

【引用格式】曹琳. 基于极限学习自编码器的水声信号目标识别方法[J]. 数字海洋与水下攻防, 2024, 7(2): 225-230.

基于极限学习自编码器的水声信号目标识别方法

曹琳

(水下测控技术重点实验室, 辽宁 大连 116013)

摘要 传统的机器学习方法在特征提取时容易受到主观经验的影响, 导致对水声信号目标的识别准确率不高。而一般深度学习算法模型较复杂, 通常具有训练耗时、计算复杂度高等缺点。极限学习自编码器具有很强的非线性处理能力, 适合针对具有非线性特点的水声信号目标的识别, 而且模型具有学习速度快, 泛化能力强等显著优势。将极限学习自编码器算法应用于水声信号目标识别中, 并与卷积神经网络、自编码器和极限学习机识别方法进行对比, 结果表明: 提出的方法对水声信号目标识别的准确率最优, 且训练时间较短。

关键词 水声信号目标识别; 极限学习自编码器; 卷积神经网络

中图分类号 TB556

文献标识码 A

文章编号 2096-5753(2024)02-0225-06

DOI 10.19838/j.issn.2096-5753.2024.02.012

A Target Recognition Method for Underwater Acoustic Signals Based on Extreme Learning Autoencoder

CAO Lin

(Underwater Test and Control Laboratory, Dalian 116013, China)

Abstract Traditional machine learning methods are easily influenced by subjective experience during feature extraction, which leads to low recognition accuracy of underwater acoustic targets. However, deep learning algorithm models are relatively complex, which usually have the disadvantages of time-consuming training and high computational complexity. Extreme learning autoencoder has strong non-linear processing ability, which is suitable for recognition of underwater acoustic signals with nonlinear characteristics. Moreover, the model has significant advantages such as fast learning speed and strong generalization ability. In this paper, the extreme learning autoencoder algorithm is applied to underwater acoustic signal recognition, and is compared with convolutional neural networks, autoencoders, and extreme learning machine recognition methods. The results show that the proposed method has the best accuracy in target recognition of underwater acoustic signals and needs shorter training time.

Key words underwater acoustic signal target recognition; extreme learning autoencoder; convolutional neural network

0 引言

水声信号的识别问题一直以来都是水声工程领域研究的重点和难点。国内外的研究人员长期以

来一直致力于研究应用机器学习理论进行水声目标识别的方法, 即在研究水声信号的非线性、非平稳等特性的基础上^[1-2], 提取了水声信号的时域特征、频域特征、时频特征等, 并利用支持

收稿日期: 2023-12-08

作者简介: 曹琳(1985-), 女, 硕士, 工程师, 主要从事水声信号处理、大数据分析挖掘研究。

向量机、人工神经网络等分类器来实现水声信号的目标分类识别。然而，传统的机器学习方法在特征提取过程中需要较高质量的数据并容易受到主观经验的影响，导致水声信号目标识别的效果不理想。

近年来，深度学习方法在处理非线性信号方面的研究取得了很大的进展，国内外学者纷纷将深度学习方法应用到水声信号目标识别中^[3-7]。深度学习是一种利用非线性信息处理技术实现多层次、监督或无监督的特征学习和转换，并进行模式分析和分类的机器学习理论和方法。可以自适应地学习水声信号中的深度特征并进行分类识别，不需要依赖大量先验知识。然而，一般的深度学习网络为了学习更深层的特征，往往具有模型训练耗时、计算复杂度高等缺点。KASUN 等人^[8]提出了极限学习机自编码器的特征表示方法，极限学习机自编码器具有非常棒的非线性处理能力，且具有模型学习的速度快、精度高、人为干扰最小等显著优势^[9-10]，利用该方法能够快速地建立水声信号与目标之间的非线性关联关系，目前尚未在公开资料中见到其应用在水声信号目标识别领域中。因此，本文将极限学习自编码器算法应用于水声信号目标识别中，以现有的实测水声信号数据为例，对输入为不同帧长的水声信号进行训练和测试，并对识别准确率和训练时间进行对比，结果表明，该方法可以有效的提升对水声信号目标的识别性能。

