

DOI:10.19651/j.cnki.emt.2211454

基于多尺度融合模型的化工故障诊断

杨晓岗 夏涛

(北京化工大学信息科学与技术学院 北京 100029)

摘要: 针对日前化工过程故障诊断中降噪效果不佳、多尺度特征未区分重要性、时序特征提取不充分等问题,本文提出了一种基于多尺度融合模型的化工故障诊断方法,该方法将注意力机制分别与软阈值方法和多尺度学习相结合,构建了多尺度深度残差收缩网络,并将提取到的多尺度空间特征送入双向门控循环单元进一步提取时序特征,相比于单通道网络,双向门控循环单元不仅能够完成对过去信息的学习,而且还能够完成对未来信息的学习,因此能够得到更多的时间关联信息。最后使用修正田纳西-伊斯曼过程数据进行验证,最终取得了95.08%的分类精度和94.76%的召回率,明显优于对比方法,证明了方法的有效性。

关键词: 故障诊断;深度学习;多尺度学习;注意力机制;融合模型

中图分类号: TP391.5 **文献标识码:** A **国家标准学科分类代码:** 520.2060

Chemical fault diagnosis based on multi-scale fusion model

Yang Xiaogang Xia Tao

(School of Information Science and Technology, Beijing University of Chemical Technology, Beijing 100029, China)

Abstract: Aiming at the problems of poor noise reduction effect, failure to distinguish the importance of multi-scale features, and insufficient extraction of temporal features in the current fault diagnosis of chemical processes, this paper proposes a chemical fault diagnosis method based on multi-scale fusion model. In this method, the attention mechanism is combined with soft threshold method and multi-scale learning respectively, and a multi-scale deep residual shrinkage network (MDRSN) is constructed. Moreover, the extracted multi-scale spatial features are sent to the bidirectional gated cyclic unit (BIGRU) to further extract temporal features. Compared with the single-channel network, BIGRU can not only complete the learning of past information, but also complete the learning of future moment information, so more temporal correlation information can be obtained. Finally, the modified Tennessee-Eastman process data were used to verify the classification accuracy of 95.08% and the recall rate of 94.76%, which was obviously better than the comparison method, and the effectiveness of the method was proved.

Keywords: fault diagnosis; deep learning; multi-scale learning; attention mechanism; fusion model

0 引言

随着科学技术的不断发展,现代化工产业在提高产品质量和生产效率、促进社会的进步方面起到的作用越来越显著,为了获得更高的生产效益,企业的生产规模越来越大,但是生产规模扩大带来效益提升的同时也使得生产过程的非线性、不确定性、不稳定性和强关联性更加突出,一旦发生生产事故^[1-2],给人民群众带来的生命财产损失也变得更严重,因此,及时对化工过程进行故障诊断和分类变得极为重要。

自20世纪故障诊断技术被提出以来,其在保障企业的生产效益以及人民群众的生命财产安全等方面起到了无可代替的作用,随着时间的推移,故障诊断技术逐渐发展出了

3个大的方向^[3],即基于机理模型的方法、基于知识的推理方法、基于数据驱动的方法。由于近年来工厂智能化的推进提供了大量的数据,使得基于数据驱动的方法得到迅猛发展,其根据采用手段的不同具体可以划分为基于统计的方法和基于机器学习的方法。常用的统计方法包括主成分分析(principal component analysis, PCA)^[4-5]、独立成分分析(independent component analysis, ICA)^[6-7]和偏最小二乘(partial least squares, PLS)^[8-9]方法等。虽然这类方法是处理工业过程数据的常用方法,但是当数据中存在高维非线性时往往无法取得满意的效果。随着计算机处理能力的提升,能够提取数据中高维特征的深度学习等方法得到了越来越多的关注,其中卷积神经网络(convolutional neural

收稿日期:2022-09-19

network, CNN)是深度学习中最具代表性的之中方法。文献[10]将卷积神经网络用于齿轮箱故障的诊断,其通过小波变换将时序信号转换为时频域信号作为输入,构建了卷积神经网络提取故障特征,卷积神经网络通常用来处理二维数据,当处理一维数据时需将一维信号转换为二维矩阵作为卷积神经网络的输入,这样不仅会提高计算成本,而且可能会忽略一维信号中的重要特征。文献[11]首先提出了使用一维卷积神经网络(one-dimensional convolutional neural network, 1D-CNN)用于心电信号的分类和异常检测,证明了对一维信号特征提取的优越性,随着相关学者的研究,一维卷积神经网络已广泛应用于心电特征提取^[12]、轴承损伤检测^[13-15]、自然语言处理^[16]等各方面。文献[17]将 1D-CNN 用到了化工过程故障诊断方面,并取得了不错的效果。文献[18]将长短时记忆网络(long short term memory, LSTM)与 CNN 结合,提出了一种 CNN-LSTM 模型用于轴承故障诊断工作,该模型综合了 LSTM 网络能够记忆时序信号的优势和 CNN 在特征提取上的能力,并取得了 99.41%的分类准确率。文献[19]将长短期记忆网络与注意力机制结合来处理小波降噪后的 TE 数据,对故障 1~故障 5 进行分类,获得了 93.05%的分类精度。但是,由于小波降噪的局限性,存在处理数据时易滤除有效信息的问题。文献[20]将 TE 过程划分为不同的子单元,并对每个子单元用时序卷积网络提取特征,然后将提取后的特征进行融合送入下一网络进行故障分类,最终使用所提出的多块时序卷积网络(multiblock temporal convolution network, MBTCN)在不包含故障 3、故障 9、故障 15 的 TE 数据上取得 95%的分类精度。虽然该法提取的特征包含时序信息,但由于使用的卷积核尺寸单一,易忽略重要特征。文献[21]利用小波变换将 TE 过程信号进行分解,并利用不同尺度的卷积核对分解后的不同频段信号进行特征提取,根据融合后的特征进行故障分类,最终使用所提出的多尺度动态自适应残差网络(multi-scale dynamic adaptive

