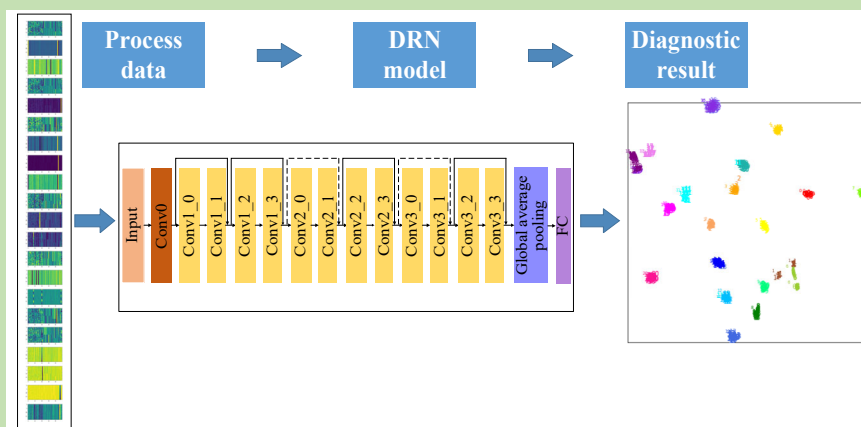


Fault diagnosis for chemical processes based on deep residual network

Lusheng ZHONG*, Xiangming XIA

College of Electrical and Automation Engineering, East China Jiaotong University, Nanchang, Jiangxi 330013, China

Abstract: A fault diagnosis method for chemical processes based on deep residual network (DRN) was proposed, which could automatically extract fault features from a large number of chemical processes operation data. The model adopted the shortcut connections to alleviate the training difficulty in the traditional deep neural network, and adopted the batch normalization (BN) method, which



could effectively alleviate the problem of vanishing/exploding gradients. The Tennessee Eastman (TE) process was used as the experimental object to evaluate the diagnostic performance of the proposed method. The proposed method and the previous TE process fault diagnosis method based on traditional deep learning model were compared. Furthermore, the effects of the number of layers, BN technology and residual structure on fault diagnosis rate were studied. Finally, the output of some layers was visualized by the t-distributed stochastic neighbor embedding (t-SNE) method. The results showed that the model achieved an average fault diagnosis rate of 94% and an average false positive rate of 0.30% for 21 working conditions, showing more excellent diagnostic performance. The two-dimensional scatter plot of the output layer showed clear clustering, which indicated that the proposed DRN model can accurately diagnose the faults.

Key words: fault diagnosis; chemical processes; deep learning; deep residual network; Tennessee Eastman process

收稿: 2019-12-23, 修回: 2020-03-07, 网络发表: 2020-03-31, Received: 2019-12-23, Revised: 2020-03-07, Published online: 2020-03-31
基金项目: 国家自然科学基金资助项目(编号: 61863012; 61263010); 江西省科技计划项目(编号: 20181BBE50020; 20161BBE50082; 20161BAB202067)
作者简介: 衷路生(1979-), 男, 江西省信丰县人, 博士, 教授, 研究方向为故障诊断与检测, E-mail: lszhongzju@163.com.

引用格式: 衷路生, 夏相明. 基于深度残差网络的化工过程故障诊断. 过程工程学报, 2020, 20(12): 1483-1490.

Zhong L S, Xia X M. Fault diagnosis for chemical processes based on deep residual network (in Chinese). Chin. J. Process Eng., 2020, 20(12): 1483-1490, DOI: 10.12034/j.issn.1009-606X.219374.

基于深度残差网络的化工过程故障诊断

袁路生*, 夏相明

华东交通大学电气与自动化工程学院, 江西 南昌 330013

摘要: 本工作提出了一种基于深度残差网络(DRN)的化工过程故障诊断方法, 可从大量化工过程运行数据中自动提取故障特征。模型采用快捷连接缓解传统深度神经网络训练困难的问题, 且使用批归一化(BN)方法, 可有效缓解梯度消失/爆炸的问题。以田纳西-伊斯曼(TE)过程为实验对象对所提方法进行诊断性能评价实验, 并与以往的基于传统深度学习模型的 TE 过程故障诊断方法进行对比, 进一步探究了模型层数、BN 技术和残差结构对故障诊断率的影响, 最后, 通过 t 分布随机邻域嵌入(t-SNE)方法对网络部分层的输出进行可视化。结果表明, 模型对 21 种工况取得了 94% 的平均故障诊断率和 0.30% 的平均假阳率, 表现出更加卓越的诊断性能。输出层的二维散点图显示了清晰的聚类, 表明所提出的 DRN 模型能够对故障进行准确诊断。

