

一种用于 Bi-LSTM 神经网络信号识别的 DO-CAB 算法

花国祥^{1,2,3}, 汤炼海³, 李伟伟^{1,3}, 李 鹏^{1,3}, 孙 炎³

(1.无锡学院 自动化学院, 江苏 无锡 214105;

2.华北电力大学 电气与电子工程学院, 北京 102206; 3.南京信息工程大学, 南京 21004)

摘要:针对双向工频通信系统(TWACS)存在上行信号识别准确率不足的问题,提出一种基于蒲公英优化(DO)算法的联合卷积神经网络(CNN)与注意力机制(AM)的双向长短时记忆(Bi-LSTM)神经网络信号识别算法,简称 DO-CAB 算法。该算法首先通过 CNN 自适应提取 TWACS 信号重要特征,然后使用 DO 算法优化 Bi-LSTM 超参数,根据优化的超参数构建网络,并引入 AM 赋予输入影响权重,以获得更好信号识别效果。实验结果表明,所提算法的识别准确率达到 92.32%,能高效、准确识别 TWACS 调制信号。

关键词:双向工频通信系统;蒲公英优化算法;双向长短时记忆网络;深度学习;信号检测

中图分类号:TP183; TN929 **文献标志码:**A **文章编号:**1002-5561(2024)06-0023-05

DOI:10.13921/j.cnki.issn1002-5561.2024.06.005

开放科学(资源服务)标识码(OSID):

DO-CAB algorithm for
Bi-LSTM neural network signal recognitionHUA Guoxiang^{1,2,3}, TANG Lianhai³, LI Weiwei^{1,3}, LI Peng^{1,3}, SUN Yan³

(1. School of Automation, Wuxi University, Wuxi Jiangsu 214105, China; 2. School of Electrical and Electronic Engineering, North China Electric Power University, Beijing 102206, China; 3. Nanjing University of Information Science & Technology, Nanjing 210044, China)

Abstract: To address the problem of insufficient recognition accuracy for uplink signals in two-way automatic communication systems(TWACS), a signal recognition algorithm based on the dandelion optimization(DO) algorithm that combines convolutional neural network(CNN) with attention mechanism(AM) and bidirectional long short-term memory(Bi-LSTM) neural networks is proposed, which is briefly referred to as the DO-CAB algorithm. The algorithm first adaptively extracts important features of TWACS signals using a CNN. It then optimizes the hyperparameters of the Bi-LSTM using the DO algorithm, constructs the network based on the optimized hyperparameters, and introduces an AM to assign influence weights to the inputs, improving the network algorithm for better signal recognition. The experimental results show that the proposed algorithm achieves a recognition accuracy of 92.32%, enabling efficient and accurate identification of TWACS modulated signals.

Key words: two-way automatic communication systems, dandelion optimization algorithm, bidirectional long short-term memory network, deep learning, signal detection

0 引言

当前,电力线通信技术主要包括电力线载波通信、电力线双向工频通信、脉冲控制载波通信、超窄带

收稿日期:2024-07-13。

基金项目:新能源电力系统全国重点实验室开放课题项目(LAPS23014)资助。

作者简介:花国祥(1985—)男,江苏淮安人,博士生,副教授,国家电网级专家人才,国家电网高级兼职培训师,浙江省电网青年岗位能手。主要研究方向为电力通信,主持了省部级项目 1 项,其它各类项目 10 余项,并参与了 2 项省部级以上的项目。

* **通信作者:**李伟伟(1982—),男,博士,讲师,主要研究方向为电力线通信和继电保护。



通信等^[1]。其中,电力线双向工频通信系统(TWACS)通过在电网中的电压和电流波形过零点处产生微弱低频畸变信号(200~600 Hz)来实现信息传输^[2]。这种畸变信号可以跨变压器传输,相较于传输高频信号的电力线载波通信,电力线 TWACS 具有灵活性高、安装与维护方便、设备成本低等优点^[3]。然而,由于该系统传输速度慢和上行信号识别方式难以适应复杂配电网信道环境,导致其发展较为缓慢。

传统的信号识别方法通常依赖于时域差分法^[4]或时频分析法^[5-7],但这些方法难以同时保证高可靠性和高实时性。为此,申东等人^[8]尝试结合自适应阻尼雅可比算法和共轭梯度算法来完成对上行链路信号的识