residual network, MSDACRN)得到了 94%的分类准确率。

虽然上述方法在化工故障诊断方面取得了不错的效果,但随着网络层数的增加,会出现模型退化和过拟合的问题,且实际化工生产过程往往存在较强的噪声干扰,既往研究多采用小波降噪来处理数据中的噪声,但即便是同一种故障类型,由于运行工况和故障程度的不同,其信号就会具备不同的特点,导致了小波降噪中的阈值难以普遍适用,因此这种方式会使得过程数据中的一部分有用信号当作噪声信号被滤除,造成滤波后的分类准确度降低。同时,由于故障信号多为低频信号,大部分卷积网络时通常使用较小的卷积核来提取这些低频故障特征,这样就造成了部分高频故障特征的丢失,使得最终无法取得较好的故障诊断结果。为了解决这些问题,本文提出了一种融合深度残差收缩网络(multi-scale deep residual networks, MDRSN)与门控循环单元(bidirectional gating recurrent unit, BIGRU)的多尺度故障诊断方法,最终通过 TE 数据集验证方法的有效性。

1 MDRSN-BIGRU 模型

1.1 模型优点

本文所提出的模型结构如图 1 所示,该模型具有如下优点:

1)在一维卷积神经网络中加入残差结构,通过恒等映射来解决随着网络层数增加而带来的特征损失,减弱了模型退化和过拟合的问题。

2)采用残差收缩块作为提取空间特征的主要结构。其将软阈值方法和注意力机制相结合,为每个通道赋予了不同权重,实现了特征提取部分与降噪部分的一体化,对不同工况下的过程数据实现自适应的降噪。

3)使用了卷积核大小分别为 1×3 、 1×5 、 1×7 的残差收缩块并行处理输入数据,将 3 个分支提取到的特征拼接,并使用注意力机制为每个特征通道赋予权重,使模型能够更好地利用不同尺度的特征,从而提高诊断精度。

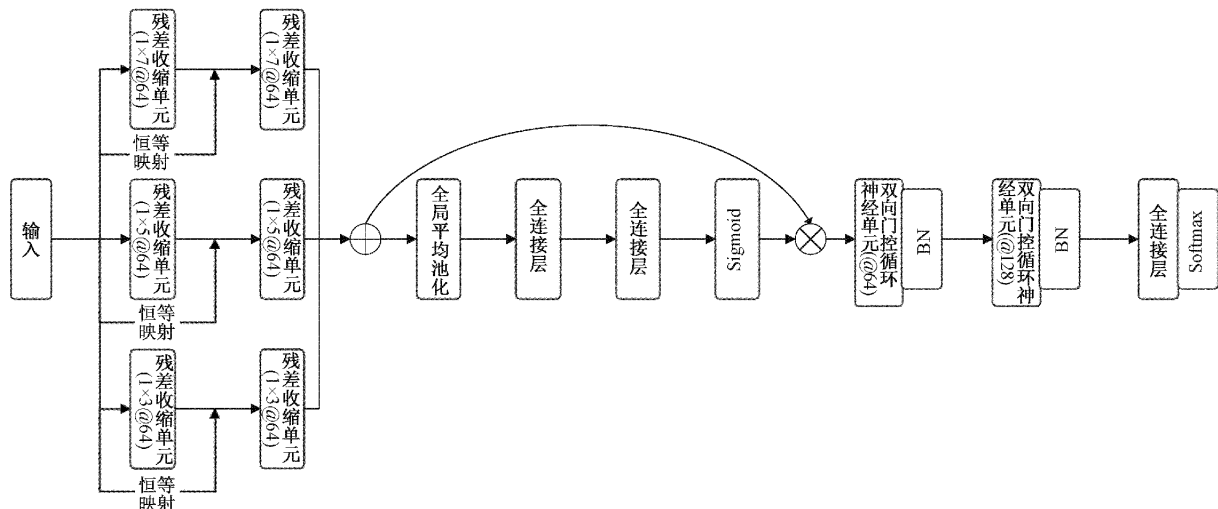


图 1 MDRSN-BIGRU

4)针对仅依靠 CNN 难以提取化工过程数据中的时序特征问题,本文将经过多尺度深度残差收缩网络后提取到的特征送入双向门控循环单元中进行进一步的特征提取,来获取深层时间关联信息。

1.2 模型组成

1)一维卷积层

一维卷积与二维卷积的主要区别在于输入维度的不同,可以直接对一维序列数据进行特征提取,对于输入信号为一维的数据来说,其避免了信号转换过程中的特征损失、提高了计算效率,其计算过程如图 2 所示。

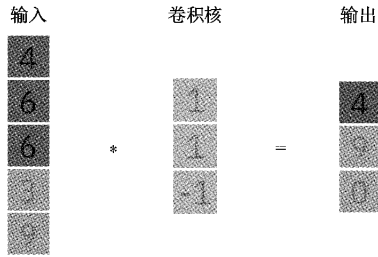


图 2 一维卷积计算过程

一维卷积操作利用一维核滤波器自上而下的计算,其计算过程如式(1)所示。

$$x_j^{(l)} = f(b_j^{(l)} + \sum_{i=1}^{N_{l-1}} conv1D(W_{ij}^{(l-1)}, x_i^{(l-1)})) \quad (1)$$

其中, $x_j^{(l)}$ 是第 l 层第 j 个神经元的输出, f 表示非线性激活函数, $b_j^{(l)}$ 表示第 l 层第 j 个神经元的偏置, $conv1D$ 表示一维卷积运算, $W_{ij}^{(l-1)}$ 表示从 $l-1$ 层的第 i 个神经元~ l 层第 j 个神经元的权重, $x_i^{(l-1)}$ 表示第 $l-1$ 层的输入。

