

面向小样本数据的水下目标识别 神经网络深层化研究

吴晏辰, 王英民

(西北工业大学 航海学院, 陕西 西安 710072)

摘 要:在面临新时期海洋工程应用领域的挑战时,可以通过利用基于深度学习的神经网络在水声工程中的实现,来达成自动化、高效性、准确性的目标。然而在面对水下目标样本匮乏、水下声环境复杂、样本信噪比差等客观问题时,深度学习也会因其自身的局限性而变得不那么灵敏。针对小样本问题,通过构建多种目标特征提取法和深层深度神经网络模型,得到了不同目标特征提取与网络模型匹配后的目标识别率与网络预测值,并通过比对实验结果,提出了通过深层神经网络深层化设计解决小样本目标识别的新思路。

关 键 词:水下目标识别;深度学习;深层神经网络设计

中图分类号:TJ630.34;TB53

文献标志码:A

文章编号:1000-2758(2022)01-0040-07

水下目标识别作为水声工程的重要技术领域,一直是海洋军事和民用的主要研究方向之一。随着近年来水下军事目标小型化、消声化,民用目标复杂化、精确化的趋势不断加深,对于水下目标识别系统的精度和灵敏度都产生了更高的要求^[1]。

与传统的利用基于先验知识的人工听音和谱图等信息进行目标识别不同,利用基于深度学习的水下目标识别系统,可以满足现代海洋作战、作业对于目标自动识别、快速识别的迫切需求^[2]。能够实现这一需求的主要原因在于深度学习提取目标声学信号中的各类特征作为基础,利用信号处理的方法,将多种信号特征通过近年来蓬勃发展的各类人工智能分类器,以实现自动识别的目的。现有研究表明,目前利用各种方法所提取的信号特征在用于水下目标识别时,在实验室验证的效果普遍较好,但在实际水下噪声环境中,应用性能却出现了明显下降^[3]。出现这种现象的主要原因在于:一是相较于实验环境,实际水下环境复杂许多,相应的样本数量与基于大数据的深度学习网络所需的量较少;二是在复杂声环境下的信号信噪比不佳,存在多目标或强干扰的

情况下,容易导致特征提取出现误差偏离;三是对于目标声信号本质的特征提取方法的研究还不够。

为了研究产生应用性差异的原因,本文选择了 2 种具有相关性且有较为明显差异的特征提取方法。作为目标应用最广泛的水声特征提取方法,梅尔倒谱系数特征提取方法(Mel-frequency cepstral coefficients, MFCC)可以将声音信号的实际频谱转化到感知频域中进行研究,从而有利于系统模拟人的感知过程,在性能和稳健性方面比较符合实际的听觉效果^[4]。而在 MFCC 基础上改进的 Gammatone 频率倒谱系数法(Gammatone frequency cepstrum coefficient, GFCC)在保有 MFCC 优点的同时,在水下复杂声场环境中,GFCC 较 MFCC 具有更优良的抗噪能力^[5]。

2006 年 Hinton 提出了以卷积神经网络(convolutional neural network, CNN)为主体的深度学习模型^[6]。CNN 具有自适应寻找最适合分类器的能力, CNN 由于采样层的存在,具有时移不变性的特征^[7]。与传统识别方法相比, CNN 通过寻找使识别效果最优的参数,使网络可以摆脱人工经验的影响,

收稿日期:2021-05-20

作者简介:吴晏辰(1989—),西北工业大学博士研究生,主要从事神经网络和水下声信号研究。

通信作者:王英民(1963—),西北工业大学教授,主要从事水声工程领域研究。e-mail:ywang@nwpu.edu.cn

对复杂的应用条件具有更好的适应能力。而基于 CNN 卷积核改进的残差神经网络(ResNet),充分利用了神经网络的函数拟合特性,对于水下目标的辐射信号而言,基于 CNN 的深度残差网络可以一定程度上获取常规特征分析方法难以发现的隐含关联,即作为特征提取器的同时,亦可作为分类器,因此非常适合处理原始水声信号。

