



智能系统学报

CAAI TRANSACTIONS ON INTELLIGENT SYSTEMS

深度残差收缩网络的多特征锅炉炉管声波信号故障识别

杨正理, 吴馥云, 陈海霞

引用本文:

杨正理, 吴馥云, 陈海霞. 深度残差收缩网络的多特征锅炉炉管声波信号故障识别[J]. 智能系统学报, 2023, 18(5): 1108–1116.

YANG Zhengli, WU Fuyun, CHEN Haixia. Fault identification of multi-feature boiler tube acoustic signal based on deep residual shrinkage network[J]. *CAAI Transactions on Intelligent Systems*, 2023, 18(5): 1108–1116.

在线阅读 View online: <https://dx.doi.org/10.11992/tis.202207012>

您可能感兴趣的其他文章

多感知兴趣区域特征融合的图像识别方法

Image recognition method based on multi-perceptual interest region feature fusion
智能系统学报. 2021, 16(2): 263–270 <https://dx.doi.org/10.11992/tis.201906032>

用于关系抽取的注意力图长短时记忆神经网络

Attention graph long short-term memory neural network for relation extraction
智能系统学报. 2021, 16(3): 518–527 <https://dx.doi.org/10.11992/tis.202008036>

基于改进的Faster RCNN面部表情检测算法

Facial expression recognition based on improved Faster RCNN
智能系统学报. 2021, 16(2): 210–217 <https://dx.doi.org/10.11992/tis.201910020>

注意力机制和Faster RCNN相结合的绝缘子识别

Insulator recognition based on attention mechanism and Faster RCNN
智能系统学报. 2020, 15(1): 92–98 <https://dx.doi.org/10.11992/tis.201907023>

层次化双注意力神经网络模型的情感分析研究

Hierarchical double-attention neural networks for sentiment classification
智能系统学报. 2020, 15(3): 460–467 <https://dx.doi.org/10.11992/tis.201812017>

深度学习的双人交互行为识别与预测算法研究

Human interaction recognition and prediction algorithm based on deep learning
智能系统学报. 2020, 15(3): 484–490 <https://dx.doi.org/10.11992/tis.201812029>

DOI: 10.11992/tis.202207012

网络出版地址: <https://kns.cnki.net/kcms2/detail/23.1538.TP.20230428.1942.004.html>

深度残差收缩网络的多特征锅炉 炉管声波信号故障识别

杨正理, 吴馥云, 陈海霞

(三江学院 机械与电气工程学院, 江苏 南京 210012)

摘要: 为了提高锅炉炉管声波信号故障识别的学习效果和识别精度, 采用特征向量并行和拼接两种融合方式构成特征层, 以及平均得分和最大值得分两种融合方式构建决策层等不同信息融合机制, 提出基于深度残差收缩网络的多特征锅炉炉管声波信号故障识别方法。首先, 考虑锅炉炉管上各声波传感器的差异性, 分别计算声波信号谱特征一阶和二阶差分构建三通道特征集作为二维网络的输入特征向量; 然后, 在卷积神经网络和双向长短时记忆网络基础上引入注意力机制构建基线模型, 并采用深度残差收缩网络对二维网络的通道权重进行优化分配, 提高模型的故障识别精度。大量实验结果表明: 采用特征向量并行融合方式构成特征层的信息融合机制是一种更有效的策略; 本文模型的识别精度得到较大程度提高, 与基线模型相比较, 未加权平均召回率提高了 4.32%。

关键词: 深度学习; 故障识别; 深度残差收缩网络; 双向长短时记忆网络; 注意力机制; 卷积神经网络; 锅炉炉管; 声波信号

中图分类号: TP391; TN912.3; TM621.2 **文献标志码:** A **文章编号:** 1673-4785(2023)05-1108-09

中文引用格式: 杨正理, 吴馥云, 陈海霞. 深度残差收缩网络的多特征锅炉炉管声波信号故障识别 [J]. 智能系统学报, 2023, 18(5): 1108-1116.

英文引用格式: YANG Zhengli, WU Fuyun, CHEN Haixia. Fault identification of multi-feature boiler tube acoustic signal based on deep residual shrinkage network[J]. CAAI transactions on intelligent systems, 2023, 18(5): 1108-1116.

Fault identification of multi-feature boiler tube acoustic signal based on deep residual shrinkage network

YANG Zhengli, WU Fuyun, CHEN Haixia

(School of Mechanical and Electrical Engineering, Sanjiang University, Nanjing 210012, China)

Abstract: To improve the learning effect and recognition accuracy in fault identification of boiler tube acoustic signals, different information fusion mechanisms were used in this study. Specifically, feature vector parallelism and splicing techniques are used to form the feature layer, while the decision layer is formed using average and maximum score approaches. A fault identification method of multifeature boiler tube acoustic signal based on a deep residual shrinkage network is proposed. First, the first-order and second-order differential characteristics of the acoustic signal spectrum were calculated separately, considering the variations of acoustic wave sensors on the boiler tube. This process constructs a three-channel feature set, which serves as the input feature vector of the two-dimensional network. Subsequently, an attentional mechanism is introduced to construct a baseline model using a combination of a convolutional neural network and a bidirectional short and long memory network. Additionally, a deep residual shrinkage network is used to optimize the allocation of channel weights within the two-dimensional network to improve the fault identification accuracy of the model. Extensive experimental results show that: it is a more effective strategy to construct the information fusion mechanism of the feature layer by using feature vector parallel fusion; compared with the baseline model, the unweighted average recall rate has increased by 4.32%, highlighting significant improvements in the recognition accuracy of the model presented in this paper.

Keywords: deep learning; fault identification; deep residual shrinkage network; bidirectional short and long memory network; attention mechanism; convolutional neural network; boiler tube; acoustic signal

收稿日期: 2022-07-11. 网络出版日期: 2023-05-04.

基金项目: 江苏省高校自然科学基金面上项目 (20KJB470029).

通信作者: 杨正理. E-mail: Zhengli-yang@163.com.