2)残差收缩块

深度残差收缩网络的主要思想是软阈值化与注意力机制的结合。通过注意力机制学习输入信号的特征重要程度,将学习到的特征权重与原信号相乘,从而将注意力集中在重要特征,来达到消除噪声的目的,由于对于不同信号学习到的权重不同,降噪阈值并不是固定不变的,而是因此阈值可以根据输入信号的特点不断变化,来达到更好的降噪效果。

软阈值化是一种数据收缩方法,当输入信号在一定阈值内时,将输入信号置为 0,当信号大于阈值时,将其减去阈值,当输入信号小于阈值时,将其加上阈值,其公式如式(2)、(3)所示。

$$y = \begin{cases} x - \tau, & x > \tau \\ 0, & -\tau \leq x \leq \tau \\ x + \tau, & x < -\tau \end{cases} \quad (2)$$

对上述公式求导可得:

$$\frac{\partial y}{\partial x} = \begin{cases} 1, & x > \tau \\ 0, & -\tau \leq x \leq \tau \\ 1, & x < -\tau \end{cases} \quad (3)$$

可以观察到,软阈值函数的导数为 0 或 1,因此,其能

有效地防止梯度爆炸和梯度消失。

注意力机制是一种权重分配机制,其通过调整权重系数来修改不同特征通道的重要程度,本文采取的注意力机制参考文献[22]提出的 SENet (squeeze-and-excitation networks),SENet 通过一个小型的子网络来学习一组权重,将学习到的权重与特征通道相乘,从而得到带有不同权重的特征图,其结构图如图 3 所示。

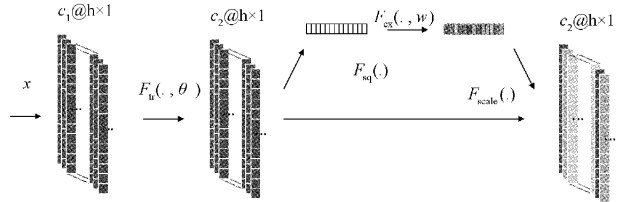


图 3 SENet 网络结构

SENet 主要由 3 个操作组成,分别是压缩、激发、缩放,在输入 x 经过一系列变换得到一个特征通道数为 c_2 的特征图后,将其沿时间维度进行特征压缩,将每个特征通道变为一个实数;激发操作则是通过学习一组参数来为每个特征通道生成权重;最后经过缩放操作即将得到的权重与先前的特征相乘,把信号按重要程度进行缩放。

本文使用的残差收缩块结构如图 4 所示,特征图 x_1 经卷积、批归一化等操作得到特征图 x_2 ,特征图 x_2 通过全局平均池化层展平,然后进入一个由两层全连接测层组成的子网络,使用 Sigmoid 函数进行归一化得到向量 α ,将 α 与展平后的信号 x_2 相乘得到软阈值,此时再将特征图减去软阈值即可得到降噪后的特征图,最终通过恒等映射将其

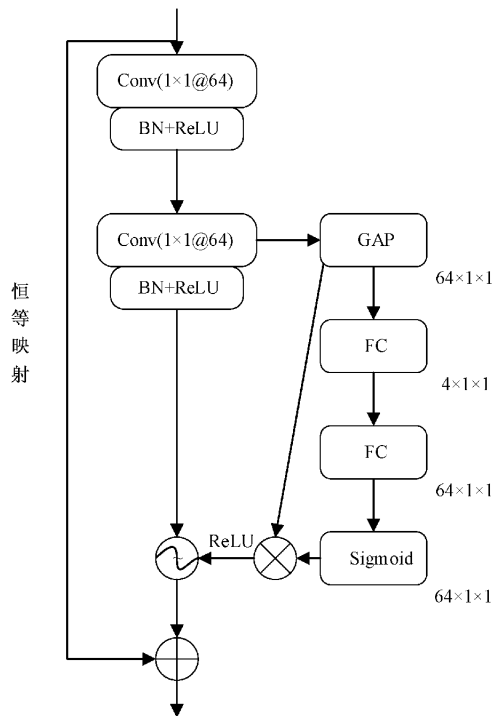


图 4 残差收缩块

与原特征图 x_1 相加,避免因网络层数过深而造成的网络退化。

3)多尺度特征注意力机制

通常卷积神经网络仅使用单一尺度的卷积核进行特征提取,而单一尺度的卷积核仅能提取固定频段的特征,近年来,虽也有专家提出使用多个卷积核来提取不同尺度的特征,然后融合不同尺度的特征最后获得输出,但这些不同尺度的卷积核的特征学习能力都是等价的,即所获得特征权重相同,而故障特征多为低频信号,因此,不加区分的使用多尺度卷积显然是不合适的,为了解决这个问题,本文将注意力机制与多尺度卷积相结合,最终得到带权重的多尺度特征。

输入信号经过 3 个具有不同尺度的残差收缩网络进行特征提取以获得 h_1, h_2, h_3 , 其中 3 个残差收缩网络中卷积核的大小分别为 $1 \times 3, 1 \times 5, 1 \times 7$, 将 3 个分支提取到的特征进行拼接获得新的特征向量 U , U 的计算公式如下所示:

$$U = \varphi[h_1, h_2, h_3] \quad (4)$$

特征向量 U 使用注意力机制为其附加权重后输入后续网络,这里使用的注意力机制与上文残差收缩块中的注意力机制相同,这里不做赘述。

4) 双向门控循环单元网络

门控循环单元网络(gated recurrent unit, GRU)是 LSTM 的一种简化变体,相比于 LSTM,它通常能以更快的计算速度达到相似的效果,GRU 隐含层结构如图 5 所示。