本文给出了一个面对水下目标识别小样本数据的基于卷积神经网络的深层神经网络构成方法。利用 MFCC 和基于 MFCC 改进的 GFCC 作为特征提取方法,构建卷积神经网络与残差神经网络形成比对网络,建立了小样本下的水下目标识别系统,并对样本在 2 种网络的较深层结构下所取得的实际识别率与网络预测值进行统计比对。发现由于特征提取方法与深度学习网络的各种组合模式对于水声信号诸如信噪比等特性的匹配程度不同,其表现出的实际识别率和网络预测值有较为明显的关系,抗噪较好的组合模式的梯度衰减、网络识别稳定性和实际识别率较抗噪较差的组合具有更佳的网络稳定性和准确率,根据这一依据,为复杂神经网络神经元的设计提供了一种思路。

1 舰船辐射噪声的特征提取

1.1 MFCC 特征提取法

作为一种常见的声音识别特征,MFCC 特征提取方法主要分以下几步:

- 1) 分帧
- 2) 求能量谱

为了得到信号能量谱,需要利用快速傅里叶变换对帧信号进行处理,其公式为

$$p(f) = |X(f)|^2 = |\text{FFT}(x(n))|^2 \quad (1)$$

式中, $x(n)$ 为输入信号。

- 3) 滤波

将 $p(f)$ 通过梅尔滤波器组,其公式为

$$E(m) = \sum_{k=0}^{N-1} (p(f) \cdot H_m(f)) \quad (2)$$

式中: N 为帧数; $H_m(f)$ 为梅尔滤波器组系数。

- 4) 对数运算

对(2)式中 $E(m)$ 求导,得到 $E'(m)$

- 5) 求倒谱

对 $E'(m)$ 求离散余弦变换,可得 MFCC 参数 $C(n)$,其公式为

$$C(n) = \sum_{k=1}^M E'(m) \cos \left[\frac{\pi(k-0.5)n}{M} \right] \quad (3)$$

式中: $n = 1, 2, \dots, p$, p 为 MFCC 的阶数; M 为滤波器的个数。

则 MFCC 特征参数的一阶以及二阶差分系数由(3)式可得,其公式分别为

$$D(n) = \frac{1}{\sqrt{\sum_{i=-k}^k i^2}} \sum_{i=-k}^k i \cdot C(n+1) \quad (4)$$

$$D'(n) = \frac{1}{\sqrt{\sum_{i=-k}^k i^2}} \sum_{i=-k}^k i \cdot D(n+1) \quad (5)$$

(3)~(5)式共同构成帧特征向量

$$\mathbf{T}_n = \{C(n), D(n), D'(n)\} \quad (6)$$

(6)式即为样本信号的 MFCC。

为了使系统的性能以及稳健性都相对符合人听觉感知效果,利用 MFCC 特征参数分析方法将声音信号频谱转换到模拟听觉感知过程的感知频域。图 1 为梅尔频率和线性频率关系图,通过关系图可以得出以下分析:当频率在 1 000 Hz 以上时,Mel 频率的分布趋势为对数增长;而频率在 1 000 Hz 以下时,Mel 频率的分布趋势为带宽 100 Hz 左右的线性分布^[8]。

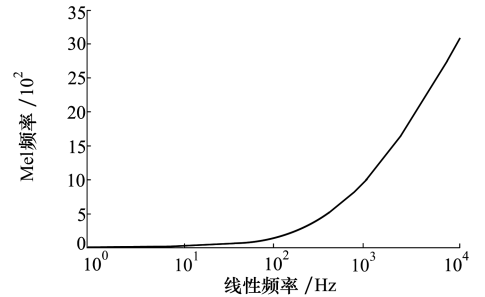


图 1 梅尔频率与线性频率的关系

利用倒谱变换进行各个滤波器组输出的幅度与能量之间的强相关性解耦,常见的处理滤波器组输出方法有对数压缩以及余弦变换,得到的 Mel 倒谱系数的方程如(7)式所示

$$c_m = \sqrt{\frac{2}{N}} \sum_{k=1}^N \ln(x_k) \cdot \cos \left(\frac{\pi m}{N} \left(k - \frac{1}{2} \right) \right) \quad (7)$$

