

DOI:10.11918/201906120

卷积神经网络多变量过程特征学习与故障诊断

陈淑梅, 余建波

(同济大学 机械与能源工程学院, 上海 201804)

摘要: 为提取复杂多变量过程的有效特征,提高故障诊断性能,提出一种基于卷积神经网络(convolutional neural network, CNN)特征学习的多变量过程故障诊断模型. 将高维过程信号归一化处理转为图像信号,多层卷积滤波器与子采样滤波器交替构成的轻量级CNN网络通过多个卷积核与图像进行卷积,采用本地连接和权重共享,滤除过程噪声和干扰信息,从而获得过程数据的高层抽象化表达. 通过Softmax层有监督的微调方式学习故障特征完成故障诊断. 利用以田纳西过程为代表的多变量非线性过程验证了模型的有效性,与经典分类器和近几年流行的深度神经网络进行对比,结果表明:将高维过程信号转为图像信号输入CNN提高了多变量过程的故障诊断精度;通过t-SNE方法对模型提取的特征进行可视化分析,说明模型强大的特征提取能力;将模型提取的特征作为传统分类器的输入时,故障识别准确率显著提升,进一步说明有效的特征提取有利于提高故障诊断的准确度和可靠性;与无监督学习方式相比,模型通过标签能获取更有效、稳定和抽象化的数据特征.

关键词: 多变量过程;故障诊断;卷积神经网络;特征学习;田纳西过程

中图分类号: TP277

文献标志码: A

文章编号: 0367-6234(2020)07-0059-09

Feature learning and fault diagnosis in multivariate process with convolutional neural network

CHEN Shumei, YU Jianbo

(School of Mechanical and Energy Engineering, Tongji University, Shanghai 201804, China)

Abstract: A multivariate process fault diagnosis model is proposed based on convolutional neural network (CNN), aiming at extracting effective features from complex multivariate processes and improving fault diagnosis performance. First, the high-dimensional process signals are normalized and then converted into images. Second, a lightweight CNN network composing of multi-layer convolution filters and sub-sampling filters is convolved with images through multiple convolution kernels, using local connections and shared weights to remove noise and interference information to obtain the high-level abstract representations of process data. Finally, a Softmax layer is used in a supervised way to implement fault diagnosis. Tennessee Eastman Process is used to verify the effectiveness of proposed model and compare the performance between the proposed model with classical classifiers and deep neural networks. The results show that the fault diagnosis accuracy is improved by converting high-dimensional process signals into images. The t-SNE visualization analysis method is used to illustrate the powerful feature extraction ability of proposed model. The features extracted by the proposed model are sent to the traditional classifiers and the accuracy of fault identification is significantly improved, which further illustrates the benefit of effective feature extraction for improving the fault diagnosis accuracy and reliability. Compared to unsupervised learning, the proposed model with the guidance of label helps to extract more efficient, stable, and abstract feature representations.

Keywords: multivariate processes; fault diagnosis; convolutional neural network; feature learning; Tennessee Eastman Process

随着制造过程变量数量增加,变量间自相关性、过程的动态性和非线性以及混合的操作增加了过程监测与故障诊断的难度^[1]. 在工业过程中应用最为广泛的是多变量统计过程(multivariate

statistical process control, MSPC)方法. 主成分分析(principle component analysis, PCA)和偏最小二乘法(partial least squares, PLS)是两种最为常见的多变量过程监测方法^[2]. 在PCA和PLS的基础上,自适应PCA、动态PCA和多尺度PCA的监测方案得以扩展,在化工过程、生化过程和半导体过程等领域的应用得到了广泛的认可^[3-5]. 其他MSPC监控方法,如独立成分分析(independent component analysis, ICA)法、典型变量分析(canonical variate

收稿日期: 2019-06-12

基金项目: 国家自然科学基金(7177173)

作者简介: 陈淑梅(1994—),女,硕士研究生;

余建波(1978—),男,教授,博士生导师

通信作者: 余建波, jbyu@tongji.edu.cn

analysis, CVA) 和机器学习模型(人工神经网络,支持向量机, K-means, 高斯混合模型, 隐马尔科夫模型等)也在过程监控中取得了良好的应用效果^[6-8]. 然而, 由于假设过程在单个最佳区域中操作, 这些传统的监控方案仍具有单峰局限性. 过程故障诊断方法的研究在过去十几年中得到了快速的发展, 如人工神经网络^[9], 支持向量机^[10]、随机森林^[11]、Fisher 判别法^[12]等.

在过程监控与故障诊断中, 高维过程信号的特征表达对故障模式的匹配和识别至关重要^[13]. 近年来, 相继开展了基于特征学习的故障检测方法的研究^[14-15]. 基于深度学习(deep Learning)的特征提取方法能自动学习大数据中的特征, 显著地提高了故障检测和诊断的准确率^[16]. 深度神经网络(deep neural networks, DNN)在计算机视觉、自然语言处理和语音识别等领域取得了巨大的成功^[17-25]. 近年来主要开发了基于 DNN 的机械故障诊断方案^[21-23]. 但是, 基于 DNN 的过程故障诊断的研究较少^[24-25].

卷积神经网络(convolutional neural network, CNN)^[26]以非监督或半监督的学习方式获取图像数据的高层抽象化表达, 克服了传统的手工提取特征人力成本大、精度不高和不可泛化的缺点. 近年来, CNN 在计算机视觉领域取得了突破性进展^[27-29]. Lee 等^[30]用卷积神经网络从半导体的传感信号中自动提取故障的内部结构特征, 开发了基于 CNN 的晶圆表面缺陷识别模型.