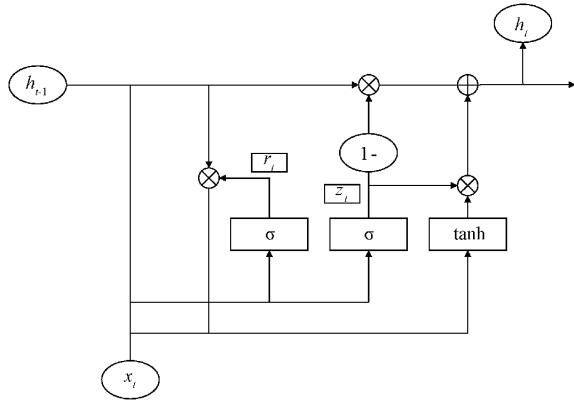


图 5 GRU 隐含结构

其中, h_{t-1} 为上一时刻的隐藏状态, x_t 为当前时刻输入信息, \tilde{h}_t 为候选隐藏状态, h_t 为传递到下一时刻的隐藏状态, r_t 为重置门,作用是将 t 时刻的输入信息 x_t 与 $t-1$ 时刻的隐藏状态 h_{t-1} 相结合, z_t 为更新门,其用于控制前一时刻的状态信息被代入到当前状态中的程度。

其计算公式如下:

$$r_t = \sigma(x_t W_r + h_{t-1} U_r + b_r) \quad (5)$$

$$z_t = \sigma(x_t W_z + h_{t-1} U_z + b_z) \quad (6)$$

$$h_t = (1 - z_t) \odot h_{t-1} + z_t \odot \tilde{h}_t \quad (7)$$

$$\tilde{h}_t = \tanh(x_t W_h + (r_t \odot h_{t-1}) U_h + b_h) \quad (8)$$

式中: $W_r, W_z, W_h, U_r, U_z, U_h$ 为权重矩阵, σ 为 sigmoid 激活函数, \odot 为向量乘积。

上述为经典 GRU 的模型结构,特征向量是根据序列由前向后单向传播的,只能捕捉当前时刻之前的历史信息,而化工过程故障诊断不仅需要考虑到故障发生之前的数据,还需考虑到故障发生之后的数据,因此,本文采用了双向门控循环单元的结构,其结构如图 6 所示。

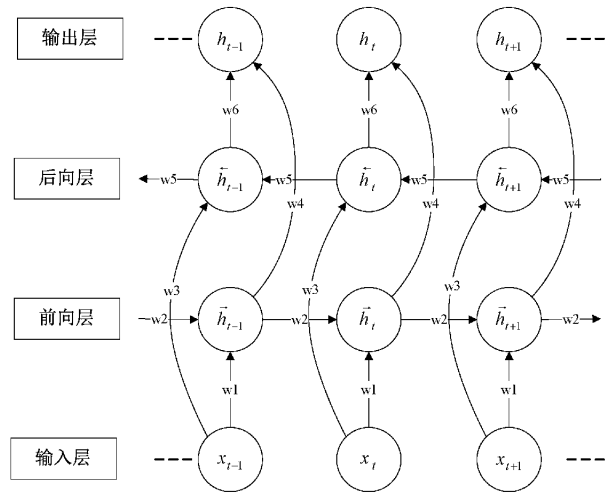


图 6 双向门控循环单元结构

图 6 中 x_t 为 t 时刻输入信号, \tilde{h}_t 为前向网络 t 时刻隐藏状态, \hat{h}_t 为后向网络 t 时刻隐藏状态, h_t 为 t 时刻输出,其计算公式如下:

$$\tilde{h}_t = GRU(x_t, \tilde{h}_{t-1}) \quad (9)$$

$$\hat{h}_t = GRU(x_t, \hat{h}_{t-1}) \quad (10)$$

$$h_t = W_h \tilde{h}_t + V_h \hat{h}_t + b_h \quad (11)$$

双向门控循环单元将同时输入给一个训练序列向前的 GRU 网络和一个训练序列向后的 GRU 网络,并将这两个网络接至同一个输出层。相比于单通道 GRU 网络,双通道 GRU 网络不仅能够完成对过去信息的学习,而且还能够完成对未来时刻信息的学习,因此能够得到更多的时间关联信息。

2 Tennessee Eastman 化工过程

2.1 数据采集

TE 过程是由田纳西州伊士曼公司的 Downs 和 Vogel^[23] 根据实际生产提出仿真模型,其仿真数据常用做化工过程故障诊断的研究。本文所用为修订后的 TE 过程,该过程共拥有 12 个操纵变量和 41 个测量变量,尽管修订版本中有 28 个过程故障类型,但为了方便与其他算法进行对比,本文仅选取了 IDV(1)~IDV(20),其工艺流程如图 7 所示。