式中: N 为滤波器的个数; m 为倒谱系数的维数。

1.2 GFCC 提取方法

图 2 为 GFCC 特征提取的主要流程。

3 改进的残差神经网络结构设计

基于 Inception 模块改进的数据池化层是由 GoogLetNet 在神经框架中首次采用的优化模块,后经过几次版本的迭代,一直到最新的 Inception-v4 和 Inception-resnet,每个版本在性能上都有一定的提升。通过建立小卷积层的串联化,当系统输出来到大卷积层时,基于堆叠替换的方法,拼接多种特征维度层面,来达到提升神经网络学习效率的目的。由于传统残差神经网络的目标是实现百万级以上的大训练样本及多 GPU 运行的复杂训练,在数据输入时,会将输入到池化层的数据通过 7×7 的大型卷积层和池化层。在面对具有时变性强的如辐射噪声等类型的时域信号时,如果有效样本有限,则很难利用大卷积层完成数据堆叠,无法有效提升网络的学习效率。因此本文针对基于 Inception 的模块特点设计了一种改进的数据池化层,这种池化层在面对小样本数据时具有较强的特征提取能力。

为了改进数据池化层,将传统的 7×7 的卷积层替代为 3 个 3×3 的小卷积层,并将每个通道数分别设置为 8,16,16。为了有效提升神经网络的特征提取速度,在每个 3×3 的小卷积层后面都加入了批量标准化和 ReLU 激活函数。传统加入单一 ReLU 激活函数的方式,会随着网络层数加深出现过拟合的问题,为了解决这个问题,给每个小卷积层都加入了一个 ReLU 激活函数,使网络对小样本特征信息的提取能力更强,之后再在 Concat 层将特征维度拼接到一起。为了避免小样本下网络层数加深可能导致训练过程中过早出现过拟合问题,将改进的残差连接结构加入到数据池化层中,再从最大池化层的输出中提取得到数据特征信息。

4 系统实验分类结果和分析

在本研究中,实地采集舰船辐射噪声样本数据 1 756 份,取自互联网的海洋生物噪声样本数据 877 份,利用高斯模拟或水下实测的其他水下噪声样本 6 410 份。每个样本都被裁剪为大小 40 帧,帧移 15 ms的块,由这些块所组成的数据集统称为样本集(epoch),当样本集过大或根据研究需要时,可以将样本集按照一定数量进行分批(batch),每批大小

(batch size)固定。在对样本集做过批处理后,将这些样本分别采用经过改进后的 GFCC 特征和 MFCC 特征进行训练,并使用深层残差神经网络进行目标分类。

在得到匹配本研究的神经元组成结构后,将 MFCC 和 GFCC 与残差网络组成的神经元分别进行全样本训练,在多次训练后,对输出求取均值后得到神经网络分类结果,见表 1~2。

表 1 Mel 频率倒谱系数特征提取后样本识别统计结果

地点	编号	类别	样本数	识别数	正确率/%
水库	A	小型船只	480	464	96.7
	B	大型船只	240	229	95.4
海洋	A	小型船只	681	564	82.8
	B	大型船只	355	286	80.6
	C	海洋生物	877	854	97.4

表 2 CFCC 特征提取后样本识别统计结果

地点	编号	类别	样本数	识别数	正确率/%
水库	A	小型船只	480	471	98.1
	B	大型船只	240	233	97.1
海洋	A	小型船只	681	602	88.4
	B	大型船只	355	315	88.7
	C	海洋生物	877	864	98.5

在对分类结果做分析时,由于检测点较多,所以仅展示前 10 个检测点在前 10 层网络下的网络预测率和实际识别率。需要说明的是由于动物叫声具有比较显著的声学特征表现^[12],考虑到对比的特征较多,工作量较大,所以在样本集中加入了动物叫声作为目标样本,以在合理范围内降低运算参数,增加研究效率。

利用 MFCC 作为特征提取方法应用在 CNN 网络中的神经网络测试结果,如图 4 所示。图中网络深度表示神经网络的层数,在实验时将特征提取在深度达到每一层时的预测结果标记在图上形成曲线。根据结果可以发现,在网络深度达到 5 层时,10 个预测点的网络预测值趋近于 100%,再将被网络贴上标签的识别样本与实际样本集进行对比,得到实际正确率,如图 5 所示。实验结果表明实际识别率在网络达到 6 层时出现了明显衰减,且在整个检测过程中实际正确率一直与网络预测值存在着一定