虽然基于 CNN 算法的机械或半导体的故障诊断方法近年来被广泛研究^[26, 30-33], 将机械或半导体产生的信号转为图像信号输入 CNN 进行故障诊断,

但是, 基于 CNN 特征学习的过程故障诊断还未见报道.

本文尝试将多变量过程信号转为图像信号, 开发基于 CNN 的过程故障诊断方案, 设计了一个轻量级的 CNN 模型, 适用于样本量不大的多变量过程的特征学习. 本文的主要贡献有以下几个方面: 1) 提出了基于 CNN 的多变量过程特征学习模型(CNN-based feature learning for multivariate process, 简称 MPCNet); 2) 将一维过程信号转为二维的图像信号, 以 MPCNet 提取的图像特征表征多变量过程状态; 3) 基于 MPCNet 提取的有效过程特征, 提出了多变量过程故障诊断模型, 显著地提高了系统的故障诊断性能.

1 MPCNet

MPCNet 是一种由多层卷积滤波器与子采样滤波器交替构成的轻量级 CNN 网络, 其特殊性表现为两个方面, 即本地连接和权重共享. MPCNet 的核心原理: 输入图像通过局部感受和权值共享的下采样方式, 映射到隐含层空间的有用信息具有某种尺度的不失真性. 在 MPCNet 中, 神经元之间的一些连接在整个层上被复制, 它们具有相同的权重和偏差.

如图 1 所示, MPCNet 共有 7 层, 包含两个卷积层, 两个池化层, 两个全连接层和一个输出层, 卷积层和池化层的堆叠构成具有深层结构的 CNN 模型. 输入图像的像素为 208, 通道为 3, 卷积层卷积核的大小为 3, 通道数为 16. 全连接层的神经元数为 128, 激活函数为 ReLu, 输出层的激活函数为 Softmax, 用于图像的分类.

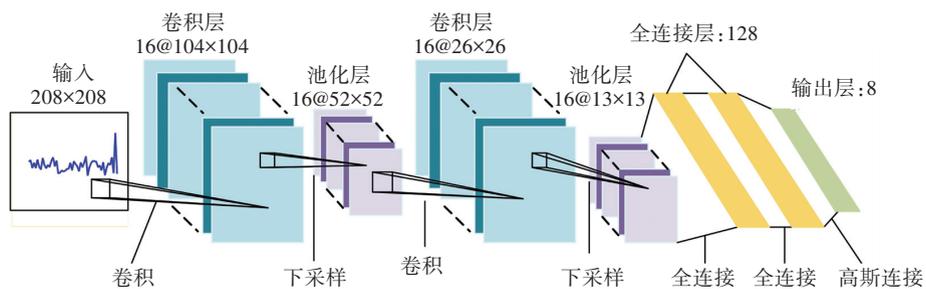


图 1 MPCNet 网络结构

Fig.1 Network structure of MPCNet

1) 卷积层. 卷积层即特征提取层, MPCNet 采用大小为 3×3 的卷积核从前一层中提取 208×208 图像的局部特征, 作为卷积层中神经元的输入, 卷积层确定了所有提取特征之间的位置关系. 传统神经网络层与层之间全连接的方式易占用计算空间, 且执行任务效率不高. 而通过特征映射, 提取输入图像的主要特征, 通过卷积核的特征提取操作, 图像由原

来的 208×208 减少至 104×104 , 实现了参数的降维, 在降低计算成本的同时也提高了准确度. 在进行卷积操作时, 图像每一维的步长均为 2. 填充图像的方法为相同填充, 即超出边界的区域使用 0 填充. 卷积计算过程如图 2 所示. 在训练过程中, 用一个卷积核 f_x 将图像中的局部区域通过特征提取和卷积核的权重 W_x 作用, 以及偏置 b_x 构成卷积层神经元的输

入 C_x , 因此卷积核也被称为滤波器. 滤波器的学习参数 (W_x 和 b_x) 可通过训练过程的迭代优化得到.

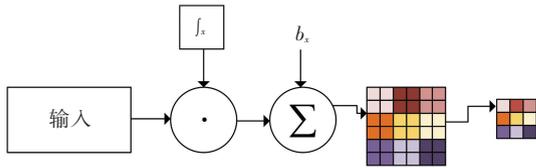


图 2 卷积和下采样过程

Fig.2 Convolution and down-sampling process

2) 池化层. 池化层即特征映射层, MPCNet 采用 2×2 的池化矩阵用以减少数据特征, 移动步长为 2, 使用全 0 填充. 受到人脑视觉皮层抑制效用的启发, 池化层在降低分辨率的同时使特征图中的重要信息得到最大化的保留, 减少主要特征发生的位移和失真. 下采样过程如图 2 所示, MPCNet 使用最大值下采样方式对图像上不同区域内的特征进行聚合统计, 得到低维的概要统计特征, 在提高性能的同时不容易产生过拟合. 其他下采样方式还包括权值可训练下采样、平均值下采样、高斯池化下采样.

3) 全连接层. MPCNet 输出层之前设置两个 128 全连接层, 将原始数据经过特征映射层提取到的隐含特征全连接, 对数据进行重新分类. 全连接层以矩阵乘法的形式对隐层特征进行空间提取整合, 在降低维度的同时, 最大化地保留有用信息. 理论上, 多个全连接层可通过激活函数实现多种非线性变化, 从复杂多变量过程的输入图像中学习数据特征的表达方法.