关键词: 故障诊断; 化工过程; 深度学习; 深度残差网络; 田纳西-伊斯曼过程

中图分类号: TQ460.6+4

文献标识码: A

文章编号: 1009-606X(2020)12-1483-08

1 前言

近年来, 化工过程安全事故时有发生, 化工过程通常具有毒性和腐蚀性强、易燃易爆的特点, 一旦发生故障, 危险性远高于其他行业。因此, 化工过程故障诊断越来越受到关注。

故障诊断技术通过对系统运行状况进行监测, 判断是否有故障发生, 同时确定故障发生位置、种类等^[1]。传统基于模型的故障诊断方法需对系统建立精确的数学模型, 而现代化工过程通常规模大、复杂度高, 建立精确的数学模型难度大。基于定性经验知识的诊断方法适用于规模小、简单系统的故障诊断, 用于复杂化工过程时, 诊断准确率较低。基于数据驱动方法对系统运行数据进行分析可有效进行故障诊断, 逐渐成为主流方法。主元分析、独立主元分析、偏最小二乘等多元统计方法已在化工过程故障诊断领域得到成功应用^[2-7]。支持向量机、人工神经网络等浅层学习的方法也逐渐被应用^[8,9], 现代化工过程通常为高度非线性过程, 而这两类数据驱动方法在面对具有非高斯分布、非线性特性的过程数据时, 难以高效提取数据中的重要特征。此外, 化工过程在运行过程中会产生海量数据, 这两类方法也存在难以处理海量数据的问题。

近年来, 深度学习的兴起吸引了大批学者的目光。深度学习是一种用含复杂结构或由多重非线性变换构成的多处理层对数据进行高层抽象的算法。深度学习擅长发现高维数据中的复杂结构, 通过足够多的转换组合, 也可学习非常复杂的函数^[10]。国际顶级期刊

《Nature》自 2015 年以来已连续多年发表关于深度学习的文章, 包括综述文章^[10]和将深度学习应用于游戏、医学、地球系统科学、聚变能科学等领域的研究成果^[11-14], 进一步掀起了对深度学习的研究热潮。目前, 主流的深度学习模型包括深度置信网络、深度自动编码器、循环神经网络、卷积神经网络及其各自的派生网络。与多元统计方法和浅层学习方法相比, 深度学习对非线性数据有着更强的特征提取能力, 且擅长处理高维海量数据。2013 年, Tamilselvan 等^[15]最早将深度学习应用于故障诊断领域。随后, 深度学习各主流模型在化工过程故障诊断中也得到初步应用。Zhang 等^[16]提出了一种基于扩展深度置信网络的故障诊断模型, 用于化工过程故障诊断。Jiang 等^[17]提出了一种基于动态稀疏堆叠自编码器的化工过程故障诊断方法, 可对动态故障数据进行分类。Xavier 等^[18]提出了一种基于长短时记忆网络的化工过程故障检测与诊断的新方法。Wu 等^[19]提出了一种基于深度卷积神经网络(Deep Convolutional Neural Network, DCNN)模型的化工过程故障诊断方法。上述基于深度学习的化工过程故障诊断方法均取得了较好的效果。然而, 这些深度学习模型存在训练困难的问题, 使故障诊断的准确率难以进一步提高。

深度残差网络(Deep Residual Network, DRN)^[20,21]是一种先进的深度学习模型, 在传统卷积神经网络的基础上生成, 不同之处在于 DRN 在网络结构中引入恒等映射快捷连接。恒等映射快捷连接可跨过多层网络直接连接到后面的隐含层, 有助于参数优化, 缓解训练困难问题。DRN 已被应用于其他对象的故障诊断, Zhang 等^[22]

提出了一种基于 DRN 的旋转机械故障诊断方法; 汪久根等^[23]提出了一种基于 DRN 的 RV(Rotate Vector)减速器故障诊断方法。以上研究都取得了非常好的诊断效果, 表明将 DRN 用于故障诊断领域潜力巨大。

本研究提出了一种基于 DRN 的化工过程故障诊断方法, 可缓解深度神经网络训练困难问题, 自动高效提取故障特征, 实现高精度故障诊断。以田纳西-伊斯曼(Tennessee Eastman, TE)化工过程为实验对象进行性能评价实验, 并对实验结果进行了分析与对比。

2 基于 DRN 的故障诊断模型

本研究提出的 DRN 故障诊断模型结构如图 1 所示。该模型由卷积层、残差块、批归一化 (Batch

Normalization, BN)层、激活函数层、全局平均池化层和全连接层组成, 共 14 层加权层。首先, 输入数据经过一个单独的卷积层, 然后经 6 个残差块及一个全局平均池化层, 最后经一个全连接层输出结果。表 1 给出了模型的内部参数。

