

## 深度学习在流程工业过程数据建模中的应用

袁小锋, 王雅琳, 阳春华, 桂卫华

(中南大学自动化学院, 湖南 长沙 410083)

**摘要:**深度学习是近年来发展的人工智能技术。相比于传统浅层学习模型,深度学习具有强大的特征表示和函数拟合能力。深度学习能够从海量数据中提取层次特征,其在流程工业过程数据驱动建模中具有较大的潜力和应用前景。首先简单介绍了深度学习的发展历程;然后,介绍了4类广泛使用的深度学习模型以及它们在流程工业过程数据建模中的应用;最后,在流程工业过程数据建模领域对深度学习进行了简要总结。

**关键词:**流程工业;深度学习;数据解析;数据建模

**中图分类号:** TP18

**文献标识码:** A

**doi:** 10.11959/j.issn.2096-6652.202012

## The application of deep learning in data-driven modeling of process industries

YUAN Xiaofeng, WANG Yalin, YANG Chunhua, GUI Weihua

School of Automation, Central South University, Changsha 410083, China

**Abstract:** Deep learning is an artificial intelligence technique developed in recent years. Compared with traditional shallow models, deep learning has a strong ability of feature representation and function fitting. It can extract hierarchical features from massive data, which has great potential for data-driven modeling in process industries. Firstly, the history of deep learning was introduced. Then, four widely used deep networks and their applications were introduced in data-driven modeling of process industries. At last, the conclusions about deep learning and its applications in process industries were given.

**Key words:** process industries, deep learning, data analytics, data modeling

### 1 引言

流程工业是我国的基础原材料工业,主要包括冶金、钢铁、化工、石油和造纸等行业。作为制造业的重要组成部分,流程工业是我国国民经济和社会发展的支柱产业。经过数十年的发展,我国流程工业的生产工艺、装备和生产过程自动化水平都得到了大幅提升。现今,我国已成为世界门类最齐

全、规模最庞大的流程工业大国。在德国“工业4.0”和美国的“再工业化”等战略背景下,我国流程工业生产正朝着智能制造的目标发展。

流程工业智能制造的关键是实现生产过程建模、控制、优化和决策的智能化。其中,过程建模是实现智能制造的基础,也是智能控制和优化的先决条件。由于现代流程工业过程的大规模化和复杂化,传统基于机理分析和专家经验的方法无法准确

收稿日期: 2020-04-11; 修回日期: 2020-05-13

通信作者: 王雅琳, ylwang@csu.edu.cn

基金项目: 国家重点研发计划基金资助项目 (No.2018YFB1701100, No.2018AAA0101603); 国家自然科学基金资助项目 (No.61703440, No.61590921); 111 引智项目 (No.B17048); 中南大学创新驱动计划基金资助项目 (No.2018CX012); 湖南省自然科学基金资助项目 (No.2018JJ3687)

**Foundation Items:** The National Key Research and Development Program of China (No.2018YFB1701100, No.2018AAA0101603), The National Natural Science Foundation of China (No.61703440, No. 61590921), The 111 Project (No.B17048), Innovation Driven Program of Central South University (No.2018CX012), The Natural Science Foundation of Hunan Province of China (No.2018JJ3687)

获取过程模型,不能满足产品质量预测、生产过程监测和故障诊断的准确性和实时性需求,这将严重导致过程控制、优化和管理决策的滞后。随着数据采集、网络通信和数据存储等技术的飞速发展,流程工业生产过程不断积累大量的数据,其中蕴含丰富的过程知识和价值信息,为数据驱动建模提供了数据基础。流程工业大数据具有高维非线性、高速多采样性、动态时序性等复杂特性,如何从生产过程中大数据中挖掘和实现数据的真实知识和价值,并指导实际过程生产,是数据驱动建模的难点。

近年来,深度学习技术在互联网、信息通信、金融等领域取得了巨大成功,并正在深刻改变社会生活和世界格局<sup>[1-2]</sup>。深度学习通过深层网络结构对低等级特征进行变换或组合来获得更高等级的抽象特征,从而学习出具有层次结构的深度特征,具有更优异的特征学习能力和获取更本质的数据特征的能力。因而,深度学习在流程工业数据建模中有着广泛的应用。

本文主要集中于深度学习在流程工业数据建模中的软测量与故障诊断研究。软测量技术被广泛用于流程工业过程中的质量指标估计和预测,主要利用工业过程数据建立难以直接测量的质量变量与易于测量的过程变量之间的数学模型。利用该模型可以实现对难以测量的变量的实时在线预测。相比于传统的机理建模,建立基于深度网络的软测量模型不需要对工艺流程和机理有明确的认识,而且相较于浅层模型,深度学习模型具有更强的复杂函数逼近能力,故而能够很好地解决过程的非线性、时变性、动态性等问题,在质量预测任务中有更高的预测精度。故障诊断技术可分为故障检测(过程监测)、故障识别和分类等。在流程工业过程故障诊断领域,深度学习有广泛的应用前景。

过程监测通过收集工业过程中的历史数据,利用各种数据分析处理方法挖掘过程变量的重要信息,提取正常操作和故障条件下的数据特征表示,进而建立过程监测模型,对流程工业生产过程进行实时监测,判断生产过程是否存在异常,从而指导生产操作,确保生产的安全、稳定运行。如今,现代工业系统愈发复杂,使用基于解析模型和传统数据驱动模型的过程监控方法得不到满意的监测效果。目前,在现代工业工厂中,不同的操作单元可测量大量反映工业过程操作状态的变量,从而给基于深度学习的过程监测提供了契机。因此基于深度

学习的过程监测可以从工业生产的大量过程数据中提取有用信息,用于判断系统的运行状态。

此外,还可以利用深度学习实现工业过程的故障诊断与分类,通过分析过程的历史数据发现设备运行中的异常,并分析故障情况,从而及时根据故障类别对设备进行维护,以保障设备的正常运行。现代工业生产中各种设备组件之间的复杂的、强耦合的相关关系以及各种复杂的环境因素使得故障的发生具有随机性、不确定性以及并发性等,传统的故障诊断方法难以解决上述问题,误诊概率较大。深度学习能够学习到数据的有用特征表示,发现数据的本质结构,学习到对任务有用的信息,因而在故障诊断中具有更高的精度。