锅炉炉管自动报警系统是当锅炉炉管在高温、高压状态下工作中发生泄露故障时, 根据安

装在锅炉炉管上的多个声波传感器检测到的声波信号判断锅炉炉管是否发生泄露故障^[1]。目前, 故障识别主要是依据声波信号的主要特征, 包括声波韵律特征和声波谱特征^[2]。

声波韵律特征包括声波基频特征和能量特征, 按照其全局和局部特征, 可以通过提取声波基频和能量的最大值、最小值、平均值及方差等构成高维声波特征集。声波谱特征包括线性频谱特征和倒谱特征, 主要反映声波信号在频域上的差异。BOU-GHAZALE 等^[3-4]的研究表明, 倒谱特征在识别声波信号特征上比线性频谱特征具有明显优势。Hsiao 等^[5-6]通过提取声波信号的基频、过零率、能量、梅尔频率倒谱系数 (Mel-frequency cepstral coefficient, MFCC) 等特征, 并计算其方差、平均值、最大值、最小值、峰度、线性回归系数等构成高维声波特征集, 然后采用深度循环神经网络 (deep recurrent neural network, DRNN) 模型在 FAU-Aibo 数据集的未加权平均召回率 (unweighted average recall, UAR) 提高了 9.3%; Han 等^[7-8]通过声波能量、共振峰、谐波噪声比等特征构建 16 维声波信号特征向量, 其中前 9 维为声波韵律特征, 后 7 维为声波谱特征, 提出了一种基于高斯核的非线性近端支持向量机模型; 胡德生等^[9-10]基于主辅网络特征融合方法, 通过计算声波信号的平均过零率、基准频率、能量、以及 MFCC 等特征, 在 IEMO-CAP 数据库上使未非加权平均精度 (unweighted average precision, UAP) 提高到 73.1%。上述文献所述的模型均采用固定信息融合方法, 将声波信号的不同特征输入到多个网络, 因而声波信号特征利用率较低, 造成模型的识别精确不高。本文充分利用声波信号的多特征性质, 对多个声波特征采用不同融合策略进行信息融合, 提高了信号特征利用效率, 从而提升模型的识别精度。

近年来, 卷积神经网络 (convolutional neural network, CNN) 和长短期记忆网络 (long short-term memory, LSTM) 以其较高的识别精度在声波信号特征识别领域得到了广泛的应用^[11-16]。而近期的研究进一步表明, 在双向长短期记忆 (bidirectional long short-term memory, BLSTM) 网络中引入注意力机制能更好地提高声波信号的识别精度^[17-20]。其主要思路是: 1) 基于当前声波信号特征与其前后声波信号之间的相关性, 采用 BLSTM 网络能够得到信号在时间上的依赖特征; 2) 基于炉管声波信号中泄漏故障信号与正常信号之间的信息特征差异, 引入注意力机制, 增加故障信号在特征模型中的权重。

在锅炉炉管声波信号故障识别系统中, 由于声波传感器之间输出的声波信号存在个体的差异性。为了降低个体差异对故障识别的影响, 往往需要对各声波信号进行差分处理后再作为 2D-CNN、3D-CNN 网络的输入特征集, 提升模型的识别精度。例如, Gao 等^[21]通过计算声音信号 MFCC 系数的一阶差分、二阶差分作为信号特征集输入到改进型 3D-CNN-LSTM 网络模型中, 使网络模型的 UAR 在 IEMCAP 和 EMO-DB 数据库上分别达到了 64.57% 和 83.15%。Sohaib 等^[22]通过计算声音信号的对数梅尔特征 (Log-Mel) 的一阶、二阶差分作为信号特征集输入到改进型 2D-CNN-LSTM 网络模型中, 使网络模型基于发声源依赖实验和基于发声源独立实验的识别精确率在 IEMOCAP 数据库上分别达到 90.02% 和 53.21%。Gupta 等^[23]通过计算语音信号 Log-Mel 的一阶、二阶差分构成 3D-Log-Mel 数据集输入到 3D-CNN-LSTM 网络模型, 使语音信号识别平均准确率在 IEMCAP 和 EMO-DB 数据库上分别达到 60.89% 和 86.07%。上述文献所述方法将信号进行差分处理后再输入到多通道卷积网络模型, 但却没有对各通道的特征输入权重进行优化。声波信号的不同阶差分对信号的区分度显然是不相同的, 输入到网络模型时应该合理分配各通道的权重, 从而提升模型的识别精度。

深度残差收缩网络 (deep residual shrinkage networks, DRSN) 引入了注意力机制和软阈值函数。在内部将软阈值化作为非线性层, 自适应地在特征学习过程滤除噪声信号, 提高有用信号特征的学习效果, 所以该模型适合给重要信号特征分配更大权重, 而降低不重要信号特征的权重。本文引入 DRSN 对声波信号的输入权重进行优化分配, 提高声波信号有效特征的输入权重, 提高声波信号识别精度。

1 锅炉炉管声波信号的有效特征

锅炉炉管在运行过程中最容易发生故障且故障后果最严重的 4 种泄露故障类型是发生在“四管”上的泄露故障, 即再热器管、水冷壁管、省煤器管和过热器管^[24]。用于故障识别的声波信号有效特征通常采用声波韵律特征和声波谱特征来描述。

1.1 声波信号韵律特征

1) 短时能量。不同类型的锅炉炉管声波信号的幅度值是不相同的。所以其声波信号的短时能量可以作为信号的特征之一; 设某声波信号为 $\{x(n)\}$, 定义加窗函数的平均短时能量为

$$E_n = \sum_{i=0}^{N-1} x(i)h(n-i) = x^2(n) * h(n) \quad (1)$$

$$h(n) = \omega(n)^2$$

式中: $\omega(n)$ 为窗函数, 可选矩形窗或汉明窗等; $h(n)$ 为单位冲激响应; $*$ 为卷积。

2) 短时平均幅度。声波信号短时能量主要用来描述信号的峰值幅度, 而声波信号的单位时间内幅度值变化情况可用短时平均幅度值 M_n 来表示:

$$M_n = \sum_{i=0}^{N-1} |x_n(i)| \quad (2)$$

3) 短时过零率。声波信号的短时过零率用来表示单位时间内声波信号穿越时间轴的次数, 即

$$Z_n = \frac{1}{2} \sum_{i=0}^{N-1} |\text{sgn}[x_n(i)] - \text{sgn}[x_n(i-1)]| \quad (3)$$

