_{A \subseteq \Theta} m(A) = 1 \quad (10)$$

对于  $\forall A \subseteq \Theta$ , 在水声噪声信号识别框架  $\Theta$  上的 Dempster 合成规则可表述为<sup>[22]</sup>

$$(m_1 \oplus m_2 \oplus m_3 \oplus m_4)(A) = \frac{1}{K_{A_1 \cap A_2 \cap A_3 \cap A_4 = A}} \sum_{A_1 \cap A_2 \cap A_3 \cap A_4 = A} m_1(A_1) \cdot m_2(A_2) \cdot m_3(A_3) \cdot m_4(A_4) \quad (11)$$

式中

$$K = \sum_{A_1 \cap A_2 \cap A_3 \cap A_4 \neq \emptyset} m_1(A_1) \cdot m_2(A_2) \cdot m_3(A_3) \cdot m_4(A_4) = 1 - \sum_{A_1 \cap A_2 \cap A_3 \cap A_4 = \emptyset} m_1(A_1) \cdot m_2(A_2) \cdot m_3(A_3) \cdot m_4(A_4) \quad (12)$$

式中,  $A_1, A_2, A_3, A_4$  分别表示当前识别框架下 4 种判别情况。

如图 4 所示,本文基于 D-S 证据理论决策级融合对目标噪声进行识别分类的具体步骤为:①对目标噪声经数据预处理获取时域包络、DEMON 谱、MFCC 数据等多类别训练数据集;②使用多类别数据训练集有监督预训练各类别数据输入情况下多层 LSTM 识别分类模型,并保存网络模型参数;③在预

测分类阶段,按照上述过程同样获取各类别测试数据,使用训练好的多层 LSTM 识别分类模型分别预测目标噪声所属类别概率  $P_a, P_b$  等;④基于 D-S 证据理论对各单一类别特征判别概率进行决策级融合判别目标类型。

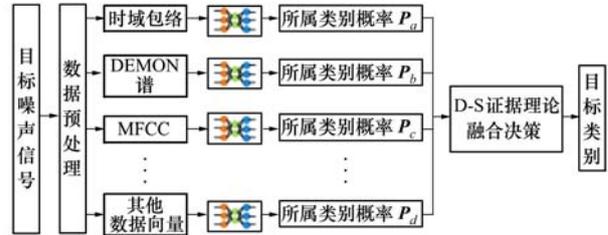


图 4 基于 D-S 证据理论决策级融合识别分类框架

此外,本文中仅以水声噪声信号的时域包络、DEMON 谱、MFCC 数据作为待融合数据信息,但实际在进行多类别特征融合判别是不仅限于上述 3 类数据,也可对其它类别数据进行联合判别。

### 1.3 模型度量方法

本文主要关注水上、水下 2 类声学信号的分类问题。使用机器学习中准确率、真正率、假正率(虚警率)、假负率<sup>[23]</sup>及声呐信号检测中检测率、漏检率、虚警率等<sup>[24]</sup>相关概念,分别定义本文使用的识别分类正确率  $P_{acc}$ , 以及真正率  $P_d$ 、假负率  $P_m$ 、假正率  $P_{fa}$ 、真负率  $P_n$  等相关概念。其中正确率  $P_{acc}$  指“对水上、水下 2 类信号识别分类正确的比例”,真正率  $P_d$  指“水下声学信号被正确识别分类的比例”,假负率  $P_m$  指“水下声学信号被错误识别分类的比例”,假正率  $P_{fa}$  指“水上声学信号被错误识别分类的比例”,真负率  $P_n$  指“水上声学信号被正确识别分类的比例”。

表 1 分类混淆矩阵

分类数目	预测为水下 $H_1$	预测为水上 $H_0$	实际总计
实际为水下 $H_1$	$N_{TP}$	$N_{FN}$	$N_{TP} + N_{FN}$
实际为水上 $H_0$	$N_{FP}$	$N_{TN}$	$N_{FP} + N_{TN}$
预测总计	$N_{TP} + N_{FP}$	$N_{FN} + N_{TN}$	$N_{TP} + N_{FN} + N_{FP} + N_{TN}$