随着深度学习技术在流程工业中的发展,深度学习在流程工业过程数据建模中具有巨大的潜力,将进一步在流程工业智能化中发挥巨大的推动作用。本文主要阐述深度学习的发展历程、典型深度学习模型及其在流程工业中的应用,并对深度学习在流程工业过程数据建模领域的工作进行简要展望。

## 2 深度学习发展历程

深度学习作为机器学习的重要发展产物,其发展历程与机器学习密不可分。在深度学习之前,机器学习领域已发展出了各种学习模型,如以主成分分析(principal component analysis, PCA)和偏最小二程(partial least squares, PLS)回归等为代表的线性模型,以支持向量机(support vector machine, SVM)、核方法等为代表的非线性模型。相比而言,传统的机器学习方法大多属于浅层学习模型,其模型结构一般不超过一个隐藏层。浅层学习模型能够有效地解决一些情况简单或条件简单的问题。但是在面临一些更复杂的问题时,浅层学习模型的函数表达能力和建模效果将受到限制,无法满足复杂场景下的建模要求。

深度学习是一种多层次结构学习模型,其通过对原始信号进行逐层非线性特征变换,将样本在原空间的特征表示变换到新的特征空间自动地学习层次化的特征表示,从而更有利于解决分类或回归等特定问题。深度学习是人工神经网络的一个重要分支,具有深度网络结构的人工神经网络是深度学习最早的网络模型。总体而言,人工神经网络的发展经历了3个阶段。

第一阶段: 人工神经网络思想起源于 1943 年的人工神经元数学模型<sup>[3]</sup>, 该模型希望能够使用计算机来模拟人的神经元反应的过程。1958 年, Rosenblatt<sup>[4]</sup>提出了感知器(perceptron)算法, 并提出利用 Hebb 学习规则或最小二乘法来训练感知器的参数, 该模型被用于解决多维数据二分类问题。1969 年, 美国数学家 Minsky<sup>[5]</sup>在其著作中证明了感知器本质上是一种线性模型, 只能处理线性分类问题, 不能处理非线性分类问题。自此之后的十几年, 神经网络的研究陷入了一段低潮时期。

第二阶段: 在 1974 年, Werbos<sup>[6]</sup>提出采用反向传播法来训练一般的人工神经网络。1986 年, 加拿大多伦多大学的 Hinton 教授等人<sup>[7]</sup>在其基础上设计了适用于多层感知器(multi-layer perceptron, MLP)的反向传播算法, 并采用 Sigmoid 激活函数进行非线性映射, 有效解决了非线性分类和学习的问题。自此, 神经网络进入第二次研究热潮。但是, 反向传播算法对具有多隐含层的神经网络的模型参数的学习效果并不好, 基于梯度下降的反向传播法很容易在训练网络参数时收敛于局部极小值。因而, 部分研究学者开始探究通过改变感知器的结构来改善网络学习的性能。1995 年, Vapnik 教授等人<sup>[8]</sup>提出了著名的支持向量机(SVM)模型, 并发展出了核技巧的 SVM 模型, 在非线性分类问题上取得了比神经网络更突出的效果。因而, 神经网络的研究又进入了一段寒冬时期。

第三阶段: 2006 年, Hinton 教授等人<sup>[9]</sup>首次提出了深度学习技术, 通过无监督逐层预训练和有监督训练微调, 能够有效地解决深层网络训练过程中的梯度消失和梯度爆炸问题。2012 年, 以卷积神经网络(convolutional neural network, CNN)为基础的 AlexNet 模型<sup>[10]</sup>在国际 ImageNet 图像识别比赛中一举夺冠, 极大地降低了图像识别的错误率。自此, 深度学习受到各个领域研究人员的关注, 并再次引领了人工智能的研究热潮。

### 3 深度学习模型及其在流程工业中的应用

#### 3.1 深度置信网络

深度置信网络(deep belief network, DBN)是最早提出的深度学习模型, 是一种无监督概率生成模型。DBN 由多个受限玻尔兹曼机(restricted Boltzmann machine, RBM)堆叠而成, 其基本结构如图 1 所示<sup>[11]</sup>。其基本组成单元——受限玻尔兹曼

机是一种由可见层和隐藏层组成的基于能量的概率图模型, 对于一组给定的状态  $(\mathbf{v}, \mathbf{h})$ , 其能量函数  $E(\mathbf{v}, \mathbf{h})$  可表示为:

$$E(\mathbf{v}, \mathbf{h}) = -\mathbf{a}^\top \mathbf{v} - \mathbf{b}^\top \mathbf{h} - \mathbf{h}^\top \mathbf{w} \mathbf{v} \quad (1)$$

其中,  $\mathbf{v}$  和  $\mathbf{a}$  分别为可见层的状态向量和偏置向量,  $\mathbf{v} = (v_1, v_2, \dots, v_i, \dots, v_m)$ ,  $\mathbf{a} = (a_1, a_2, \dots, a_i, \dots, a_m)$ ,  $\mathbf{h}$  和  $\mathbf{b}$  分别为隐藏层的状态向量和偏置向量,  $\mathbf{h} = (h_1, h_2, \dots, h_j, \dots, h_m)$ ,  $\mathbf{b} = (b_1, b_2, \dots, b_j, \dots, b_m)$ ,  $\mathbf{w}$  为隐藏层和可见层之间的权值矩阵,  $\mathbf{w} = \{w_{ij}\}$ 。

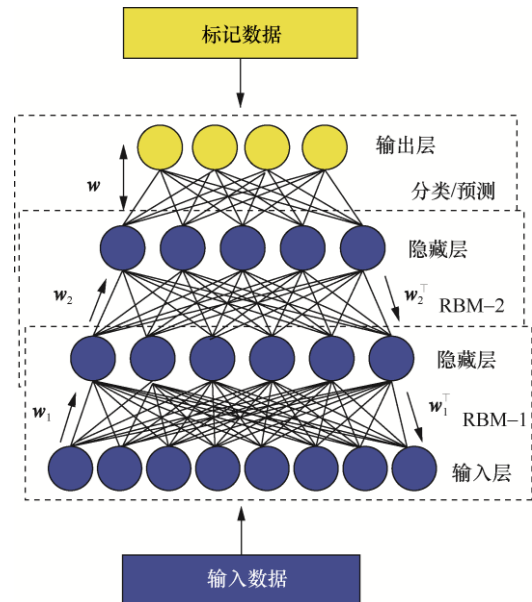


图 1 DBN 的基本结构

由上述能量函数可给出状态  $(\mathbf{v}, \mathbf{h})$  的联合概率分布:

$$P(\mathbf{v}, \mathbf{h}) = \frac{e^{-E_\theta(\mathbf{v}, \mathbf{h})}}{Z_\theta} \quad (2)$$

$$Z_\theta = \sum_{\mathbf{v}, \mathbf{h}} e^{-E_\theta(\mathbf{v}, \mathbf{h})} \quad (3)$$