花国祥,汤炼海,李伟伟,等:一种用于 Bi-LSTM 神经网络信号识别的 DO-CAB 算法

别,但该方法识别成功率不高。LIU X 等人^[9]则采用差分技术提取上行信号特征,并将提取出的工频信号特征输入支持向量机(SVM)进行信号识别,然而,这种方法的分类准确性受到差分提取特征质量的影响。如果差分技术未能准确地捕捉到信号的关键特性,或者在处理过程中引入了噪声,那么即使是最优化的 SVM 模型也可能无法达到理想的分类效果。

针对上述问题,本文利用深度学习在数据特征提取方面的强大能力,提出一种基于蒲公英优化(DO)算法的联合卷积神经网络(CNN)与注意力机制(AM)的双向长短时记忆(Bi-LSTM)神经网络信号识别算法,简称为 DO-CAB 算法。

1 DO-CAB 算法

本文设计的 DO-CAB 算法流程如图 1 所示。信号识别原理如下:首先,通过电流传感器采集原始信号,这些信号包含了来自发送端的含有电流畸变的调制信号和无畸变的非调制信号。接着,对这些不同类型的信号进行标记。为了构建训练集,本文向信号中加入噪声,而实际从居民台区收集到的信号则作为最终的测试集。训练集部分用于 DO-CAB 算法的训练。训练具体过程为:

1)利用 CNN 对输入信号进行特征提取,然后将 CNN

提取出的特征作为 Bi-LSTM 的输入。

2)由于 Bi-LSTM 在处理复杂的序列数据时可能存在一定的局限性,本文在 DO-CAB 算法中加入了 AM,以更高效地提取深层次的特征信息。

3)考虑到 DO-CAB 算法初始参数设定的不确定性,例如 Dropout 概率、Bi-LSTM 隐藏层单元数目、输入权重的学习率因子、偏置学习率因子及 AM 的通道数,同时考虑到自然优化过程的效率问题,本文引入了 DO 算法来进行高效的参数优化过程,从而加速 DO-CAB 算法的收敛速度并提升其性能。

