

基于深度学习的反应器过程故障诊断方法研究

唐琪, 郭武, 刘怀海, 刘建平, 陆相印

湖南应用技术学院, 湖南 常德 415100

DOI: 10.61369/TACS.2026020040

摘 要 : 在工业反应器运行过程中, 高维过程信号的特征提取效果与噪声抑制水平, 会直接影响故障检测与诊断的整体性能。为解决这一问题, 本文设计一种新型多通道时序双向长短期记忆-门控循环单元神经网络 (MTBiLSTM-GRU), 用于高维过程信号的深度特征学习。首先采用小波变换对高维过程信号进行多频带特征提取, 分离不同频率分量信息; 其次通过快速傅里叶变换实现时域信号到频域信号的转换, 简化信号特征呈现形式; 最后利用 MTBiLSTM-GRU 网络从多尺度过程信号中学习具有区分度的时频联合特征。在 Tennessee Eastman 数据集上开展多组故障分类方法对比实验, 结果表明: 小波变换降噪条件下模型分类准确率达到 90.71%, 当选用 haar 小波基且降噪等级设置为 3 时, 模型分类准确率可提升至 96.1%。此外, 在行业内公认难以分类的第 3 类、第 9 类与第 15 类故障数据上, 本文模型分别实现 99.50%、98.63%、97.38% 的高分类精度。

关 键 词 : 故障检测; 时间序列提取; 信息融合

Research on Reactor Process Fault Diagnosis Method Based on Deep Learning

Tang Qi, Guo Wu, Liu Huaihai, Liu Jianping, Lu Xiangyin

Hunan University of Applied Technology, Changde, Hunan 415100

Abstract : In the operation of industrial reactors, the effectiveness of feature extraction and noise suppression of high-dimensional process signals directly affect the overall performance of fault detection and diagnosis. To address this issue, this paper designs a novel multi-channel temporal bidirectional Long Short-Term Memory-Gated Recurrent Unit neural network (MTBiLSTM-GRU) for deep feature learning of high-dimensional process signals. Firstly, wavelet transform is adopted to perform multi-band feature extraction on high-dimensional process signals to separate information of different frequency components; secondly, fast Fourier transform is used to realize the conversion from time-domain signals to frequency-domain signals, simplifying the presentation form of signal features; finally, the MTBiLSTM-GRU network is utilized to learn discriminative time-frequency joint features from multi-scale process signals. Comparative experiments on multiple fault classification methods are conducted on the Tennessee Eastman dataset. The results indicate that under wavelet transform denoising conditions, the model achieves a classification accuracy of 90.71%, and when the Haar wave is selected with a denoising level set to 3, the classification accuracy can be improved to 96.1%. Furthermore, on the industry-recognized difficult-to-classify Faults 3, 9, and 15, the proposed model achieves high classification accuracies of 99.50%, 98.63%, and 97.38%, respectively.

Keywords : fault detection; time series extraction; information fusion

引言

现代化工制造流程具有高度复杂性, 需要对多项工艺变量进行同步监测, 这就要求配备高效可靠的监测技术, 及时识别过程变量中的异常状态, 尤其是带有噪声干扰的过程信号。由此可见, 过程故障诊断本质上是依托传感器信号实现的多分类问题。传统故障诊断技术高度依赖信号预处理操作与人工诊断经验, 仅能完成浅层故障特征提取, 傅里叶变换作为平稳信号处理的常用手段^[1], 可完成时域到频域的信号转换。但工业现场的实际信号大多呈现非平稳、非线性特性, 无法适配传统方法。小波变换能够在时频域内对动态信号进行处理, 通过将信号分解为不同频带分量实现多分辨率分析, 精准表征信号局部时频特征^[2], 同时可由粗到细逐步观测信号, 实现原始信号噪声抑制。

深度学习技术已在图像识别、语音识别、机器翻译等领域取得显著突破。卷积神经网络 (CNN) 作为深度学习核心技术之一, 由 Krizhevsky 等人^[3]首次应用于图像识别领域, 可通过多层卷积运算自动完成特征提取。递归神经网络 (RNN) 同样是深度学习的重要