的差值,其差值在 6 层时出现最小值,约为 25% 左右。

将基于 MFCC 改进的 GFCC 特征提取法应用于 CNN 网络中,其网络预测值和实际识别率如图 6 和 7 所示,可以发现在网络深度达到 8 层时,网络预测值趋近于 100%,与实际识别率相对比,其差值为 10% 左右,其实际识别率在深度达到 9 层时出现了明显衰减。

通过研究,对产生这种现象的原因做了如下几点推测,首先,神经网络由于其计算方式不是线性

的,在小样本情况下,每一次样本输入进入卷积核后,通过反复卷积-池化-反向传播的计算会产生多个带有不同标签的结果,只有在样本大小匹配网络结构的情况下,网络对权重的选择才会逐渐趋向固定,输出值才会趋向于唯一。另外由于梯度弥散问题,当神经网络深度达到一定值时,梯度衰减过快会导致卷积层内的部分卷积核出现无效化从而出现“卷积核死亡”的问题,可能正是因为这种线性导致了在某层出现的识别率快速衰减的结果。

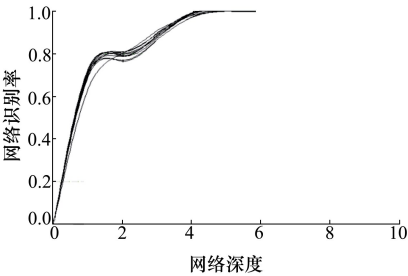


图 4 基于 Mel 频率倒谱系数特征提取的卷积神经网络预测值

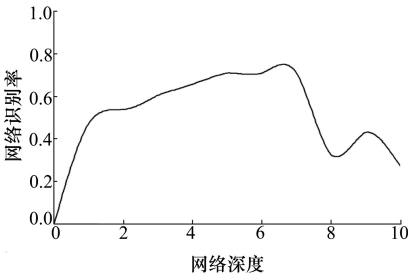


图 5 基于 Mel 频率倒谱系数特征提取的卷积神经网络实际识别率

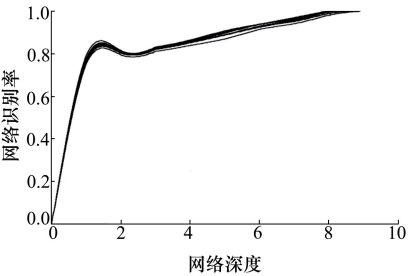


图 6 基于 Gammatone 频率倒谱系数特征提取的卷积神经网络预测值

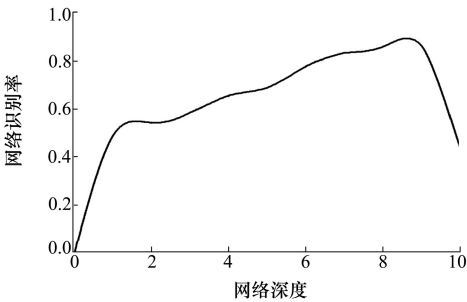


图 7 基于 Gammatone 频率倒谱系数特征提取的卷积神经网络实际识别率

因此,在研究对象为小样本的客观基础上,很容易得到一种直观判断神经元构成的标准,也就是利用输出结果平稳性、神经元死亡深度来判断特征提取法和网络关于本研究的匹配程度,从而决定最终神经元的组成结构。

在文献[11]中发现 2015 年研发成功的残差神经网络(ResNet)具有比较好的抗噪抗弥散性能,且常被用于(相对)小样本问题的处理上,因此尝试了几种特征提取与残差神经网络的融合,同 CNN 网络一样,本文将前 10 个检测点的 10 层网络识别数据分别展现如下。

首先是将 MFCC 与 ResNet 进行融合,其网络预测值和实际识别率如图 8 和图 9 所示,由预测结果图可以看出,当网络深度为 7 层时,ResNet 网络预测值趋近 100%,而与实际识别率相对比,其差值最小为 20% 左右,实际识别率在深度达到 7 层时出现了明显的衰减。