显然, 锅炉炉管声波信号的高、低频部分具有较高和较低的短时过零率, 可以用来表现声波信号的振动情况, 所以能作为区分声波信号特征之一。

1.2 声波信号谱特征

声波信号的倒谱特征对信号特征的区分能力

$$H_m(k) = \begin{cases} 0, & k < f(m-1) \\ \frac{2(k-f(m-1))}{(f(m+1)-f(m-1)(f(m)-f(m-1)))}, & f(m-1) \leq k < f(m) \\ \frac{2(f(m+1)-k)}{(f(m+1)-f(m-1)(f(m)-f(m-1)))}, & f(m) \leq k < f(m+1) \\ 0, & k \geq f(m+1) \end{cases} \quad (6)$$

⑤ 计算三角滤波器输出的对数能量, 并对对数能量进行离散余弦变换, 得到声波信号 MFCC 系数:

$$s(m) = \ln \left(\sum_{k=0}^{N-1} |X_a(k)|^2 H_m(k) \right), \quad 0 \leq m \leq M \quad (7)$$

$$C(n) = \left(\sum_{m=0}^{N-1} s(m) \cos \left(\frac{n\pi(m-0.5)}{M} \right) \right), \quad n = 1, 2, \dots, L \quad (8)$$

2) LFCC。LFCC 计算过程与 MFCC 相同, 只是其滤波器组的频率需要按照线性频率进行分布;

3) LPCC。LPCC 采用线性预测分析方法获得声波信号的倒谱系数^[26]。其计算过程如下:

① 采用线性预测分析方法得到分段声波信号的全极点模型, 即

$$H(z) = \frac{G}{1 - \sum_{i=1}^p a_i z^{-i}} \quad (9)$$

② 声波信号的激励模型为

$$U(z) = E(z)G(z) = \frac{A_u}{1-z^{-1}} \cdot \frac{1}{(1-e^{-cT_s})^2} \quad (10)$$

③ 模型的输入和输出关系可用表示为

明显优于线性谱特征。倒谱特征常用 MFCC、线性频率倒谱系数 (linear frequency cepstral coefficient, LFCC)、线性预测倒谱系数 (linear predictive cepstral coefficient, LPCC) 来表示。

1) MFCC。MFCC 可有效表征声波信号的共振情况。在声波信号故障识别领域, MFCC 已经成为表征声波信号特征应用最广泛的参数^[25]。声波信号的 MFCC 的计算过程如下:

① 采用高频滤波器对声波信号进行预处理, 滤除声波信号中的高频部分:

$$H_z = 1 - \mu z^{-1} \quad (4)$$

② 按照段长为 25 ms、移动段长为 10 ms 对声波信号进行分段处理, 并给各段声波信号加汉明窗;

③ 对各分段声波信号进行傅里叶变换, 并计算各分段声波信号的能量分布:

$$X_a(k) = \sum_{i=0}^{N-1} x(i) e^{-j \frac{2\pi k i}{N}}, \quad 0 \leq k \leq N \quad (5)$$

④ 采用梅尔尺度三角滤波器对能量谱作平滑处理:

$$s(n) = \sum_{k=1}^p a_k s(n-k) + G u(n) \quad (11)$$

④ 假设某段声波信号的第 n 个采样点的模型输出为 $\tilde{s}(n)$, 且 $\tilde{s}(n)$ 能够用该段声波信号前面 p 个模型输出线性表示, 即

$$\tilde{s}(n) \approx a_1 s(n-1) + a_2 s(n-2) + \dots + a_p s(n-p) \quad (12)$$

式中: a_1, a_2, \dots, a_p 为常数, 称为线性预测系数。

⑤ 从而, LPCC 可表示为

$$c(n) = \begin{cases} 0, & n < 0 \\ \ln G, & n = 0 \\ a_n + \sum_{k=1}^{n-1} \left(\frac{k}{n} \right) c(k) a_{n-k}, & 0 < n \leq p \\ \sum_{k=n-p}^{n-1} \left(\frac{k}{n} \right) c(k) a_{n-k}, & n > p \end{cases} \quad (13)$$

2 锅炉炉管声波信号识别模型构建

2.1 基线模型

在 CNN 和 BLSTM 网络模型基础上, 引入注意力机制, 通过增强有效声波信号的权重, 降低无效声波信号的权重构建基线模型, 如图 1 所示。

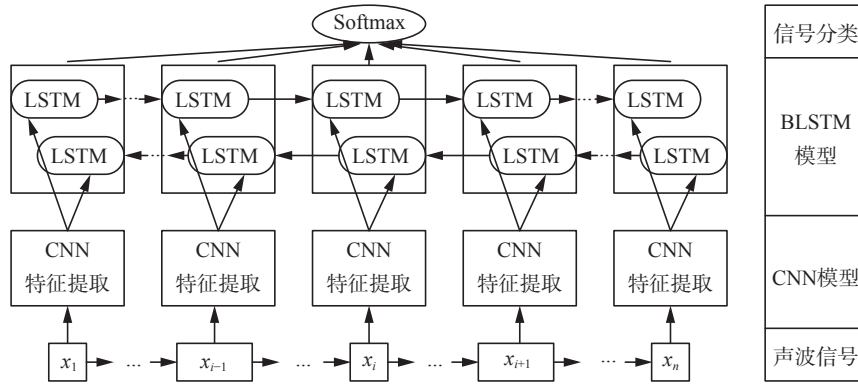


图 1 CNN-BLSTM 模型结构

Fig. 1 Structure of CNN-BLSTM

1) CNN 模型。首先, 将声波信号的韵律特征数据集作为 1D-CNN 网络的输入, 1D-CNN 网络设置 2 个卷积层, 卷积核大小为 5, 数量定义为 128, 步长为 1, 激活函数采用 ReLU。设置池化大小分别为 5 和 3 的两个池化层; 然后, 计算声波信号谱特征的一阶、二阶差分构成 3 个通道的数据集作为 2D-CNN 的输入; 2D-CNN 网络设置 3 个卷积层, 卷积核大小分别为 5×5 、 5×5 、 3×3 , 数量定义为 128, 步长为 1, 激活函数采用 ReLU。设置池化大小为 2×2 的 3 个池化层;

2) BLSTM 模型。当前声波信号的信号特征与其前、后声波信号的特征信息均相关。所以在处理声波信号序列时, 需要采用两个独立的 BLSTM 网络从前、后两个方向分别对声波信号进行处理。所以, BLSTM 网络在 t 时刻的隐藏状态输出结果为

$$\mathbf{h}_t = \begin{bmatrix} \vec{\mathbf{h}}_t; \overleftarrow{\mathbf{h}}_t \end{bmatrix} \quad (14)$$