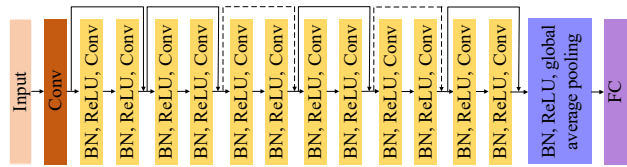


图 1 DRN 故障诊断模型结构
Fig.1 Architecture of the DRN fault diagnostic model

表 1 DRN 故障诊断模型参数
Table 1 Parameters of the DRN fault diagnostic model

Layer name	Layer type	Filter	Stride	Output size
Input	Input	—	—	20×50×1
Conv0	Conv	3×3, 16	1	20×50×16
Conv1_x	[BN - ReLU - Conv] ×2	[3×3, 16] ×2	[1×1] ×2	20×50×16
Conv2_x	[BN - ReLU - Conv] ×2	[3×3, 32] ×2	[1×2] ×2	10×25×32
Conv3_x	[BN - ReLU - Conv] ×2	[3×3, 64] ×2	[1×2] ×2	5×13×64
—	BN-ReLU	—	—	5×13×64
—	Average pooling	—	—	1×1×64
FC	FC	—	—	1×21

2.1 卷积层

卷积层的主要目的是从输入特征图中提取特征, 一般可表达为

$$Y_{i,j,o} = b_o + \sum_m \sum_n \sum_k X_{i-m,j-n,k} W_{m,n,k,o} \tag{1}$$

其中, X 为输入特征图, W 为卷积核, b 为偏置, Y 为输出特征图, i, j, k 分别为输入特征图的高度、宽度及通道索引, m, n, o 分别为卷积核的高度、宽度及输出特征图的通道索引。在本模型中, 所有卷积层均使用 3×3 大小的卷积核。

2.2 激活函数层

模型使用了新型的修正线性单元(Rectified Linear Units, ReLU)激活函数。函数表达式为

$$f(x) = \max\{0, x\} = \begin{cases} x, & x \geq 0 \\ 0, & x < 0 \end{cases} \tag{2}$$

2.3 BN 层

在训练神经网络时会出现内部协方差漂移问题, 导

致网络需要重新学习新的分布, 从而减缓训练速度。BN 是用于解决内部协方差漂移问题的一种归一化方法, 可表示为

$$\mu = \frac{1}{m} \sum_{i=1}^m x_i \tag{3}$$

$$\sigma^2 = \frac{1}{m} \sum_{i=1}^m (x_i - \mu)^2 \tag{4}$$

$$\hat{x}_i = \frac{x_i - \mu}{\sqrt{\sigma^2 + \varepsilon}} \tag{5}$$

$$y_i = \gamma \hat{x}_i + \beta \tag{6}$$

式中, x_i 和 y_i 分别为 BN 层的输入和输出, m 为批量大小, μ 为批量均值, σ 为批量标准差, ε 为一个接近 0 的常数, \hat{x}_i 为输入归一化后的值, γ 和 β 为待训练的参数。在网络中引入 BN 层, 可在取得相同精度的情况下减少训练步骤, 还可缓解梯度消失/爆炸问题, 提高模型泛化能力。

2.4 残差块

残差块是 DRN 的核心部分, 分为原始版^[20]和改进版^[21]两个版本, 结构如图 2 所示。图中右侧带箭头的线条称为快捷连接, 执行恒等映射。本模型采用的是改进版残差块。改进版与原始版的组成成分相同, 不同之处在于对 BN 层和 ReLU 层的位置进行了调整, 将它们置于卷积层之前, 这样可使输入跨过多层网络直接连接到输出, 更利于参数训练。假设多个残差块堆叠, 则从第 i 个残差块到第 j 个残差块的信息向前传递, 可表示为

$$x_j^{\text{out}} = x_i^{\text{in}} + \sum_{k=i}^j F(x_k^{\text{in}}, \{W_k\}) \quad (7)$$

其中, x_k^{in} , x_k^{out} 分别为第 k 个残差块的输入和输出, W_k 为第 k 个残差块中的权重, F 为残差映射。在误差反向传播过程中, 网络优化的梯度为

$$\frac{\partial L}{\partial x_i^{\text{in}}} = \frac{\partial L}{\partial x_j^{\text{out}}} \frac{\partial x_j^{\text{out}}}{\partial x_i^{\text{in}}} = \frac{\partial L}{\partial x_j^{\text{out}}} \left[1 + \frac{\partial}{\partial x_i^{\text{in}}} \sum_{k=i}^j F(x_k^{\text{in}}, \{W_k\}) \right] \quad (8)$$