4)收敛效果的评估指标是 DO 算法的适应度值,该值是通过训练过程中的五折交叉验证获得的准确率来确定的。经过迭代优化后,输出针对当前训练集的最佳信号识别模型。

5)最后,将实际信号输入此最佳模型以完成测试阶段。

通过上述流程,DO-CAB 算法能够实现对信号的精确识别。

2 测试与分析

本文旨在提出一种能够在复杂噪声和不同背景电流条件下精确识别电力线调制信号的算法。为了充分利用采集的数据并使算法具备良好的泛化能力,本

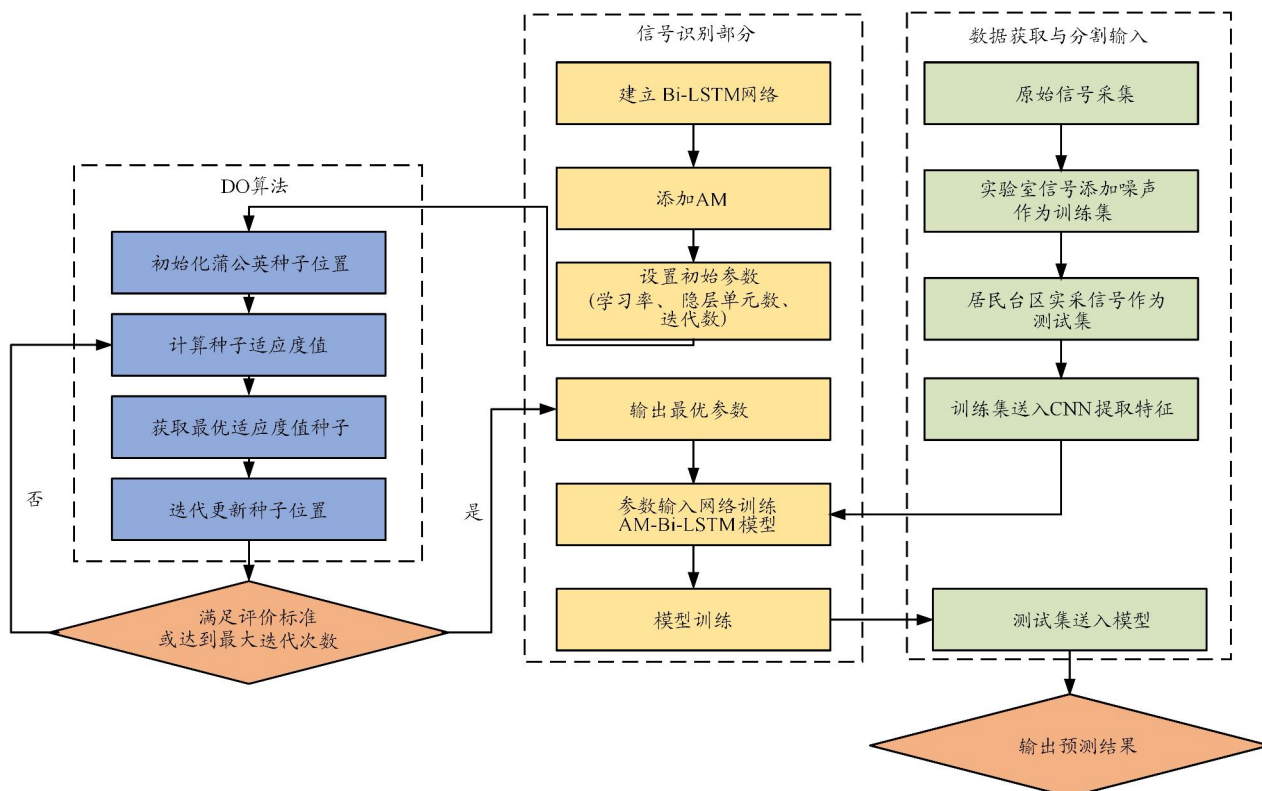


图 1 DO-CAB 算法流程图

文在实验室和居民台区采集了电力线调制信号的时间序列数据,并在实验室的理想信号中添加了高斯白噪声。这些数据被输入到本文提出的 DO-CAB 算法中进行训练,以输出最优模型参数,从而实现最佳的信号识别效果。

本次实验的硬件环境配置为 NVIDIA GTX 1050 Ti 显卡与 Intel i5-7300HQ CPU。算法的搭建与训练是在 Matlab 2023b 环境下完成的。在训练过程中,本文选择 Adam 优化器来调整学习率,初始学习率设置为 0.001,权重衰减系数为 0.01,批量大小为 40,整个训练过程共进行了 50 轮迭代。

2.1 实验数据采集

电力线 TWACS 信号畸变主要发生在过零点前 30° 附近。因此,在解调端,以过零点前 30° 为中心,使用 STM32 采集持续时间为 $150\ \mu\text{s}$ 的信号,采样频率为 1.5 MHz。为了满足后续处理的实时性要求,采集 100 个点作为样本。实验室采集环境中,发送端和接收端距离较近,并且只通过电灯连接。根据欧姆定律,电流与电压成正比,因此电流波形与电压波形相似。

实验总共采集了 180 000 条数据。然而,由于调制端发送信号间隔为 5 s,并且发送的数据为 EA、9B、0A、01 和 2F,所以调制信号仅占全部信号的 7.2%,即包含 12 960 条“1”信号的调制信号。为了保证 2 种信号的数量接近,仅选取含有调制信号的 50 个周期作为原始数据集,其中 90% 用于训练集,其余作为验证集。

本文通过在实验室采集的低噪声信号中添加上述电力线噪声来模拟实际环境中接收到的信号,并完成信号识别算法的研究。由于该数据集中噪声的采样频率为 0.4 MHz,因此需要对其进行插值处理,使其采样频率提升至 1.5 MHz,然后将预处理后的噪声信号作为干扰源加入实验室采集的信号中。

2.2 算法模拟信号识别结果评估

本文分别采用 CNN2、ResNet、DenseNet、卷积层尝试神经网络 (CLDNN)^[9]、DO-CAB 算法进行对比实验。实验中,将高斯白噪声加入到实验室采集的原始调制信号中,设定了不同信噪比下的含噪信号,以模拟实际传输中的信号情况。信噪比范围设置为 $-20\sim 18\ \text{dB}$,并观察了不同信噪比下各算法的检测准确率。不同网络算法的参数设置如表 1 所示。