其中, 参数集合  $\theta = \{\mathbf{v}, \mathbf{h}, \mathbf{a}, \mathbf{b}, \mathbf{w}\}$ ,  $Z_\theta$  为  $\theta$  的配分函数,  $E_\theta(\mathbf{v}, \mathbf{h})$  为  $\theta$  的能量函数, 其与能量函数  $E(\mathbf{v}, \mathbf{h})$  在数学计算上并无本质不同。

由此, 可计算得到对应  $\mathbf{v}$  和  $\mathbf{h}$  的边缘分布(似然函数):

$$P(\mathbf{v}) = \sum_{\mathbf{h}} P(\mathbf{v}, \mathbf{h}) = \frac{\sum_{\mathbf{h}} e^{-E(\mathbf{v}, \mathbf{h})}}{Z} \quad (4)$$

$$P(\mathbf{h}) = \sum_{\mathbf{v}} P(\mathbf{v}, \mathbf{h}) = \frac{\sum_{\mathbf{v}} e^{-E(\mathbf{v}, \mathbf{h})}}{Z} \quad (5)$$

因此,在给定可见层或隐藏层的神经元状态时,可计算对应隐藏层或可见层的神经元被激活(即取值为1)的概率:

$$P(h_j=1|\mathbf{v})=f(b_j+\sum_{i=1}^m v_i w_{ij}) \quad (6)$$

$$P(v_i=1|\mathbf{v})=f(a_i+\sum_{j=1}^n h_j w_{ij}) \quad (7)$$

其中,  $f(\cdot)$  为激活函数。

最后,采用最大化似然函数的方法求解得到 RBM 的参数,使得在该参数下得到的概率分布尽可能与样本的分布一致。为简化该计算过程,优化过程使用了对比散度(contrastive divergence, CD)算法。其原理是:通过隐含层对样本数据进行特征提取,并对可视层和隐含层神经元进行吉布斯(Gibbs)采样。在 CD- $k$  算法中,参数  $k$  为执行步数,训练 RBM 时  $k$  的值通常设为 1, RBM 执行一步对比散度计算后的权值增量为:

$$\Delta w_{ij}=P(h_j=1|\mathbf{v}^{(0)})v_i^{(0)}-P(h_j=1|\mathbf{v}^{(1)})v_i^{(1)} \quad (8)$$

$$\Delta b_j=P(h_j=1|\mathbf{v}^{(0)})-P(h_j=1|\mathbf{v}^{(1)}) \quad (9)$$

$$\Delta a_i=v_i^{(0)}-v_i^{(1)} \quad (10)$$

由输入层、隐藏层和输出层组成 DBN 模型的训练可分为无监督逐层预训练和有监督微调 2 个步骤。在无监督逐层预训练阶段,主要通过逐层贪婪训练方法对多个受限玻尔兹曼机进行逐层训练,单个 RBM 训练时的参数更新过程如式(8)~式(10)所示。首先从输入层引入原始输入数据,由输入层和第一隐藏层组成第一个受限玻尔兹曼机 RBM-1,通过最小化 RBM-1 的可视层观测数据的能量函数对其进行预训练。RBM-1 预训练完成后,再将其隐藏层的输出作为 RBM-2 的输入,继续对 RBM-2 进行预训练。以此类推,可逐层对整个 DBN 进行预训练。在整个预训练中,只用了输入数据,而未涉及数据标签,因而该过程被称为无监督预训练。预训练完成后,在 DBN 模型顶层加上特定任务的输出层,将预训练阶段的权重参数作为整个网络的初始参数,以数据标签信息最小化为目标,通过反向传播算法可对整个网络参数进行微调。

通过无监督逐层预训练和有监督微调,DBN 可对原始输入数据的深层抽象特征进行学习,其在流程工业过程软测量、过程监测和故障诊断等方面有广泛应用。例如,Shang 等人<sup>[12]</sup>建立了基于 DBN

的软测量模型,其分析了深度学习模型相较于传统机器学习模型在流程工业过程软测量建模方面的优势,并将其用于原油蒸馏装置中柴油 95% 分馏点的预测,实验结果证明 DBN 能够取得比传统方法更好的预测效果。刘瑞兰等人<sup>[13]</sup>建立了基于深度置信网络的精对苯二甲酸(pure terephthalic acid, PTA)工业生产过程中 4-CBA 预测的软测量模型。康岩等人<sup>[14]</sup>提出了基于 DBN-ELM 的球磨料位软测量方法。Liu 等人<sup>[15]</sup>提出了基于 DBN 的燃烧系统氧含量预测方法,其通过设计多层 DBN 结构,提取燃烧系统火焰图像的多层特征描述来估计氧含量。Yuan 等人<sup>[16]</sup>建立了基于 DBN 的冶金烧结过程的烧结成品矿成分含量预测模型,其提出的深度置信网络模型通过多层特征提取,挖掘出数据的本质特征,利用远超浅层网络的复函数逼近能力较准确地完成了基于复杂工业数据的质量预测建模;在故障检测与诊断方面,Zhang 等人<sup>[19]</sup>提出了基于深度置信网络的化工过程故障诊断模型。针对高炉冶炼系统复杂性特点,结合深度学习在处理复杂分布数据、提取特征等方面的优势,赵辉等人<sup>[18]</sup>提出了一种基于深度置信网络模型的高炉炉况分类方法。葛强强<sup>[19]</sup>针对目前多维时间分类方法没有考虑多维时间序列变量之间重要性联系以及不同样本矩阵约减后维数不完全对应的问题,将共同主成分分析(common principle component analysis, CPCA)和动态时间规整(dynamic time warping, DTW)方法引入 DBN 模型,提出了基于 CPCA-DTW-DBN 的多维时间序列故障诊断方法,在公开数据集和田纳西-伊斯曼过程故障案例中取得较好的分类效果。Wang 等人<sup>[11]</sup>提出了一种扩充 DBN 模型的化工过程故障分类模型,在预训练阶段通过逐层加入原始数据信息,可以更有效地提高特征提取和故障分类预测能力。