表 1 给出了本识别分类问题的分类混淆矩阵<sup>[22]</sup>。其中,  $N_{TP}$  为“真正例”,即判别为水下类别中正确的数目;  $N_{FP}$  为“假正例”,即判别为水下类别中错误的信号数目;  $N_{FN}$  为“假反例”,即判别为水上类别中错误的信号数目;  $N_{TN}$  为“真反例”,即判别为水上类别中正确的信号数目,设分类目标总数为  $N$ ,则有

$$N = N_{TP} + N_{FP} + N_{TN} + N_{FN} \quad (13)$$

$$P_{acc} = \frac{N_{TP} + N_{TN}}{N} \quad (14)$$

$$P_d(H_1 | H_1) = \frac{N_{TP}}{N_{TP} + N_{FN}} \quad (15)$$

$$P_m(H_0 | H_1) = \frac{N_{FN}}{N_{TP} + N_{FN}} \quad (16)$$

$$P_{fa}(H_1 | H_0) = \frac{N_{FP}}{N_{FP} + N_{TN}} \quad (17)$$

$$P_n(H_0 | H_0) = \frac{N_{TN}}{N_{FP} + N_{TN}} \quad (18)$$

另外,使用“ROC 曲线”(receiver operation characteristic)衡量识别分类模型性能<sup>[25]</sup>,ROC 曲线横坐标为假正率,纵坐标为真正率。通常,若一个模型的 ROC 曲线被另一个模型的曲线完全“包住”,则后者性能优于前者,另外也可使用 ROC 曲线下的面积,即 AUC(area under roc curve)评价模型优劣。

## 2 模型验证

本文首先使用样本库数据对上述多类别特征融合判别方法进行了测试,之后使用港池试验数据对上述方法进行了验证,并对比研究了多类别特征融合判别与单一类别特征识别分类的效果差异。

### 2.1 样本库测试

样本库包含 65 000 条噪声信号,其中水上类别 42 000 条,水下类别 23 000 条。随机选择 2 类目标各 4/5 的样本组成训练集。按照图 3 所示过程,本文使用了在水声信号处理中常用的时域包络、

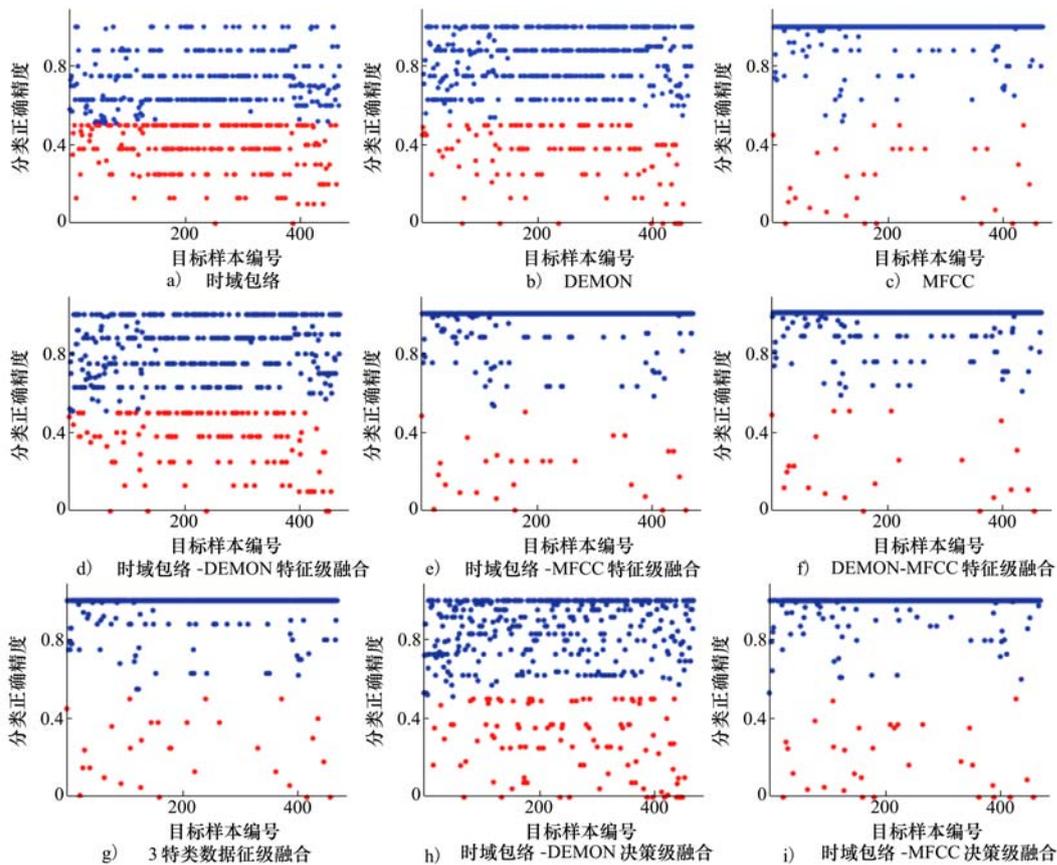
DEMON 谱、MFCC 系数作为待融合类别数据,分别通过有监督预训练多层 LSTM 模型,自主学习提取目标噪声特征子集,之后,采用串联特征融合运算符融合特征子集形成融合特征,并训练后端识别分类模型,相关参数设置见表 2。在前期研究中,在上述 3 类输入数据条件下,采用“9 层网络-100 个 LSTM 单元节点数”的组合均取得了相对较好的训练效果。同时,为降低这 3 类特征子集特征点数、幅值大小等因素差异造成得影响,设置特征子集点数均为 390,并在进行特征融合之前对各特征子集幅值作归一化处理。

