

航行工况失配条件下的深度神经网络 水声目标识别方法

王海涛, 晋安其, 杨爽, 曾向阳

(西北工业大学 航海学院, 陕西 西安 710072)

摘要:舰船的运行工况会对舰船辐射噪声产生很大影响,即使同一艘船行驶在同一片海域,不同的运行工况也会产生不同的辐射噪声,从而影响目标识别的准确性。特别是在工况失配的情况下,识别结果的正确率会大大降低。针对这一问题,提出了基于知识蒸馏的智能水下声学目标识别方法,以提高识别精度。使用听觉特征作为系统的输入,利用知识蒸馏学习不同工况条件下目标特征的内在联系。教师网络由大量现有工况数据训练而成,用于辅助学生网络(由少量工况数据训练而成)解决不同情况下的工况失配问题。测试使用了4种不同工况下的船舶辐射噪声数据集。结果表明,所提出的方法在各种工况失配问题上的表现都优于其他方法,这也体现了其在工程问题上的智能性和实用性。

关键词:舰船辐射噪声;知识蒸馏;水声目标识别;工况失配

中图分类号:TB566

文献标志码:A

文章编号:1000-2758(2024)06-1039-08

被动声呐识别系统由于其隐蔽性在军事和民用任务中发挥着重要作用。基于被动声呐的水下声学目标识别(UATR)方法从船舶辐射噪声中提取特征,并通过分类将这些特征映射到目标属性,从而实现目标识别^[1]。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,但对目标噪声信号进行分析处理,能够有效提高识别分类的正确率,并降低虚警^[2]。水声目标识别技术在实际应用中面临诸多困难,如缺乏大量数据类别信息、训练与测试不匹配等。其中,目标的工况条件也是导致训练与测试不匹配,从而影响识别效果的重要因素之一^[3]。

舰船噪声主要由机械噪声、螺旋桨噪声和水动力噪声组成,会随着舰船运行工况的变化而变化^[4]。当舰船处于不同的运行状态,如匀速直行、转弯、加速和减速时,会产生不同的目标辐射噪声,影响目标识别的准确性。目标的各种运行条件及其

状态变化反映在舰船噪声的变化中,因此也反映在信号的时域和频域中。

在过去几十年中,用于海洋工程的深度学习技术发展迅速,智能UATR系统受到了广泛关注^[5-6]。智能UATR系统的应用场景多种多样,如智能船舶、无人智能探测平台和无人潜航器(UUV)。特征提取和分类器设计是UATR系统的关键组成部分^[7]。近年来,许多学者在深度神经网络的帮助下将深度学习应用于UATR的任务,利用深度学习网络的简单结构和拟合能力进行特征优化和目标识别^[8-9]。例如,基于CNN-DNN模型的UATR模型使用开始层提取特征,使用最后一层进行分类和识别^[10]。随着对人耳听觉机制的深入研究,一些基于人耳听觉感知的特征也被应用于UATR,基于听觉特征^[11]直接使用卷积神经网络进行分类。

结合时域和频域分析的优势,时频分析提供了大量关于时频变化的信息^[12]。根据现有文献,基于时频的特征是大多数深度学习技术的首选^[13-14]。Feng等^[15]使用Mel滤波器组(FBank)作为输入特征,首次将变压器网络引入UATR领域,并取得了比DeepShip数据集中几个CNN网络更好的结果。联

收稿日期:2023-09-26

基金项目:国家自然科学基金面上项目(52271351)资助

作者简介:王海涛(1986—),副教授

通信作者:曾向阳(1974—),教授 e-mail:zengxy@nwpu.edu.cn

合波形和时频表示的特征^[16]也用于 UATR。通过具有深度相互学习的联合模型实现高效的 UATR。使用短时傅里叶变换 (STFT) 作为输入, AMNet^[17] 通过结合多分支网络和注意力机制, 在多个数据集上取得了成功。与传统技术相比, 深度神经网络成功地识别了基于时频的特征, 因为它们包含更详细的信息^[18-20]。