### 3.2 堆栈自编码器

深度学习策略在 DBN 上取得成功后,该思想被进一步延伸到其他深度网络。通过将 DBN 结构中的基本 RBM 模块替换成自编码器(autoencoder, AE)单元,构成了堆栈自编码器(stacked autoencoder, SAE)深度网络。图 2 分别给出了 AE 和 SAE<sup>[20]</sup>的基本结构。其中,AE 是单隐含层神经网络,且输出层重构其自身输入变量。在自编码器中,输入层到隐藏层部分称为编码器,隐藏层到输出层部分称为解码器。编码器将输入数据  $x$  由非线性函数  $f$  编码为隐层特征  $h$ ,解码器利用非线性函数  $\tilde{f}$  将  $h$  解码为  $\tilde{x}$ 。



因此, 自编码器的目标函数为:

$$J(\mathbf{W}, \mathbf{b}, \tilde{\mathbf{W}}, \tilde{\mathbf{b}}) = \frac{1}{2N} \sum_{i=1}^N \|\tilde{f}(f(x_i)) - x_i\|_2^2 = \frac{1}{2N} \sum_{i=1}^N \|\tilde{x}_i - x_i\|_2^2 \quad (11)$$

其中,  $\mathbf{W}$ 、 $\mathbf{b}$  分别为编码器的权值矩阵和偏置向量;  $\tilde{\mathbf{W}}$ 、 $\tilde{\mathbf{b}}$  分别为解码器的权值矩阵和偏置向量;  $N$  是样本总数,  $x_i$  和  $\tilde{x}_i$  分别为样本值和自编码器的输出值。AE 以最小化输出层对其输入数据的重构误差为目标, 通过反向传播算法进行网络参数训练和学习。因而, AE 是一个无监督特征自学习网络。类似于 DBN, SAE 深度网络的训练也分为无监督逐层预训练和有监督微调 2 个步骤。在无监督逐层预训练阶段, 网络从底层第一个 AE1 开始训练, 训练完后将其隐藏层的输出作为 AE2 的输入, 并对 AE2 进行训练。以此类推, 逐层训练, 直至最顶层 AEK 训练完成。预训练完成后, 各 AE 的编码器部分被保留在 SAE 网络中, 然后利用数据标签对整个网络进行有监督微调。

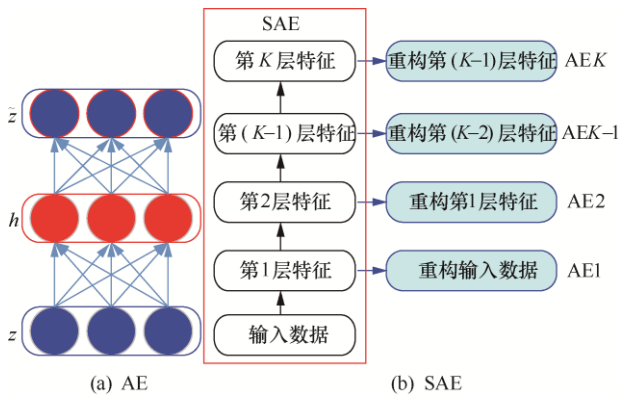


图2 AE 和 SAE 的基本结构

堆栈自编码器模型不依赖于过程标签数据, 能够通过无监督逐层预训练获取海量输入数据的层次特征表示, 其也被广泛用于流程工业过程数据建模。例如, 针对原始 SAE 预训练学习无法学习与质量变量相关的深度特征, Yuan 等人<sup>[21]</sup>提出了变量加权 SAE、混合变量加权 SAE<sup>[22]</sup>和堆栈质量驱动编码器<sup>[23]</sup>等模型的软测量方法, 将输出的标签信息以不同形式引入网络预训练过程, 可引导网络学习到与输出指标相关的层次特征, 从而提高软测量模型的预测精度。邱禹等人<sup>[24]</sup>提出了基于 SAE 深度学习的多输出自适应软测量建模方法。在过程监测与故

障诊断建模中, Kong 等人<sup>[25]</sup>提出了基于正则化自编码器的数据深度特征表示, 并用于工业过程监测。Yan 等人<sup>[26]</sup>提出了教师和监督双重堆叠自动编码器模型, 用于提取与质量相关的特征学习和工业过程故障监测。Yan 等人<sup>[27]</sup>提出用带标签的自编码器对神经网络进行监督, 并结合  $k$  近邻算法进行可视化工业过程监测。蒋立<sup>[28]</sup>提出了一种鲁棒自编码器模型及其故障检测方法以及一种基于反向传导的贡献度计算及其故障识别方法, 可提高模型的重构误差对故障的灵敏度, 且能避免对故障来源的错误识别。针对实际炼铁过程, Zhang 等人<sup>[29]</sup>提出了一种基于去噪自编码器的监控方法, 该方法考虑了 2 台不同的主热风炉之间切换引起的峰值扰动, 并将所提方法在冷炉故障对应的数据上进行了验证。Zhang 等人<sup>[30]</sup>提出了一种堆叠去噪自编码器和  $k$  近邻规则的非线性过程监测方法, 该方法利用堆叠去噪自编码器对非线性过程数据进行建模, 并自动提取关键特征, 然后通过堆叠自编码器将原始非线性空间映射到特征空间和剩余空间, 通过引入  $k$  近邻算法, 构造了上述空间中的 2 个新的统计量, 并在一个非线性数值系统和 TE 基准过程中验证了该方法的有效性。