MPCNet 在前向传播阶段通过卷积核对输入图像进行卷积操作而获得局部特征映射, 通过反向传播方式更新网络参数. 输入层通过特征映射与内核卷积, 在激活函数的非线性转化下形成输出特征映射. 卷积核可与多个输入映射组合形成对应的输出映射, 计算过程如下:

$$x_j^l = f\left(\sum_{i \in M_j} x_i^{l-1} * k_{ij}^l + b_j^l\right). \quad (1)$$

式中: M_j 为输入特征映射, l 为 MPCNet 网络的第 l 层, k 为 3×3 矩阵的卷积核, f 为 ReLU 函数, b_j^l 为每个特征映射在输出时的偏差. 卷积层通过不同的卷积核 k 与输入映射卷积形成特定的输出映射.

子采样层提取输入映射的下采样特征, N 个输入映射对应 N 个输出特征, 输出特征的维度可能

较小. 下采样计算过程如下:

$$S_j^l = \max(a_i^l), \quad i \in R_j^l.$$

式中: R_j^l 为 MPCNet 网络第 l 层的第 j 个池化域, i 为池化域内特征的索引值, a_i^l 为池化域内索引值为 i 的特征值. 对输入特征映射中不同的 2×2 模块中取最大值, 得到的输出特征映射在两个空间维度上均缩小至 $1/2$.

在 MPCNet 的末端为具有 8 个神经元的全连接层用于执行 8 个过程故障类分类任务, 输出层前的级联特征映射作为输出层的输入, 记特征向量为 f_v , 输出为

$$O = f(b_o + w_o f_v).$$

式中: b_o 和 w_o 分别为偏置向量和特征向量. 在反向传播过程中, MPCNet 使用梯度下降法优化学习参数 k_{ij}^l , b_j^l 和 b_o . 为避免模型过拟合, 本文将数据集分为训练集、验证集和测试集. 在训练过程中, 每隔 200 步使用验证集对训练的模型进行预测. 初始阶段模型在验证集上的训练误差随着训练集的误差下降而下降, 当超过一定的训练步数后, 验证集上的误差不再下降, 此时提前终止训练模型, 防止网络结构过拟合. 由于在整个输入映射中都有卷积核的参与, 所以 MPCNet 模型中的连接数量远远多于权重数量. 与一般的前馈反向传播网络相比, 卷积层的共享权重减少了网络参数的数量, 增大了每个权重对输出过程的影响力, 同时也降低了误差梯度为零的速度, 使得 MPCNet 深度模型对复杂多变量过程进行特征学习时更为容易.

2 基于 MPCNet 的过程故障诊断方法

本文通过图像处理的方法将原始一维过程向量转为图像信号, 从而可为 MPCNet 提供图像数据输入, 构建具有多层结构的 MPCNet 模型 (参数见表 1) 对输入图像进行特征提取, 提取到的特征进一步输入至 Softmax 分类器中, 以有监督的方式学习与过程状态匹配的显著性特征. 过程的异常状态可通过各类控制图 (比如基于局部和全局主成分分析法 (Local and global principle component analysis, LGPCA) 构造的 T^2 和 Q^2 统计值加以判别^[34]). 对于识别为异常状态的样本, 则进一步进行故障诊断, 判断样本所属的故障类别.

表 1 MPCNet 的模型参数

Tab.1 Parameters of MPCNet model

输入/连接层	输入图像	卷积层	卷积层	全连接层	全连接层	输出层
参数/激活参数	208×208×3	3-16	3-16	128/ReLU	128/ReLU	Softmax

基于 MPCNet 模型的复杂多变量过程的特征学习过程如图 3 所示。

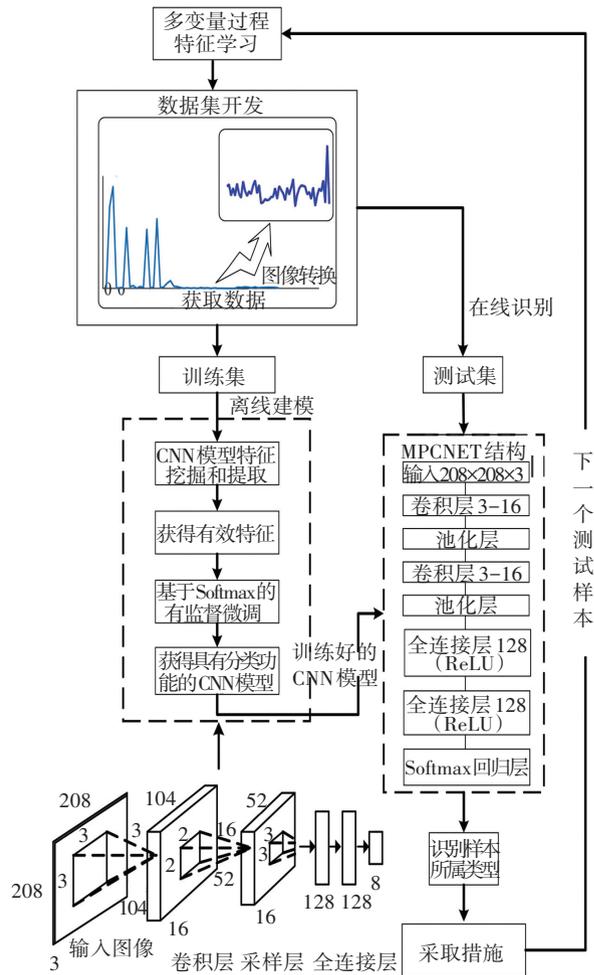


图 3 基于 MPCNet 的复杂多变量过程特征学习与故障诊断
Fig.3 Feature learning and fault diagnosis in complex multivariate process based on MPCNet