1 模型算法原理

1.1 极限学习机算法原理

2004 年，南洋理工大学的黄广斌教授提出了极限学习机 (ELM) 算法^[11]，该算法能够改善传统基于梯度的单隐层网络中存在的训练速度慢、局部最小值、学习率选择敏感等问题^[12-16]。ELM 算法的主要思想是只要激活函数可以无穷地可微，输入权值和隐藏层偏置值就能随机设置。一旦在训练开始这些参数被赋予随机的值，那么隐藏层的输出矩阵 H 就保持不变了，网络的训练相当于只是简单的求取矩阵的广义逆而无需反复的迭代。因此 ELM 不仅能达到最小的训练误差，使

得训练速度极快，而且能达到输出权值的最小范数，因此有着较强的泛化能力。ELM 的网络结构如图 1 所示。

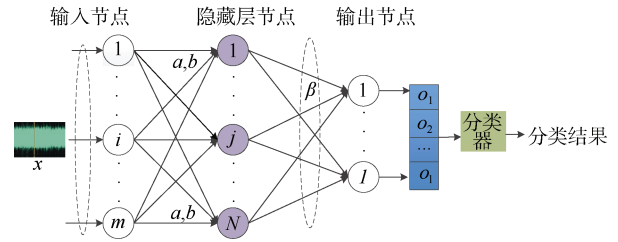


图 1 ELM 模型结构图
Fig. 1 ELM model structure diagram

极限学习机建立于单隐层神经网络基础之上，对于 N 个不同样本 $\{(x_i, o_i) | x_i \in \mathbf{R}^m, o_i \in \mathbf{R}^l\}_{i=1}^N$ ，一个隐层神经元节点的数目为 N ，激励函数为 $g(x)$ 的 ELM 网络数学模型可表示为

$$\sum_{i=1}^{\tilde{N}} \beta_i g_i(x_j) = \sum_{i=1}^{\tilde{N}} \beta_i g(a_i \cdot x_j + b_j) = o_j, j=1, \dots, N$$

式中: $a_i = [a_{i1}, a_{i2}, \dots, a_{im}]^T$ 表示连接第 i 个隐层节点的输入权, b_i 表示第 i 个隐层节点的偏差, $\beta_i = [\beta_{i1}, \beta_{i2}, \dots, \beta_{in}]^T$ 表示连接第 i 个隐层节点的输出权, o_j 表示第 j 个样本对应的 ELM 网络输出, 其矩阵形式为: $H\beta = O$, 其中 H 表示隐藏层的输出矩阵, 即 $H = [h_1, h_2, \dots, h_N]$ 。模型训练的关键取决于在极短的时间内确定网络的结构以及网络的各项参数: $W = (a, b, \beta)$, 即

$$\begin{aligned} \underset{W=(a,b,\beta)}{\operatorname{argmin}} E(W) &= \underset{W=(a,b,\beta)}{\operatorname{argmin}} \|\epsilon\|^2 \\ \text{s.t.} \sum_{i=1}^{\tilde{N}} \beta_i g(a_i \cdot x_j + b_j) - o_j &= \epsilon_j, j=1, \dots, N \end{aligned}$$

黄广斌教授通过理论推导证明: 给 ELM 网络的输入权值 a 随机赋值, 将 b_i 全部赋值为 0, 只需隐层节点足够多 (最高为学习样本总数), 隐层输出矩阵总为满秩矩阵, ELM 有精确解, 可直接求解输出权值矩阵 β 。

1.2 自编码器算法原理

自编码器 (AE) 是一种无监督学习算法, 由编码器部分和解码器部分构成^[17], 编码器可以将输入数据进行压缩, 到隐藏层之后重新解压并映射回输出层, 如图 2 所示。