式中: $\vec{\mathbf{h}}_t = \vec{\mathbf{L}}(x_t, \vec{\mathbf{h}}_{t-1})$ 表示 BLSTM 网络对前向信号进行处理在 t 时刻的隐藏状态输出结果; $\overleftarrow{\mathbf{h}}_t = \overleftarrow{\mathbf{L}}(x_t, \overleftarrow{\mathbf{h}}_{t-1})$ 表示 BLSTM 网络对后向信号进行处理在 t 时刻的隐藏状态输出结果; 符号 \rightarrow 、 \leftarrow 分别表示前向、后向。

BLSTM 网络的最终隐藏层输出可表示为

$$\mathbf{H} = (h_1, h_2, \dots, h_t) \quad (15)$$

式中: $\mathbf{H} \in \mathbf{R}^{t \times d}$; t 为声波信号的分段数; d 为 BLSTM 网络输出最终隐藏层的大小。

3) 注意力机制。将 BLSTM 网络输出的最终隐藏层作为注意力机制的输入数据集。

$$e_i = \tanh(W_j h_i + b_i) \quad (16)$$

$$a_i = \exp(e_i) / \sum_t \exp(e_i) \quad (17)$$

$$h'_i = a_i h_i \quad (18)$$

式中: a_i 为注意力机制赋予给有效声波信号的权重; h'_i 为有效声波信号加权后的输出值。

在 CNN 模型基础上构建引入了注意力机制

的 BLSTM 网络。注意力机制的主要作用是合理分配输入通道的权重, 对包含故障特征的声波信号分配较大的权重。锅炉炉管声波信号特征与炉内温度、炉管内压力、设备损耗、泄漏故障位置和程度等多种因素有关。注意力机制能够使模型更有效地提取声波信号本身特征及其耦合的炉管内压力信息、炉管老化程度、设备损耗特征等。并在很大程度上避免在机器学习中出现过拟合现象。

2.2 信息融合

本文所述的信息融合是指综合不同模型的优势, 达到不同模型间提优补缺的作用, 主要包括特征层融合和决策层融合^[27]。

特征层融合是指将原始样本分别输入到多个深度学习网络得到多个降维特征向量, 由多个降维特征向量融合得到一个特征向量的方法, 有特征向量并行和特征向量拼接两种方式。特征向量并行是将多个同维的特征向量进行叠加得到新的同维特征向量, 新特征向量的维数没有增加, 但却增加了每一维向量的特征信息; 特征向量拼接是将多个同维度或不同维度的特征向量串联拼接而得到新的特征向量, 新特征向量的维数增加, 但每一维向量的特征信息量却没有增加。

决策层融合是指采用代数方法对多个模型网络的识别结果进行融合。多个模型网络的识别结果之间相互独立, 识别结果一般用一个预测评分来表示, 常见的决策层融合方式就是对多个预测评分求平均值、最大值等。

2.3 深度残差收缩网络

在深度残差网络中引用软阈值化作为非线性层。软阈值化的本质是通过设计合适的滤波器对信号噪声进行滤波。在深度残差网络基础上应用软阈值化构建 DRSN, 提高网络模型对含噪数据或复杂数据的学习效果和识别精度。用 \mathbf{x} 表示输入特征向量, 用 \mathbf{y} 表示输出特征向量, 用 τ 表示一

个正参数阈值, 软阈值在网络模型中的作用可表示为

$$y = \begin{cases} x - \tau, & x > \tau \\ 0, & -\tau \leq x \leq \tau \\ x + \tau, & x < -\tau \end{cases} \quad (19)$$

当软阈值的输出对输入的导数为 1 或 0 时, 即

$$\frac{\partial y}{\partial x} = \begin{cases} 1, & x > \tau \\ 0, & -\tau \leq x \leq \tau \\ 1, & x < -\tau \end{cases} \quad (20)$$

可有效防止信号经软阈值处理后出现梯度消失和爆炸等异常问题。

将多个具有不同通道阈值的残差收缩模块进行叠加构成 DRSN, 如图 2 所示。

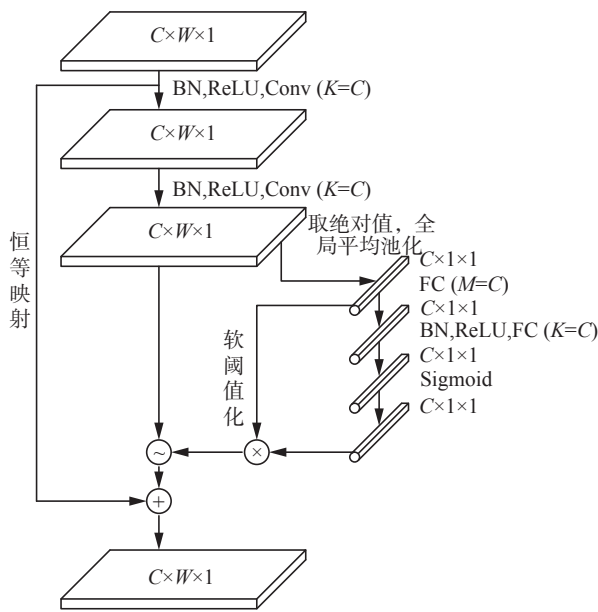


图 2 阈值独立型残差收缩模块

Fig. 2 Residual shrinkage building unit with channel-wise thresholds

图 2 中的每个通道均对应一个独立阈值, 所以构成 DRSN 的各残差收缩模块又称为阈值独立型残差收缩模块 (residual shrinkage building unit with channel-wise thresholds, RSBUCW)。通过对特征向量中每一个元素求绝对值, 并利用全局平均池化方法将特征向量映射为一个一维向量作为一个两层全连接网络的输入, 其中全连接网络的第二层神经元个数等于模型输入特征映射的特征通道个数, 全连接网络的输入缩放在 (0,1) 之间, 缩放参数表示为

$$\sigma_c = \frac{1}{1 + e^{-z_c}} \quad (21)$$

式中: z_c 表示第 c 通道的输出特征; σ_c 表示第 c 通道的缩放参数。

第 c 通道的阈值计算公式为

$$\tau_c = \sigma_c \cdot f_{\text{average}} |X_{i,j,c}| \quad (22)$$

式中: τ_c 表示第 c 通道阈值; $X_{i,j,c}$ 表示特征向量 X 第 c 通道下特征向量, 坐标为 (i, j) ; f_{average} 表示求均值。