式中, L 为损失函数, $\frac{\partial L}{\partial x_j^{\text{out}}}$ 项保证了低层网络都能接收到这个梯度, 缓解了训练困难问题, $1 + \frac{\partial}{\partial x_i^{\text{in}}} \sum_{k=i}^j F(x_k^{\text{in}}, \{W_k\})$ 项不小于 1, 避免了梯度消失。

在图 1 的 DRN 故障诊断模型的残差块中, 实线快捷连接表示输入通道数和输出通道数相同时的连接, 虚线表示不同时的连接。对于通道数不同的快捷连接, 采用的方法是先进行平均池化, 再使用零填充的方法使通道数匹配。

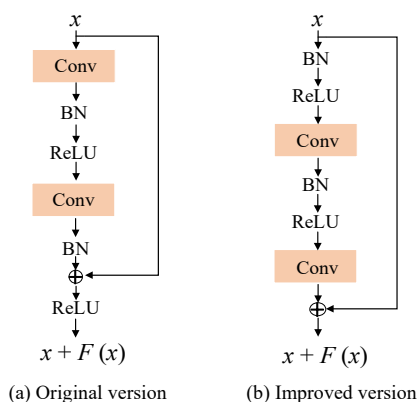


图 2 残差块示意图

Fig.2 Illustrations of residual building blocks

2.5 全局平均池化层

在模型中引入全局平均池化层, 可用全局信息并直接实现降维, 对整个网络在结构上起到了正则化作用, 提高了泛化能力。

2.6 全连接层

DRN 模型的最后一层是全连接层。在全连接层之后采用 Softmax 函数作为分类器评估输入数据属于各类工况的概率。Softmax 函数的输出可视为概率分布, 所以一般只在输出层中作为分类器使用。Softmax 函数的表达式为

$$y_j = \frac{e^{z_j}}{\sum_{k=1}^K e^{z_k}}, j=1, \dots, K \quad (9)$$

式中, z_j 和 z_k 分别是 Softmax 函数的第 j, k 个输入, y_j 可认为是输入数据属于第 j 类工况的概率, K 是工况类别数。

3 仿真实验

选取 TE 实验平台产生的 TE 数据作为数据源验证 DRN 故障诊断方法的性能。

3.1 TE 过程

TE 过程是根据现实中某化工过程建立的模型, 被广泛用于评价过程监测方法的性能^[24]。采用 TE 过程的修订版^[25]作为仿真平台。修订版包括 53 个过程变量, 28 种故障类型, 选取前 20 种故障和正常工况共 21 类工况进行研究。在 53 个过程变量中, 3 个变量(压缩机循环阀、汽提塔水流阀、搅拌速率)在运行过程中为恒定值, 可删去。正常工况数据和故障数据可从 MATLAB 的 Simulink 仿真模块中采集得到, 采样频率为 100 个数据/h。从模拟器中采集了每种工况数据 8000 个, 按 6:2:2 的比例划分训练集、验证集和测试集。然后, 对所有数据进行归一化处理, 公式为

$$Z(x) = \frac{x - \bar{x}}{S(x)} \quad (10)$$

式中 x 为原始数据, $Z(x)$ 为归一化后的数据, \bar{x} 为原始数据的均值, $S(x)$ 为标准差。为提取过程数据在时空域上的故障特征, 取 12 min 内采集的 20 个数据作为一个样本, 因此, 数据被转换成 20×50 的矩阵。

3.2 模型训练

本研究在开源机器学习软件库 TensorFlow 上实现该 DRN 模型。在超参数设置上, 将批尺寸设为 128, 权重衰减因子设为 0.0002, 迭代次数设为 250; 学习率采取动态调整的方式: 初始学习率设为 0.1, 并设定训练完 80 次后学习率衰减为 0.01, 训练完 160 次后学习率衰减为 0.001。

3.3 结果讨论

3.3.1 性能对比

用故障诊断率(Fault Diagnosis Rate, FDR)和假阳率(False Positive Rate, FPR)这两个指标评价模型的故障诊断性能。表 2 为定义的第 i 类工况的混淆矩阵。

表 2 第 i 类工况的混淆矩阵
Table 2 Confusion matrix of type i working conditions

	Number of samples in type i working conditions (actual)	Number of samples in the other conditions (actual)
Number of samples in type i working conditions (predicted)	TP	FP
Number of samples in the other conditions (predicted)	FN	TN

FDR 和 FPR 可定义为

$$FDR = \frac{TP}{TP + FN}$$
 (11)

$$FPR = \frac{FP}{FP + TN}$$
 (12)