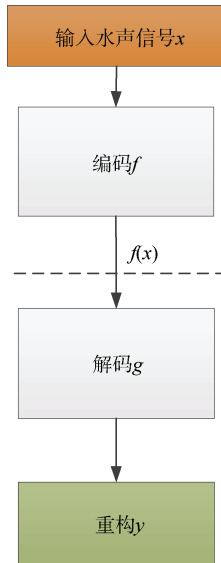


图 2 自编码器模型结构图

Fig. 2 Autoencoder model structure diagram

式中, 编码函数为: $f(x) = \beta_f(\mathbf{w}x + b)$, β_f 为编码器的激活函数, 通常为 sigmoid 激活函数, \mathbf{w} 为输入层和隐藏层之间的权重。解码函数为

$$g(f(x)) = \beta_g(\mathbf{w}^T f(x) + c)$$

式中, β_g 为解码器的激活函数, \mathbf{w}^T 为隐藏层与输出层间的权重。由于自编码器是一种无需标记的无监督特征学习方法^[18], 可以降低数据的维度并能够

有效的学习水声信号数据中的深层特征, 从而提取鲁棒的声纹特征。

1.3 极限学习自编码器算法原理

极限学习机作为一种单隐层神经网络, 具有较高的分类准确度和较强的非线性数据拟合能力。在将极限学习机用于水声信号的目标识别时发现, 对于小样本识别问题, 极限学习机的泛化能力相对较差, 此外, 由于输入权值采用随机赋值方式, 导致极限学习机性能不稳定。自编码器在不需标签数据的情况下经过训练可以将输入数据压缩为潜在空间表示, 从潜在的空间表示中重构输入数据, 这样做的好处是可以学习水声信号中的非线性特征。

因此, 本文考虑结合极限学习机和自编码器各自的优点, 将极限学习自编码器 (ELM-AE) 应用到水声信号目标识别问题中, ELM-AE 模型结构如图 3 所示。该模型像 ELM 一样, 也由输入层、隐含层、输出层组成, 对于单隐层神经网络, ELM-AE 模型将输入数据作为输出数据, 其目标是输入 X 的重构误差, 这样可以学习到原始水声信号数据高级的特征表示。此外, ELM-AE 选择随机生成的隐藏节点的权值和偏差正交, 不仅能提高模型的训练速度, 还能获得全局最优解^[19-21]。

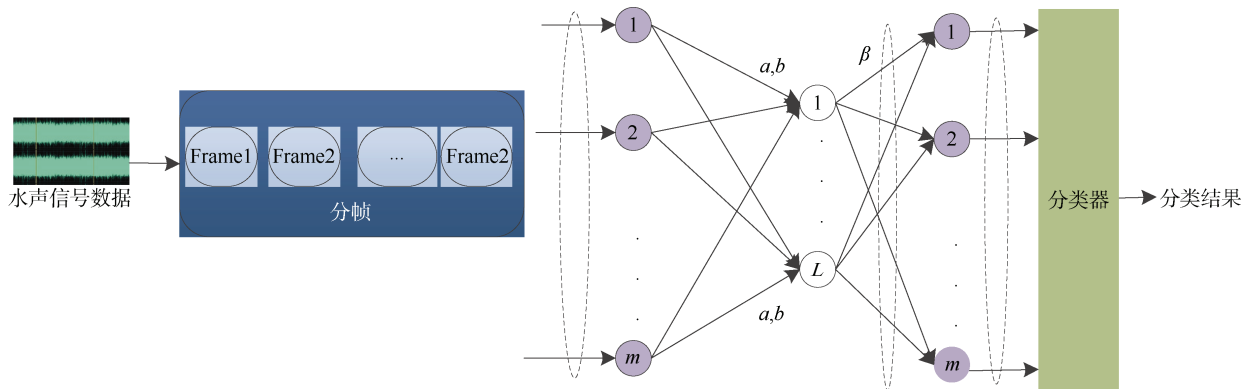


图 3 ELM-AE 模型结构图

Fig. 3 ELM-AE model structure diagram

ELM-AE 的输出可以用如下表达式表示:

$$x_j' = \sum_{i=1}^L \beta_i g(a_i \cdot x_j + b_i), j = 1, 2, \dots, N$$

式中: a_i 和 b_i 与 ELM 网络表示的意思相同, β_i 输出层权值的计算与 ELM 相似, 公式为^[22]

$$\beta = \left(\frac{1}{C} + \mathbf{H}^T \mathbf{H} \right)^{-1} \mathbf{H}^T X$$