与一般故障诊断方法不同,本文将多变量高维过程信号转换为图像作为 MPCNet 的输入. CNN 在图像识别中应用广泛,但收敛速度慢,对数据进行预处理可防止梯度弥散. 本文对数据进行归一化预处理,以加快模型收敛. 如图 3 所示,在一维过程向量转化为图像的过程中,为防止模型梯度爆炸,首先对训练数据进行归一化处理,用 matlab 绘图工具为每个过程信号样本生成一张图片,构成样本图片的训练集和测试集. 同时为提高模型的运算速度,在绘图时固定图片的大小和背景,并将图片的像素设为 208×208 . MPCNet 采用大小为 3×3 的卷积核从前一层中提取图像的局部特征,通过卷积核的特征提取操作,图像由原来的 208×208 减少至 104×104 . 下采样过程用 2×2 的池化矩阵对图像进一步进行参数降维,使图像缩减至原来的 $1/4$. 卷积层和下采样层

构成特征映射平面, MPCNet 采用两个特征映射平面提取输入图像的主要特征,实现了参数的降维,在降低计算成本的同时也提高了准确度. 后一个特征平面的卷积核尺寸和池化矩阵的大小与前面相同. 在 MPCNet 输出层前设置两个大小为 128 的全连接层,将原始数据经过特征映射层提取到的隐含特征全连接,作为分类层的输入. MPCNet 通过 Softmax 层有监督的微调方式学习故障特征,与无监督学习方式相比,有监督的学习方式能通过提供有标签的训练样本数据提取更有效、稳定和抽象化的数据特征.

基于 MPCNet 的多变量过程故障识别的过程包含离线建模和在线识别两个过程. 离线建模过程包含以下 4 个步骤:

步骤 1 收集各类过程样本,构建训练集;

步骤 2 对训练集样本进行归一化处理,并将过程信号生成图像,作为 MPCNet 的输入;

步骤 3 MPCNet 从过程图像中提取故障模式特征;

步骤 4 MPCNet 将提取到的数据特征输入分类器,以有监督的学习方式调整和优化网络参数.

在线故障诊断过程包含以下 5 个步骤:

步骤 1 待测试样本进行归一化处理;

步骤 2 将过程信号转化为图像信号;

步骤 3 将图像输入到 MPCNet 产生过程特征;

步骤 4 输入特征到故障识别器,输出过程故障识别结果;

步骤 5 根据故障诊断结果制定过程维护决策.

3 实验结果分析

田纳西过程是复杂多变量过程故障诊断应用最为典型的一种化工生产仿真系统,具有 5 个操作单元,分别为反应器、冷凝器、循环压缩机、分离器和汽提塔. TEP 有 52 个变量,其中包括 12 个操纵变量和 40 个测量变量. 在实际的工业过程中,TEP 常见的故障模式有 21 种,本文选取 8 种故障模式进行研究,如表 2 所示. 在实验中,通常以 3 min 的采样间隔时间收集数据,构造训练集和测试集. 训练集包含 500 个正常样本和 10 080 个故障样本,其中每种故障类型的样本数为 480. 在测试集中,正常样本和每种故障类型的样本数均为 960 个,每种故障类型在第 161 个样本开始引入异常. 本文选取的 TEP 故障类型为 4、5、11、13、14、16、19 和 21 共 8 种故障类型,如表 2 所示. 每个样本的数据维度为 52. 每种故障的训练集和测试集的样本数量分别为 480 和 800.

表 2 TEP 故障描述

Tab.2 Description of faults in TEP

故障	描述	类型
故障 4	反应器冷却水入口温度变化	阶跃
故障 5	冷凝器冷却水入口温度变化	阶跃
故障 11	反应器冷却水入口温度变化	随机
故障 13	反应动力学参数变化	缓慢漂移
故障 14	反应器冷却水阀	阀粘滞
故障 16	未知	未知
故障 19	未知	未知
故障 21	管道 4 的值固定在稳态位置	固定位置

3.1 TEP 故障诊断

在选取故障类别时,首先用常见的故障诊断方法(BPN, SDAE 和 SVM)测试 21 种故障数据的识别率. 故障类别依次增加,当故障类别增加到 8 类时,上述方法难以区分本文选取的 8 种故障类型. 本文选取的 8 种故障类型的数据特征具有相似性,

在空间上表现为部分重叠,一般分类方法难以对这 8 种故障类型的数据进行区分. 本文利用 CNN 强大的特征学习能力提取这 8 种故障类型的内部结构特征,提高模型的故障识别率. 表 3 为 MPCNet 故障类型识别结果的混淆矩阵. 由于 8 种故障的数据特征在空间上交叠, MPCNet 在对每种故障类型进行识别时均有一定程度的“类别混淆”情况. 除了在对故障 4 和故障 5 进行识别时, MPCNet 将一小部分样本混淆为故障 16 和故障 19 外,在对其他类型的故障进行识别时, MPCNet 基本能正确识别出特征模式所属的故障类别. 图 4 为 MPCNet 对 TEP 故障类型的识别结果,横坐标上的 8 类样本对应纵坐标上的故障 4 到故障 21. 图中点的纵坐标代表对应样本的分类结果,点分布得越密集,每类样本属于对应类别的可能性越大. 如图 4 所示,每类样本的分类结果大都集中于对应类别的位置,有少数样本分布在其他类别. 因此,每种故障类型均有少数的“误分类”现象,但大多数的样本能被正确归类到所属的故障类别中.