分支,但受递归、加权等运算影响,易出现梯度消失或梯度爆炸问题,无法有效处理长时序依赖关系。Shahnazari 等人^[4]将 RNN 用于执行器与传感器故障检测,Sun 等人^[5]提出贝叶斯递归神经网络并应用于田纳西-伊斯曼过程(TEP)故障诊断,但其对故障 3、9、15 的检测率仅为 5%、5%、7.12%,远低于其他故障类型。Kang 等人^[6]对比人工神经网络(ANN)与 RNN 在化工过程故障诊断中的性能,结果显示 RNN 在故障分类方面表现更优,但对故障 9 的分类效果较差,且无法区分故障 15。相较于 TEP 其他方法,RNN 可提升诊断精度,却仍难以改善故障 3、9、15 的识别效果。相关研究表明,LSTM、GRU 等改进型循环网络能够进一步提升 TEP 故障分类精度。Han 等人^[7]通过优化 LSTM 网络隐藏层节点数改进 TEP 故障诊断效果,故障 3、9 诊断准确率超过 90%,但故障 15 准确率仅为 47%。Zhang 和 Xiong^[8]采用 GRU 模型提升 TEP 故障诊断精度,故障 3、9 识别效果良好,故障 15 精度在 12%–70% 区间浮动。Somayeh Mirzaei 等人^[9]对比基础 LSTM 与 GRU 的故障诊断性能及分类机制,为化工过程故障诊断模型选型提供参考依据。

长短期记忆网络(LSTM)属于特殊的 RNN 结构,内部包含 LSTM 模块,可实现不定时长信息记忆,有效捕捉时序延迟关联信息。深度学习在其他领域的关键问题解决中也得到广泛应用。门控循环单元(GRU)是 LSTM 的高效变体,结构更简洁且性能表现优异。GRU 本质是去除输出门的 LSTM 网络,每个时间步都会将记忆单元信息完整写入整体网络,仅包含重置门与更新门两个门控结构,将遗忘门与输入门整合为更新门,同时合并隐藏状态与隐藏层输出。双向长短期记忆网络(BiLSTM)由前向 LSTM 与后向 LSTM 组合而成,LSTM 仅能提取上文信息,而 BiLSTM 可同时感知上下文信息,特征提取更全面。

基于上述分析,本文将 GRU 与 BiLSTM 进行融合,提出一种深度学习故障诊断方法,即基于 Multichannel Temporal BiLSTM-GRU 的时序提取模型,用于化工过程时序信号特征学习。该网络以处理时序信号的 LSTM 为基础架构,增设多通道输入层适配多尺度过程信号。在时序信号输入深层网络前,先通过小波变换完成多尺度分解,使 MTBiLSTM-GRU 能够从多通道学习不同尺度高级特征;再利用快速傅里叶变换将处理后的数据转为频域信号,获取不同故障源的鉴别特征;最终通过 MTBiLSTM-GRU 完成多传感器数据融合,全面表征设备运行状态,提升整体分类性能。

本文的主要贡献如下:

- (1) 提出基于 Multichannel Temporal BiLSTM-GRU 的时序提取模型(MTBiLSTM-GRU),实现时频域特征权重融合,有效提升故障诊断分类精度。
- (2) 针对 TE 数据集高维信号噪声问题,采用小波变换将高噪声信号转为低噪声方波信号,结合快速傅里叶变换提取强辨别力频域信息,降低噪声对模型精度的影响,优化故障分类效果。
- (3) 结合 GRU 时序语义关联捕捉能力与 BiLSTM 空间特征提取优势,弥补单一网络缺陷,使模型学习更全面的时频特征。

一、相关理论

(一) BiLSTM/GRU

长短期记忆网络(LSTM)是循环神经网络(RNN)的改进形式,传统 RNN 在实际应用中存在梯度消失、梯度爆炸及长距离依赖建模能力弱等问题,LSTM 针对上述缺陷进行优化。LSTM 主体结构结构与 RNN 相近,核心改进是在隐藏层 h 中增设遗忘门、输入门、输出门三个门控结构,并新增细胞状态(cell state)存储关键信息。LSTM 隐藏层结构原理如下图所示, $f(t)$ 、 $i(t)$ 、 $o(t)$ 分别表示 t 时刻遗忘门、输入门、输出门的值, $a(t)$ 表示 t 时刻对 $h(t-1)$ 和 $x(t)$ 的初步特征提取。