然而, 在目前的研究中, 使用的训练集和测试集是同源的, 但在实际场景中, 获取的数据样本覆盖了不同的目标工况条件, 容易导致测试数据与训练数据不匹配, 影响实际应用效果。受知识蒸馏 (knowledge distillation, KD)^[21] 的启发, 提出了一种用于工况失配的智能水声目标识别方法——UAKD。Mel 谱^[22] 作为系统的特征输入, 采用大量现有工况数据训练教师模型, 以辅助学生模型的训练。教师模型是一个强大的学习者, 可以通过将其学到的知识转移到学习中相对较弱的学生模型以提高学生模型的表现。学生模型的训练仅包含来自其他工况条件的少量数据。UAKD 允许学生模型学习更多信息, 例如不同工况条件之间的特征相关性, 以帮助解决工况失配问题。对多个工况失配案例进行了实验, 将

所提出的 UAKD 与几种基准方法进行比较。本文使用的数据包 4 种工作条件: 恒速直行、加速、减速和转弯。所提模型方法在所有工况失配案例中都取得了最佳结果, 证明了 UAKD 在工程问题上的有用性和智能性。

1 用于工况失配的水声目标识别模型

1.1 模型总体框架

水声目标的工作条件及其变化状态将反映在目标辐射噪声的变化中, 这也将反映时域和频域的差异。然而, 线谱结构与连续谱结构仍有一些相似之处, 这为不匹配工况条件下识别提供了依据, 以此为基础, 提出了一种智能 UATR 模型, 以提高工况失配下的识别性能。该模型的核心是利用对某一工作条件下目标辐射噪声数据的充分了解, 辅助完成少量其他工况下目标辐射噪声的识别任务。智能模型实现的方法是知识蒸馏。智能识别模型的框架见图 1。

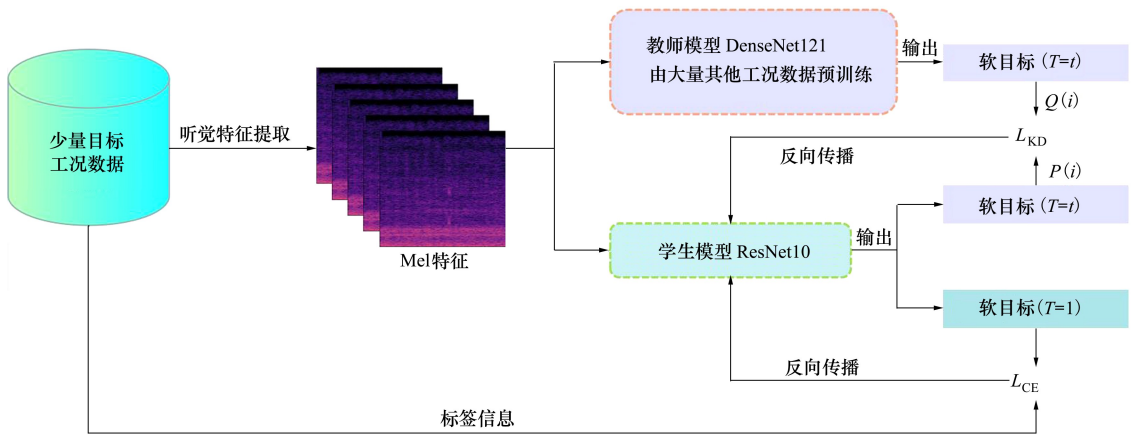


图 1 UAKD 模型结构图

在任务中, 假设一个工作条件的数据集是足够的, 而另一个工作条件的数据集是稀缺的。所有工作条件数据在输入模型之前都需要提取听觉特征。使用足够的工作条件数据训练教师网络并获得预训练教师模型。预训练教师模型可以对当前工作状态数据进行分类。之后, 将另一个工作条件的数据分别输入到预训练教师模型和学生模型中, 以获得它们的输出。软目标的定义如(1)式所示。

$$p(z_i, T) = \frac{\exp(z_i/T)}{\sum_j \exp(z_j/T)} \quad (1)$$

式中: z_i 表示第 i 类的输出; T 表示控制每个软目标重要性的温度参数。学生模型要学习教师模型的分类知识, 需要最小化 2 个网络输出的概率分布之间的距离。KL 散度用于度量 2 个模型输出之间的概率分布差异。假设 $P(i)$ 是学生模型的概率输出, $Q(i)$ 是预训练教师模型的概率输出。 $P(i)$ 和 $Q(i)$