由于各通道阈值由各样本自身的特征自适应确定, 所以 DRSN 对含噪信号处理同样具有良好效果。当样本信号中的噪声较小时, 经学习和训练后所得的阈值非常接近于 0, 从而使软阈值化对训练模型几乎不会造成不良影响。本文通过计算声波信号的谱特征, 并经一阶差分和二阶差分后得到 3 个通道的特征数据集, 在输入二维网络之前, 利用 DRSN 得到各通道的输入权重。经上述方法处理后, 每组声波信号的训练样本均能获得独立的通道权重, 并根据各训练样本的信号特征对通道权重进行调整, 从而得到具有独立通道权重的 CNN, 提升模型的识别效果。

2.4 DRSN 多特征融合模型

锅炉炉管声波信号故障识别系统采用 DRSN 多特征融合模型 (DRSN with multi-feature fusion, DRSN-MF)。该模型包含一个一维网络 (1D-CNN-BLSTM-attention) 和一个二维网络 (2D-CNN-BLSTM-attention); 两个网络分别都引入注意力机制和 Dropout 层。注意力机制的主要作用是通过提高有效声波信号的权重来提高声波信号故障识别效果; Dropout 层的主要作用是有效缓解模型发生过拟合, 提升模型的泛化能力, 在一定程度上达到正则化的效果。二维网络比一维网络多了一个维度, 模型利用这一维度实现样本数据的多通道输入。声波信号的韵律特征数据作为一维网络的输入; 声波信号的谱特征作为二维网络的输入, 在输入到二维网络前, 首先计算声波信号谱特征的一阶和二阶差分, 形成 3 通道特征数据集, 再通过 DRSN 对二维网络的 3 个通道权重进行合理分配, 最后再将特征数据集输入到二维网络。MFCC、LFCC 和 LPCC 均采用这种方式输入数据。

不同类型的特征数据集经过一维或二维深度学习网络后, 其输出结果为对应的降维特征向量。在特征层融合中, 降维特征向量采用并行和拼接方法进行信息融合后作为全连接层的输入, 输出结果采用 Softmax 函数对声波信号进行识别, 如图 3(a) 所示。在决策层融合中, 降维特征数据集先通过全连接层, 再采用 Softmax 函数得到各类降维特征的分类预测得分, 最后采用求平均值和最大值的代数组合规则输出声波信号的识别结果, 如图 3(b) 所示。