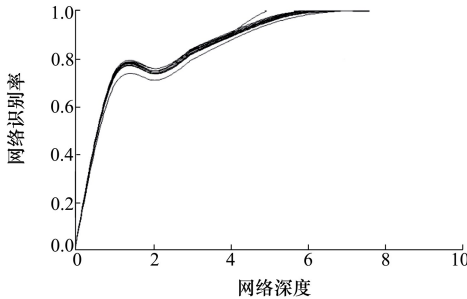


图 8 基于梅尔频率倒谱系数特征提取的残差网络预测值

将 GFCC 与 ResNet 融合,其网络预测值和实际识别率如图 10 和图 11 所示,由预测结果图可以看出,该网络在深度为 10 时虽然接近 100%,但仍未到饱和程度,而与实际识别率相对比,其差值为 10% 左右,为了研究该网络的衰减层数,将该网络的深度加深到 15 层,可以明显看出,基于 GFCC 的残差网

络在 12 层时出现了衰减,其实际识别率差值小 于 10%。

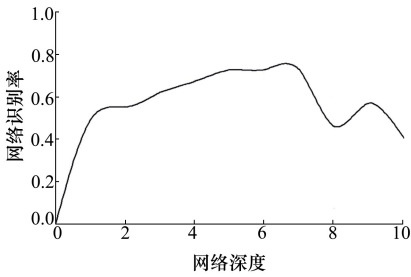


图 9 基于 Mel 频率倒谱系数特征提取的残差网络实际识别率

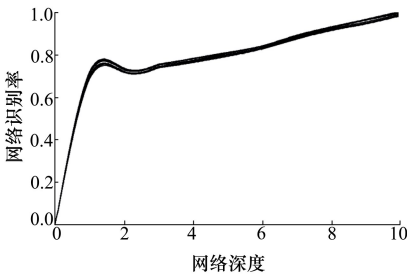


图 10 基于 Gammatone 频率倒谱系数特征提取的网络预测值

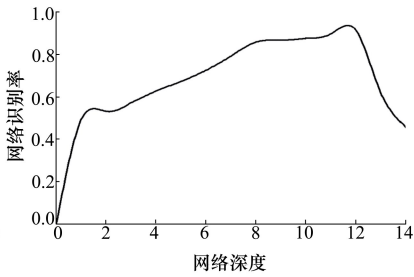


图 11 基于 Gammatone 频率倒谱系数特征提取的实际识别率

5 结 论

通过研究发现,面向水下目标识别的深度神经网络,其识别稳定性与识别准确率与目标特征提取方法、所选用神经网络模型之间有密切的联系。与常见的浅层水下目标识别神经网络系统相比,利用如改良的残差神经网络等具有较强抗网络退化功能的网络结构,同时选择匹配的特征提取法所组成的

网络,可以有效增加网络的深度,提高小样本利用率,增强系统对于多种目标在水下复杂环境中的识别效率。在使用 GFCC 作为特征提取方法的残差神经网络中,深层网络的识别率和系统稳定性都显著强于使用其他多种组合。这说明,对于基于小样本的水下目标识别系统,在选择特征提取方法和神经网络模型时,应当择优考虑抗噪强,梯度衰减弱,具有强抗退化的种类,以达到网络深层化,以及充分利用有限样本内的各类目标信息的目的。

参考文献:

[1] 方尔正,黄志浩,桂晨阳,等. 水面水下目标识别技术的现状与挑战[J]. 国防科技工业, 2020(7): 66-68
FANG Erzheng, HUANG Zhihao, GUI Chenyang, et al. Current situation and challenges of surface and underwater target recognition technology[J]. National Defense Technology Industry, 2020(7): 66-68 (in Chinese)

[2] 李兰瑞,李鹏,刘天宇,等. 水声信号检测与识别技术研究现状[J]. 通信技术, 2020, 53(12): 2904-2907
LI Lanrui, LI Peng, LIU Tianyu, et al. Research status of underwater acoustic signal detection and recognition technology[J]. Communications Technology, 2020, 53(12): 2904-2907 (in Chinese)

