



基于动态注意力深度迁移网络的高炉铁水硅含量在线预测方法

蒋珂 蒋朝辉 谢永芳 潘冬 桂卫华

Online Prediction Method for Silicon Content of Molten Iron in Blast Furnace Based on Dynamic Attention Deep Transfer Network

JIANG Ke, JIANG Zhao-Hui, XIE Yong-Fang, PAN Dong, GUI Wei-Hua

在线阅读 View online: <https://doi.org/10.16383/j.aas.c210524>

您可能感兴趣的其他文章

[集成自编码与PCA的高炉多元铁水质量随机权神经网络建模](#)

Autoencoder and PCA Based RVFLNs Modeling for Multivariate Molten Iron Quality in Blast Furnace Ironmaking

自动化学报. 2018, 44(10): 1799–1811 <https://doi.org/10.16383/j.aas.2018.c170299>

[高炉铁水质量鲁棒正则化随机权神经网络建模](#)

Robust Regularized RVFLNs Modeling of Molten Iron Quality in Blast Furnace Ironmaking

自动化学报. 2020, 46(4): 721–733 <https://doi.org/10.16383/j.aas.2018.c170670>

[基于多参数灵敏度分析与遗传优化的铁水质量无模型自适应控制](#)

Model Free Adaptive Control of Molten Iron Quality Based on Multi-parameter Sensitivity Analysis and GA Optimization

自动化学报. 2021, 47(11): 2600–2613 <https://doi.org/10.16383/j.aas.c180741>

[基于注意力机制的协同卷积动态推荐网络](#)

Attention-based Collaborative Convolutional Dynamic Network for Recommendation

自动化学报. 2021, 47(10): 2438–2448 <https://doi.org/10.16383/j.aas.c190820>

[一种迁移学习和可变形卷积深度学习的蝴蝶检测算法](#)

A Butterfly Detection Algorithm Based on Transfer Learning and Deformable Convolution Deep Learning

自动化学报. 2019, 45(9): 1772–1782 <https://doi.org/10.16383/j.aas.c190104>

[基于迁移学习的细粒度实体分类方法的研究](#)

Fine-grained Entity Type Classification Based on Transfer Learning

自动化学报. 2020, 46(8): 1759–1766 <https://doi.org/10.16383/j.aas.c190041>

基于动态注意力深度迁移网络的高炉铁水硅含量在线预测方法

蒋珂¹ 蒋朝辉^{1,2} 谢永芳¹ 潘冬¹ 桂卫华¹

摘要 铁水硅含量是反映高炉冶炼过程中热状态变化的灵敏指示剂, 但无法实时在线检测, 造成铁水质量调控盲目. 为此, 提出一种基于动态注意力深度迁移网络 (Attention deep transfer network, ADTNet) 的高炉铁水硅含量在线预测方法. 首先, 针对传统深度网络静态建模思路无法准确描述过程变量与铁水硅含量之间的关系, 提出一种基于注意力机制模块的输入过程变量与输出硅含量之间的动态关系描述方法; 其次, 为降低硅含量预测模型训练时对标签数据的依赖, 考虑到铁水温度与硅含量数据之间的正相关性, 利用小时级硅含量标签数据微调基于分钟级铁水温度数据预训练好的深度模型的结构, 进而提高基于动态注意力深度迁移网络的硅含量预测精度; 同时, 为增强预测网络的可解释性, 实时给出了基于动态注意力机制模块计算的每个样本各过程变量对铁水硅含量的贡献度; 最后, 基于某钢铁厂 2 号高炉的工业实验, 验证了该方法的准确性、有效性和先进性.

关键词 高炉炼铁, 铁水硅含量, 深度网络, 迁移学习, 动态注意力机制, 预测

引用格式 蒋珂, 蒋朝辉, 谢永芳, 潘冬, 桂卫华. 基于动态注意力深度迁移网络的高炉铁水硅含量在线预测方法. 自动化学报, 2023, 49(5): 949–963

DOI 10.16383/j.aas.c210524

Online Prediction Method for Silicon Content of Molten Iron in Blast Furnace Based on Dynamic Attention Deep Transfer Network

JIANG Ke¹ JIANG Zhao-Hui^{1,2} XIE Yong-Fang¹ PAN Dong¹ GUI Wei-Hua¹

Abstract The molten iron silicon content, which can reflect the thermal state in blast furnace hearth during the ironmaking process, is difficult to detect in real time. This will cause blind adjustment to the quality of the molten iron. Hence, this paper proposes a data-driven model for the online prediction of the silicon content based on dynamic attention deep transfer network (ADTNet). First, considering that the deep network cannot accurately describe the relationship between process variables and the silicon content based on static modeling process, a dynamic attention module is designed to describe the dynamic relationship between the inputs and outputs. Then, to reduce the dependence of labeled silicon content samples during model training process, an online prediction model is established based on minute-level temperature data after considering the positive correlation between molten iron temperature and silicon content data. Subsequently, using the labeled silicon content data to fine tune the parameters of the well-trained deep model and improve the silicon content prediction performance based on dynamic attention deep transfer network. Furthermore, to enhance the interpret-ability of the deep black box model, the contribution of process variables of each sample to the silicon content is presented based on the dynamic attention module. Finally, industrial experiments of No. 2 blast furnace in a steel plant verify the accuracy, effectiveness and advancement of the proposed method.