DRN 模型在训练集和验证集上的损失曲线图和 FDR 曲线如图 3(a)和 3(b)所示。可以看出, 训练约 100 次后, 模型的损失和 FDR 逐渐达到收敛。模型在测试集上的诊断结果如图 4 所示, 图中显示了 21 种工况的诊断情况。图中“0”代表正常工况,“1”代表故障 01, 等等; 每种工况均用不同颜色表示, 每个数据显示为一个带数字标签的点, 显示的数字标签(纵坐标)为模型预测结果, 所带颜色(横坐标)对应实际所属标签。

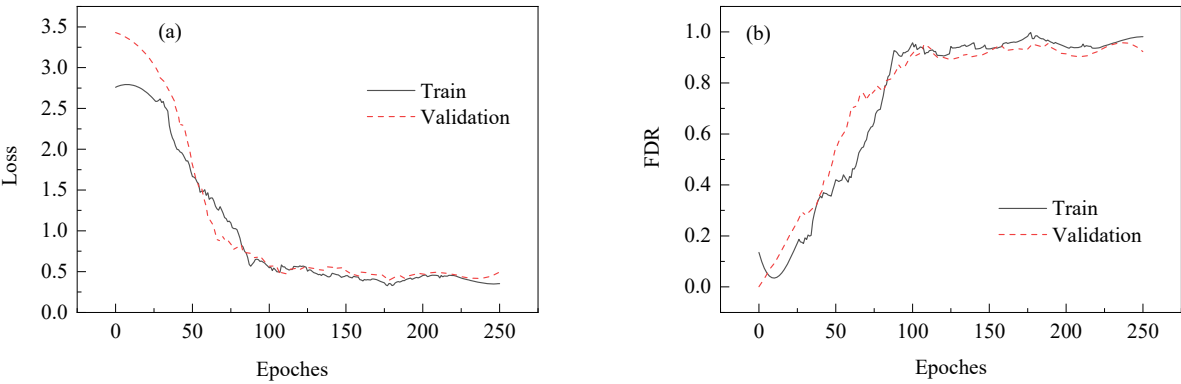


图 3 训练集和验证集上的损失曲线和 FDR 曲线
Fig.3 Loss and FDR curves on training set and validation set

表 3 在测试集上的 FDR 和 FPR 值

Table 3 The values of FDR and FPR on the testing set

Type of working condition	FDR/%	FPR/%
Normal	100	0.00
Fault 01	98	0.04
Fault 02	99	0.03
Fault 03	82	0.18
Fault 04	100	0.12
Fault 05	96	1.26
Fault 06	99	0.01
Fault 07	100	0.00
Fault 08	95	0.00
Fault 09	87	0.34
Fault 10	95	0.04
Fault 11	95	0.01
Fault 12	96	0.82
Fault 13	98	0.25
Fault 14	100	0.04
Fault 15	70	1.18
Fault 16	70	1.76
Fault 17	98	0.07
Fault 18	90	0.10
Fault 19	99	0.04
Fault 20	99	0.01
Average	94	0.30

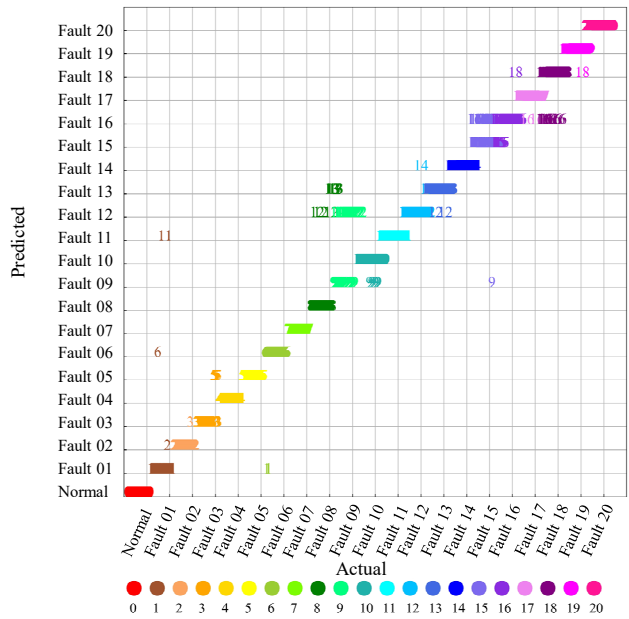


图 4 测试集上的诊断结果
Fig.4 Diagnostic results on testing set

为提高结果的可靠性,采用 5 折交叉验证,在测试集上的结果如表 3 所示,表中结果均为 5 个测试结果的平均值。模型的 21 种工况的平均 FDR 和 FPR 分别为 94%和 0.30%。基于同一数据集的诊断方法中,已有多 种基于深度学习的经典方法提出,并取得了很好的结果。为验证所提方法的性能,将上述结果与基于 DCNN 的方法^[19]在测试数据上的结果进行比较,结果如图 5 所