2 模型实验与结果分析

2.1 实验环境配置与数据样本分布

本文实验的运行环境为 ubuntu16.04 操作系

统, GPU: A6000, 内存: 512G, 编程语言: python3.5.3。

将现有的实测船舶水声音频信号数据整理成 A、B 两类, 其中每类目标包括若干种不同的航行工况, 样本分布情况如表 1 所示, 每个时序样本的时间长度为 3 s, 采样率为 28 000 Hz, 样本幅值经归一化处理在 [-1, 1] 范围内, 在实验过程中选取数据集中的 80% 作为训练集, 20% 作为测试集。

表 1 水声信号数据样本分布情况
Table 1 Sample distribution of underwater acoustic signals

数据类型	工况种类	样本个数
A	14 种	531
B	21 种	606

2.2 模型训练与结果分析

本文中的 ELM-AE 网络模型结构设置为 128-64-32-64-128-40, 如表 2 所示, 其中 128-64-32-64-128 为 AE 网络的模型结构, 编码块的结构为 128-64, 32 为中间隐层节点, 64-128 为解码块的结构, 此外, ELM 结构的隐层节点个数为 40。

表 2 ELM-AE 网络模型结构
Table 2 The ELM-AE network model structure

网络结构	ELM 隐层节点个数
128-64-32-64-128-40	40
128-64-32-64-128-40	40
128-64-32-64-128-40	40
128-64-32-64-128-40	40
128-64-32-64-128-40	40

在模型的训练阶段, 首先将水声信号数据送入 AE 网络以实现水声信号潜在特征的提取, 误差函数采用均方根误差, 然后利用 ELM 网络对 AE 的参数进行优化, ELM-AE 模型的最大迭代次数设置为 100, 学习率设置为 0.001, 激活函数采用 ReLu, 当模型经过 100 次的迭代之后, 得 ELM-AE 模型的训练损失曲线, 如图 4 所示, 从图中可以看到, ELM-AE 模型在训练时拟合效果不错, 并未出现过拟合和欠拟合的现象。

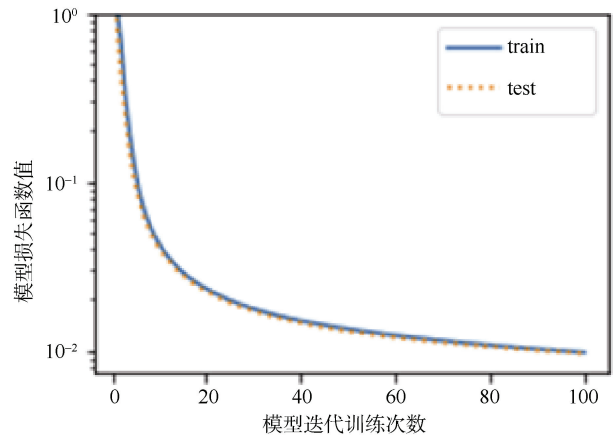


图 4 ELM-AE 模型训练损失曲线图

Fig. 4 Training loss curve diagram of ELM-AE model

此外, 本文分别将水声信号的测试数据送入到训练好的 ELM、AE、CNN 和 ELM-AE 模型中进行识别性能验证。

表 3 列出了不同帧长的原始水声信号作为 ELM 的输入信号, ELM 的网络结构包括输入层、一个隐藏层 (隐藏层节点个数设置如表 3 所示) 和输出层, 用 ELM 模型进行分类识别的识别效果。