上述算法在不同信噪比下的识别准确率如图 2 所示。可以看出,当信噪比低于 $-8\ \text{dB}$ 时,所有算法的识别准确率都较低;随着信噪比的提高,各算法的识别准确率显著增加。当信噪比大于 $-2\ \text{dB}$ 后,所有网

表 1 不同网络算法参数设置

网络算法	卷积层	全连接
CNN2	(256, 2×8), (128, 2×8), (64, 2×8), (64, 2×8)	100, 2
ResNet	(256, 1×3), (256, 2×3), (80, 1×3), (80, 1×3)	100, 2
DenseNet	(256, 1×3), (256, 2×3), (80, 1×3), (80, 1×3)	100, 2
CLDNN	(256, 1×3), (256, 2×3), (80, 1×3), (80, 1×3)	100, 2
DO-CAB	(256, 1×3), (256, 2×3), (64, 1×3), (32, 1×3)	100, 2

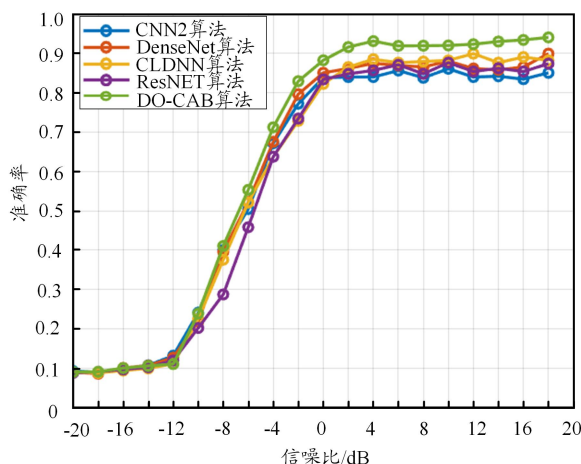


图 2 不同网络算法识别准确率

络都能够有效地识别出调制信号,并且随着信噪比的进一步提高,网络性能继续改善。在这一整体趋势下,经过优化后的 DO-CAB 算法识别准确率高於其它网络算法。具体如下:CNN2 算法、ResNet 算法、DenseNet 算法、CLDNN 算法的最高识别准确率分别为 85.06%、89.93%、89.07%、87.39%,而本文提出的 DO-CAB 算法的最高识别准确率达到 94.07%。

5 种网络算法的参数量和平均识别准确率如表 2 所示。可以看出,相较于其它 4 种网络算法,本文提出的 DO-CAB 算法的参数量最少,并且在本数据集上的平均识别准确率最高,比其它算法高 3.28%~5.32%。由于本文算法通过 DO 算法对注意力层的注意力通道数

表 2 5 种网络算法参数量和平均识别准确率对比

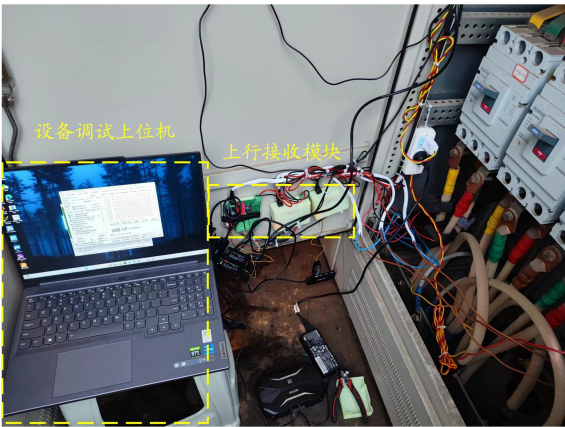
网络算法	参数量/个	平均识别准确率/%
CNN2	739 330	57.75
ResNet	492 642	59.01
DenseNet	534 882	58.62
CLDNN	2 086 942	56.96
DO-CAB	478 556	62.28