表 2 模型相关参数

输入数据	LSTM 输入节点数	LSTM 层数	LSTM 单元节点数	特征子集维数
时域数据	30 000	9	100	390
MFCC	3 861	9	100	390
DEMON	3 276	9	100	390

选取样本库中 466 条信号进行测试,根据样本时长进行多次判别,基于判别统计情况给出最终判别分类结果。

图 5 给出了基于统计值的各样本的分类正确概



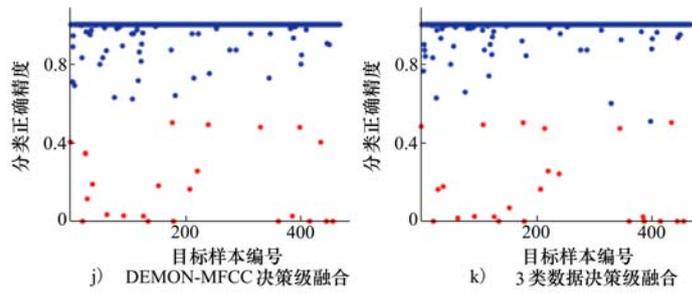


图 5 各样本识别分类正确概率

率,其中,横坐标为各测试样本编号,纵坐标为分类正确概率,另外,蓝色点表示分类正确,红色点表示分类错误。从图中可以看出,①单一类别特征判决情况下,采用 MFCC 作输入数据,样本总体分类正确概率优于其他两类;②多类别特征融合判别的样本总体分类正确概率优于单一类别特征判决情况。