之间的距离测量如(2)式所示。

$$L_{KD} = \sum_i P(i) \lg \frac{P(i)}{Q(i)} \quad (2)$$

除了鼓励学生模型学习预训练教师网络的分类知识外,还需要对学生模型在少量目标工况分类任务上进行微调。交叉熵损失函数用于计算学生网络输出概率分布与实数标签之间的误差,定义为

$$L_{CE} = -\frac{1}{N} \sum_{j=1}^N \sum_{k=1}^K t_{jk} \lg(P_{jk}) \quad (3)$$

式中: N 表示批大小数; K 表示类数; t_{jk} 表示样本 j 属于类别 k 的标签信息; P_{jk} 表示样本 j 属于类别 k 的预测概率。智能识别模型的最终损失函数可以定义为

$$L = \alpha L_{KD} + (1 - \alpha) L_{CE} \quad (4)$$

式中, α 是用于在 L_{KD} 和 L_{CE} 之间进行权衡的超参数。

反向传播算法会根据损失函数 L 的值计算每个参数的梯度,然后通过优化器调整参数,使 L 的值最小化,从而实现识别系统的训练。将教师模型和学生模型之间的分类知识相匹配是一种有效的知识转移方法。

1.2 Mel 特征

教师模型和学生模型都使用 Mel 听觉特征作为输入。人耳对音高的感知与声波的频率不呈正比,而是遵循一种对数关系,可以用 Mel 频率标度更好地近似。这种关系可以通过(5)式在数学上表示

$$f_{Mel} = 2595 \times \lg \left(1 + \frac{f}{700} \right) \quad (5)$$

式中: f 是实际频率; f_{Mel} 是以 Mel 表示的感知频率。在 Mel 频谱图的计算中,使用 Mel 滤波器组对频谱图进行加权,可以将原始频谱图的高维特征映射到低维 Mel 频谱图表示,实现音频信息的压缩。Mel 谱将频带均匀地分布在 Mel 音阶上,使其更符合人类听觉系统的感知特征。

1.3 教师模型和学生模型

教师模型是层数较多的 DenseNet121。与其他 CNN 网络不同, DenseNet 是密集连接的,网络的每一层都连接到前一层的所有网络。每个密集块由三部分组成:批量归一化(BN)、ReLU 激活函数和 3D 卷积(即大小 $3 \times 3 \times 3$)。BN 减少了内部协变量的偏移,不仅加快了训练速度,而且提高了准确性。ReLU 降低了梯度消失的概率。在每个密集块之间设置过渡层,如图 2 所示。过渡层由用于整合数据的 BN、用于保持原始特征图大小的 $(1 \times 1 \times 1)$ Conv

和最大池化组成,以减少特征维度。DenseNet 的每一层都连接到前一层的所有网络,允许直接访问输入特征并通过通道上的连接重用具有更好鲁棒性的特征,此外,教师网络在密集连接结构中增加了注意力机制,如图 3 所示。

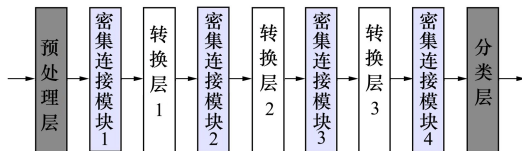


图 2 教师模型 DenseNet121 模型

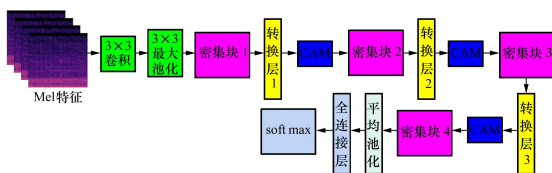


图 3 加入通道注意力机制的 DenseNet121 模型

学生模型是层数较少的轻量级 ResNet10。此外,学生网络在残差结构中增加了注意力机制,以进一步提高学生模型的性能。学生模型的具体结构如图 4 所示。MFCC 特征被输入到 7×7 卷积层和 3×3 最大池化层中,用于浅层特征提取。深度提取模块由 4 个残差块组成,每个残差块包括 2 个 3×3 卷积层和 1 个注意力块。注意力块将在 1.4 节中详细介绍。全连接层和平均池化层组成分类模块,按软目标 ($T=1$) 分类。表 1 显示了学生模型每一层的精确参数设置。从表 1 可以看出,浅层特征提取模块通过卷积和池化操作压缩特征图的形状,同时增加了特征图的通道。