本文基于 CNN 模型在特征提取方面的优势和 BLSTM 在时间连续上的识别优势, 将 CNN 和 BLSTM 进行融合得到 CNN-BLSTM 模型用于锅

锅炉管声波信号故障识别, 该模型能够提取丰富的信号特征, 提高信号特征的利用效率, 并对时

间连续的行为动作具有较强的识别性, 因此具有较高的信号识别率。

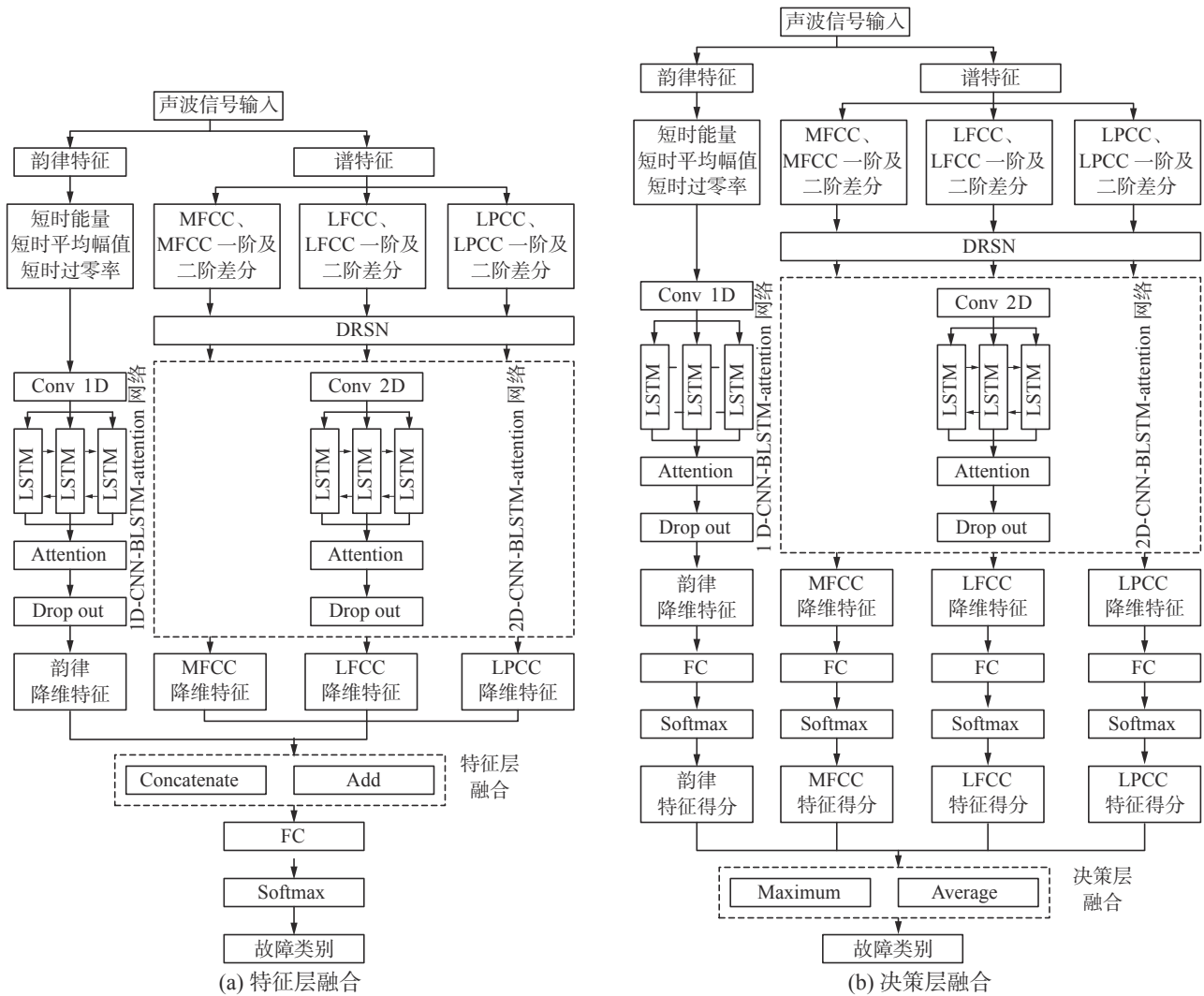


图 3 DRSN-MF 结构图

Fig. 3 DRSN-MF structure

3 模型实验及结果分析

3.1 锅炉炉管声波信号原始样本数据集

选择某典型锅炉炉管声波自动报警系统的历史数据作为本文的原始实验样本数据集。该样本数据集经过一定的数据筛选、排异处理, 共 12 600 条数据, 包括经多种类型、不同时段、多个声波传感器采样得到的再热器管、水冷壁管、省煤器管、过热器管泄漏故障数据, 以及无泄露故障数据。数据集采样率为 16 kHz, 采用 16 bit 量化。

3.2 故障识别关键评价指标

锅炉炉管声波信号故障识别是一个典型的多分类任务, 评价识别效果的主要指标有准确率、精确率、召回率和 F_1 值^[28]。对于一个二分类模型, 模型输入样本可以分为正样本和负样本; 模型输出结果可分为 4 种情况: 1) 真阳性 T_P , 即预测结果

和实际样本均为正样本; 2) 假阳性 F_P , 即预测结果为正样本, 实际样本为负样本; 3) 真阴性 T_N , 即预测结果和实际样本均为负样本; 4) 假阴性 F_N , 即预测结果为负样本, 实际样本为正样本。

1) 准确率。准确率为预测结果正确的数量与全部样本的比值:

$$A = \frac{T_P + T_N}{T_P + T_N + F_P + F_N} \quad (23)$$

2) 精确率。精确率为真阳性样本数量与所有正样本数量的比值。

$$R = \frac{T_P}{T_P + F_N} \quad (24)$$

3) 召回率。召回率为真阳性样本数量与所有预测结果为正样本数量的比值。

$$P = \frac{T_P}{T_P + F_P} \quad (25)$$

4) F_1 值。 F_1 值为算数平均数除以几何平均数,且越大越好,本文的 F_1 值就是精确率和召回率的调和均值。识别模型中,希望准确率和召回率都高,但是一般情况下,当准确率越高时、召回率就越低,反之亦然。 F_1 值的计算公式可表示为

$$F_1 = 2 \times \frac{P \times R}{P + R} \quad (26)$$

多分类任务中,计算所有分类结果的精确率 R 和召回率 P 可分别构成一个 n 维向量 \mathbf{R} 和 \mathbf{P} ,即

$$\mathbf{R} = (R_1, R_2, \dots, R_i, \dots, R_n) \quad (27)$$

$$\mathbf{P} = (P_1, P_2, \dots, P_i, \dots, P_n) \quad (28)$$

式中: n 表示多分类模型输出的分类类别数; R_i 、 P_i 分别表示第 i 个分类类别的精确率和召回率。

对多分类模型输出分类类别的精确率求平均值就得到分类类别的 $UAP(\hat{R})$;对多分类模型输出分类类别的召回率求平均值就得到的分类类别的 $UAR(\hat{P})$,即

$$\hat{R} = \sum_{i=1}^n R_i, \quad \hat{P} = \sum_{i=1}^n P_i \quad (29)$$

3.3 模型参数设置

本文所有实验在搭载第 12 代 Inter(R)Core(TM) i5-12500H 标压处理器的 HUAWEI MateBook 14 s 上运行,基于 TensorFlow 构建深度学习框架。模型识别效果的评价指标采用 UAR、UAP、准确率和 F_1 值。按照 25 ms 帧长和 10 ms 移动帧长的定义,将原始样本数据集的声波信号划分成 200 帧的声波段,训练样本数和测试样本数的比例为 4:1,所有实验结果取 10 次平均值以消除随机误差。设置模型的学习率为 10^{-4} ,衰减率为 10^{-6} ,学习最高迭代次数为 150。

3.4 实验结果及分析

1) 信息融合策略实验。

对声波信号的 4 种信息融合方式进行实验,实验结果如表 1 所示。

表 1 声波信号不同信息融合方式实验结果
Table 1 Experimental results of different information fusion methods of acoustic signals %

融合策略	准确率	精确率	召回率	F_1 值
特征层融合(Concatenate)	81.47	80.78	80.21	81.34
特征层融合(Add)	84.95	85.20	85.11	84.79
决策层融合(Average)	80.03	81.23	80.06	80.12
决策层融合(Maximun)	64.25	70.26	63.42	66.18

从表 1 中数据可见,声波信号采用特征层融合策略比决策层融合的效果更好。这一实验结果表明,先将声波信号的特征进行整合,再进行分类的信号识别效果更好。另外表中数据还表明:

信号在特征层融合时,采用并行方式的效果比拼接方式好;信号在决策层融合时,采用求平均值的效果比求最大值方式好,这主要是因为求最大值方式容易出现过拟合现象。

2) RSBU-CW 模块数量设置实验。

多个 RSBU-CW 模块进行叠加构成 DRSN,对构成模块的数量对模型的识别效果进行实验验证。模型的信息融合在特征层进行,各特征向量融合方法采用并行方式。当 RSBU-CW 模块数量分别为 1、2、3 时,模型的识别效果 UAP 分别为 84.65%、81.37%、79.29%。由此可见,构成 DRSN 模型的 RSBU-CW 模块为 1 时,模型的识别效果就达到最佳状态。所以模型构建时,不必过多追求 RSBU-CW 模块的叠加数量,原因是 DRSN 模型的输入特征数据量较小,过多的 RSBU-CW 模块叠加反而容易造成 DRSN 模型过拟合,使其识别效果降低。

3) 模型训练过程实验。

对模型的训练过程进行实验,设置模型的最大迭代次数为 100 次,并限定当模型连续 40 次迭代,其模型准确率没有提升时主动中断训练过程。得到的迭代次数-准确率变化曲线如图 4 所示。