表 3 基于混淆矩阵的 MPCNet 故障识别结果

Tab.3 Confusion matrix for fault diagnosis results based on MPCNet

故障	4	5	11	13	14	16	19	21
4	86.87	0	0.71	0.35	0	11.70	0.35	0
5	0	78.50	2.05	0.34	1.02	2.73	12.29	3.07
11	3.14	2.79	84.32	1.05	1.05	0.70	5.23	1.74
13	0.35	1.38	0.69	92.73	0.69	0.69	2.08	1.38
14	0	1.34	0	1.00	90.27	0	4.70	2.68
16	4.35	0.72	0.72	1.09	0.72	90.22	1.45	0.72
19	0.35	10.6	3.18	1.77	1.77	2.12	74.56	5.65
21	0.35	4.21	0.35	1.05	0	1.05	4.21	88.77

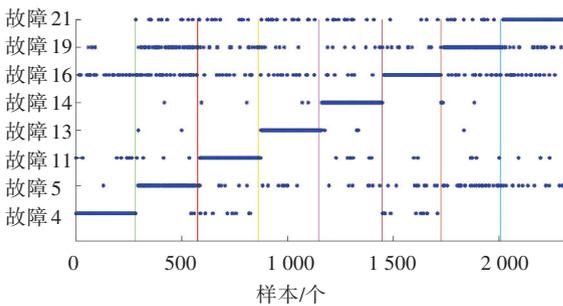


图 4 基于 MPCNet 特征学习的 TEP 故障分类结果

Fig.4 Fault classification for TEP based on MPCNet feature learning

为进一步验证基于特征提取的 MPCNet 模型对数据分类的有效性,本文比较了传统分类器

(分别为 SDAE、BPN、SVM_{LK}、SVM_{RBF}、LVQ、KNN 和 DT 分类器)对 8 种故障的类别识别性能. SDAE 和 BPN 均采用“52-70-45-8”的网络结构,激活函数均为 Sigmoid; SVM_{LK} 和 SVM_{RBF} 采用的核函数分别为线性核 (linear kernel, LK) 和径向基函数 (radial basic function, RBF); LVQ 的隐藏层大小为 100. 表 4 为 MPCNet 模型和一般机器学习方法对所选取的 8 种故障类型识别准确率的比较结果. 除故障 19 外,基于特征学习的 MPCNet 模型对故障类型的识别率均显著高于其他分类器,且 MPCNet 模型的故障类别平均识别率高于其他机器学习方法,显著地提高了识别器的故障诊断能力.

表 4 MPCNet 与典型识别器的故障识别率比较

Tab.4 Comparison of fault recognition rates between MPCNet and conventional classifiers

%

分类器	故障								Average
	4	5	11	13	14	16	19	21	
MPCNET	86.88	90.27	78.50	84.30	92.73	90.22	74.56	88.77	85.78
SDAE	79.92	50.34	66.89	73.75	83.87	81.27	85.20	54.80	72.01
BPN	79.92	84.56	69.84	76.68	83.87	80.94	68.95	65.48	76.28
SVM _{LK}	80.30	40.93	22.95	64.78	34.76	80.93	45.13	63.70	54.19
SVM _{RBF}	79.92	89.26	72.79	76.41	84.22	82.94	83.75	81.14	81.30
LVQ	74.62	42.95	27.54	68.77	79.57	75.25	56.32	60.50	60.69
KNN	69.32	40.27	25.90	73.09	83.87	54.18	37.91	58.72	55.41
DT	78.79	74.83	70.16	75.42	84.95	69.90	63.18	79.36	74.57

3.2 特征可视化分析

为进一步说明 MPCNet 模型对高维数据的特征提取性能,本文通过 t-SNE 方法^[35]将 MPCNet 模型第二层池化层提取到的图像特征进行二维平面可视化,如图 5 所示。

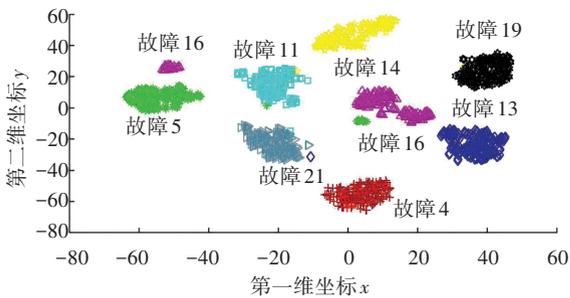


图 5 MPCNet 模型对故障数据的特征可视化

Fig.5 Feature visualization for faulty data using MPCNe

表 5 基于 MPCNet 提取特征的典型识别器识别率

Tab.5 Fault recognition rates of conventional classifiers based on features extracted by MPCNet

%

分类器	故障								Average
	4	5	11	13	14	16	19	21	
BPN	86.17	80.89	85.71	94.12	94.97	87.32	85.51	89.47	88.02
SVM _{LK}	89.36	77.47	86.41	95.16	96.31	85.14	74.56	85.61	86.25
KNN	87.23	70.99	70.03	84.78	84.90	77.17	55.83	69.82	75.09

3.3 性能对比分析

为了进一步证明 MPCNet 模型提取特征的有效性,本文进行了五折交叉实验,对本文提出的 MPCNet

MPCNet 模型强大的特征提取能力使得每种故障类型的数据特征在空间上具有良好的聚簇性和可分性,这显著提高分类器对故障数据的分类精度。

本文将 MPCNet 提取到图像特征作为传统分类器 KNN、BPN 和 SVM 的输入,得到的识别率如表 5 所示。比较表 4 与 5 可知,这些分类器的故障识别率均有大幅度提升,表明有效地过程特征提取有助于准确率更高、消耗更低的识别器的构建,在降低经验风险最小化的同时,也降低了模型的复杂度。MPCNet 以分层提取的方式获得数据特征的抽象化表达,这种结构化特征提高了传统分类器对数据类别属性的理解,表现为故障识别准确率的显著提升。