Key words Blast furnace ironmaking process, silicon content of molten iron, deep network, transfer learning, dynamic attention mechanism, prediction

Citation Jiang Ke, Jiang Zhao-Hui, Xie Yong-Fang, Pan Dong, Gui Wei-Hua. Online prediction method for silicon content of molten iron in blast furnace based on dynamic attention deep transfer network. *Acta Automatica Sinica*, 2023, 49(5): 949–963

收稿日期 2021-06-10 录用日期 2021-11-02

Manuscript received June 10, 2021; accepted November 2, 2021
国家自然科学基金 (61773406, 61725306, 61290325), 国家重大科研仪器研制项目 (61927803), 中南大学研究生自主探索创新项目 (2021zzts0183), 湖南省研究生科研创新项目 (CX20210242) 资助

Supported by National Natural Science Foundation of China (61773406, 61725306, 61290325), National Major Scientific Research Equipment of China (61927803), Independent Exploration and Innovation Project for Postgraduate of Central South University (2021zzts0183), and Hunan Provincial Innovation Foundation for Postgraduate (CX20210242)

本文责任编辑 陈胜勇

高炉炼铁是钢铁流程中的上游和关键工序, 是铁素物质流转换的核心关键单元, 是钢铁制造过程中能耗最大和生产成本最高的环节^[1-2]. 炼铁过程中的主体设备高炉是气体、液体和固体三相共存的竖

Recommended by Associate Editor CHEN Sheng-Yong

1. 中南大学自动化学院 长沙 410000 2. 鹏城实验室 深圳 518000

1. School of Automation, Central South University, Changsha 410000 2. Peng Cheng Laboratory, Shenzhen 518000

式反应器,具有连续鼓风、周期性加料和出铁、高温高压高尘等特征,其内部连续发生复杂的物理化学反应,并最终生产出熔融铁水^[3]。钢铁工业要实现真正的高质量发展,至少要过“五关”,更要斩“六降”,“五关”即产能关、布局关、绿色关、质量关和效率关,“六降”即降预期、降杠杆、降排放、降人员、降风险和降成本。为了满足这种趋势的发展,生产过程中关键性能指标的实时监控,在工业界和学术界得到了广泛的关注^[4-6]。

铁水硅含量是高炉冶炼过程中表征铁水质量的重要指标,也是反映炉缸内部热状态和炉况的灵敏指示剂。铁水硅含量偏低,说明铁水物理热不足,炉缸热储备不够,容易造成炉缸冻结事故。铁水硅含量偏高,有利于去除铁水中的有害元素磷和硫,但是过高的铁水硅含量又会使生铁变硬变脆,收得率降低且易引起喷溅。另外,铁水中过高的硅也意味着炉缸煤气流过于旺盛,易导致悬料、崩料等炉况故障及能源的浪费^[7]。通常,铁水硅含量需要根据原燃料品位和铁水品质的要求,控制在一个合理的范围内,如本文研究的中国西南某钢铁厂需要控制在0.3%~0.6%。目前,高炉现场操作者主要通过人工采样后离线化验的方式获取硅含量值,但在撇渣器处取铁水样本的过程具有一定的危险性,化验过程需要昂贵的仪器和人工成本,且化验的数据不具有时效性,使得铁水质量信息不能及时反馈,造成高炉铁水质量精细化调控困难。因此,实现铁水硅含量的实时在线预测,对评估高炉状态、提升铁水质量、降低高炉能耗和保障高炉顺行具有重要的意义。高炉炼铁过程具有强非线性、非稳态、大时滞等特点,涉及气、液、固三相及其耦合,是一个极其复杂的物理化学反应过程,因此难以建立精确的机理模型描述冶炼过程。通过简化部分条件建立的机理模型,在一定程度上为认知冶炼过程及炉内现象起到了积极作用,但建模条件要求严格,模型所需参数无法准确获取,建立的静态机理模型无法应用于原燃料品位波动频繁的国内高炉炼铁过程^[8-10]。随着集散控制系统和工业互联网成功运用以来,高炉冶炼系统已积累了海量能反映冶炼过程的知识和数据,且基于数据驱动的方法不需要详细准确的先验知识,仅仅依靠过程运行中的大量数据就能实现对铁水硅含量的在线实时预测,因此,基于数据驱动的铁水硅含量在线预测方法已经成为当前研究领域的热点方向^[11-13]。