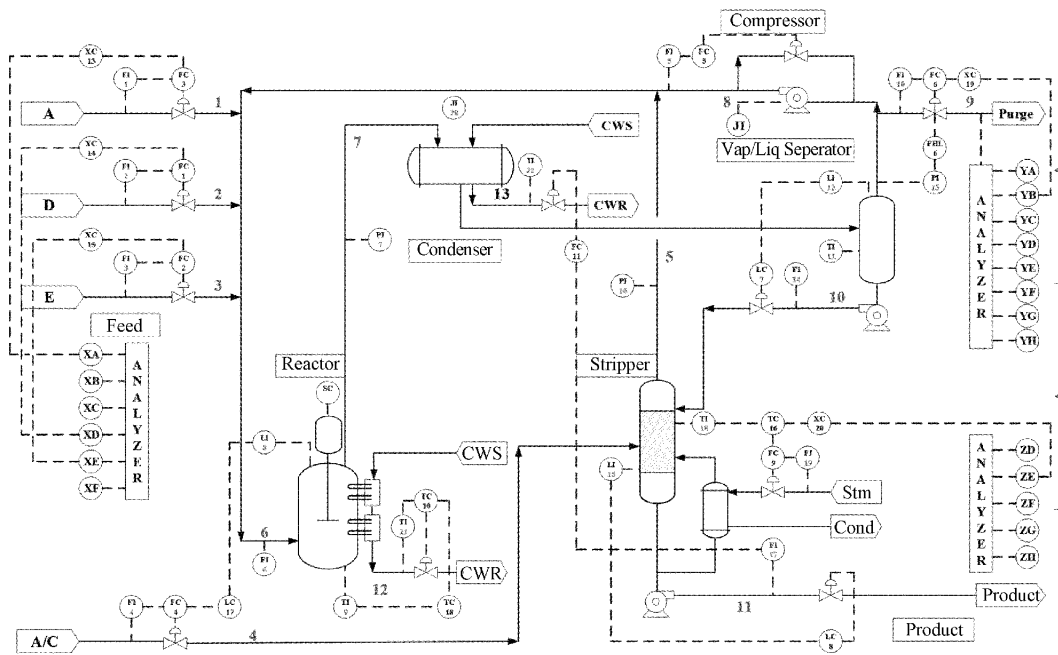


图 7 TE 工艺流程

仿真数据采样周期为 3 min, 采集正常样本时, 模拟器在正常状态下运行 500 h, 共采集到 10 000 个正常样本, 采集正常样本时, 模拟器先正常运行 10 h, 然后引入相应的故障, 再运行 10 h, 一次采集 200 个样本, 在不同的初始状态下 20 种故障各采集 10 次, 共采集到 40 000 个故障样本。选择 70% 数据作为训练集, 其余数据作为测试集。

2.2 数据处理

1) 数据归一化

在 TE 数据等化工数据中, 其采集的不同特征具有不同的量纲和单位, 为了解决不同数据间的可比性, 加快网络收敛, 本文采用 z-score 归一化处理数据, 其公式如下:

$$Z(x) = \frac{x - \bar{x}}{S(x)} \quad (12)$$

其中, x 为原始数据, \bar{x} 为样本均值, $S(x)$ 为样本标准差。

2) 标签平滑

将归一化后的数据转换为 one-hot 编码, 本文在进行 one-hot 编码转换时使用了标签平滑技术, 其计算公式如下:

$$Y = \begin{cases} 1 - \alpha, & i = target \\ \frac{\alpha}{K}, & i \neq target \end{cases} \quad (13)$$

传统 one-hot 编码中, 通常将训练数据中标签向量的目标类别概率应为 1, 非目标类别概率应为 0, 这种编码方式会在使用交叉熵函数作为损失函数时, 使得到的预测向量中目标类别的值趋于无穷大, 非目标类别趋于无穷小, 而过大的差值会使得模型缺乏适应性, 在样本数据不能覆

盖全部情况时导致过拟合, 在加入标签平滑后, 使得样本的标签不再是非 1 即 0, 能够避免对于正确标签的过于自信。

3 仿真实验与结果分析

本文所用网络模型的主要结构如图 3 所示, 由于数据量较多, 本文在每个卷积层后加入了最大池化层, 池化大小为 2, 同时为了降低过拟合提高泛化能力, 本文在 softmax 层前交替加入了两层 dropout 与全连接层, 其中 dropout 层的大小为 0.3, 全连接层的节点数分别为 128 和 64, 在卷积层和全连接层加入了 L2 正则化机制, 正则参数为 0.01, 在模型编译阶段, 使用了对学习率具有更强约束作用的 Nadam 优化器, 初始学习率设置为 0.001, 使用交叉熵函数作为损失函数。

为了反映所提出的模型在化工过程故障诊断中的有效性, 本文选取了 1D-CNN、GRU、CNN-GRU、MBTCN、MSDACRN 与本文所提出的 MDRSN-BIGRU 网络进行对比, 表 1 中列出了对照模型的详细架构。

表 2 显示了 MDRSN-BIGRU 与其他方法在 TE 过程上的故障诊断精度 (Precision) 和召回率 (Recall), 从表 2 可以看出, 从加权平均精度和召回率两个指标可以看出, 基于 MDRSN-BIGRU 的故障诊断方法效果明显好于其他方法, 其加权平均精度达到了 95.08%, 是所有模型中最高的, 在比较模型中, 仅经过卷积特征提取的 1D-CNN 表现最差, 其加权平均精度只有 80.87%, 相比于 1D-CNN, 具备长时间历史信息记忆能力的 GRU 网络表现效果要好, 说明合理的时序特征在进行故障分类时也具备十分优

表 1 对照模型架构

模型	架构
1D-CNN	Conv(64)-Conv(64)-Conv(128)-FC(300)-FC(21)
GRU	GRU(128)-GRU(128)-FC(300)-FC(21)
CNN-GRU	Conv(64)-Conv(64)-Conv(128)-GRU(128)-GRU(128)-FC(300)-FC(21)
MBTCN	4×(Temporal residual block×3)-Concatenation(400)-FC(21)
MSDACRN	Wide Conv(128)-MSDAVC(64)-MSDAVC(128)-FC(100)-FC(21)