与深度学习方法(DBN 和 SDAE)^[36]、最近几年流行的 CNN 结构(Vgg16、Vgg19、Resnet50)^[37]和轻量级 CNN 模型(Mobilenet^[38])进行故障识别率的比较,结果见表 6。

表 6 MPCNet 与典型深度神经网络的五折交叉对比实验

Tab.6 Comparison of five-cross validation between MPCNet and conventional deep network models

%

分类器	故障								Average
	4	5	11	13	14	16	19	21	
MPCNet	85.20	91.19	85.48	85.56	92.06	96.41	67.02	82.99	85.74
	84.18	80.21	72.43	83.96	91.53	92.82	80.32	87.05	84.06
	85.71	86.98	78.38	83.33	93.12	91.79	72.34	83.94	84.45
	85.20	80.21	71.89	84.95	87.23	90.26	65.78	87.56	81.64
	83.16	81.25	84.32	87.63	86.70	86.60	62.03	83.94	81.96
									83.57
DBN	46.52	48.53	20.51	53.52	28.98	15.34	0	0	26.68
	50.75	53.89	0	33.51	63.13	0	0	34.92	29.53
	51.55	41.62	11.98	37.81	64.10	0	0	0	25.88
	51.81	49.01	12.43	52.76	0	0	0	69.61	29.45
	56.44	53.51	6.06	25.81	0	0	0	87.20	28.63
									28.03
SDAE	46.52	50.49	47.69	46.95	48.86	23.81	44.38	30.41	42.39
	48.75	59.59	31.55	46.91	40.22	42.70	61.06	43.39	46.77
	51.03	43.15	29.94	38.81	33.33	56.78	38.83	47.69	42.45
	51.30	50.00	25.95	51.26	40.31	66.01	23.63	43.09	43.94
	53.37	54.59	24.75	48.39	42.39	39.47	73.97	39.34	47.03
									44.52
Vgg16	79.08	73.58	45.70	62.03	84.58	81.03	42.55	64.43	66.62
	83.67	69.27	56.22	57.75	78.61	83.59	42.02	63.21	66.79
	87.76	73.44	57.30	60.22	82.09	82.05	42.05	60.62	68.25
	79.08	69.79	52.43	54.84	77.50	85.64	45.45	61.66	65.80
	84.69	67.19	56.22	68.82	74.00	82.99	46.52	61.14	67.70
									67.03
Vgg19	79.59	67.17	45.16	63.64	83.58	81.54	43.62	64.43	67.22
	83.16	70.31	54.05	60.43	79.60	83.08	48.94	69.95	68.69
	87.76	73.96	50.81	61.29	80.60	81.54	39.36	64.77	67.51
	80.10	72.40	52.97	54.84	78.50	85.64	43.32	62.18	66.24
	82.65	69.79	52.97	70.43	74.50	82.99	41.18	65.80	67.54
									67.44
Resnet50	79.59	65.80	42.47	52.94	82.09	81.03	40.96	61.45	63.28
	83.67	60.42	48.65	52.94	75.12	82.56	47.34	63.21	64.24
	87.24	64.06	52.43	51.08	80.60	82.05	39.89	64.77	65.27
	78.57	61.98	49.19	48.92	74.00	85.64	41.71	59.07	62.39
	81.63	63.54	47.57	57.53	43.50	82.47	41.18	61.66	63.63
									63.76
Mobilenet	81.82	56.69	79.07	87.41	86.03	70.80	65.65	80.29	75.97
	68.18	67.72	83.72	90.91	91.11	81.02	61.07	86.76	78.81
	59.09	68.50	80.62	85.31	91.85	81.75	50.38	88.24	75.72
	61.90	68.25	81.40	85.31	89.63	81.02	61.07	81.62	76.28
	76.19	64.29	81.40	86.62	88.89	85.40	66.41	80.15	78.67
									77.09

本文提出的 MPCNet 的平均故障识别率显著高于其他分类器,说明 MPCNet 模型具有从图像中提取更高级特征的能力,这些对原始数据的抽象化表征使模型能迅速捕捉到数据所属类别标签间的区别,从而提高对数据的分类精度. DBN 和 SDAE 分

类器识别率非常低,无法应用于 TEP 故障诊断. Vgg16、Vgg19 与 Resnet50 模型具有相似的平均故障识别率. Mobilenet 的故障识别率与本文提出的 MPCNet 最为相似. Vgg16、Vgg19、Resnet50 和 Mobilenet 是近几年较为流行的深度较大的 CNN 网

络,网络模型使用的卷积核较小,瓶颈层和模块化的使用降低了计算量,从而提高了模型的运算效率.但几种 CNN 模型对 TEP 故障类型的识别率不高,原因在于网络训练所采用的数据集过小,不足以支持庞大网络进行正常的训练.因此在训练样本集不大的情况下,使用较小的 CNN 网络识别 TEP 故障类型往往能取得更为理想的结果.通过 CNN 中的局部连接、对权值共享和下采样的参数进行约束,可实现“较小”的 CNN 结构. MPCNet 采用大小为 3×3 的卷积核、 2×2 的池化矩阵和大小为 128 的全连接层,减少了网络参数,使网络训练过程容易,不易过学习,这种轻量级的 CNN 模型更适合用于小样本数据量的过程诊断. CNN 常用于处理图像数据^[17,28,38],用于计算机视觉设计的 CNN 具有庞大的网络结构,如 AlexNet,参数共计 62 378 344 个. MPCNet 的参数量为 2 085 072,与大型网络相比, MPCNet 明显具有较小的网络结构,适用于样本量不大的多变量过程的故障诊断.