同时,使用“样本均方误差(MSE)”作为衡量样本总体分类效果的标准,其计算公式可表述为

$$\delta_{MSE} = \frac{1}{N} \sum_{i=1}^N (p_i - p'_i)^2 \quad (19)$$

其含义为样本测试正确概率与期望概率(100%)的偏差,同样,也可使用“均方根误差(RMSE)”作为衡量标准,其计算公式可表述为

$$\delta_{RMSE} = \sqrt{\delta_{MSE}} \quad (20)$$

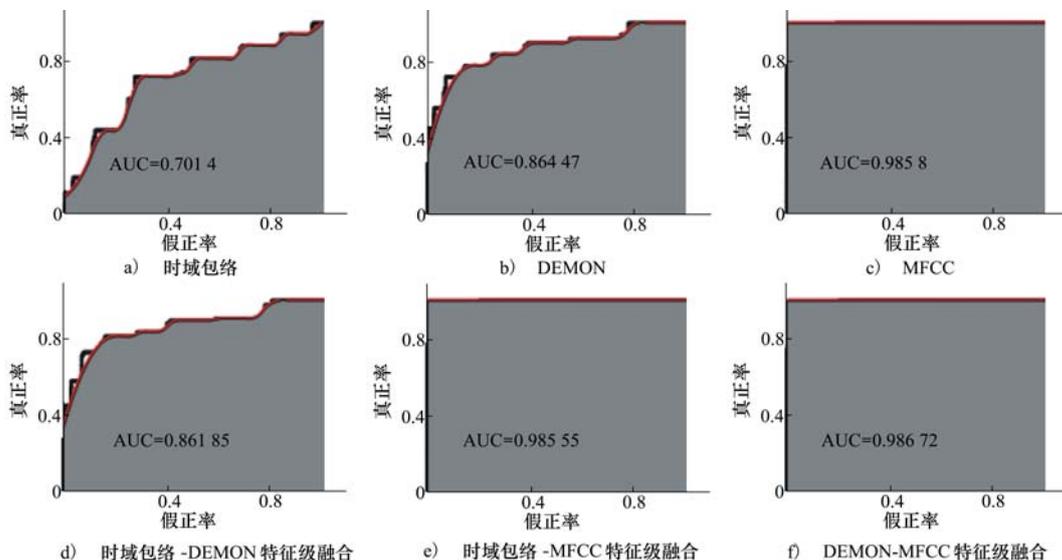
表 3 给出了使用样本库数据测试时,单一类别数据判别及多类别特征融合判别的  $\delta_{RMSE}$  及  $\delta_{MSE}$  对比情况,表中结果同样表明多类别特征融合判别的样本总体分类正确概率优于单一类别特征判决情况。

表 3 测试 RMSE 及 MSE 对比情况

判别效果	单一类别数据判别			特征级融合判别				决策级融合判别			
	时域包络	DEMON	MFCC	时域-DEMON	时域-MFCC	DEMON-MFCC	时域-DEMON-MFCC	时域-DEMON	时域-MFCC	DEMON-MFCC	时域-DEMON-MFCC
$\delta_{RMSE}$	10.14	8.05	4.81	8.22	4.64	4.23	4.59	8.29	4.91	4.19	4.31
$\delta_{MSE}$	102.93	64.82	23.20	67.63	21.53	17.93	21.23	68.8	24.16	17.6	18.59

图 6 样本库测试 ROC 曲线给出了与图 5 相对应的样本库数据测试的 ROC 曲线,从图中 AUC 面

积可看出多类别特征融合判别模型较单一类别特征判别模型判别效果更好。



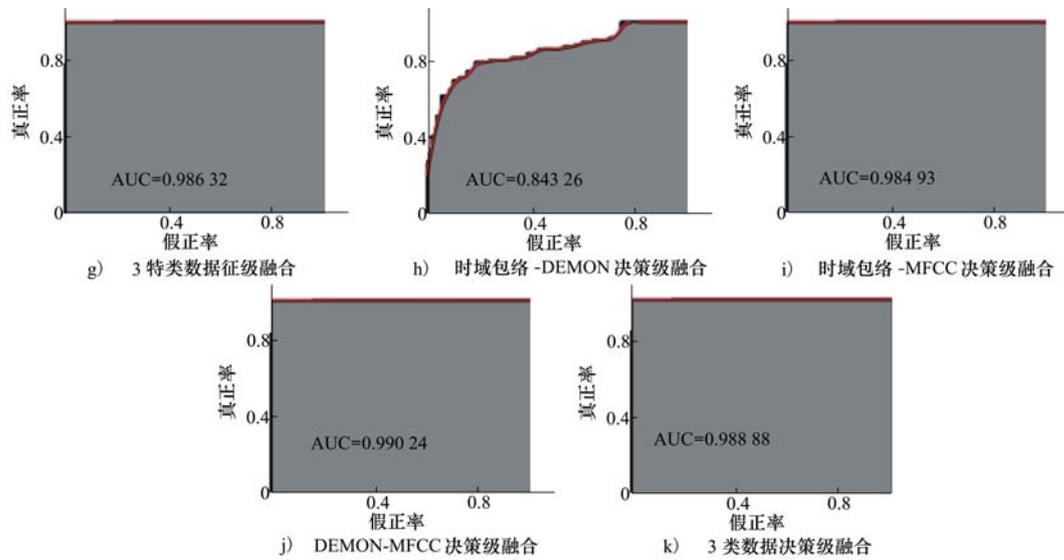


图 6 样本库测试 ROC 曲线

表 4 给出了使用样本库数据测试的分类混淆矩阵,表中依次分别为使用单一类别特征判别与多类别特征融合判别情况下的测试结果,根据公式(13)~(18),计算各度量指标,详情见表 5,其结果表明:①当采用多类别特征融合方法进行判别时,模型判别的正确率、真正率和假正率(虚警率)等指标均优