示。基于 DRN 的方法对故障 09、故障 15、故障 16 的 诊断效果显著优于基于 DCNN 的方法,对其他工况类 型的诊断效果与基于 DCNN 的方法相差不大或略有降 低。从平均诊断效果看,基于 DRN 的方法对 20 种故障 取得了 94%的平均 FDR 和 0.30%的平均 FPR,优于基 于 DCNN 方法的平均 FDR(88.2%)和平均 FPR(0.5%)。

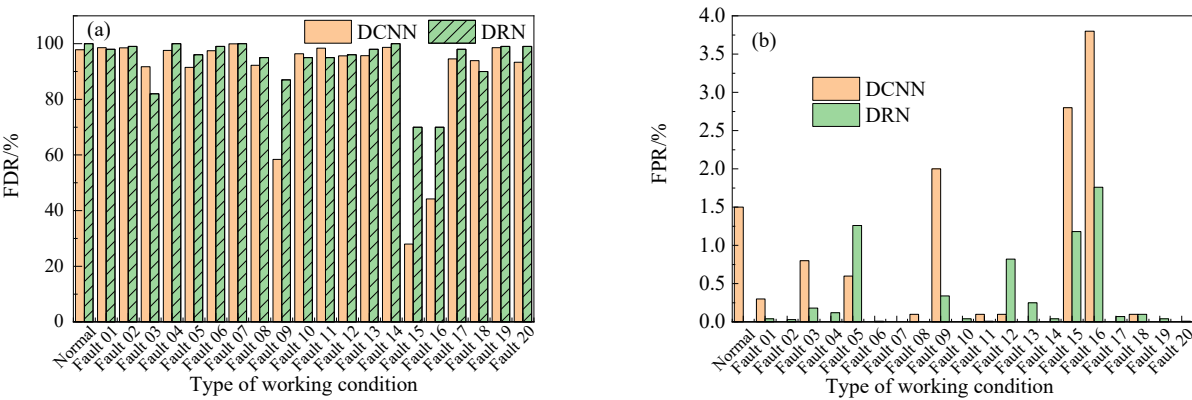


图 5 DRN 和 DCNN 诊断效果比较
Fig.5 Comparison of diagnostic results based on DRN and DCNN

表 4 给出了 DCNN 模型和 DRN 模型的浮点运算量 (Floating Point Operations, FLOPs)、模型每次迭代的训练时间和每个样本矩阵的测试时间。FLOPs 可由以下公式 计算得到:

$$\text{FLOPs}_{\text{conv}} = (2C_{\text{in}}k^2 - 1)HWC_{\text{out}} \tag{13}$$

$$\text{FLOPs}_{\text{fc}} = 2IO \tag{14}$$

$$\text{FLOPs}_{\text{avg_pooling}} = HWC_{\text{out}} \tag{15}$$

式中, $\text{FLOPs}_{\text{conv}}$, FLOPs_{fc} 和 $\text{FLOPs}_{\text{avg_pooling}}$ 分别为卷积层(不带偏置项)、全连接层(带偏置项)和全局平均池化层的 FLOPs, k 为卷积核边长, C_{in} 和 C_{out} 分别为卷积层输入和输出的通道数, H 和 W 分别为卷积层输出特征图的高度和宽度, I 和 O 分别为全连接层输入和输出的维度。DRN 模型的 FLOPs、训练时间和测试时间均小于 DCNN 模型, 实时性更好。总体而言, 基于 DRN 的方法表现出更优越的性能。

表 4 DCNN 和 DRN 模型的 FLOPs、训练时间和测试时间
Table 4 FLOPs, training times and testing times of DCNN and DRN models

Model	FLOPs	Training time/(epoch/s)	Testing time/ms
DCNN	7.8×10^7	30	1.5
DRN	5.1×10^7	4	0.4

3.3.2 网络层数、BN、残差结构对 FDR 的影响

对比 DCNN 和 DRN 模型的结构, DRN 模型的层数更多, 且使用了 BN 技术和残差结构。为进一步探究网络层数、BN、残差结构对 FDR 的影响, 设计了 A, B, C, D 四个基于卷积神经网络的模型, 并将这四个模型对 20 种故障的平均 FDR 与 DCNN 和 DRN 模型进行对比, 结果如表 5 所示。