4 结 论

1) 本文提出的基于 CNN 的多变量过程故障诊断方法,将过程信号转为图像信号,利用 MPCNet 从图像中提取故障特征向量进行模式识别,克服了传统人工提取特征主观性大、依赖性强和效率低等缺点.

2) 与经典的分类器相比,基于 MPCNet 提取的特征进行故障模式识别可以得到更理想的结果,充分证明了 MPCNet 强大的特征学习能力.

3) 本文为以 CNN 为代表的深度学习方法在多变量过程故障诊断的应用提供了较为充分的参考依据,也为将来开发更有效的多变量监测方案奠定了基础.

下一步工作可用 MPCNet 提取故障数据特征,并结合多变量统计过程方法开发多变量过程的过程监控方法.

参 考 文 献

- [1] YU Jie. Localized fisher discriminant analysis based complex process monitoring [J]. *AIChE Journal*, 2011, 57(7): 1817. DOI: 10.1002/aic.12392
- [2] 李建军, 韦志辉, 张正军. 多专家的 PCA 边缘检测模型 [J]. *哈尔滨工业大学学报*, 2012, 44(11): 92
LI Jianjun, WEI Zhihui, ZHANG Zhengjun. Edge detection model based on multi-expert and principal component analysis [J]. *Journal of Harbin Institute of Technology*, 2012, 44(11): 92
- [3] NOMIKOS P, MACGREGOR J F. Monitoring batch processes using multiway principal component analysis [J]. *AIChE Journal*, 1994, 40(8): 1361. DOI: 10.1002/aic.690400809
- [4] NOMIKOS P, MACGREGOR J F. Multivariate SPC charts for monitoring batch processes [J]. *Technometrics*, 1995, 37(1): 41. DOI: 10.1080/00401706.1995.10485888
- [5] CHERRY G A, QIN S J. Multiblock principal component analysis based on a combined index for semiconductor fault detection and diagnosis [J]. *IEEE Transactions on Semiconductor Manufacturing*, 2006, 19(2): 159. DOI: 10.1109/tsm.2006.873524
- [6] YU Jie, QIN S J. Multimode process monitoring with Bayesian inference-based finite Gaussian mixture models [J]. *AIChE Journal*, 2018, 54(7): 1811. DOI: 10.1002/aic.11515
- [7] YU Jianbo. Hidden Markov models combining local and global information for nonlinear and multimodal process monitoring [J]. *Journal of Process Control*, 2010, 20(3): 344. DOI: 10.1016/j.jprocont.2009.12.002
- [8] 吴海曦, 余忠华, 张浩, 等. 面向熔融沉积成型的 3D 打印机故障声发射监控方法 [J]. *浙江大学学报(工学版)*, 2016, 50(1): 78
WU Haixi, YU Zhonghua, ZHANG Hao, et al. Method for monitoring of FDM 3D printer failure based on acoustic emission [J]. *Journal of Zhejiang University (Engineering Science)*, 2016, 50(1): 78
- [9] 徐文, 王大忠. 结合遗传算法的人工神经网络在电力变压器故障诊断中的应用 [J]. *中国电机工程学报*, 1997, 17(2): 109
XU Wen, WANG Dazhong. Application of artificial neural network combined with genetic algorithm in fault diagnosis of power transformer [J]. *Proceedings of the Chinese Society for Electrical Engineering*, 1997, 17(2): 109. DOI: 10.13334/j.0258-8013.psee.1997.02.009
- [10] GAO Xin, HOU Jian. An improved SVM integrated GS-PCA fault diagnosis approach of Tennessee Eastman Process [J]. *Neurocomputing*, 2016, 174: 906. DOI: 10.1016/j.neucom.2015.10.018
- [11] QIAO Zijian, LEI Yaguo, LIN Jing, et al. An adaptive unsaturated bistable stochastic resonance method and its application in mechanical fault diagnosis [J]. *Mechanical Systems and Signal Processing*, 2017, 84: 731. DOI: 10.1016/j.ymsp.2016.08.030
- [12] CHIANG L H, KOTANCHEK M E, KORDON A K. Fault diagnosis based on Fisher discriminant analysis and support vector machines [J]. *Computers and Chemical Engineering*, 2004, 28(8): 1389. DOI: 10.1016/j.compchemeng.2003.10.002
- [13] BENGIO Y, COURVILLE A, VINCENT P. Representation learning: a review and new perspectives [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2013, 35(8): 1798. DOI: 10.1109/tpami.2013.50
- [14] WANG Cong, GAN Meng, ZHU Chang'an. Fault feature extraction of rolling element bearings based on wavelet packet transform and sparse representation theory [J]. *Journal of Intelligent Manufacturing*, 2018, 29(4): 937. DOI: 10.1007/s10845-015-1153-2
- [15] 雷翠红, 邹平华. 供热管网的二级 BP 神经网络泄漏故障诊断 [J]. *哈尔滨工业大学学报*, 2011, 43(2): 75
LEI Cuihong, ZOU Pinghua. Two-stage BP neural network leakage fault diagnosis of heating networks [J]. *Journal of Harbin Institute of Technology*, 2011, 43(2): 75
- [16] TANG Peng, PENG Kaixiang, ZHANG Kai, et al. A deep belief network-based fault detection method for nonlinear processes [J]. *IFAC-Papers OnLine*, 2018, 51(24): 9. DOI: 10.1016/j.ifacol.2018.09.522
- [17] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks [C]// *International Conference on Neural Information processing System*. Lake Tahoe: Associate Inc, 2012: 1097