表 2 6 种模型的精度与召回率

方法	1D-CNN		GRU		CNN-GRU		MBTCN		MSDACRN		MDRSN-BIGRU	
	P	R	P	R	P	R	P	R	P	R	P	R
Normal	0.613	0.423	0.787	0.585	0.807	0.939	0.873	0.901	0.880	0.911	0.812	0.828
Fault 1	0.960	0.854	0.856	0.932	1.000	0.961	1.000	1.000	1.000	1.000	1.000	0.961
Fault 2	1.000	0.718	1.000	0.933	1.000	0.980	1.000	1.000	1.000	0.961	1.000	0.953
Fault 3	0.327	0.884	0.789	0.605	0.826	0.876	0.864	0.860	0.842	0.821	0.538	0.845
Fault 4	0.741	0.444	0.740	0.638	0.917	0.884	0.899	0.845	0.795	0.816	0.877	0.785
Fault 5	0.753	0.845	0.855	0.879	0.938	1.000	0.943	1.000	0.960	0.971	0.972	0.967
Fault 6	1.000	0.708	1.000	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Fault 7	0.831	0.847	0.796	0.941	0.953	0.961	0.934	0.946	0.904	0.976	0.909	1.000
Fault 8	0.947	0.979	0.988	1.000	1.000	1.000	1.000	0.962	1.000	0.957	1.000	1.000
Fault 9	0.510	0.456	0.465	0.717	0.812	0.876	0.733	0.913	0.631	0.839	0.782	0.822
Fault 10	0.553	0.834	0.731	0.756	0.875	0.959	0.861	0.900	0.889	0.963	0.803	0.970
Fault 11	0.929	0.924	0.976	0.947	0.995	0.964	0.993	0.941	0.995	0.961	0.973	0.949
Fault 12	0.691	0.991	0.999	0.979	1.000	1.000	1.000	0.966	0.980	0.978	1.000	0.970
Fault 13	0.986	0.651	1.000	0.767	1.000	0.953	1.000	0.967	1.000	0.942	1.000	0.930
Fault 14	0.966	0.996	0.987	1.000	0.962	1.000	1.000	0.978	0.959	0.995	0.982	1.000
Fault 15	0.502	0.463	0.517	0.652	0.929	0.793	0.800	0.813	0.840	0.727	0.863	0.730
Fault 16	0.845	0.623	0.675	0.812	0.997	0.903	0.879	0.826	0.994	0.896	1.000	0.854
Fault 17	0.877	0.917	0.971	0.971	0.987	0.988	0.973	0.997	0.987	0.986	0.958	0.930
Fault 18	1.000	0.988	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Fault 19	0.998	0.536	1.000	0.642	1.000	0.864	0.957	0.798	0.994	0.831	1.000	0.723
Fault 20	1.000	0.826	0.886	0.912	1.000	1.000	0.955	0.982	0.996	0.986	1.000	0.957
avg	0.809	0.754	0.857	0.839	0.926	0.912	0.936	0.933	0.935	0.929	0.951	0.948

秀的潜力,与前面两种网络相比,1D-CNN 与 GRU 的串联网络不仅能够提取非线性高维特征,还能够提取时序特征,其加权平均精度达到了 92.59%,这表明能够同时提取空间特征和时序特征的模型具有更好的故障诊断性能。MBTCN 与 MSDACRN 在上述基础上分别加入了分块机制与多尺度机制,提高了特征提取能力,因此取得了更好的效果。而本文提出的 MDRSN-GRU 网络不仅能够同时提取数据的多尺度空间特征和时序特征,还通过将注意力机制与软阈值相结合的方法自适应的消除噪声,从最终结果来看,无论是各分类故障精度和召回率还是总体的精度和召回率,本文所提出的模型都优于其他模型。

图 8 与 9 是 MDRSN-BIGRU 网络与其对比网络在测试集上的准确率和损失率变化曲线,由图可见,在大约 10

次迭代后,MDRSN-BIGRU 的精确度曲线开始大幅领先其他网络,并最终在 30 次迭代后稳定在 0.95 附近。

为了可视化并且进一步理解网络学习到的特征,本文采用了 t 分布-随机近邻嵌入算法对 6 种模型最后一层的输出进行了特征可视化,结果如图 10 所示,图 10(a)为 1D-CNN 最后一层的输出特征,可以看出 1D-CNN 等有许多数据点都是重合的,这表示 1D-CNN 的故障分类能力不足,未能成功分出这些故障,从图 10(b)、(c)可以看出,GRU 与 CNN-GRU 相比于 1D-CNN,故障点减少了很多,但还有部分故障点重叠,从图 10(d)、(e)可以看出,加入分块机制或多尺度机制后的方法相较于上述网络具有更好的分类效果,而图 10(d) MDRSN-BIGRU 模型输出的特征则能够很清楚区分大多数故障,仅有很少一部分重叠在一