对比模型 A, C 表明, 单纯地增加网络层数, FDR 值降低, 这是因为网络越深, 参数优化越困难; 对比模型 A, B 表明, 使用 BN 技术对提升 FDR 有微小的帮助, 且在实验中发现模型 B 在训练早期就能快速收敛; 对比模型 C, D 表明, 使用残差结构可使 FDR 有较大幅度的提升。因此, 在模型构建过程中, 将 BN 技术和残差结构结合使用会对性能提升有较大帮助。

表 5 不同模型对 20 种故障的平均 FDR
Table 5 The average FDR of 20 faults for different models

Model	The number of layer	Whether the BN was used	Whether residual building blocks were used	FDR/%
A	5	No	No	87.9
B	5	Yes	No	88.1
C	14	No	No	85.4
D	14	No	Yes	90.7
DCNN	5	No	No	88.2
DRN	14	Yes	Yes	93.3

3.3.3 可视化分析

采用 t 分布随机邻域嵌入(t-Distributed Stochastic Neighbor Embedding, t-SNE)方法, 对某些网络层输出进行可视化。选取模型的输入层、Conv0 层、Conv1_3 层、Conv2_3 层、Conv3_3 层和全连接层, 将这些网络层的高维输出特征用 t-SNE 方法嵌入到二维散点图中显示结

果, 如图 6 所示。

最初, 输入层的数据样本以混杂状态显示。经一个卷积层和几个残差块后, 样本逐渐聚集到相应的类中。最后的全连接层的二维散点图显示了清晰的样本聚类, 表明 DRN 模型对故障诊断任务的有效性。

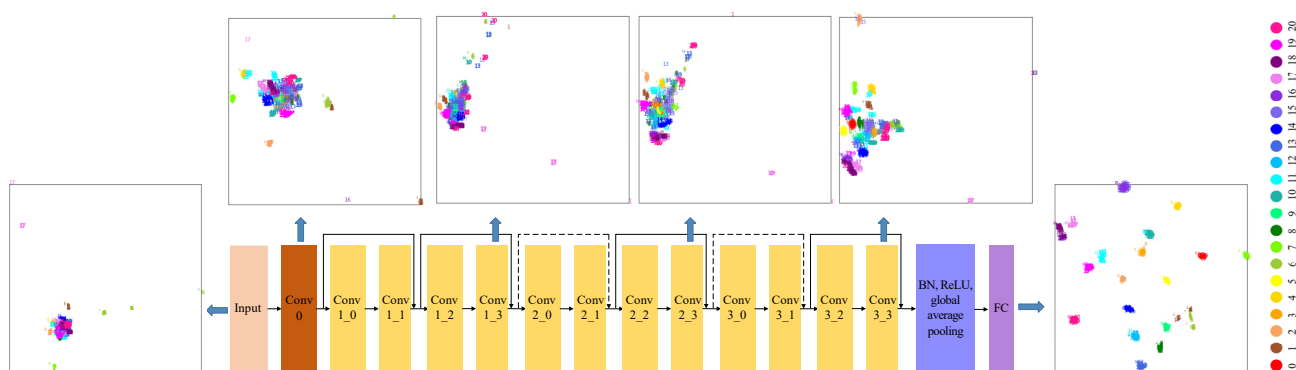


图 6 t-SNE 可视化
Fig.6 t-SNE visualizations

4 结 论

以深度残差网络(DRN)模型为主要方法对化工过程进行故障诊断, 并以田纳西-伊斯曼(TE)化工过程为案例进行了性能评价实验, 讨论了模型层数、批归一化(BN)技术、残差结构对模型的故障诊断率的影响, 得到以下结论:

(1) 基于 DRN 的故障诊断方法对 21 类工况的平均故障诊断率(FDR)达到了 94%, 平均假阳率(FPR)低至 0.30%, 优于基于传统深度学习模型的故障诊断方法。