### 3.3 卷积神经网络

CNN 是一种前馈型神经网络, 该网络由卷积层和池化层交替堆叠而成。卷积神经网络是目前深度学习领域研究较广泛的模型之一, 其在图像处理方面有突出的优势。图 3 给出了卷积神经网络的结构<sup>[31]</sup>。在卷积层中, 通过多个卷积核在输入数据中进行有规律的平移滑动, 获得其低维表示的特征图。卷积核所处的输入数据的各个位置被称为感受野 (receptive field) 区域, 其类似于视觉皮层细胞的感受野。卷积核的数量等于输出特征图的数量。假设输入特征图的宽和高分别为  $w_1$  和  $h_1$ ,  $n$  个卷积核的宽和高均为  $f$ , 那么将得到  $n$  个输出特征图的宽  $w_2$  和高  $h_2$  为:

$$w_2 = (w_1 - f + 2p) / s + 1 \quad (12)$$

$$h_2 = (h_1 - f + 2p) / s + 1 \quad (13)$$

其中,  $p$  为在输入特征图周围的零填充数,  $s$  为卷积核的步长。卷积运算是卷积核权值和感受野对应元素的加权和, 输出特征图在坐标  $(i, j)$  的值定义为:

$$y_{ij} = \sum_{k=0}^{f-1} \sum_{l=0}^{f-1} w_{kl} \times x_{(k+ixs)(l+jxs)}, (0 \leq i \leq w_2, 0 \leq j \leq h_2) \quad (14)$$

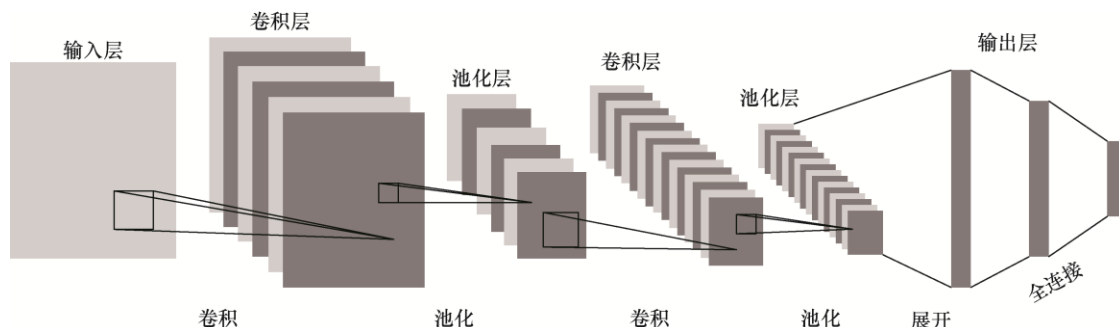


图3 卷积神经网络结构

其中,  $w_{kl}$  为卷积核在坐标  $(w, l)$  处的权重,  $x_{(k+i \times s)(l+j \times s)}$  为输入特征图在坐标  $(k+i \times s, l+j \times s)$  处的值,  $w_2$ 、 $h_2$  分别为输出特征图的宽和高。

在卷积层进行特征提取后,输出的特征图将被传递至池化层进行特征选择和信息过滤。池化层包含预设定的池化函数,其功能是将特征图中单个点的结果替换为其相邻区域的特征图统计量。常见的池化操作有最大池化、平均池化等。池化操作没有任何参数,并且可以对池化层的输入特征进行进一步降维,因此池化层有利于降低网络的参数规模。将多个卷积层和池化层进行交替堆叠后,可构建深度 CNN 结构。CNN 最高隐藏层输出的特征将被展开,并经过全连接层到达网络输出层,用于特定建模任务。不同于 DBN 和 SAE 等全连接神经网络, CNN 具有强大的局部特征学习能力,因而在局部特征学习方面有着比其他算法更高的优越性。

相较而言, CNN 在大型二维图像处理研究领域取得了巨大的成功,因而, CNN 也被用于流程工业过程图像处理相关研究。例如, Horn 等人<sup>[32]</sup>将卷积神经网络应用于泡沫浮选过程中基于图像的软测量建模。但是,目前针对流程工业中过程变量数据的 CNN 建模较少,其主要原因在于流程工业过程的采样数据是以向量形式存在的,而不是二维数据形式。然而,通过对一维数据进行二维动态扩充, CNN 能够有效地对流程工业过程数据进行局部动态特征学习。例如,易令等人<sup>[33]</sup>利用 CNN 对国内某炼油厂的原油总氢物性进行回归预测。Wang 等人<sup>[34]</sup>提出了将优先脉冲响应和 CNN 结合的有限冲激响应卷积神经网络用于动态过程软测量建模,并将该模型用于仿真算例和化工过程,结果显示有限冲激响应卷积神经网络能给出与质量相关的变量

的最佳预测精度和最易解释的趋势。Wu 等人<sup>[35]</sup>通过数据动态扩充,提出了基于 CNN 模型的化工过程故障诊断与分类方法。苏堪裂<sup>[36]</sup>研究了基于卷积神经网络的化工过程故障诊断。Lee 等人<sup>[37]</sup>提出了基于 CNN 网络的半导体生产过程故障分类与诊断方法,这种方法使得第一卷积层的输出与原始数据的结构相关联,由此能够定位表示过程故障和时间信息,在一个化学气相沉积过程的实验中,该方法要优于其他深度学习模型。

### 3.4 循环神经网络

循环神经网络 (recurrent neural network, RNN) 是一类以序列数据为输入,在序列的演进方向进行递归且所有节点按链式连接的神经网络。图 4 (a) 给出了 RNN 模型的基本结构。不同于前述 DBN、SAE、CNN 等模型, RNN 模型是一种时间序列模型,模型中样本之间具有时间关联性。而 DBN、SAE、CNN 等模型的基本假设是数据之间不存在时间相关性,即所有样本服从独立同分布假设。因此,在 RNN 结构中,每个神经元不但可以接收当前时刻的信息,还可以接收其前面时刻的信息,形成具有环路的网络结构。假设当前时刻  $t$  的输入为  $x_{(t)}$ ,前一时刻的隐状态为  $h_{(t-1)}$ ,则可以计算出当前时刻 RNN 单元的隐状态为:

$$h_{(t)} = \tanh(Wx_{(t)} + Vh_{(t-1)} + b) \quad (15)$$

其中,  $W$  和  $V$  为网络的参数矩阵,  $b$  为偏置向量,  $\tanh(\cdot)$  为非线性激活函数。

由隐状态可得到当前时刻  $t$  的质量变量预测值:

$$y_{(t)} = Uh_{(t)} + b_y \quad (16)$$

其中,  $U$  为输出层的权值参数矩阵,  $b_y$  为输出层的偏置向量。

由于时间上的关联性, RNN 参数的训练与前述

深度网络训练方式不同,RNN 采用基于时间的反向传播算法来学习,即按照时间的逆序把误差一步步往前传递。RNN 是一种具有短期记忆能力的神经网络,其无法解决长期依赖问题,即当输入序列较长时,容易产生梯度爆炸或梯度消失问题<sup>[38]</sup>。为了有效缓解该问题,在 RNN 的基本结构中引入了记忆细胞和门控机制,用于对历史信息进行存储以及选择性记忆。典型的基于门控机制的改进 RNN 模型包括长短期记忆(long short-term memory, LSTM)网络<sup>[39]</sup>和门控循环单元(gated recurrent unit, GRU)<sup>[40]</sup>等,图4(b)为基本 LSTM 单元的结构。 $f$ 、 $i$ 、 $o$  分别表示 LSTM 单元的遗忘门、输入门和输出门。假设当前时刻的输入为  $x_{(t)}$ , 前一时刻的隐状态为  $h_{(t-1)}$ , 前一时刻记忆单元的输出为  $c_{(t-1)}$ , 则遗忘门、输入门和输出门的输出为:

$$f_{(t)} = \sigma(W_{fx}x_{(t)} + W_{fh}h_{(t-1)} + b_f) \quad (17)$$

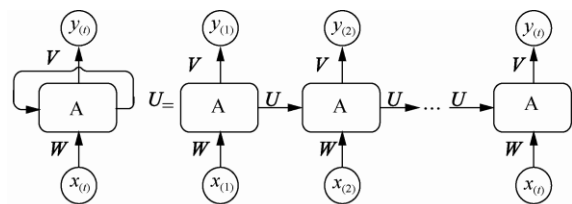
$$i_{(t)} = \sigma(W_{ix}x_{(t)} + W_{ih}h_{(t-1)} + b_i) \quad (18)$$

$$o_{(t)} = \sigma(W_{ox}x_{(t)} + W_{oh}h_{(t-1)} + b_o) \quad (19)$$

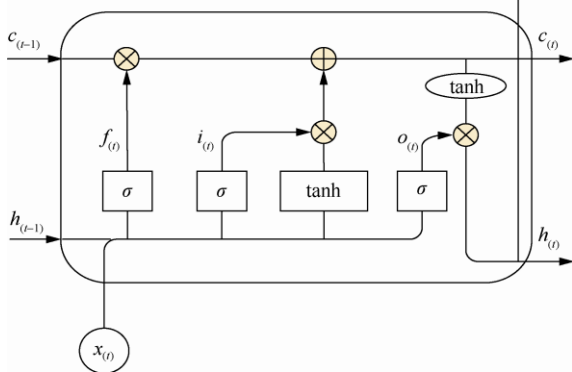
其中,  $W_{*x}$  和  $W_{*h}$  分别为 3 个控制门与输入信息和前一时刻隐状态之间的连接权值矩阵,  $b_f$ 、 $b_i$ 、 $b_o$  为相应的偏置向量,  $\sigma(\cdot)$  为 Sigmoid 激活函数。当前时刻的候选输入信息为:

$$\tilde{c}_{(t)} = \tanh(W_{cx}x_{(t)} + W_{ch}h_{(t-1)} + b_c) \quad (20)$$

其中,  $b_c$  为当前时刻的偏置向量。



(a) RNN模型结构



(b) 基本LSTM单元结构

图4 RNN 模型和 LSTM 单元的结构

由此可更新 LSTM 单元的记忆细胞,并得到当前时刻的隐状态输出:

$$c_{(t)} = f_{(t)} \cdot c_{(t-1)} + i_{(t)} \cdot \tilde{c}_{(t)} \quad (21)$$

$$h_{(t)} = o_{(t)} \cdot \tanh(c_{(t)}) \quad (22)$$

由于复杂物理化学反应机理、反馈控制和动态噪声等因素,流程工业过程本身具有自然的动态特征属性。因而,通过连续采样获得的过程数据具有较强的时间关联性。为了描述过程时序动态特性,基于 RNN 的深度学习模型也被广泛用于流程工业过程数据建模。最开始, Su 等人<sup>[41]</sup>将并行方法和 RNN 结合的外部循环网络用于炼油厂污水生物处理及催化重整装置; Chen 等人<sup>[42]</sup>将 RNN 用于生物量浓度的在线预测,网络的输入变量为进料量、液体体积和溶解氧浓度,试验结果的精度为 $\pm 11\%$ ; Zhao 等人<sup>[43]</sup>针对焦炉煤气运行优化问题,提出了基于高斯过程的长短期记忆网络,实现了对煤气实时流量和煤气柜液位的实时预测。近几年,随着经济的快速发展,人们对生产过程的经济、高效、安全提出了更高的要求,使得 LSTM 在流程工业软测量建模中得到了更加广泛的应用。例如 Fu 等人<sup>[44]</sup>针对影响石灰石/石膏湿法烟气脱硫效率的主要因素,采用 LSTM 建立了脱硫效率预测模型; Han 等人<sup>[45]</sup>为了准确分析和评价乙烯生产装置在工业过程中的能源效率状况,提出了一种基于双向长短期记忆(bidirectional long short-term memory, Bi LSTM)网络的产量预测方法; Yuan 等人提出了基于有监督 LSTM 网络<sup>[46]</sup>和变量注意力机制 LSTM 网络模型<sup>[47]</sup>的非线性动态过程软测量建模方法。Wang 等人<sup>[48]</sup>将 LSTM 用于批次过程每个阶段的长时间序列中,用于提取与质量相关的隐藏特征,并对产品质量进行预测建模。Sun 等人<sup>[49]</sup>提出了基于 DBN 和 RNN 的概率序列网络,并将其用于复杂过程数据软测量建模。Kataria 等人<sup>[50]</sup>成功地将一种基于 RNN 的软测量模型用于反应精馏塔的监测和控制系统。

## 4 结束语

大数据时代的到来使得数据驱动建模在流程工业智能工厂建设中发挥着至关重要的作用。深度学习作为大数据处理的核心技术之一,能够从海量过程数据中挖掘和学习到有用的层次特征表示,从而有效地提高数据驱动模型的性能。本文首先简单回顾了深度学习的发展历程,然后介绍了目前 4 种