花国祥,汤炼海,李伟伟,等:一种用于 Bi-LSTM 神经网络信号识别的 DO-CAB 算法

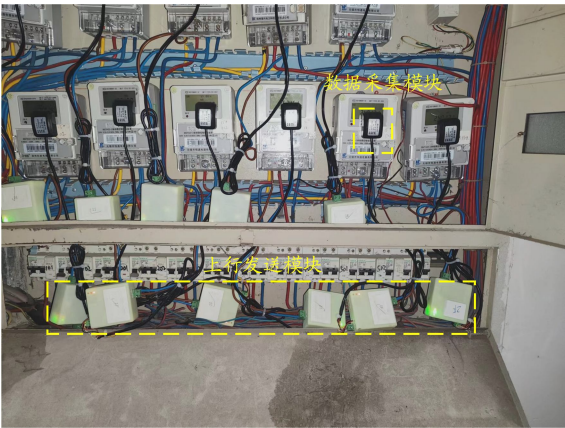
及 Bi-LSTM 的隐藏层单元数量进行迭代优化,避免了人工选择参数的盲目性和陷入局部最优的问题,同时还能防止因参数过多而导致的过拟合现象。

2.3 算法实际测试

在完成上述模拟信号识别后,将设备的发送端安装在某居民单元楼内,发送端会发送对应的电表数据,接收端则安装在分支箱中。由于不同用户的负载位于不同的相位上,因此存在 3 条不同的信道。本文在不同时间点进行信号发送与采集,以便获得真实条件下不同时间段内的负载对信道的干扰情况。安装环境如图 3 所示。



(a) 接收端设备安装图



(b) 发送端设备安装图

图 3 居民负载安装示意图

由于居民台区的负载与是否为工作日或休息时段密切相关,因此本文分别在周一、周三和周六这 3 个典型日期的早上 9 点、下午 3 点和晚上 7 点采集了电流数据用于测试。采集到的实验室电流信号对比图如图 4 所示。其中,信号 2 代表实验室电流信号;信号 3、10、11 代表周六不同时间点的数据;信号 1、5、9 代表周一不同时间点的数据;其余信号则代表周三不同时间点的数据。从图中可以看出,在不同时间段和相

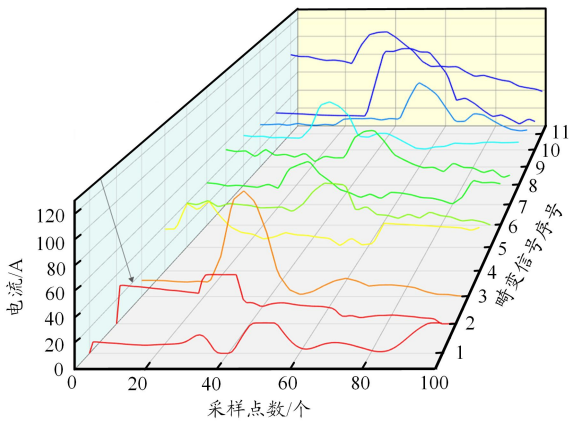


图 4 实测含畸变信号

位条件下,负载和噪声导致的畸变信号具有多样的形态。此外,在采集到的非畸变信号中还包含了背景信号,这使得使用传统方法来区分它们变得非常困难。

将上文所述的训练好的算法分别用于识别实测数据,使用 4 个二分类指标进行评价,得到不同算法的测试结果如表 3 所示。可以看出,时域差分法的受试者工作特征曲线下的面积(ROC-AUC)低于 0.6,表明其分类性能仅略高于随机水平,尤其在识别正例方面表现不佳。相比之下,时频分析法和 CNN2 算法虽然在识别正例方面具有较高的准确率,但整体识别准确率偏低,容易将无畸变信号误判为含有畸变的信号,导致在不同时间和负载条件下难以准确识别信号。与此相反,ResNet 算法和 DenseNet 算法表现出较高的整体准确率。特别是本文提出的 DO-CAB 算法,识别准确率达到了 92.32%,并且由于其较高的特异性和灵敏度,误报的概率远低于其它算法,从而显著提高了系统的可靠性。

为了进一步验证 DO-CAB 算法的识别能力,在相同的实验数据集基础上,设置了包含不同不均平衡比例的样本集,对不同算法进行性能测试。在测试集中,畸

表 3 算法实际测试结果

算法	ROC-AUC	准确率/%	灵敏度	特异性
时域差分	0.594 5	62.5	0.442	0.747
时频分析法	0.774 8	73.5	0.973 5	0.576
CNN2	0.792 9	75.63	0.975 8	0.61
ResNet	0.889 8	87.61	0.958 3	0.821 3
DenseNet	0.893 5	88.94	0.914	0.873
DO-CAB	0.925 5	92.32	0.936 7	0.914 2