于单一类别特征判别方法;②在对特征进行两两融合判别时,识别分类结果易受单一特征判别影响,如表中时域-DEMON 特征级融合及决策级融合判别结果相对较差,但对上述 3 类数据同时进行融合判别时,结果较为稳定,受单一特征判别影响较小。

表 4 样本库测试分类混淆矩阵

判别结果	判别方式	预测为水下	预测为水上	实际总计	
实际为水下	单一特征	时域包络	61	24	85
		DEMON	66	19	
		MFCC	80	5	
	特征级融合	时域-DEMON	69	16	
		时域-MFCC	79	6	
		DEMON-MFCC	79	6	
		时域-DEMON-MFCC	79	6	
	决策级融合	时域-DEMON	67	18	
		时域-MFCC	80	5	
		DEMON-MFCC	79	6	
实际为水上	单一特征	时域包络	103	278	381
		DEMON	91	290	
		MFCC	26	355	
	特征级融合	时域-DEMON	99	282	
		时域-MFCC	22	359	
		DEMON-MFCC	17	364	
		时域-DEMON-MFCC	24	357	
	决策级融合	时域-DEMON	82	299	
		时域-MFCC	26	355	
		DEMON-MFCC	17	364	
		时域-DEMON-MFCC	17	364	

表 5 样本库测试各度量指标

判别方式	度量指标	$P_{acc}$	$P_d$	$P_m$	$P_{fa}$	$P_n$
单一特征	时域包络	72.7	71.8	28.2	27.0	73.0
	DEMON	76.4	77.6	22.4	23.9	76.1
	MFCC	93.3	94.1	5.9	6.8	93.2
特征级融合	时域-DEMON	75.3	81.2	18.8	26.0	74.0
	时域-MFCC	94.0	92.9	7.1	5.8	94.2
	DEMON-MFCC	95.1	92.9	7.1	4.5	95.5
	时域-DEMON-MFCC	93.6	92.9	7.1	6.3	93.7
决策级融合	时域-DEMON	78.5	78.8	21.2	21.5	78.5
	时域-MFCC	93.3	94.1	5.9	6.8	93.2
	DEMON-MFCC	95.1	92.9	7.1	4.5	95.5
	时域-DEMON-MFCC	95.3	94.1	5.9	4.5	95.5

2.2 港池试验数据验证

使用港池试验数据对本文模型进行验证,港池试验示意图如图 7 所示。试验过程中,使用试验船吊放宽带声源播放水声信号,声源深度约 7 m,声学传感器与声源位同一深度,距声源直线距离约为 100 m。采集播放信号的 1/3 倍频程谱级约为 80~85 dB@1 000 Hz。试验期间,实际探测水上类别的信号 39 批次,水下类别的信号 32 批次。表 6 给出了使用港池试验数据测试的分类混淆矩阵,其中,采

用单一 DEMON 特征进行判别为现场试验结果。

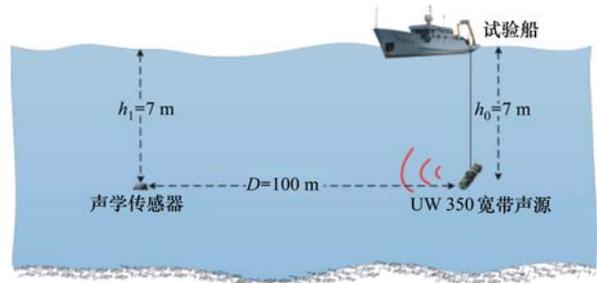


图 7 港池试验示意图

表 6 港池试验数据分类混淆矩阵

判别结果	判别方式	预测为水下	预测为水上	实际总计
实际为水下	单一特征			
	时域包络	21	11	
	DEMON(现场试验结果)	23	9	
	MFCC	26	6	
	特征级融合			
	时域-DEMON	22	10	
	时域-MFCC	26	6	
	DEMON-MFCC	26	6	32
	时域-DEMON-MFCC	26	6	
	决策级融合			
时域-DEMON	22	10		
时域-MFCC	25	7		
DEMON-MFCC	27	5		
时域-DEMON-MFCC	27	5		
实际为水上	单一特征			
	时域包络	9	30	
	DEMON(试验结果)	9	31	
	MFCC	7	32	
	特征级融合			
	时域-DEMON	9	30	
	时域-MFCC	7	32	
	DEMON-MFCC	5	34	39
	时域-DEMON-MFCC	6	33	
	决策级融合			
时域-DEMON	9	30		
时域-MFCC	8	31		
DEMON-MFCC	6	33		
时域-DEMON-MFCC	7	32		