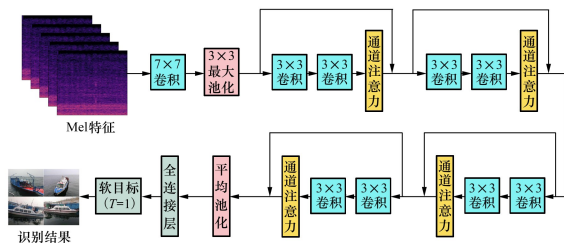


图 4 学生模型 ResNet10 模型

在深度提取模块中,第一残差块不会改变通道数或特征图的形式。下采样操作在表 1 的后 3 个残差块中实现,它们压缩输入特征图的形状并增加通道维度以提取尽可能多的有效特征。在 UAKD 中,学生模型降低了模型的复杂性,这对于提高识别效

率是非常必要的。

表 1 学生模型具体参数

层名称	输出尺寸	学生模型
Me1	128×44×1	
卷积	64×22×64	7×7, 64, 步长 2
最大池化	32×11×64	3×3, 步长 2
残差块	32×11×64	$\begin{bmatrix} 3 \times 3 \text{ 卷积, } 64 \\ 3 \times 3 \text{ 卷积, } 64 \\ \text{通道注意力层} \end{bmatrix} \times 1$
残差块	16×6×128	$\begin{bmatrix} 3 \times 3 \text{ 卷积, } 128 \\ 3 \times 3 \text{ 卷积, } 128 \\ \text{通道注意力层} \end{bmatrix} \times 1$
残差块	8×3×256	$\begin{bmatrix} 3 \times 3 \text{ 卷积, } 256 \\ 3 \times 3 \text{ 卷积, } 256 \\ \text{通道注意力层} \end{bmatrix} \times 1$
残差块	4×2×512	$\begin{bmatrix} 3 \times 3 \text{ 卷积, } 512 \\ 3 \times 3 \text{ 卷积, } 512 \\ \text{通道注意力层} \end{bmatrix} \times 1$
平均池化	1×1×512	4×2
分类层	4	全连接, 软目标 ($T=1$)

1.4 通道注意力机制

注意力机制有能力优先考虑高权重的重要信息,而忽略低权重的不相关信息。它还可以连续调整权重,从而可以在各种上下文中选择重要信息,并提高网络模型的可扩展性和鲁棒性。本文在教师和学生网络中引入通道注意力机制来优化网络,使网络充分获取特征中的有效信息,增加卷积核的感知范围。通道注意力层的结构如图 5 所示。

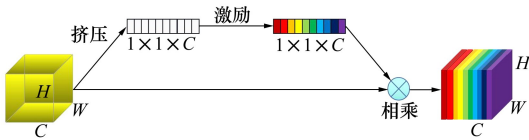


图 5 通道注意力结构

在通道维度上增加注意力机制的通道注意力机制(CAM)的主要动作是“挤压”和“激励”。特征图中每个特征通道的相关性通过自动学习确定,然后利用该重要性为每个特征赋予权重值,以便神经网络专注于特定的特征通道。增强对当前任务有帮助的特征图通道,并抑制不太有用的功能通道。如图 4 所示,在输入 CAM 之前,特征图每个通道的重要性是相同的,而在通过 CAM 后,每个特征通道的相关性用不同方框颜色表示,这导致神经网络专注于

具有高权重值的选择通道。

CAM 通过对卷积特征通道之间的相互依赖关系进行建模来改进网络表示。首先,通过挤压压缩操作生成 $1 \times 1 \times c$ 通道描述符,将全局信息压缩到上述通道描述符中,以便输入层使用。这里使用全局平均池化

$$y_c = F_{sq}(m_c) = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W m_c(i, j) \quad (6)$$

式中: y_c 表示经过全局池化处理后的输出,下标 c 表示通道序号; $F_{sq}(\cdot)$ 表示对特征矩阵的挤压压缩操作; m_c 表示第 c 个通道的特征矩阵; H, W 表示每个通道中特征矩阵的高与宽; $m_c(i, j)$ 表示第 c 个通道特征矩阵的第 i, j 个元素。