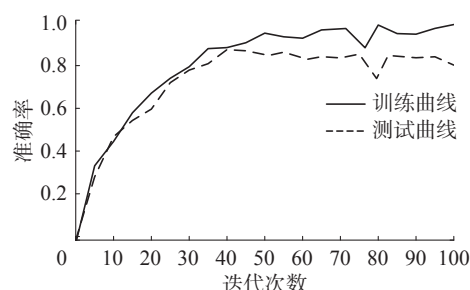


图 4 迭代次数变化时模型准确率变化曲线

Fig. 4 Model accuracy curves when the number of iterations changes

图 4 中的曲线表明,训练没有达到最大设定迭代次数训练就被中断,这是因为样本数据较少的缘故。同时,当训练次数达到 50 次左右时,模型就开始收敛,这个现象表明本文所构建的模型是实用且有效的。

4) 样本数据的混淆矩阵。

由学习框架得到原始样本数据的混淆矩阵如图 5 所示。从图 5 中数据可以看出,本文所构建的模型对再热器管、水冷壁管、省煤器管、过热器管泄漏故障,以及无泄露故障信号均有较高的召回率,这说明本文所构建的模型能够有效区分锅炉炉管是否发生泄漏故障,并能较好地地区分泄漏故障的类型。

真实标签	再热器管故障	0.92	0.06	0.04	0	0.02
	水冷壁管故障	0.04	0.85	0.04	0.03	0.03
	省煤器管故障	0.01	0.03	0.81	0.04	0.05
	过热器管故障	0.04	0	0.02	0.90	0.01
	无泄漏故障	0.03	0.01	0.01	0.03	0.95
	再热器管故障	水冷壁管故障	省煤器管故障	过热器管故障	无泄漏故障	
	预测标签					

图 5 原始样本的混淆矩阵

Fig. 5 The confusion matrix of the original sample

5) 各类模型识别效果对比。

本文在 CNN-BLSTM 模型的基础上引入注意力机制作为基线模型, 基线模型的信息融合在特征层采用特征向量并行融合方式。基线模型与本文所构建的 DRSN-MF 模型的识别效果如表 2。

表 2 基线模型与 DRSN-MF 模型识别效果对比

Table 2 Comparison of recognition effect between baseline model and DRSN-MF model %

模型	准确率	精确率	召回率	F_1 值
基线模型	80.42	80.96	80.51	80.49
DRSN-MF	85.05	84.38	84.83	84.61

如表 2 中数据所示, 本文所构建的 DRSN-MF 模型在准确率、精确率、召回率和 F_1 值 4 个模型识别效果评价标准上均优于基线模型。为进一步验证本文引入 DRSN, 在特征层进行信息融合, 以及采用特征向量并行方式的模型的优势, 与其他文献中所述的基于 CNN、LSTM 及注意力机制所构建模型的识别效果进行比较的结果如表 3 所示。

表 3 DRSN-MF 模型与其他模型 UAR 对比

Table 3 Comparison of UAR between DRSN_MF and other models %

模型	UAR
DCNN-LSTM ^[26]	80.47
3D ACRNN ^[20]	82.41
DCNN-DTPM ^[5]	80.27
CNN-LSTM ^[14]	80.39
LSTM-self-Attention ^[19]	81.24
DRSN-MF	86.82

表 3 中的数据表明, DRSN 能够有效提高基于 CNN 和 LSTM 深度学习的声波信号故障识别精度; 同时, 在特征层进行信息融合, 以及采用特征向量并行方式能够提高模型的识别效果。

4 结束语

本文在 CNN-BLSTM 模型的基础上引入注意力机制作为基线模型, 构建了引入 DRSN, 在特征层进行信息融合, 以及采用特征向量并行方式识别的模型, 对锅炉炉管声波信号的故障进行识别。

利用原始样本数据对本文所构建的 DRSN-MF 模型进行了大量的实验验证, 并将本文模型与其他文献所述的模型进行对比, 验证了本文构建模型的有效性, 具有良好的推广价值和实用性能。

参考文献:

- [1] 何俊松. 炉管泄漏自动报警装置的技术改造 [J]. 广西电力, 2018, 41(1): 63–66.
HE Junsong. Technical modification of automatic warning system of boiler pipe leakage[J]. Guangxi electric power, 2018, 41(1): 63–66.
- [2] OH H, JUNG J H, JEON B C, et al. Scalable and unsupervised feature engineering using vibration-imaging and deep learning for rotor system diagnosis[J]. IEEE transactions on industrial electronics, 2018, 65(4): 3539–3549.
- [3] BOU-GHAZALE S E, HANSEN J H L. A comparative study of traditional and newly proposed features for recognition of speech under stress[J]. IEEE transactions on speech and audio processing, 2000, 8(4): 429–442.
- [4] LAUKKA P, NEIBERG D, FORSELL M, et al. Expression of affect in spontaneous speech: acoustic correlates and automatic detection of irritation and resignation[J]. Computer speech & language, 2011, 25(1): 84–104.
- [5] HSIAO P W, CHEN C P. Effective attention mechanism in dynamic models for speech emotion recognition[C]//2018 IEEE International Conference on Acoustics, Speech and Signal Processing. Calgary: IEEE, 2018: 2526–2530.
- [6] LIU Ruonan, YANG Boyuan, ZIO E, et al. Artificial intelligence for fault diagnosis of rotating machinery: a review[J]. Mechanical systems and signal processing, 2018, 108: 33–47.
- [7] HAN Zhiyan, WANG Jian. Speech emotion recognition based on Gaussian kernel nonlinear proximal support vector machine[C]//2017 Chinese Automation Congress. Jinan: IEEE, 2018: 2513–2516.
- [8] LI Yifan, LIANG Xihui, LIN Jianhui, et al. Train axle bearing fault detection using a feature selection scheme based multi-scale morphological filter[J]. Mechanical systems and signal processing, 2018, 101: 435–448.
- [9] 胡德生, 张雪英, 张静, 等. 基于主辅网络特征融合的语音情感识别 [J]. 太原理工大学学报, 2021, 52(5): 769–774.
HU Desheng, ZHANG Xueying, ZHANG Jing, et al. Feature fusion based on main-auxiliary network for speech emotion recognition[J]. Journal of Taiyuan University of Technology, 2021, 52(5): 769–774.
- [10] KIM J, SAUROUS R A. Emotion recognition from human speech using temporal information and deep learning[C]//Proc Interspeech 2018. Hyderabad: ISCA, 2018: 937–940.
- [11] 胡康, 何思宇, 左敏, 等. 基于 CNN-BLSTM 的化妆品违法违规分类模型 [J]. 智能系统学报, 2021, 16(6): 1151–1157.
HU Kang, HE Siyu, ZUO Min, et al. Classification model for judging illegal and irregular behavior for cosmetics based on CNN-BLSTM[J]. CAAI transactions on intelligent systems, 2021, 16(6): 1151–1157.
- [12] 刘威, 王薪予, 刘光伟, 等. 融合关系特征的半监督图像分类方法研究 [J]. 智能系统学报, 2022, 17(5): 886–899.