变信号与非畸变信号数量相等(畸变信号:非畸变信号=1:1)。本文使用不同比例的训练样本集完成识别模型的训练,并利用统一的测试集来评估各算法的表现。由于时域差分法与时频分析法不属于机器学习算法,其性能不受训练集比例的影响,因此不在表中列出。最终的识别性能对比结果如表4所示。可以看出,改变畸变信号与非畸变信号的比例会导致不同算法的分类准确率发生变化,特别是当样本比例下降时,某些算法的准确度显著降低。当不平衡比例从1:2变化到1:15时,CNN2、ResNet和DenseNet的准确率分别下降了0.13、0.11和0.08,这表明数据不平衡对这些算法的识别性能有较大影响。相比之下,本文提出的算法即使在不平衡比例达到1:15的情况下也能保持较高的分类准确率,说明该算法对不平衡样本比例的鲁棒性强,显示出其优越性。

表4 算法实际测试结果

训练集畸变信号:非畸变信号 (比例)	测试集样本数量	不同算法的分类准确率			
		CNN2	ResNet	DenseNet	DO-CAB
300:600(1:2)	1 000	0.763	0.881	0.897	0.923
200:600(1:3)	1 000	0.754	0.872	0.891	0.917
120:600(1:5)	1 000	0.727	0.857	0.877	0.908
60:600(1:10)	1 000	0.683	0.811	0.842	0.893
40:600(1:15)	1 000	0.632	0.772	0.813	0.864

3 结束语

本文提出的 DO-CAB 算法通过实验室环境试验和实际数据验证,在工频通信上行信号识别方面展现了一定的实用价值。虽然实际信号中的噪声存在大量不确定性,导致识别结果与实验室试验结果存在一定

差异,但该算法在居民台区负载下仍能以较高准确率进行识别,表明其具有较强的泛化性能。然而,该算法也存在一定的局限性:目前的实测数据仅来自一个小区负载,并未涉及工业负载下的信号识别实验,因此后续将对此类型条件下的算法性能进行进一步测试。此外,识别准确率还有较大的提升空间,未来将研究算法在不同环境下的自适应参数调节,以实现不同负载和背景噪声条件下对工频调制信号的精确识别。

参考文献:

- [1] YAACOUB J P A, FERNANDEZ J H, NOURA H N, et al. Security of power line communication systems: issues, limitations and existing solutions[J]. Computer Science Review, 2021(39): 100331-1-100331-21.
- [2] 赵必游,林声锦,甄冲,等. 电力线工频通信抗干扰技术[J]. 电力系统及其自动化学报, 2016, 28(S1): 163-168.
- [3] 庄元强, 申敏. 改进深度神经网络的电力线脉冲噪声抑制与补偿算法[J]. 光通信技术, 2023, 47(4): 67-72.
- [4] MAK T S, REED L D. TWACS, a new viable two-way automatic communication system for distribution networks. part i: outbound communication[J]. IEEE Transactions on Power Apparatus & Systems, 1982 (8): 2941-2949.
- [5] 张玉恒,吴启晖,王金龙. 基于时频加窗短时傅里叶变换的 LFM 干扰抑制[J]. 电子与信息学报, 2007(6): 1361-1364.
- [6] SHENG M, IBM T J, WATSON C J. Modeling heterogeneous network traffic in wavelet domain[J]. IEEE/ACM Transactions on Networking, 2001, 9(5): 634-649.
- [7] CRISTAN A C, WALDEN A T. Multitaper power spectrum estimation and thresholding: wavelet packets versus wavelets[J]. IEEE Transactions on Signal Processing, 2002, 50(12): 2976-2986.
- [8] 申东,赵丹,李强,等. 低复杂度的大规模 MIMO 上行链路软输出信号检测[J]. 计算机应用研究, 2021, 38(5): 1524-1528.
- [9] LIU X, YANG D, GAMAL A E. Deep neural network architectures for modulation classification [C]//IEEE. Proceedings of 2017 51st Asilomar Conference on Signals, Systems, and Computers. Pacific Grove: IEEE, 2017: 915-919.