然后,每个通道通过依赖于通道的门限机制学习特定样本,该机制利用全局信息学习,只强调信息量最大的特征,忽略信息量较小的特征,本文使用中间嵌入 ReLU 函数的 Sigmoid 来限制模型的复杂性并帮助训练。此外,门控机制由 2 个全连接层(FC)组成的瓶颈层进行参数化,即 W_1 用于降维, W_2 用于维数增量

$$n = F_{ex}(y, W) = \sigma(g(y, W)) = \sigma(W_2 \delta(W_1 y)) \quad (7)$$

式中: n 是激励操作后的通道权重向量; $F_{ex}(\cdot)$ 表示对通道信息的激励操作; y 表示挤压压缩操作的输出; W 表示全连接层的权重矩阵; $\sigma(\cdot)$ 表示 Sigmoid 激活函数,用于将通道权重规范化到 0 到 1 之间; $g(\cdot)$ 表示对 y 及 W 的门控; $\delta(\cdot)$ 表示 ReLU 激活函数,用于引入非线性特性; W_1, W_2 表示 2 个全连接层的权重矩阵。

$$\tilde{M}_c = F_{scale}(m_c, n_c) = n_c m_c \quad (8)$$

式中: \tilde{M}_c 表示经过挤压 - 激励操作后的输出; $F_{scale}(\cdot)$ 表示对特征的缩放处理; m_c, n_c 表示第 c 个通道的挤压压缩操作与激励操作的输出。

CAM 模块自动学习成分权重矩阵,找出每个特征通道在船舶辐射噪声数据通道维度中的重要性,从而获得通道级水声信号的特性。

2 实验和讨论

2.1 数据集和工况失配案例

本文测试实验中使用的数据源于自测湖试数据集,该数据集于 2018 年 9 月采集自丹江口湖。数据

集中包含 4 类目标,为 4 种不同类型的船只。湖试过程中利用 2 个 8 元线阵采集目标辐射噪声信号,并记录了船只的 GPS 坐标。如图 6 所示,噪声采集过程中,目标从五角星指定的起始位置出发,沿着数字 1~8 的路线行驶,最后回到五角星位置,每类目标均航行 3 圈。识别测试实验以其中 1 个线阵列的 1 个通道的数据为研究对象,采样频率为 48 kHz。通过观察目标每一圈的轨迹图,可以确定目标直行和转弯的轨迹,坐标和时刻点一一对应。因此,从轨迹图中找出目标直行和转弯轨迹对应的 GPS 坐标,并得到坐标对应的时刻点,再据此从原始数据中剪切出直行工况和转弯工况对应的数据。

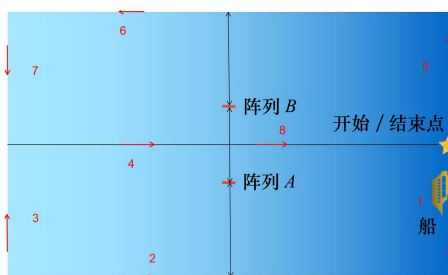


图 6 数据采集

阵元的坐标是已知的,以阵元为坐标系计算出目标在某个时刻点对应的速率,绘制出 4 类目标每圈的速率图。从上升和下降趋势中,可以确定目标是否处于加速或减速状态,确定其对应的时刻点,然后从原始数据中截出加速工况和减速工况对应的数据。最终获得包含 4 种工况的样本,其总数为 5 780,每条样本长度为 1 s。

当训练集和测试集不一致时,就会在一定程度上造成训练测试不匹配问题。从数据数理统计分析角度看,训练数据与测试数据的匹配程度是影响最终识别正确率指标的关键因素。当训练数据与测试数据对应的工况条件不匹配时,其变化会影响目标特征的变化,最终使得测试数据基于训练数据的分布发生变化。测试集的数据不能很好地拟合训练集下分类器训练的数据分布,使分类性能明显降低,导致正确识别率降低,而实际情况只能从少量样本数据中获取,很难覆盖所有可能的操作条件,训练集有限,因此也会对分类结果产生影响。