- [18] GERS F A, SCHMIDHUBER J. LSTM recurrent networks learn simple context free and context sensitive languages [J]. IEEE Transactions on Neural Networks, 2001, 12 (6): 1333. DOI: 10.1109/72.963769
- [19] GRAVES A, MOHAMED A R, HINTON G. Speech recognition with deep recurrent neural networks [C]// IEEE International Conference on Acoustics, Speech and Signal Processing. Vancouver: IEEE, 2013: 6645
- [20] 赵光权, 葛强强, 刘小勇, 等. 基于 DBN 的故障特征提取及诊断方法研究 [J]. 仪器仪表学报, 2016, 37(9): 1946
ZHAO Guangquan, GE Qiangqiang, LIU Xiaoyong, et al. Fault feature extraction and diagnosis method based on deep belief network [J]. Chinese Journal of Science Instrument, 2016, 37(9): 1946
- [21] 孙文珺, 邵思羽, 严如强. 基于稀疏自动编码深度神经网络的感应电动机故障诊断 [J]. 机械工程学报, 2016, 52(9): 65
SUN Wenjun, SHAO Siyu, YAN Ruqiang. Induction motor fault diagnosis based on deep neural network of sparse auto-encoder [J]. Journal of Mechanical Engineering, 2016, 52(9): 65
- [22] 王丽华, 谢阳阳, 周子贤, 等. 基于卷积神经网络的异步电机故障诊断 [J]. 振动, 测试与诊断, 2017, 37(6): 1208
WANG Lihua, XIE Yangyang, ZHOU Zixian, et al. Motor fault diagnosis based on convolutional neural networks [J]. Journal of Vibration, Measurement & Diagnosis, 2017, 37(6): 1208. DOI: 10.16450/j.cnki.issn.1004-6801.2017.06.021
- [23] JIA Feng, LEI Yaguo, GUO Liang, et al. A neural network constructed by deep learning technique and its application to intelligent fault diagnosis of machines [J]. Neurocomputing, 2018, 272: 619. DOI: 10.1016/j.neucom.2017.07.032
- [24] ZHANG Zhanpeng, ZHAO Jinsong. A deep belief network based fault diagnosis model for complex chemical processes [J]. Computers & Chemical Engineering, 2017, 107: 395. DOI: 10.1016/j.compchemeng.2017.02.041
- [25] JIANG Peng, HU Zhixin, LIU Jun, et al. Fault diagnosis based on chemical sensor data with an active deep neural network [J]. Sensors (Switzerland), 2016, 16 (10): 1695. DOI: 10.3390/s16101695
- [26] CHEN Zhiqiang, LI Chuan, SANCHEZ R V. Gearbox fault identification and classification with convolutional neural networks [J]. Shock and Vibration, 2015, 2015: 390134. DOI: 10.1155/2015/390134
- [27] LECUN Y, BOSER B, DENKER J S, et al. Backpropagation applied to handwritten zip code recognition [J]. Neural Computation, 1989, 1(4): 541. DOI: 10.1162/neco.1989.1.4.541
- [28] JI Shuiwang, XU Wei, YANG Ming, et al. 3D convolutional neural networks for human action recognition [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2013, 35(1): 221. DOI: 10.1109/TPAMI.2012.59
- [29] LEE H, GROSSE R, RANGANATH R, et al. Convolutional deep belief networks for scalable unsupervised learning of hierarchical representations [C]// Proceedings of 26th Annual International Conference on Machine Learning. Montreal: ACM, 2009: 609
- [30] LEE K B, CHEON S, KIM C O. A convolutional neural network for fault classification and diagnosis in semiconductor manufacturing processes [J]. IEEE Transactions on Semiconductor Manufacturing, 2017, 30(2): 135. DOI: 10.1109/tsm.2017.2676245
- [31] NCE T, KIRANYAZ S, EREN L, et al. Real-time motor fault detection by 1-D convolutional neural networks [J]. IEEE Transactions on Industrial Electronics, 2016, 63(11): 7067. DOI: 10.1109/tie.2016.2582729
- [32] JANSSENS O, SLAVKOVIKJ V, VERVISCH B, et al. Convolutional neural network based fault detection for rotating machinery [J]. Journal of Sound and Vibration, 2016, 377: 331. DOI: 10.1016/j.jsv.2016.05.027
- [33] ZHANG Wei, LI Chuanhao, PENG Gaoliang, et al. A deep convolutional neural network with new training methods for bearing fault diagnosis under noisy environment and different working load [J]. Mechanical Systems and Signal Processing, 2018, 100: 439. DOI: 10.1016/j.ymsp.2017.06.022
- [34] YU Jianbo. Local and global principal component analysis for process monitoring [J]. Journal of Process Control, 2012, 22(7): 1358. DOI: 10.1016/j.jprocont.2012.06.008
- [35] MAATEN L, HINTON G. Visualizing data using t-SNE [J]. Journal of Machine Learning Research, 2008, 9: 2579
- [36] ERHAN D, BENGIO Y, COURVILLE A, et al. Visualizing higher-layer features of a deep network [J]. University of Montreal, 2009, 1341(3): 1
- [37] WANG Guan, SUN Yu, WANG Jianxin. Automatic image-based plant disease severity estimation using deep learning [J]. Computational Intelligence and Neuroscience, 2017, 2017: 1. DOI: 10.1155/2017/2917536
- [38] HOWARD A G, ZHU Menglong, CHEN Bo, et al. MobileNets: efficient convolutional neural networks for mobile vision applications [DB/OL]. (2017-04-17) [2019-06-01]. <https://arxiv.org/abs/1704.04861>

(编辑 杨波)