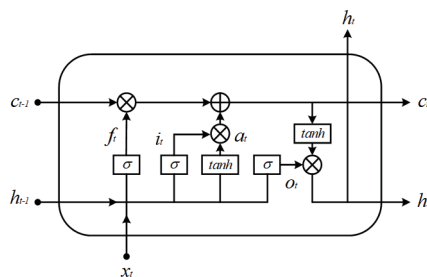


图 1 LSTM 单元结构图

具体的计算过程为:

$$f(t) = \sigma(W_f h_{t-1} + U_f x_t + b_f)$$

$$i(t) = \sigma(W_i h_{t-1} + U_i x_t + b_i)$$

$$o(t) = \sigma(W_o h_{t-1} + U_o x_t + b_o)$$

$$a(t) = \tanh(W_a h_{t-1} + U_a x_t + b_a)$$

其中, x_t 表示 t 时刻的输入, h_{t-1} 表示 $t-1$ 时刻隐层状态; W_f 、 W_i 、 W_o 和 W_a 分别为遗忘门、输入门、输出门、特征提取针对 h_{t-1} 的权重系数; U_f 、 U_i 、 U_o 和 U_a 分别为对应结构针对 x_t 的权重系数; b_f 、 b_i 、 b_o 和 b_a 为对应结构偏置值; $\tanh$ 为双曲正切激活函数, σ 为 Sigmoid 激活函数。

$$\tanh(x) = \frac{1 - e^{-2x}}{1 + e^{-2x}}$$

$$\sigma(x) = \frac{1}{1 + e^{-x}}$$

遗忘门与输入门运算结果作用于 $c(t-1)$, 生成 t 时刻细胞状态 $c(t)$, 用公式表示为:

$$c(t) = c(t-1) \odot f(t) + i(t) \odot a(t)$$

其中, $\odot$ 为 Hadamard 积。 t 时刻隐藏层状态 $h(t)$ 由输出门 $o(t)$ 与当前细胞状态 $c(t)$ 共同计算得到:

$$h(t) = o(t) \odot \tanh(c(t))$$

双向长短期记忆网络(BiLSTM)包含两个独立 LSTM 结构,输入序列分别以正序、逆序输入网络进行特征提取,将两组输出特征向量拼接后作为最终特征表达。BiLSTM 可使 t 时刻特

征同时包含历史与未来时序信息，实验证明其特征提取效率与性能优于单向 LSTM。

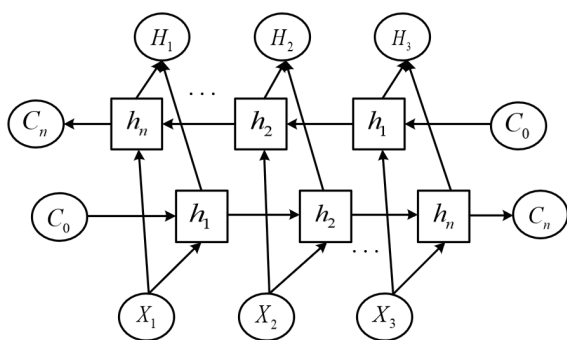


图2 BiLSTM 结构图

GRU 是 LSTM 的轻量化变体，主要用于解决长时序记忆与反向传播梯度异常问题。GRU 与 LSTM 在多数场景下性能接近，但 GRU 结构更简单、计算量更小、易于实现：

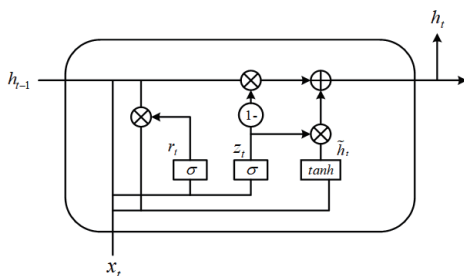


图3 GRU 单元结 + 构图

具体的计算过程为：

$$z(t) = \sigma(x_t W_{xz} + h_{t-1} W_{hz} + b_z)$$

$$r(t) = \sigma(x_t W_{xr} + h_{t-1} W_{hr} + b_r)$$

$$\tilde{h}(t) = \tanh(x_t W_{hx} + r_t \odot h_{t-1} W_{hh} + b_h)$$

$$h(t) = (1 - z_t) \odot h_{t-1} + z_t \odot \tilde{h}_t$$

其中， $\odot$ 为 Hadamard 积； $z(t)$ 为更新门， $r(t)$ 为重置门，

$\tilde{h}(t)$ 为候选隐藏层状态， $h(t)$ 为最终隐藏状态； W_{hz} 、 W_{hr} 和 W_{hh} 分别表示更新门、重置门和特征提取过程中 h_{t-1} 的权重系数； W_{xz} 、 W_{xr} 和 W_{hx} 分别为对应结构针对 x_t 的权重系数； b_z 、 b_r 和 b_h 分别表示更新门、重置门和特征提取过程中的偏置值。