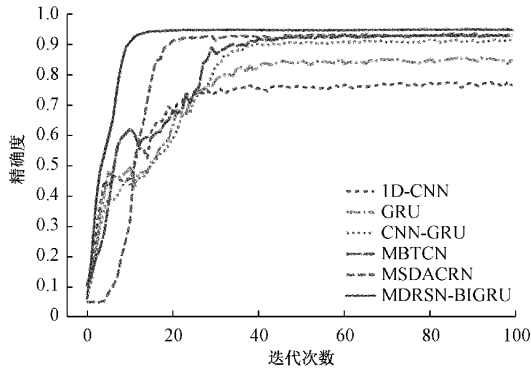


图 8 精确度曲线

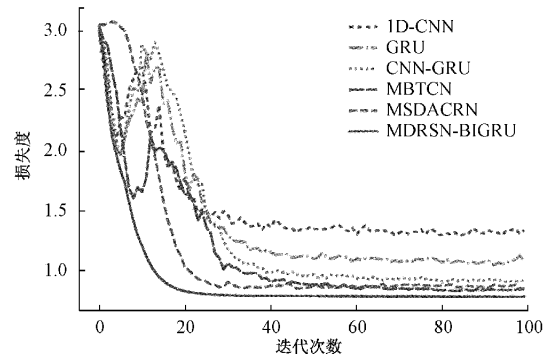


图 9 损失函数曲线

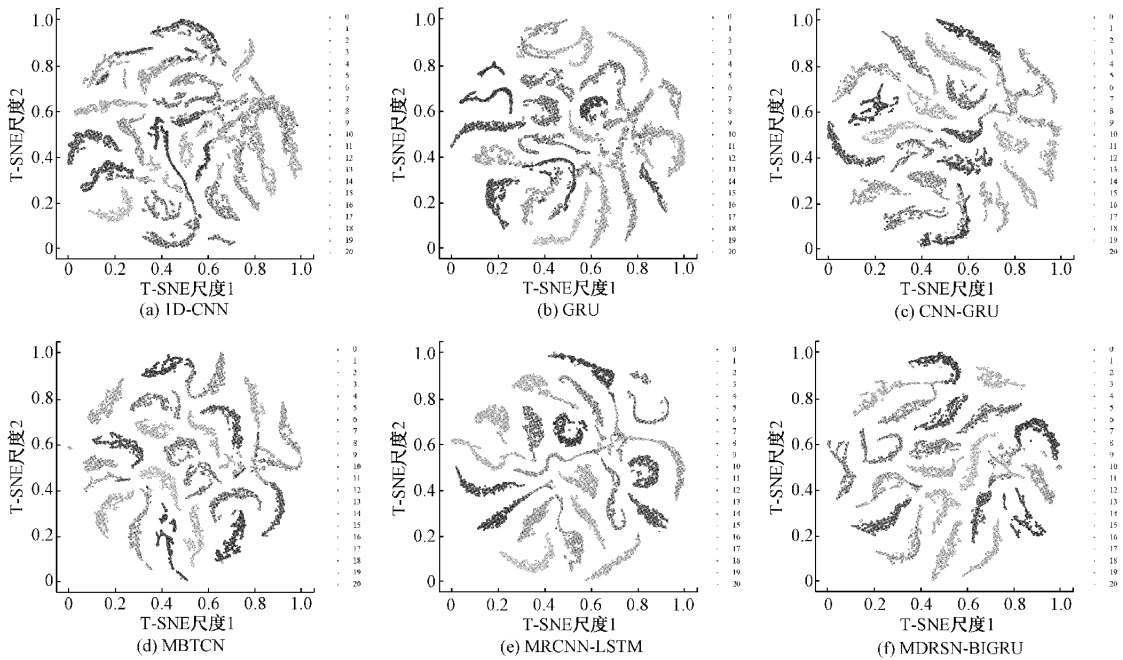
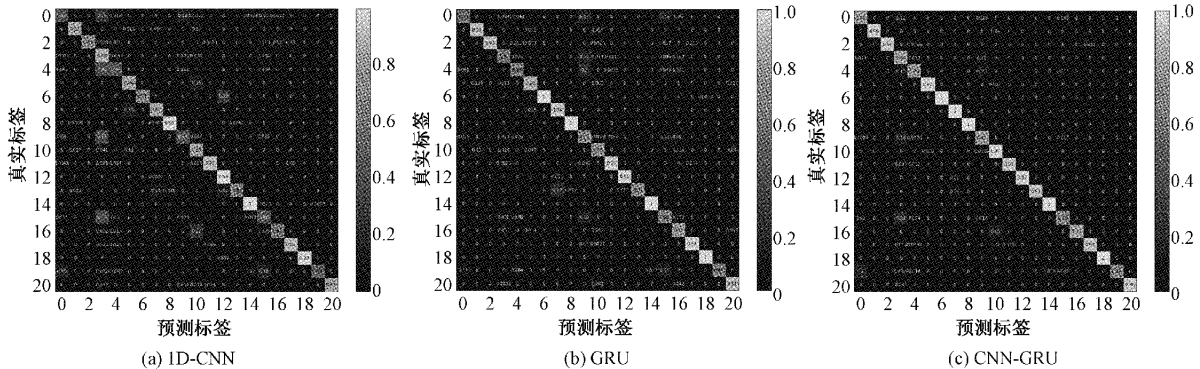


图 10 6 种模型的 t-SNE 特征可视化

起,这表明 MDRSN-BIGRU 具有更好的故障分类能力,证明了模型的整体有效性。

为了分析所提出的方法在混合故障分类中的性能,本文展示了上述 6 种方法的混淆矩阵,如图 11 所示,图中横坐标为预测标签,纵坐标为真实标签,可以看出,1D-CNN 仅有 3 种故障分类正确率较高,GRU 有 5 种故障分类率正

确率较高,CNN-GRU 则大部分分类正确率都很高,仅有 7 种较低,MBTCN 与 MSDACRN 则只有故障 3、4、15、19 这些微小故障分类正确率未达 90%,而本文提出的 MDRSN-BIGRU 相比于其他模型对故障的正确分类明显增多,仅对于较难区分故障 15 和故障 19 有略低的正确率,但相比与其他模型这几类故障的正确率也有较大的提升。