表 3 ELM 模型识别结果
Table 3 Recognition results of ELM model

数据长度/s	ELM 隐层节点个数	训练时间/s	识别率/%
0.5	40	8.7	76.68
1	40	8.9	87.12
3	20	9.5	88.40
3	40	10.2	90.04
3	60	12.6	89.29
5	40	12.9	87.22
5	60	13.5	86.56
10	40	14.2	85.74

通过上述表格以及大量的实验可以得知: 选择长度为 3 s 的水声信号测试数据作为极限学习机的输入最为合适, 且 ELM 隐藏层节点个数设置 40 的水声信号目标的识别准确率最高, 达到 90.04%, 同时, 模型的训练速度很快, 训练时间达到 10.2 s。

表 4 中列出应用 AE 对水声信号特征学习并进行分类识别的结果。

表 4 AE 模型识别结果

Table 4 Recognition results of AE model

数据长度/s	网络结构	训练时间/s	识别率/%
0.5	128-64-32-64-128	9.3	75.24
1	128-64-32-64-128	9.8	76.33
3	128-64-32-64-128	11.4	86.67
5	128-64-32-64-128	13.1	80.12
10	128-64-32-64-128	16.5	78.75

在输入水声信号为 3 s 时长的测试数据时, 应用 AE 模型的最优识别结果为 86.67%, 而应用极限学习机分类识别结果为 90.04%。结果表明, 极限学习机分类识别性能优于 AE, 模型的训练速度略高于 AE 网络。

表 5 中列出应用 CNN 对水声信号特征学习并进行分类识别的结果。

表 5 CNN 模型识别结果

Table 5 Recognition results of CNN model

数据长度/s	网络结构	训练时间/s	识别率/%
0.5	64C-2P-128C-2P-512C-100C	52.9	90.24
1	64C-2P-128C-2P-512C-100C	55.8	92.33
3	64C-2P-128C-2S-512C-100C	65.7	95.67
5	64C-2P-128C-2P-512C-100C	69.2	87.82
10	64C-2P-128C-2P-512C-100C	73.5	89.75

在上表卷积网络结构中, C 和 P 分别代表卷积层和池化层, 两者前面的数字代表卷积层或者池化层特征图的个数。在池化层中, 第 1 个池化层选择最大值池化, 第 2 个池化层选择均值池化。在输入水声信号为 3 s 时长的情况下, 应用 CNN 的最优识别结果为 95.67%, 优于 ELM 的识别率 90.04% 和 AE 的识别率 86.67%。结果表明, 虽然 CNN 的识别率较高, 但是它的训练速度远低于 ELM 和 AE。

表 6 中列出应用 ELM-AE 对水声信号特征学习并进行分类识别的结果。

表 6 ELM-AE 模型识别结果

Table 6 Recognition results of ELM-AE model

数据长度/s	网络结构	ELM 隐层节点个数	训练时间/s	识别率/%
0.5	128-64-32-64-128-40	40	11.2	78.21
1	128-64-32-64-128-40	40	12.4	89.58
3	128-64-32-64-128-40	40	13.6	96.25
5	128-64-32-64-128-40	40	14.8	88.81
10	128-64-32-64-128-40	40	20.2	87.46

在输入测试数据时长为 3 s 情况下, ELM 隐层节点个数为 40 时, 应用 ELM-AE 分类识别结果为 96.25%, 高于 ELM、AE 和 CNN 网络, 模型的训练时间为 13.6 s, 远高于 CNN 网络。

3 结束语

本文通过利用含有不同工况的 A、B 两类船舶水声信号数据集作为实验对象, 将极限学习自编码器应用到水声信号的目标识别中, 并与 ELM、AE 和 CNN 模型的识别性能进行对比。实验结果表明, 在选择输入信号时长为 3 s 的情况下, 本文提出方法的识别准确率为 96.25%, 高于 ELM、AE 和 CNN, 模型的训练速度较快, 远高于 CNN 网络。能够有