GRU 仅设置两个门控结构，将 LSTM 的输入门与遗忘门合并为更新门 $z(t)$ ，用于控制历史信息传递至当前时刻的比例；另一门控为重置门 $r(t)$ ，用于控制历史信息遗忘程度。总体而言，LSTM 与 GRU 均通过门控函数保留关键特征，避免长时序传播中的特征丢失。GRU 相较于 LSTM 减少一个门控结构，参数量更少、训练速度更快，具体应用需结合实际场景选择。

（二）基于 Multichannel Temporal BiLSTM\GRU 的时间序列提取模型

本文提出的模型处理流程分为数据预处理与故障诊断两个阶段：

1、数据预处理：首先对时序数据进行标准化处理，将数值缩放至 $[-1, 1]$ 区间，减少后续计算量；随后采用小波变换对标准化数据降噪，生成高识别率方波信号，提升故障检测能力；接着通过快速傅里叶变换将时域信号转为频域信号，进一步优化分类精

度；最后在数据输入模型前再次标准化，最大限度降低模型计算负荷。

2、故障诊断：预处理完成后，数据分别输入 BiLSTM 与 GRU 模型进行多维度特征提取，从动态过程信号的时频域中学习高级有效特征；最终输入分类器完成分类并输出检测结果：

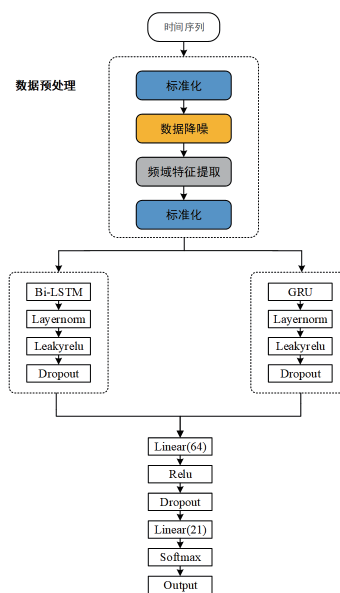


图4 Multichannel Temporal BiLSTM\GRU 模型结构图

二、实验结果与分析

（一）数据预处理

数据噪声会严重影响模型训练效果，不仅增加数据量与计算开销，占用计算机内存资源，还会增大计算误差。尤其是迭代寻优类线性算法，大量噪声会大幅降低收敛速度，甚至破坏模型训练精度。

常用数据降噪方法包括 Butterworth 滤波、FIR 滤波、移动平均滤波、中值滤波、维纳滤波、自适应滤波、小波变换等。本文对七种降噪方法分别进行处理，并统计模型分类结果。分类结果如表1所示。

表1 不同数据降噪预处理方式结果对比

数据处理方式	ACC
Butterworth 滤波	0.7392
FIR	0.7573
移动平均滤波	0.8643
中值滤波	0.8722
维纳滤波	0.8321
自适应滤波	0.8532
小波变换	0.9071

由表1可知，移动平均滤波、中值滤波、维纳滤波、自适应滤波四种方法分类准确率约为 85%，七种方法中小波变换降噪效果最优，模型分类准确率达到 90.71%，证明小波变换更适用于本文模型的信号预处理环节。

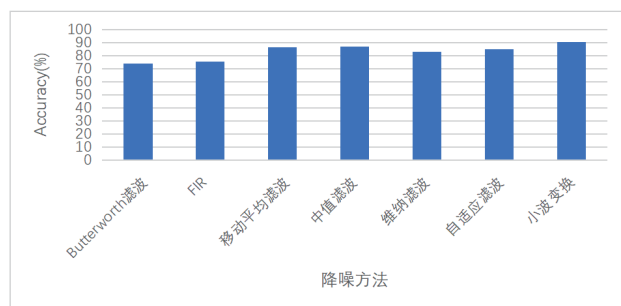


图5 不同数据降噪预处理方式结果对比

(二) 与其他论文的实验结果对比

将本文 MTBiLSTM-GRU 模型与 BiLSTM、GRU、MC1-DCNN、SimCLR-TS 方法进行 21 类故障分类对比, 结果如表 2 所示。在不考虑正常样本 Case0 的条件下, 本文模型准确率表现突出; 针对行业内难分类的第 3 类、第 9 类、第 15 类故障数据, 模型实现高精度分类, 充分证明所提模型性能优于现有多数研究方法。