较为流行的深度学习基本模型,并对各个模型在流程工业过程数据建模中的应用进行了简要概述。

深度学习作为机器学习的一个新领域,现在仍处于蓬勃发展时期。由于深度学习具有强大的数据处理和特征学习能力,其在流程工业过程数据建模中取得了一定的成效。但是,深度学习在流程工业过程这一特定领域的应用还有许多问题值得研究者们进一步深入探索和研究。例如,流程工业过程往往具有缓慢时变特性,其操作条件和数据模式也随着过程运行而不断发生改变。深度学习模型的性能在过程时变特性下将一步步退化,如何有效维持并提高过程时变特性下深度学习的建模精度和性能,是未来深度学习在流程工业研究领域的一个难点。此外,由于流程工业过程数据的动态时序性,数据样本之间存在较强的时间相关性。因而,如何建立有效的流程工业过程时间序列模型将是未来深度学习在本领域的重点研究方向之一。

## 参考文献:

- [1] 张钹. 人工智能进入后深度学习时代[J]. 智能科学与技术学报, 2019, 1(1): 4-6.  
ZHANG B. Artificial intelligence is entering the post deep-learning era[J]. Chinese Journal of Intelligent Science and Technology, 2019, 1(1): 4-6.
- [2] 刘远平, 宋昱锴, 张小燕, 等. 基于深度学习的髋关节应力分布算法研究[J]. 智能科学与技术学报, 2019, 1(3): 260-268.  
LIU Y P, SONG Y K, ZHANG X Y, et al. Research on hip joint stress distribution algorithms based on deep learning[J]. Chinese Journal of Intelligent Science and Technology, 2019, 1(3): 260-268.
- [3] MCCULLOCH W S, PITTS W. A logical calculus of the ideas immanent in nervous activity[J]. Bulletin of Mathematical Biophysics, 1943, 52(1-2): 99-115.
- [4] ROSENBLATT F. The perceptron: a probabilistic model for information storage and organization in the brain[J]. Psychological Review, 1958, 65(6): 386-408.
- [5] MINSKY M, PAPERT S. Perceptrons an introduction to computational geometry[M]. Cambridge: MIT Press, 1969.
- [6] WERBOS P. Beyond regression: new tools for prediction and analysis in the behavioral sciences[D]. Boston: Harvard University, 1974.
- [7] RUMMELHART D E, HINTON G E, WILLIAMS R J. Learning internal representations by error propagation[J]. Nature, 1986, 323(2): 318-362.
- [8] CORTES C, WAPNIK V. Support-vector networks[J]. Machine Learning, 1995, 20(3): 273-297.
- [9] HINTON G E, OSINDERO S, THE Y W. A fast learning algorithm for deep belief nets[J]. Neural Computation, 2006, 18(7): 1527-1554.
- [10] KRIZHEVSKY A, SUTSKEVER I, HINTON G. Imagenet classification with deep convolutional neural networks[C]//Advances in Neural Information Processing Systems. [S.l.:s.n.], 2012: 1097-1105.
- [11] WANG Y L, PAN Z F, YANG C H, et al. A novel deep learning based fault diagnosis approach for chemical process with extended deep belief network[J]. ISA Transactions, 2020, 96: 457-467.
- [12] SHANG C, YANG F, HUANG D X, et al. Data-driven soft sensor development based on deep learning technique[J]. Journal of Process Control, 2014, 24(3): 223-233.
- [13] 刘瑞兰, 毛佳敏. 基于深度置信网络的 4-CBA 软测量建模[J]. 计算机工程与应用, 2017, 53(6): 227-230.  
LIU R L, MAO J M. Soft sensor modeling of 4-CBA based on deep belief networks[J]. Computer Engineering and Applications, 2017, 53(6): 227-230.
- [14] 康岩, 卢慕超, 阎高伟. 基于 DBN-ELM 的球磨机料位软测量方法研究[J]. 仪表技术与传感器, 2015(4): 73-75, 92.  
KANG Y, LU M C, YAN G W. Soft sensor for ball mill fill level based on DBN-ELM model[J]. Instrument Technique and Sensor, 2015(4): 73-75, 92.
- [15] LIU Y, FAN Y, CHEN J H. Flame images for oxygen content prediction of combustion systems using DBN[J]. Energy & Fuels, 2017, 31(8): 8776-8783.
- [16] YUAN Z, WANG B, LIANG K, et al. In application of deep belief network in prediction of secondary chemical components of sinter[CI]/The 13th IEEE Conference on Industrial Electronics and Applications (ICIEA). Piscataway: IEEE Press, 2018: 2746-2751.
- [17] 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.
- [18] 赵辉, 赵德涛, 岳有军, 等. 基于深度置信网络的高炉炉况故障分类方法的研究[J]. 铸造技术, 2018, 39(5): 1028-1032.  
ZHAO H, ZHAO D T, YUE Y J, et al. Research on fault classification of blast furnace condition based on deep belief network[J]. Foundry Technology, 2018, 39(5): 1028-1032.
- [19] 葛强强. 基于深度置信网络的数据驱动故障诊断方法研究[D]. 哈尔滨: 哈尔滨工业大学, 2016.  
GE Q Q. Research on data driven fault diagnosis method based on deep belief network[D]. Harbin: Harbin Institute of Technology, 2016.
- [20] YUAN X F, GU Y J, WANG Y L, et al. A deep supervised learning framework for data-driven soft sensor modeling of industrial processes[J]. IEEE Transactions on Neural Networks and Learning Systems, 2019: 1-10.
- [21] YUAN X F, HUANG B, WANG Y L, et al. Deep learning based feature representation and its application for soft sensor modeling with variable-wise weighted SAE[J]. IEEE Transactions on Industrial Informatics, 2018, 14(7): 3235-3243.
- [22] YUAN X F, OU C, WANG Y L, et al. Deep quality-related feature extraction for soft sensing modeling: a deep learning approach with hybrid VW-SAE[J]. Neurocomputing, 2019.
- [23] YUAN X F, ZHOU J, HUANG B, et al. Hierarchical quality-relevant feature representation for soft sensor modeling: a novel deep learning strategy[J]. IEEE Transactions on Industrial Informatics, 2020, 16(6): 3721-3730.
- [24] 邱禹, 刘乙奇, 吴菁, 等. 基于深层神经网络的多输出自适应软测量建模[J]. 化工学报, 2018, 69(7): 3101-3113.  
QIU Y, LIU Y Q, WU J, et al. A self-adaptive multi-output soft sensor modeling based on deep neural network[J]. CIESC Journal, 2018, 69(7): 3101-3113.
- [25] KONG D H, YAN X F. Industrial process deep feature representation by regularization strategy autoencoders for process monitoring[J]. Measurement Science Technology, 2019, 31(2): 025104.
- [26] YAN S F, YAN X F. Design teacher and supervised dual stacked auto-encoders for quality-relevant fault detection in industrial process[J]. Applied Soft Computing, 2019, 81: 105526.
- [27] YAN S F, YUAN X F. Using labeled autoencoder to supervise neural network combined with k-nearest neighbor for visual industrial process monitoring[J]. Industrial & Engineering Chemistry Research, 2019, 58(23): 9952-9958.
- [28] 蒋立. 基于自编码器模型的非线性过程监测[D]. 杭州: 浙江大学, 2018.