[3] 陈阳. 水声阵列信号处理理论及实验研究[D]. 哈尔滨: 哈尔滨工程大学, 2010: 10-18
CHEN Yang. Theoretical and experimental study on underwater acoustic array processing[D]. Harbin: Harbin Engineering University, 2010: 10-18 (in Chinese)

[4] 吴姚振. 水下目标辐射噪声生理感知特征的提取方法研究[D]. 西安: 西北工业大学, 2014
WU Yaozhen. Studies on physiological perception feature extraction methods in underwater target-radiated noise[D]. Xi'an: Northwestern Polytechnical University, 2014 (in Chinese)

[5] 吴晏辰,王英民. 基于 Gammatone 频率倒谱系数的舰船辐射噪声分析[J]. 水下无人系统学报, 2021, 29(1): 60-64
WU Yanchen, WANG Yingmin. Ship-radiated noise analysis based on the Gammatone frequency cepstrum coefficient[J]. Journal of Unmanned Undersea Systems, 2021, 29(1): 60-64 (in Chinese)

[6] Hinton G E, Salakhutdinov R R. Reducing the dimensionality of data with neural networks[J]. Science, 2006, 313(5786): 504-507

[7] 王强,曾向阳. 深度学习方法及其在水下目标识别中的应用[C]//中国声学学会水声学分会 2015 年学术会议, 2015
WANG Qiang, ZENG Xiangyang. Deep learning methods and their applications in underwater targets recognition[C]//Proceedings of the 2015 Academic Conference of the Hydroacoustics Branch of the Acoustical Society of China, 2015 (in Chinese)

[8] 程锦盛,杜选民,周胜增,等. 基于目标 MFCC 特征的监督学习方法在被动声纳目标识别中的应用研究[J]. 舰船科学

技术, 2018, 40(17): 116-121

CHENG Jinsheng, DU Xuanmin, ZHOU Shengzeng, et al. Application of the MFCC feature based supervised learning method in passive sonar target recognition[J]. Ship Science and Technology, 2018, 40(17): 116-121 (in Chinese)

[9] 王继霞. 基于卷积神经网络的瓶状垃圾检测研究[D]. 哈尔滨: 哈尔滨理工大学, 2020

WANG Jixia. Research on bottle waste detection based on convolution neural network[D]. Harbin: Harbin University of Science and Technology, 2020 (in Chinese)

[10] 徐及, 黄兆琼, 李琛, 等. 深度学习在水下目标被动识别中的应用进展[J]. 信号处理, 2019, 35(9): 1460-1475

XU Ji, HUANG Zhaoqiong, LI Chen, et al. Advances in underwater target passive recognition using deep learning[J]. Journal of Signal Processing, 2019, 35(9): 1460-1475 (in Chinese)

[11] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition, 2016

[12] 周卓, 梁红, 杨长生, 等. 基于海豚听觉系统模型的水下目标识别[J]. 水下无人系统学报, 2021, 29(2): 147-152

ZHOU Zhuo, LIANG Hong, YANG Changsheng, et al. Underwater target identification based on dolphin auditory system model [J]. Journal of Unmanned Undersea Systems, 2021, 29(2): 147-152 (in Chinese)

A research on underwater target recognition neural network for small samples

WU Yanchen, WANG Yingmin

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

Abstract: In the face of the challenges in the field of marine engineering applications in the new era, the goal of automation, high efficiency and accuracy can be achieved by using deep learning-based neural networks in hydroacoustic engineering. However, in the face of objective problems such as the lack of underwater target samples, the complex underwater sound environment, and the poor sample signal-to-noise ratio, the deep learning also becomes less sensitive due to its own limitations. In this paper, by constructing a variety of target feature extraction methods and a deep neural network model, we obtain the target recognition rate network prediction value after matched different target feature extraction with neural network model. Through comparing experimental results, a new idea of solving small sample target identification through deep neural network deep design is proposed.

Keywords: underwater target identification; deep learning; deep neural network design

引用格式: 吴晏辰, 王英民. 面向小样本数据的水下目标识别神经网络深层化研究[J]. 西北工业大学学报, 2022, 40(1): 40-46

WU Yanchen, WANG Yingmin. A research on underwater target recognition neural network for small samples[J]. Journal of Northwestern Polytechnical University, 2022, 40(1): 40-46 (in Chinese)