效地对水声信号目标进行识别。

本文的不足之处在于所使用的水声信号数据集样本量较少, 数据种类不够丰富, 使得模型结果的可靠性降低, 下一步需增加水声数据的种类和数量, 以提高模型的泛化能力。

参考文献

- [1] 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): 954-965.
- [2] 张奇, 笄良龙, 王超, 等. 基于深度学习的水声被动目标识别研究综述[J]. 电子与信息学报, 2023, 45(11): 4190-4202.
- [3] 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.
- [4] 王升贵, 胡桥, 陈迎亮, 等. 基于深度学习的水下目标识别方法研究[J]. 舰船科学技术, 2020, 42(12): 141-145.
- [5] 刘峰, 罗再磊, 沈同圣, 等. 时频谱图和数据增强的水声信号深度学习目标识别方法[J]. 应用声学, 2021, 40(4): 518-524.
- [6] 薛灵芝, 曾向阳. 动态水声环境中的 SE_ResNet 模型目标识别方法[J]. 哈尔滨工程大学学报, 2023, 44(6): 939-946.
- [7] 杨宏晖, 李江涛, 申昇, 等. 用于水下目标识别的无监督特征选择算法[J]. 舰船科学技术, 2017, 39(12): 92-94.
- [8] KASUN L L C, YANG Y, HUANG G B, et al. Dimension reduction with extreme learning machine[J]. IEEE Transactions on Image Processing, 2016, 25(8): 3906-3918.
- [9] 刘楠, 刘翠玲, 徐金阳, 等. 基于极限学习机自编码算法的近红外光谱模型传递的研究[J]. 食品安全质量检测学报, 2023, 14(5): 30-36.
- [10] ZHANG J, YIN J, ZHANG Q, et al. Robust sound event classification with bilinear multi-column ELM-AE and two-stage ensemble learning[J]. EURASIP Journal on Audio, Speech, and Music Processing, 2017, 2017(1): 1-11.
- [11] HUANG G B, ZHU Q Y, SIEW C K. Extreme learning machine: theory and applications[J]. Neurocomputing, 2006, 70(1-3): 489-501.
- [12] HUANG G, SONG S J, GUPTA J N D, et al. Semi-supervised and unsupervised extreme learning machines[J]. IEEE Transactions on Cybernetics, 2017, 44(12): 2405-2417.
- [13] HUANG G, HUANG B B, SONG S J, et al. Trends in extreme learning machines: a review[J]. Neural Networks, 2015, 61: 32-48.
- [14] 温柔. 基于极限学习机的目标检测方法研究[D]. 西安: 西安电子科技大学, 2021.
- [15] 王雪松, 赵季娟, 程玉虎. 基于极限学习机的无监督领域适应分类器[J]. 控制与决策, 2020, 35(4): 861-868.
- [16] 熊檬. 基于极限学习机与深度学习理论的滚动轴承故障诊断研究[D]. 无锡: 江南大学, 2020.
- [17] 贾文娟, 张煜东. 自编码器理论与方法综述[J]. 计算机系统应用, 2018, 27(5): 1-9.
- [18] 来杰, 王晓丹, 向前, 等. 自编码器及其应用综述[J]. 通信学报, 2021, 42(9): 218-228.
- [19] 许鸿奎, 姜彤彤, 李鑫, 等. 结合降噪自编码与极限学习机的 LTE 上行干扰分析[J]. 吉林大学学报: 工学版, 2022, 52(1): 195-203.
- [20] 陈媛, 陈晓云. 流形极限学习机自编码特征表示[J]. 计算机工程与应用, 2020, 56(17): 150-154.
- [21] 张卫东, 路皓翔, 甘博瑞, 等. 基于栈式自编码融合极限学习机的药品鉴别[J]. 计算机工程与设计, 2019, 40(2): 545-549.
- [22] 程蓉, 白艳萍, 胡红萍, 等. 含类信息的极限学习机自动编码器特征学习方法[J]. 电子测量技术, 2022, 45(16): 71-79.

(责任编辑: 肖楚楚)