表 7 给出了不同判别方式对应的分类指标,其结果同样表明,在港池试验数据条件下,采用多类别特征融合判别效果优于单一类别特征判别效果,且

对上述 3 类数据同时进行融合判别时,结果较为稳定。

表 7 港池试验数据测试各度量指标

判别方式	度量指标	$P_{acc}$	$P_d$	$P_m$	$P_{fa}$	$P_n$
单一特征	时域包络	71.8	65.6	34.4	23.1	76.9
	DEMON(现场试验结果)	76.1	71.9	28.1	23.1	79.5
	MFCC	81.7	81.3	18.8	17.9	82.1
特征级融合	时域-DEMON	73.2	68.8	31.3	23.1	76.9
	时域-MFCC	81.7	81.3	18.8	17.9	82.1
	DEMON-MFCC	84.5	81.3	18.8	12.8	87.2
	时域-DEMON-MFCC	83.1	81.3	18.8	15.4	84.6
决策级融合	时域-DEMON	73.2	68.8	31.3	23.1	76.9
	时域-MFCC	78.9	78.1	21.9	20.5	79.5
	DEMON-MFCC	84.5	84.4	15.6	15.4	84.6
	时域-DEMON-MFCC	83.1	84.4	15.6	17.9	82.1

### 3 结 论

为解决水声目标噪声识别过程中识别分类正确率相对较低、虚警高等问题,本文采用长短时记忆网络,建立了多层 LSTM 水声信号特征提取模型,自动学习提取声学信号时域包络、DEMON 线谱、梅尔倒谱系数等信息的数据特征,构建多类别特征子集。在此基础上,建立了基于多类别特征子集的特征级融合识别分类模型和基于 D-S 证据理论的决策级融合识别分类模型,利用样本库数据和港池试验数据对上述融合识别分类模型进行了测试,对比多类别特征融合判别与单一类别特征识别分类效果的差异,得到如下结论:

1) 多类别特征融合判别模型的识别分类效果总体优于单一类别特征判别模型。在对特征进行两两融合判别时,识别分类结果易受单一特征判别影响,但对文中时域包络、DEMON 谱、MFCC 3 类数据

同时进行融合判别时,结果较为稳定,受单一特征判别影响较小,因此,增加融合数据类型,有助于提高判别稳定性,降低错误判别结果的影响。

2) 较单一类别特征判别方法,多类别特征融合判别可有效提高识别分类的正确率,降低虚警率,但特征级和决策级 2 种融合方法优劣程度及融合判别实时性问题仍需作进一步研究。

3) 当融合的局部单一类别数据判别模型相同时,决策级融合判别更能有效利用单一类别数据判别模型的优化结果,对各输入数据条件下局部判别模型的限制较小,模型优化更为简便。相比之下,特征级融合判别模型受各类别特征子集维数、数量等参数影响较大,具体而言,可分性强的特征子集,对融合判别结果会产生良性影响,而可分性差的特征子集,对融合判别结果会产生不良影响。本文仅针对这两种融合判别模型进行了研究介绍,但由于未开展优化选择相关研究,因此二者优化后判别效果对比情况需进一步研究探讨。

### 参考文献:

- [1] 王逸林. 希尔伯特黄变换在矢量信号处理中的应用研究[D]. 哈尔滨: 哈尔滨工程大学, 2006  
WANG Yinlin. Research on the Application of Hilbert-Huang Transformation to Vector Signal Processing[D]. Harbin: Harbin Engineering University, 2006 (in Chinese)
- [2] 李秀坤, 谢磊, 秦宇. 应用希尔伯特黄变换的水下目标特征提取[J]. 哈尔滨工程大学学报, 2009, 30(5): 542-546  
LI Xiukun, XIE Lei, QIN Yu. Underwater Target Feature Extraction Using Hilbert-Huang Transform[J]. Journal of Harbin Engineering University, 2009, 30(5): 542-546 (in Chinese)
- [3] 李新欣. 船舶及鲸类声信号特征提取和分类识别研究[D]. 哈尔滨: 哈尔滨工程大学, 2012