工况失配包含 6 个案例,分别以 2 种不同动工况条件构成训练集和测试集。如表 2 所示,将这 6 个工况失配案例标记为 1~6,其中直线-转弯表示

训练集为直线工况,测试集为转弯工况。

表 2 工况失配案例的编号设置

工况失配	训练集	测试集	编号
直线-加速	直线	加速	1
直线-减速	直线	减速	2
直线-转弯	直线	转弯	3
加速-减速	加速	减速	4
加速-转弯	加速	转弯	5
减速-转弯	减速	转弯	6

2.2 实验设置

在实验中,使用 Ubuntu 18.04.1 x64 操作系统,带有 Intel(R) 酷睿(TM) i9-9920X CPU @ 3.50 GHz 和 NVIDIA GeForce GTX 2080 Ti。采用 PyTorch 深度学习框架来实现实验,有 100 个 epoch,学习率为 0.000 1,批量大小为 4。利用自适应估计(Adam)优化器,将 L2 正则化设置为设置为 4×10^{-5} ,根据梯度信息更新神经网络模型的参数。这种方法有效地优化了模型,在实验中取得了更好的结果。

使用来自某一工况条件的所有数据预训练教师模型。然后使用来自其他工况条件的少量数据来训练学生模型,而预训练的教师模型则有助于学生模型的训练。在测试过程中,来自其他工况条件的剩余数据用于测试模型的识别性能。用于训练和测试的样本数定义为样本比例 k 。当训练涉及 72 个样本,测试涉及 1 364 个样本时, $k=0.05$ 。

2.3 实验结果与分析

采用正确识别率和 F1 分数评估系统的性能,以对比识别实验结果。

为了证明所提 UAKD 的通用性和智能性,研究了 UAKD 在 6 种不同工况案例中的效果,并将其与各种先进的深度学习方法进行了比较,包括:ResNet18、DenseNet121 和 MobileNetV2 方法。此外,还研究了 k 值对识别精度的影响。

在训练样本中加入少量目标工况样本作为先验信息,训练样本包括某一工况样本的全部数据和少量目标工况数据,测试样本为目标工况的剩余数据。其中,训练集和测试集中的目标样本之比为 0.05。表 3 显示了 4 种模型在 6 种工况不匹配工作条件下的正确识别率和 F1 分数。可以看出,UAKD 模型的正确识别率最高,表明 UAKD 模型在工况失配问题上具有较好的性能,在面对不同工况时具有更好的泛化能力,对工况具有更强的鲁棒性。

在不匹配工况条件下,使用少量目标工况数据

进行知识提炼可以显著提高本文模型的性能。6 种工况失配案例中,UAKD 模型在案例 1 工况下的精度提升最为显著。但是,从 UAKD 模型在 6 种不匹配工况下的精度来看,案例 1 中的精度最低,即缓解直线-加速失配的任务是最困难的,这说明直线工况与加速工况下辐射噪声差距较大。

表 3 识别实验结果

案例编号	模型方法	正确识别率/%	F1 分数/%
1	ResNet18	85.8	85.7
	DenseNet121	85.0	85.3
	MobileNetV2	82.5	82.2
	UAKD	94.2	94.1
2	ResNet18	85.7	85.8
	DenseNet121	90.6	90.6
	MobileNetV2	79.3	79.1
	UAKD	96.1	96.1
3	ResNet18	87.1	87.0
	DenseNet121	91.1	91.3
	MobileNetV2	85.9	85.6
	UAKD	95.1	95.1
4	ResNet18	86.6	86.8
	DenseNet121	88.7	88.5
	MobileNetV2	81.1	81.1
	UAKD	94.7	94.8
5	ResNet18	92.1	91.9
	DenseNet121	92.2	92.1
	MobileNetV2	89.2	88.9
	UAKD	96.5	96.5
6	ResNet18	89.8	89.9
	DenseNet121	89.7	89.7
	MobileNetV2	84.6	84.8
	UAKD	95.6	95.6

图 7~9 显示了样本比例对 UAKD 模型精度的影响。可以看出,随着 k 的增加,UAKD 模型在 3 种案例中性能都有所提高。当 k 为 0.1 时,模型的精度最高。此外,在面对工况失配问题时,UAKD 模型可以在目标样本较少的情况下取得良好的识别效