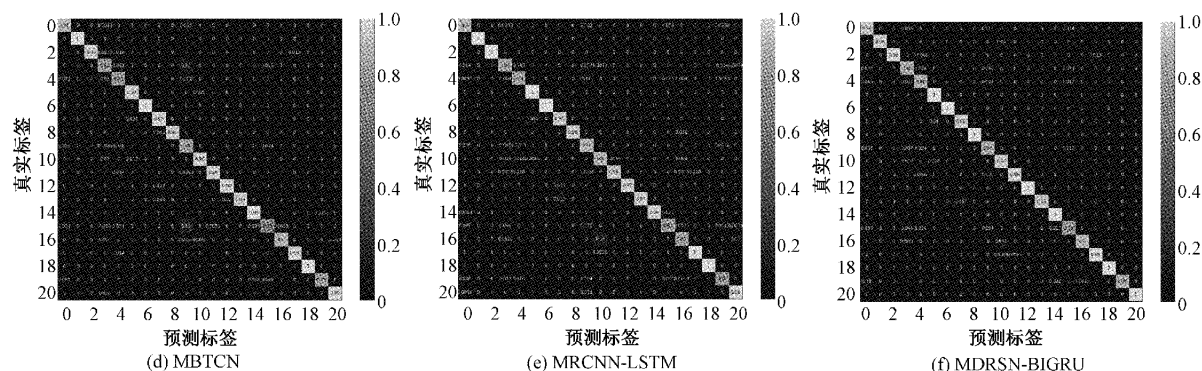


图 11 6 种模型混淆矩阵

4 结 论

本文提出了一种基于 MDRSN-BIGRU 融合模型的故障诊断方法,该方法引入残差收缩结构与多尺度卷积核,将注意力机制与软阈值相结合,在提取高维多尺度空间特征的同时自适应的将注意力集中到更加有用的特征信息上,以达到增强有效信息,减弱噪声信息的目的,之后将提取到的特征输入 BIGRU 网络,进一步提取时序特征,且由于 BIGRU 具有的双向通道特性,不仅能够完成对过去信息的学习,而且还能完成对未来时刻信息的学习,因此能够学习到更有效的特征。通过 MDRSN 与 BIGRU 的融合,使得网络不仅能够学习到空间特征和时序特征,还能在学习时降低噪声信号的干扰,最终获得了 95.08% 的分类精度和 94.76% 的召回率,证明了该模型的有效性。但是,该方法与大部分有监督算法存在相同的弊端,即需要大量的标记样本,且化工生产过程正常样本数据数量远大于故障样本数据数量,因此,如何从有限的中提取更加优秀的特征,以及如何处理不平衡数据是接下来研究的重点。

参考文献

[1] 马艳雪. 机械故障诊断基础研究“何去何从”[J]. 环球市场信息导报, 2017(37): 90-91.

[2] 黄志坚, 高立新, 廖一凡. 机械设备振动故障监测与诊断[M]. 北京: 化学工业出版社, 2010.

[3] 旷天亮. 基于深度神经网络的复杂化工过程故障检测研究[D]. 广州: 华南理工大学, 2018.

[4] 曹鹏晋. 基于 PCA-CNN 的光伏电站设备故障诊断研究[J]. 科学技术创新, 2021(30): 27-29.

[5] 张博峰. 基于 PCA 与 BP 神经网络的滚动轴承故障诊断[J]. 内燃机与配件, 2021(11): 131-132.

[6] 刘明光. 基于独立成分分析的故障诊断方法研究[D]. 武汉: 武汉理工大学, 2019.

[7] 田立勇, 张一轶. 基于局部均值分解与快速独立成分分析的潜水泵故障诊断[J]. 计量学报, 2020, 41(5): 585-591.

[8] 李坚铎. 基于改进的偏最小二乘法和贡献图的故障诊断方法[D]. 锦州: 渤海大学, 2021.

[9] YIN S, ZHU X, KAYNAK O. Improved pls focused on key-performance-indicator-related fault diagnosis [J]. IEEE Transactions on Industrial Electronics, 2015, 62(3): 1651-1658.

[10] 彭成, 李凤娟, 蒋金元. 基于深度学习的故障诊断与预测方法综述[J]. 现代电子技术, 2022, 45(3): 111-120.

[11] KIRANYAZ S, INCE T, GABBOUJ M. Real-time patient-specific eeg classification by 1-d convolutional neural networks[J]. IEEE Transactions on Biomedical Engineering, 2016, 63(3): 664-675.

[12] ÖZAL Y, ULAS B B, RAJENDRA A. A deep convolutional neural network model for automated identification of abnormal EEG signals[J]. Neural Computing and Applications, 2018, 32(20): 15857-15868.

[13] 王琦, 邓林峰, 赵荣珍. 基于改进一维卷积神经网络的滚动轴承故障识别[J]. 振动与冲击, 2022, 41(3): 216-223.

[14] 陈彦龙, 张培林, 徐超, 等. 基于 DCT 和 EMD 的滚动轴承故障诊断[J]. 电子测量技术, 2012, 35(2): 121-125.

[15] 杨永灿, 刘韬, 柳小勤, 等. 基于注意力机制的一维卷积神经网络行星齿轮箱故障诊断[J]. 机械与电子, 2021, 39(10): 3-8.

[16] 陈郑溟, 冯翱, 何嘉. 基于一维卷积混合神经网络的文本情感分类[J]. 计算机应用, 2019, 39(7): 1936-1941.

[17] MALEK S, MELGANI F, BAZI Y. One-dimensional convolutional neural networks for spectroscopic signal regression [J]. Journal of Chemometrics, 2018, 32(5): 2977-2994.

[18] 陈保家, 陈学力, 沈保明, 等. CNN-LSTM 深度神经网络在滚动轴承故障诊断中的应用[J]. 西安交通大学学报, 2021, 55(6): 28-36.

- [19] 唐颖川, 黄姣茹, 钱富才. 基于深度学习与注意力机制的化工故障分类[J]. 电子测量技术, 2022, 45(4): 168-174.
- [20] HE Y, SHI H, TAN S, et al. Multiblock temporal convolution network-based temporal-correlated feature learning for fault diagnosis of multivariate processes[J]. Journal of the Taiwan Institute of Chemical Engineers, 2021, 122(2): 78-84.
- [21] YU J, ZHANG C, WANG S. Multichannel one-dimensional convolutional neural network-based feature learning for fault diagnosis of industrial processes[J]. Neural Computing and Applications, 2020, 33(8): 3085-3104.
- [22] JIE H, LI S, GANG S, et al. Squeeze-and-excitation networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 42(8): 2011-2023.
- [23] DOWNS J J, VOGEL F E. A plant-wide industrial process control problem[J]. Computers & Chemical Engineering, 1993, 17(3): 245-255.

作者简介

杨晓岗(通信作者), 硕士, 主要研究方向为故障诊断等。

E-mail: xgyangwork@163.com

夏涛, 副教授, 主要研究方向为故障诊断、半实物仿真等。

E-mail: 969316743@qq.com