表2 与其他论文的实验结果对比

Case	MTBiLSTM-GRU	BiLSTM	GRU	MC1-DCNN [10]	SimCLR-TS [11]
Case 0					55.12
Case 1	94.63	96.25	91.25	100.00	96.29
Case 2	98.13	98.63	94.63	100.00	100.00
Case 3	99.50	96.38	99.38	83.48	25.57
Case 4	78.50	91.88	76.88	99.22	96.57
Case 5	97.00	96.88	96.00	90.40	83.86
Case 6	93.38	70.88	100.00	93.10	99.14
Case 7	97.38	97.75	85.38	100.00	100.00
Case 8	96.74	92.13	62.63	99.27	86.71

Case 9	98.63	87.38	76.25	74.60	29.43
Case 10	99.13	96.38	100.00	93.75	91.71
Case 11	100.00	100.00	91.50	95.45	99.29
Case 12	98.63	94.50	100.00	100.00	95.14
Case 13	100.00	88.75	81.25	99.14	57.86
Case 14	97.13	91.00	97.00	100.00	99.57
Case 15	97.38	87.00	85.25	73.50	2.71
Case 16	93.25	89.50	94.75	100.00	98.57
Case 17	100.00	96.50	93.88	100.00	100.00
Case 18	86.38	91.88	90.38	98.43	99.00
Case 19	99.75	100.00	98.00	100.00	100.00
Case 20	98.75	95.38	94.38	94.12	99.14
Case 21	93.88	95.88	90.75	70.59	81.71
Average	96.10	93.09	90.45	94.08	81.43

三、结论

本文提出一种面向工业过程故障诊断的 MTBiLSTM-GRU 新型 LSTM 模型。首先通过小波变换依据时频信息对信号降噪处理, 生成高识别率方波信号; 随后采用快速傅里叶变换完成时域到频域的信号转换, 简化信号特征; 最后依托 MTBiLSTM-GRU 多通道深度学习框架, 适配多通道信号的多分辨率输入过程数据, 通过隐含层学习时频域典型特征并完成多通道信号融合。

在 Tennessee Eastman 数据集上的测试结果表明, 该方法表现出优异的故障诊断性能, 相较于其他深度学习模型, MTBiLSTM-GRU 实现更高的故障分类精度, 数据集整体识别准确率达到 96.1%。尤其针对第 3 类、第 9 类、第 15 类难分类故障, 模型分别实现 99.50%、98.63%、97.38% 的高精度识别, 可为工业反应器过程故障诊断提供有效技术支撑。

参考文献

- [1]Faradisa, I. S., et al. (2016). Identification of phonocardiogram signal based on STFT and Marquart Lavenberg Backpropagation. 2016 International Seminar on Intelligent Technology and Its Applications (ISITIA).
- [2]Daubechies, I. (1990). "The wavelet transform, time-frequency localization and signal analysis." IEEE Transactions on Information Theory 36(5): 961-1005.
- [3]Krizhevsky A, Sutskever I, Hinton GE (2017) ImageNet classification with deep convolutional neural networks. Commun ACM 60(6):84 - 90
- [4]Shahnazari, H. (2020). "Fault diagnosis of nonlinear systems using recurrent neural networks." Chemical Engineering Research and Design 153: 233-245.
- [5]Sun, W., et al. (2020). "Fault detection and identification using Bayesian recurrent neural networks." Computers & Chemical Engineering 141: 106991.
- [6]Kang, J.-L. (2020). "Visualization analysis for fault diagnosis in chemical processes using recurrent neural networks." Journal of the Taiwan Institute of Chemical Engineers 112: 137-151.
- [7]Han, Y., et al. (2020). "An optimized long short-term memory network based fault diagnosis model for chemical processes." Journal of Process Control 92: 161-168.
- [8]Zhang, R. and Z. Xiong (2019). Recurrent Neural Network Model with Self-Attention Mechanism for Fault Detection and Diagnosis. 2019 Chinese Automation Congress (CAC).
- [9]Mirzaei, S., et al. (2022). "A comparative study on long short-term memory and gated recurrent unit neural networks in fault diagnosis for chemical processes using visualization." Journal of the Taiwan Institute of Chemical Engineers 130.
- [10]Yu, J., et al. (2020). "Multichannel one-dimensional convolutional neural network-based feature learning for fault diagnosis of industrial processes." Neural Computing and Applications 33(8): 3085-3104.
- [11]Pöppelbaum, J., et al. (2022). "Contrastive learning based self-supervised time-series analysis." Applied Soft Computing 117.