(2) BN 技术和残差结构结合使用有助于提升模型性能。

(3) 通过 t 分布随机邻域嵌入(t-SNE)方法将 DRN 模型内部的高维输出特征降至二维, 得到的二维散点图表明 DRN 模型对故障诊断任务的有效性。

参考文献

- [1] 周东华, 刘洋, 何潇. 闭环系统故障诊断技术综述 [J]. 自动化学报, 2013, 39(11): 1933-1943.
Zhou D H, Liu Y, He X. Review on fault diagnosis techniques for closed-loop systems [J]. Acta Automatica Sinica, 2013, 39(11): 1933-1943.
- [2] Wise B M, Veltkamp D J, Davis B, et al. Principal component analysis for monitoring the West Valley liquid fed ceramic melter [J]. Waste Management, 1988, 88: 811-818.
- [3] Ge Z Q, Yang C J, Song Z H. Improved kernel PCA-based monitoring approach for nonlinear processes [J]. Chemical Engineering Science, 2009, 64(9): 2245-2255.
- [4] Kano M, Tanaka S, Hasebe S, et al. Monitoring independent components for fault detection [J]. AIChE Journal, 2003, 49(4): 969-976.
- [5] Ge Z Q, Xie L, Kruger U, et al. Local ICA for multivariate statistical fault diagnosis in systems with unknown signal and error distributions [J]. AIChE Journal, 2012, 58(8): 2357-2372.
- [6] Kresta J V, Macgregor J F, Marlin T E. Multivariate statistical monitoring of process operating performance [J]. The Canadian Journal of Chemical Engineering, 1991, 69(1): 35-47.
- [7] Wen Q J, Ge Z Q, Song Z H. Nonlinear dynamic process monitoring based on kernel partial least squares [C]//2012 American Control Conference (ACC). Montréal: IEEE, 2012: 6650-6654.
- [8] 陈剑雪, 侍洪波. 基于支持向量机的化工过程故障诊断 [J]. 华东理工大学学报, 2004, 30(3): 315-317.
Chen J X, Shi H B. SVM-based fault diagnosis for chemical process [J]. Journal of East China University of Science and Technology, 2004, 30(3): 315-317.
- [9] Rad M A A, Yazdanpanah M J. Designing supervised local neural network classifiers based on EM clustering for fault diagnosis of Tennessee Eastman process [J]. Chemometrics and Intelligent Laboratory Systems, 2015, 146: 149-157.
- [10] LeCun Y, Bengio Y, Hinton G. Deep learning [J]. Nature, 2015, 521(7553): 436-444.
- [11] Silver D, Huang A, Maddison C J, et al. Mastering the game of Go with deep neural networks and tree search [J]. Nature, 2016, 529(7587): 484-489.
- [12] Esteva A, Kuprel B, Novoa R A, et al. Dermatologist-level classification of skin cancer with deep neural networks [J]. Nature, 2017, 542(7639): 115-118.
- [13] Reichstein M, Camps-Valls G, Stevens B, et al. Deep learning and process understanding for data-driven earth system science [J]. Nature, 2019, 566(7743): 195-204.
- [14] Kates-Harbeck J, Svyatkovskiy A, Tang W. Predicting disruptive

- instabilities in controlled fusion plasmas through deep learning [J]. *Nature*, 2019, 568(7753): 526–531.
- [15] Tamilselvan P, Wang P. Failure diagnosis using deep belief learning based health state classification [J]. *Reliability Engineering & System Safety*, 2013, 115: 124–135.
- [16] Zhang Z P, Zhao J S. A deep belief network based fault diagnosis model for complex chemical processes [J]. *Computers & Chemical Engineering*, 2017, 107: 395–407.
- [17] Jiang L, Ge Z Q, Song Z H. Semi-supervised fault classification based on dynamic sparse stacked auto-encoders model [J]. *Chemometrics and Intelligent Laboratory Systems*, 2017, 168: 72–83.
- [18] Xavier G M, de Seixas J M. Fault detection and diagnosis in a chemical process using long short-term memory recurrent neural network [C]//2018 International Joint Conference on Neural Networks (IJCNN). Rio: IEEE, 2018: 1–8.
- [19] Wu H, Zhao J S. Deep convolutional neural network model based chemical process fault diagnosis [J]. *Computers & Chemical Engineering*, 2018, 115: 185–197.
- [20] He K M, Zhang X Y, Ren S Q, et al. Deep residual learning for image recognition [C]//The IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas: IEEE Computer Society, 2016: 770–778.
- [21] He K M, Zhang X Y, Ren S Q, et al. Identity mappings in deep residual networks [C]//European Conference on Computer Vision. Amsterdam: Springer, 2016: 630–645.
- [22] Zhang W, Li X, Ding Q. Deep residual learning-based fault diagnosis method for rotating machinery [J]. *ISA Transactions*, 2018, 95: 295–305.
- [23] 汪久根, 柯梁亮. 基于残差网络的 RV 减速器故障诊断 [J]. *机械工程学报*, 2019, 55(3): 73–80.
- Wang J G, Ke L L. Fault diagnosis for RV reducer based on residual network [J]. *Journal of Mechanical Engineering*, 2019, 55(3): 73–80.
- [24] Downs J J, Vogel E F. A plant-wide industrial process control problem [J]. *Computers and Chemical Engineering*, 1993, 17(3): 245–255.
- [25] Bathelt A, Ricker N L, Jelali M. Revision of the Tennessee Eastman process model [J]. *IFAC-PapersOnLine*, 2015, 48(8): 309–314.