- JIANG L. Nonlinear process monitoring based on auto-encoder model[D]. Hangzhou: Zhejiang University, 2018.
- [29] ZHANG T S, WANG W, YE H, et al. Fault detection for ironmaking process based on stacked denoising autoencoders[C]//The 2016 American Control Conference (ACC). [S.l.:s.n.], 2016.
- [30] ZHANG Z H, JIANG T, LI S H, et al. Automated feature learning for nonlinear process monitoring, an approach using stacked denoising autoencoder and k-nearest neighbor rule[J]. Journal of Process Control, 2018, 64: 49-61.
- [31] LECUN Y, BOTTOU L, BENGIO Y, et al. Gradient-based learning applied to document recognition[J]. Proceedings of the IEEE, 1998, 86(11): 2278-2324.
- [32] HORN Z C, AURET L, MCCOY J T, et al. Performance of convolutional neural networks for feature extraction in froth flotation sensing[J]. IFAC Papersonline, 2017, 50(2): 13-18.
- [33] 易令, 吕忠元, 丁进良, 等. 面向原油总氢物性预测的数据扩增预处理方法[J]. 控制与决策, 2018, 33(12): 2153-2160.
- YI L, LYU Z Y, DING J L, et al. Data pretreatment approach for crude oil hydrogen properties prediction[J]. Control and Decision, 2018, 33(12): 2153-2160.
- [34] WANG K C, SHANG C, LIU L, et al. Dynamic soft sensor development based on convolutional neural networks[J]. Industrial & Engineering Chemistry Research, 2019, 58(26): 11521-11531.
- [35] WU H, ZHAO J S. Deep convolutional neural network model based chemical process fault diagnosis[J]. Computers & Chemical Engineering, 2018: 115.
- [36] 苏堪裂. 基于卷积神经网络的化工过程故障诊断研究[D]. 广州: 华南理工大学, 2019.
- SU K L. Fault diagnosis of chemical process based on convolutional neural network research[D]. Guangzhou: South China University of Technology, 2019.
- [37] 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): 135-142.
- [38] BENGIO Y, SIMARD P, FRASCONI P. Learning long-term dependencies with gradient descent is difficult[J]. IEEE Transactions on Neural Networks, 1994, 5(2): 157-166.
- [39] HOCHREITER S, SCHMIDHUBER J. Long short-term memory[J]. Neural Computation, 1997, 9(8): 1735-1780.
- [40] CHO K, WAN MERRIENBOER B, GULCEHRE C, et al. Learning phrase representations using RNN encoder-decoder for statistical machine translation[C]//The 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP). [S.l.: s.n.], 2014: 1724-1734.
- [41] SU H T, MCAVOY T J, WERBOS P. Long-term predictions of chemical processes using recurrent neural networks: a parallel training approach[J]. Industrial Engineering Chemistry Research, 1992, 31(5): 1338-1352.
- [42] CHEN L Z, NGUANG S K, LI X M, et al. Soft sensors for on-line biomass measurements[J]. Bioprocess and Biosystems Engineering, 2004, 26(3): 191-195.
- [43] ZHAO J, LIU Q L, WANG W, et al. Hybrid neural prediction and optimized adjustment for coke oven gas system in steel industry[J]. IEEE Transactions on Neural Networks & Learning Systems, 2012, 23(3): 439-450.
- [44] FU J G, XIAO H, WANG T, et al. Prediction model of desulfurization efficiency of coal-fired power plants based on long short-term memory neural network[C]//The 2019 International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData). [S.l.:s.n.], 2019: 40-45.
- [45] HAN Y M, ZHOU R D, GENG Z Q, et al. Production prediction modeling of industrial processes based on Bi-LSTM[C]//The 34th Youth Academic Annual Conference of Chinese Association of Automation. [S.l.:s.n.], 2019: 285-289.
- [46] YUAN X F, LI L, WANG Y L. Nonlinear dynamic soft sensor modeling with supervised long short-term memory network[J]. IEEE Transactions on Industrial Informatics, 2020, 16(5): 3168-3176.
- [47] YUAN X F, LI L, WANG Y L, et al. Deep learning for quality prediction of nonlinear dynamic process with variable attention-based long short-term memory network[J]. The Canadian Journal of Chemical Engineering, 2019.
- [48] WANG K, GOPALUNI B, CHEN J H, et al. Deep learning of complex batch process data and its application on quality prediction[J]. IEEE Transactions on Industrial Informatics, 2018: 1.
- [49] SUN Q Q, GE Z Q. Probabilistic sequential network for deep learning of complex process data and soft sensor application[J]. IEEE Transactions on Industrial Informatics, 2018, 15(5): 2700-2709.
- [50] KATARIA G, SINGH K. Recurrent neural network based soft sensor for monitoring and controlling a reactive distillation column[J]. Chemical Product and Process Modeling, 2017, 13(3).

#### [作者简介]



袁小锋(1988—), 男, 博士, 中南大学自动化学院副教授, 主要研究方向为工业大数据与人工智能、深度学习与模式识别、过程监测与软测量建模等。



王雅琳(1973—), 女, 博士, 中南大学自动化学院教授、副院长、博士生导师, 主要研究方向为大数据分析处理、智能集成建模、过程监测与故障诊断等。



阳春华(1965—), 女, 博士, 中南大学自动化学院教授、院长、博士生导师, 主要研究方向为复杂工业过程建模与优化、分析检测与自动化装置、智能化系统等。



桂卫华(1950—), 男, 中国工程院院士, 中南大学自动化学院教授, 主要研究方向为工业大系统递阶和分散控制理论及应用、复杂工业过程建模、优化与控制应用和知识自动化。