- LI Xinxin. Research on Feature Extraction and Classification of Ship Noise and Whale Sound[D]. Harbin: Harbin Engineering University, 2012 (in Chinese)
- [4] 韩雪. 基于听觉特征的水中目标辐射噪声特征提取[D]. 哈尔滨: 哈尔滨工程大学, 2013  
HAN Xue. Feature Extraction of Underwater Target Based on Auditory Features[D]. Harbin: Harbin Engineering University, 2013 (in Chinese)
- [5] 石超雄, 李钢虎, 何会会, 等. 基于提升小波变换的 MFCC 在目标识别中的应用[J]. 声学技术, 2014, 33(4):372-375  
SHI Chaoxiong, LI Ganghu, HE Huihui, et al. Application of the Lifting Wavelet Transform Based MFCC in Target Identification [J]. Technical Acoustics, 2014, 33(4): 372-375 (in Chinese)
- [6] 程玉胜, 王易川, 史广智, 等. 基于现代信号处理技术的舰船噪声信号 DEMON 分析[J]. 声学技术, 2006, 25(1): 276-281  
CHENG Yusheng, WANG Yichuan, SHI Guangzhi, et al. DEMON Analysis of Underwater Target Radiation Noise Based on Modern Signal Processing[J]. Technical Acoustics, 2006, 25(1): 276-281 (in Chinese)
- [7] 葛青. 水下目标识别中的数据融合技术[D]. 哈尔滨: 哈尔滨工程大学, 2008  
GE Qing. The Information Fusion in Underwater Target Recognition[D]. Harbin: Harbin Engineering University, 2008(in Chinese)
- [8] 王杨. 水下目标辐射噪声多维特征分析技术[D]. 哈尔滨: 哈尔滨工程大学, 2012  
WANG Yang. Underwater Target Multidimensional Radiation Analysis of Characteristics of the Technology[D]. Harbin: Harbin Engineering University, 2012 (in Chinese)
- [9] 张大伟, 章新华, 付留芳, 等. 基于听觉感知与卷积神经网络识别舰船目标[J]. 声学技术, 2015, 34(6):181-184  
ZHANG Dawei, ZHANG Xinhua, Fu Liufang, et al. Recognition of Ships Based on Auditory Sense and Convolutional Neural Networks[J]. Technical Acoustics, 2015, 34(6): 181-184 (in Chinese)
- [10] 卢安安. 基于深度学习方法的水下声音目标识别研究[D]. 哈尔滨: 哈尔滨工程大学, 2017  
LU Anan. Underwater Acoustic Classification Based on Deep Learning[D]. Harbin: Harbin Engineering University, 2017 (in Chinese)
- [11] VALDENEGRO-TORO M. Improving Sonar Image Patch Matching via Deep Learning[C]//2017 European Conference on Mobile Robots, Paris, 2017
- [12] WILLIAMS D P. Underwater Target Classification in Synthetic Aperture Sonar Imagery Using Deep Convolutional Neural Networks[C]//International Conference on Pattern Recognition 2016, Mexico, 2016:2497-2502
- [13] 王强, 曾向阳. 深度学习方法及其在水下目标识别中的应用[J]. 声学技术, 2015, 34(2):138~140  
WANG Qiang, ZENG Xiangyang. Deep Learning Methods and Their Applications in Underwater Targets in Recognition[J]. Technical Acoustics, 2015, 34(2):138-140 (in Chinese)
- [14] KAMAL S, MOHAMMED S K, PILLAI P R S, et al. Deep learning Architectures for Underwater Target Recognition[C]//Ocean Electronics, 2013:48-54
- [15] 杨宏晖, 申昇, 姚晓辉, 等. 用于水声目标特征学习与识别的混合正则化深度置信网络[J]. 西北工业大学学报, 2017, 35(2): 220-225  
YANG Honghui, SHEN Sheng, YAO Xiaohui, et al. Underwater Acoustic Target Feature Learning and Recognition Using Hybrid Regularization Deep Belief Network[J]. Journal of Northwestern Polytechnical University, 2017, 35(2): 220-225 (in Chinese)
- [16] 陈越超, 徐晓男, 姚晓辉, 等. 基于降噪自编码的水中目标识别方法[J]. 声学与电子工程, 2018, 1: 30-33  
CHEN Yuechao, XU Xiaonan, YAO Xiaohui, et al. Underwater Target Recognition method based on Denoising Auto-Encoder [J]. Acoustics and Electronics Engineering, 2018, 1: 30-33 (in Chinese)
- [17] 刘同明, 夏祖勋, 谢洪成. 数据融合技术及其应用[M]. 北京: 国防工业出版社, 1998  
LIU Tongming, XIA Zuxun, XIE Hongcheng. Technology and Application of Data Fusion Techniques[M]. Beijing: National Defense Industry Press, 1998 (in Chinese)
- [18] KLAUS Greff, RUPESH K, SRIVASTAVA, et al. LSTM: a Search Space Odyssey[J]. IEEE Trans on Neural Networks and Learning Systems, 2017, 28(10): 2222-2232
- [19] ALEX Graves. Learning Precise Timing with LSTM Recurrent Networks[J]. Journal of Machine Learning Research, 2002(3): 115-143