果。当 k 为 0.025 时,UAKD 模型仍能保持较高的精度。

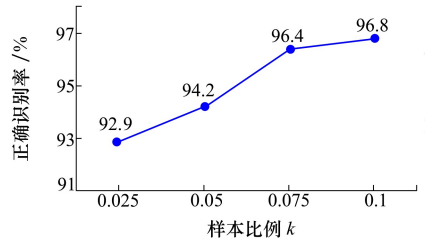


图 7 案例 1 中样本比例 k 对正确识别率的影响

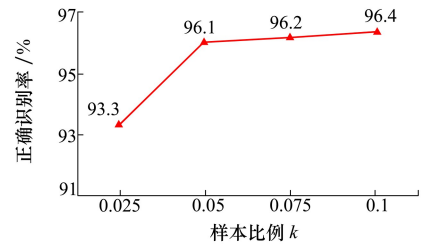


图 8 案例 2 中样本比例 k 对正确识别率的影响

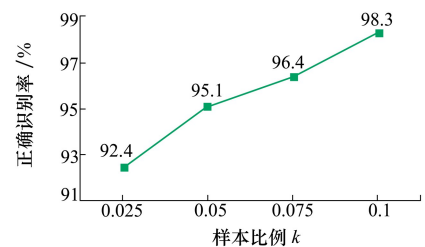


图 9 案例 3 中样本比例 k 对正确识别率的影响

3 结论

针对水声目标识别中的工况失配问题,提出一个针对工况失配的智能 UATR 模型——UAKD。首先,在输入模型之前,需要提取所有工况条件数据以获取听觉特征。其次,使用足够的工况条件数据训练教师网络并获得预先训练的教师模型。最后,将另一组工况条件数据分别输入到预训练教师模型和学生模型中。较小的学生网络更适合实际部署。采用教师网络辅助学生网络,利用通道注意力增强模型对通道特征的敏感度,学习对工况条件影响较小、适应不同工况失配场景的深度特征。

6 个工况失配案例中的实验表明,UAKD 优于其他先进方法,证明了 UAKD 的有效性,为其在实际工程中的应用提供了可行的基础。未来,计划进一步探索 UAKD 在更复杂的数据或任务上的表现。

参考文献:

- [1] 王强. 基于深度学习理论的水声目标识别方法研究[D]. 西安: 西北工业大学, 2018
WANG Qiang. Research on hydroacoustic target recognition method based on deep learning theory[D]. Xi'an: Northwestern Polytechnical University, 2018 (in Chinese)
- [2] 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
ZHANG Shaokang, WANG Chao, SUN Qindong. Classification technique for hydroacoustic target noise recognition based on multi-category feature fusion[J]. Journal of Northwestern Polytechnical University, 2020, 38(2): 366-376 (in Chinese)
- [3] JIN Anqi, ZENG Xiangyang. A novel deep learning method for underwater target recognition based on res-dense convolutional neural network with attention mechanism[J]. Journal of Marine Science and Engineering, 2023, 11(1): 69
- [4] MCINTYRE Duncan, LEE Walfred, FROUIN-MOUY Héloïse, et al. Influence of propellers and operating conditions on underwater radiated noise from coastal ferry vessels[J]. Ocean Engineering, 2021, 232: 109075
- [5] REN Jiawei, XIE Yuan, ZHANG Xiaowei, et al. UALF: a learnable front-end for intelligent underwater acoustic classification system[J]. Ocean Engineering, 2021, 264: 112394
- [6] GAO Miao, SHI Guoyou. Ship collision avoidance anthropomorphic decision-making for structured learning based on AIS with Seq-CGAN[J]. Ocean Engineering, 2020, 217: 107922
- [7] DING Yuwei. Review on passive sonar target recognition[J]. Technical Acoustics, 2004, 4: 253-257
- [8] KAMAL Suraj, MOHAMMED K Shameer, PILLAI P R Saseendran, et al. Deep learning architectures for underwater target recognition[C]//Proceedings of 2013 Ocean Electronics, Kochi, India, 2013
- [9] FERGUSON L Eric, RAMKRISHNAN Rish, WILLIAMS B Stefan, et al. Convolutional neural networks for passive monitoring of a shallow water environment using a single sensor[C]//Proceedings of 2017 IEEE International Conference on Acoustics, Speech and Signal Processing, New Orleans, LA, USA, 2017
- [10] HU Gang, WANG Kejun, PENG Yuan, et al. Deep learning methods for underwater target feature extraction and recognition[J]. Computational Intelligence and Neuroscience, 2018, 2018: 1214301
- [11] LI Junhao, 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
- [12] WANG Shuguang, ZENG Xiangyang. Robust underwater noise targets classification using auditory inspired time-frequency analysis[J]. Applied Acoustics, 2014, 78: 68-76
- [13] ZHENG Yunliang, GONG Qiyong, ZHANG Shufang. Time-frequency feature-based underwater target detection with deep neural network in shallow sea[J]. Journal of Physics: Conference Series, 2021, 1756: 012006
- [14] MUHAMMAD Irfan, ZHENG Jiangbin, ALI Shahid, et al. DeepShip: an underwater acoustic benchmark dataset and a separable convolution based autoencoder for classification[J]. Expert Systems with Applications, 2021, 183: 115270
- [15] FENG Sheng, ZHU Xiaoqian. A transformer-based deep learning network for underwater acoustic target recognition[J]. IEEE Geoscience and Remote Sensing Letters, 2022, 19: 1505805
- [16] ZHU Yunan, WANG Biao, ZHANG Youwen, et al. Convolutional neural network based filter bank multicarrier system for underwater acoustic communications[J]. Applied Acoustics, 2021, 177: 107920
- [17] WANG Biao, ZHANG Wei, ZHU Yunan, et al. An underwater acoustic target recognition method based on AMNet[J]. IEEE Geoscience and Remote Sensing Letters, 2023, 20: 1-5
- [18] MUHAMMAD Khishe. DRW-AE: a deep recurrent-wavelet autoencoder for underwater target recognition[J]. IEEE Journal of Oceanic Engineering, 2022, 47(4): 1083-1098
- [19] SUN Qinggang, WANG Kejun. Underwater single-channel acoustic signal multitarget recognition using convolutional neural networks[J]. The Journal of the Acoustical Society of America, 2022, 151(3): 2245-2254
- [20] ZHANG Wen, LIN Bin, YAN Yulin, et al. Multi-features fusion for underwater acoustic target recognition based on convolution recurrent neural networks[C]//2022 8th International Conference on Big Data and Information Analytics, Guiyang, China, 2022
- [21] BERGMANN Paul, FAUSER Michael, SATTLEGGER David, et al. Uninformed students: student-teacher anomaly detection

with discriminative latent embeddings[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 2020

[22] ZHANG Lanyue, WU Di, HAN Xue, et al. Feature extraction of underwater target signal using Mel frequency cepstrum coefficients based on acoustic vector sensor[J]. Journal of Sensors, 2016, 2016: 7864213

Underwater acoustic target recognition under working conditions mismatch

WANG Haitao, JIN Anqi, YANG Shuang, ZENG Xiangyang

(School of Marine Science and Technology, Northwestern Polytechnical University, Xi'an 710072, China)

Abstract: The working conditions of the ship will have a great impact on the radiated noise of the ship. Even if the same ship is traveling in the same sea area, different working conditions will produce different radiated noise, thus affecting the accuracy of target recognition. Especially in the case of working condition mismatch, the correct rate of the recognition results will be greatly reduced. To address this problem, an intelligent underwater acoustic target recognition method based on knowledge distillation is proposed to improve the recognition accuracy. Auditory features are used as inputs to the system, and knowledge distillation is utilized to learn the intrinsic connection of target features under different working conditions. The teacher network, trained from a large amount of existing working condition data, is used to assist the student network (trained from a small amount of working condition data) to solve the working condition mismatch problem under different conditions. Tests were conducted using ship radiated noise datasets under four working conditions. The results show that the proposed method outperforms the other methods in all kinds of working condition mismatch problems, which demonstrates its intelligence and practicality in engineering problems.

Keywords: ship radiated noise; knowledge distillation; underwater acoustic target recognition; working condition mismatch

引用格式:王海涛, 晋安其, 杨爽, 等. 航行工况失配条件下的深度神经网络水声目标识别方法[J]. 西北工业大学学报, 2024, 42(6): 1039-1046

WANG Haitao, JIN Anqi, YANG Shuang, et al. Underwater acoustic target recognition under working conditions mismatch [J]. Journal of Northwestern Polytechnical University, 2024, 42(6): 1039-1046 (in Chinese)