- LIU Wei, WANG Xinyu, LIU Guangwei, et al. Semi-supervised image classification method fused with relational features[J]. CAAI transactions on intelligent systems, 2022, 17(5): 886–899.
- [13] 苏丽, 孙雨鑫, 苑守正. 基于深度学习的实例分割研究综述 [J]. 智能系统学报, 2022, 17(1): 16–31.
SU Li, SUN Yuxin, YUAN Shouzheng. A survey of instance segmentation research based on deep learning[J]. CAAI transactions on intelligent systems, 2022, 17(1): 16–31.
- [14] 胡海华, 韩国军, 张孝谊. 基于卷积神经网络的闪存信道检测技术研究 [J]. 智能系统学报, 2021, 16(6): 1090–1097.
HU Haihua, HAN Guojun, ZHANG Xiaoyi. Research on flash memory channel detection technology based on convolutional neural network[J]. CAAI transactions on intelligent systems, 2021, 16(6): 1090–1097.
- [15] 潘国壮. 基于实时陆空通话情感识别的管制员疲劳状态快速监测研究 [D]. 南京: 南京航空航天大学, 2020.
PAN Guozhuang. Research on rapid monitoring of air traffic controller's fatigue state based on real-time emotion recognition of land-air communication[D]. Nanjing: Nanjing University of Aeronautics and Astronautics, 2020.
- [16] FARHOUDI Z, SETAYESHI S. Fusion of deep learning features with mixture of brain emotional learning for audio-visual emotion recognition[J]. *Speech communication*, 2021, 127: 92–103.
- [17] 李瑞航, 吴红兰, 孙有朝, 等. 基于深度残差收缩网络多特征融合语音情感识别 [J]. 数据采集与处理, 2022, 37(3): 542–554.
LI Ruihang, WU Honglan, SUN Youchao, et al. Speech emotion recognition base on multi-feature fusion of deep residual shrinkage network[J]. Journal of data acquisition and processing, 2022, 37(3): 542–554.
- [18] DANGOL R, ALSADOON A, PRASAD P W C, et al. Speech emotion recognition using convolutional neural network and long-short term memory[J]. Multimedia tools and applications, 2020, 79(43/44): 32917–32934.
- [19] 胡婷婷, 冯亚琴, 沈凌洁, 等. 基于注意力机制的 LSTM 语音情感主要特征选择 [J]. 声学技术, 2019, 38(4): 414–421.
HU Tingting, FENG Yaqin, SHEN Lingjie, et al. The salient feature selection by attention mechanism based LSTM in speech emotion recognition[J]. Technical acoustics, 2019, 38(4): 414–421.
- [20] 姜特, 陈志刚, 万永菁. 基于注意力机制的多任务 3D CNN-BLSTM 情感语音识别 [J]. 华东理工大学学报(自然科学版), 2022, 48(4): 534–542.
JIANG Te, CHEN Zhigang, WAN Yongjing. Multi-task learning 3D CNN-BLSTM with attention mechanism for speech emotion recognition[J]. Journal of East China University of Science and Technology(natural science edition), 2022, 48(4): 534–542.
- [21] GAO Zehai, MA Cunbao, SONG Dong, et al. Deep quantum inspired neural network with application to aircraft fuel system fault diagnosis[J]. *Neurocomputing*, 2017, 238: 13–23.
- [22] SOHAIB M, KIM C H, KIM J M. A hybrid feature model and deep-learning-based bearing fault diagnosis[J]. *Sensors (Basel, Switzerland)*, 2017, 17(12): 2876.
- [23] GUPTA A, ANJUM, GUPTA S, et al. InstaCovNet-19: A deep learning classification model for the detection of COVID-19 patients using Chest X-ray[J]. Applied soft computing, 2020, 99: 1–13.
- [24] 陈捷, 陆云. 2016 年度上海并网发电厂发电锅炉“四管”泄漏统计及案例分析 [J]. 电力与能源, 2017, 38(2): 170–175.
CHEN Jie, LU Yun. Statistics and case analysis of Shanghai grid plant power boiler four-tube leakages in 2016[J]. Power & energy, 2017, 38(2): 170–175.
- [25] WANG Tao, PEI Yu, XIAO Huiheng, et al. Detection of small gas leaks based on neural networks and D-S evidential theory using ultrasonics[J]. *Insight-non-destructive testing and condition monitoring*, 2014, 56(4): 189–194.
- [26] 宁方立, 韩鹏程, 段爽, 等. 基于改进 CNN 的阀门泄漏超声信号识别方法 [J]. 北京邮电大学学报, 2020, 43(3): 38–44.
NING Fangli, HAN Pengcheng, DUAN Shuang, et al. Identification method of valve leakage ultrasonic signal based on improved CNN[J]. *Journal of Beijing University of Posts and Telecommunications*, 2020, 43(3): 38–44.
- [27] ZHAO Minghang, ZHONG Shisheng, FU Xuyun, et al. Deep residual shrinkage networks for fault diagnosis[J]. *IEEE transactions on industrial informatics*, 2020, 16(7): 4681–4690.
- [28] 卢锦玲, 郭鲁豫. 基于改进深度残差收缩网络的电力系统暂态稳定评估 [J]. 电工技术学报, 2021, 36(11): 2233–2244.
LU Jinling, GUO Luyu. Power system transient stability assessment based on improved deep residual shrinkage network[J]. *Transactions of China electrotechnical society*, 2021, 36(11): 2233–2244.

作者简介:



杨正理, 副教授, 主要研究方向为复杂系统与计算智能、软件工程。发表学术论文 63 篇。



吴馥云, 讲师, 主要研究方向为新能源接入电网控制技术。发表学术论文 16 篇。



陈海霞, 副教授, 主要研究方向为海量信息处理的计算模型、自动推理。发表学术论文 33 篇。