- [20] WOJCIECH Zaremba. Recurrent Neural Network Regularization[J/OL]. (2015-12-19) [2019-03-01]. <https://arxiv.org/pdf/1409.2329v5.pdf>
- [21] NITISH Srivastava, GEOFFREY Hinton, ALEX Krizhevsky, et al. Dropout: a Simple Way to Prevent Neural Networks from Overfitting[J]. Journal of Machine Learning Research, 2014(15): 1929-1958
- [22] 夏佩伦. 目标跟踪与信息融合[M]. 北京:国防工业出版社, 2010  
XIA Peilun. Target Tracking and Information Fusion[M]. Beijing: National Defense Industry Press, 2010 (in Chinese)
- [23] PETER Flach. 机器学习[M]. 北京:人民邮电出版社, 2016  
PETER Flach. Machine Learning[M]. Beijing: Posts & Telecom Press, 2016 (in Chinese)
- [24] 田坦, 刘国枝, 孙大军. 声纳技术[M]. 哈尔滨:哈尔滨工程大学出版社, 1999  
TIAN Tan, LIU Guozhi, SUN Dajun. Sonar Technology[M]. Harbin, Harbin Engineering University Press, 1999 (in Chinese)
- [25] 周志华. 机器学习[M]. 北京:清华大学出版社, 2016  
ZHOU Zhihua. Machine Learning[M]. Beijing: Tsinghua University Press, 2016 (in Chinese)

## Underwater Target Noise Recognition and Classification Technology based on Multi-Classes Feature Fusion

ZHANG Shaokang<sup>1,2,3</sup>, WANG Chao<sup>1,2</sup>, SUN Qindong<sup>1,2</sup>

(1.Navy Submarine Academy, Qingdao 266000, China;  
2.National Laboratory for Marine Science and Technology, Qingdao 266000, China;  
3.Joint Operations College, National Defense University, Shijiazhuang 050000, China)

**Abstract:** As one of the main signal sources of underwater acoustic target recognition, the target noise signal is difficult to characterize the characteristics of the target from clearly comparing with the multi-sensor detection technology, which may lead to lower recognition rate and higher false alarm rate and seriously restricts the function of underwater acoustic detection system. In order to solve this problem, a multi-layers LSTM underwater acoustic target noise feature extraction model is established by using the long short term memory network. The information features such as time-domain envelope of target noise, Demon line spectrum and Mel frequency cepstrum coefficient are extracted, and a subset of multi-classes features is constructed. On this basis, the feature level fusion recognition and classification model based on the multi-classes features subset and the decision level fusion recognition and classification model based on the D-S evidence theory are established, and the above-mentioned models are tested by using the sample database. The difference of classification result between the multi-classes feature fusion and the single class feature recognition classification is compared, and the above model is tested and verified by using the relevant test data of port basin verification experiment. The correlation results show that the present intelligent recognition and classification method of underwater target noise based on the multi-classes feature fusion is more robust, and the recognition rate and false alarm rate of underwater target are better than those of single category feature discrimination method.

**Keywords:** underwater acoustic target recognition; underwater acoustic target noise; multi-class feature fusion; feature level fusion recognition method; decision level fusion recognition method