随着人工智能技术的快速发展和应用,大量的基于数据驱动的模式被用来在线预报铁水硅含量,如支持向量回归机(Support vector regression, SVR)^[14]、神经网络^[15]、T-S(Takagi-Sugeno)模糊模型^[16]等。文献[14]基于非线性自回归模型提出了一

种鲁棒的最小二乘支持向量机模型,同时预测多个铁水质量性能指标。文献[15]融合多次有放回重采样下建立的神经网络模型,实现了硅含量数值和可信度二维预报。文献[16]提出了一种基于贝叶斯块结构稀疏的T-S模糊模型预测高炉铁水硅含量。尽管这些模型在高炉冶炼过程的关键性能指标的预测方面都取得了积极的进展,但还有一些问题没有被解决,仍有进一步提高的空间。首先,这些模型大都是浅层学习模型,其表达能力和泛化能力有限,难以以为复杂的高炉冶炼过程提供深层次的非线性描述,在非平稳炉况时模型的性能波动较大;其次,这些模型都是监督学习模式,对铁水硅含量标签数据的依赖较强,硅含量标签以人工采样、离线化验的方式获取,使得采集大量带标签的样本是困难的。

深度网络能在一定程度上解决上述铁水硅含量在线预测面临的两个问题。首先,通过堆叠多层隐含层使得模型能逐层地提取抽象的特征表示,来描述复杂的高炉冶炼过程;其次,模型的构建分为无监督的预训练和有监督的微调两阶段,先利用无标签的快速采样过程变量数据预训练深度网络,可以更好地挖掘数据中隐含分布特征,进而降低对标签硅含量数据依赖^[17]。由于深度网络的特点,在很多领域都取得了出色效果,如自然语言处理、文本识别和计算机视觉^[18-20]等。在工业过程中,深度学习也被引入工业过程关键性能指标在线估计。文献[21]利用深度网络,预测高炉铁水硅含量短期内的变化趋势。文献[22]堆叠多个受限玻尔兹曼机构成深度网络,实现了美国田纳西州伊士曼化工过程状态分类。深度网络建模优势的引入,使得工业过程的关键性能指标预测的精度取得了一定的提升。但是,直接建立基于深度网络的铁水硅含量在线预测模型,还存在一定的问题。首先,训练好的深度网络无差别地逐层提取输入样本的各维过程变量抽象特征表示,来完成硅含量在线预测任务,实际对于高炉冶炼过程,特别是国内入炉矿源频繁变动的情况,影响硅含量的过程变量的重要性呈现的是一种动态的变化规律,因此深度网络静态的建模思路无法完整地描述动态的高炉冶炼过程;其次,基于数据驱动的深度黑箱模型虽然展示了较佳的性能,但建模的过程难以被理解和信任,且模型缺乏一定的可解释性;再次,深度模型的训练对标签的硅含量数据量有一定的要求,而高成本的化验过程导致获取实时的硅含量标签样本是困难的。基于上述考虑,本文提出了基于动态注意力深度迁移网络(Attention deep transfer network, ADTNet)的高炉铁水硅含量在线预测方法。首先,提出了一种轻量化的动态注意力机制模块,将该模块嵌入在深度去噪自编码器(Denoising autoencoders, DAE)网络的前端,能实

时地计算输入样本的各过程变量注意力分数, 使得深度网络能有差别地动态提取与硅含量相关的过程变量的抽象特征, 且该模块还能给出对应的过程变量对硅含量预测值的贡献度, 在一定程度上增强黑箱模型的可解释性. 此外, 为了减少模型对硅含量标签数据的依赖, 本文介绍和引入了前期工作中研发的高炉铁口铁水温度红外视觉检测系统, 迁移该系统获取的实时的铁水温度数据训练好的深度网络模型到铁水硅含量在线预测任务中, 进一步提高铁水硅含量在线预测的性能^[23-24].

1 高炉冶炼过程及问题描述

1.1 过程描述

高炉炼铁过程如图 1 所示, 固体燃料(焦炭、煤粉等)、含铁原料(烧结矿、球团矿和块矿)和溶剂(白云石、石灰石和锰矿等)按一定配比从高炉炉顶分批下料, 同时由热风炉产生高温热风, 并通过高炉四周的风口平台向高炉内部鼓入高温热风及煤粉, 与焦炭发生燃烧反应生成一氧化碳和氢气等高温还原性气体. 下降的炉料和上升高温煤气流相遇, 先后发生传热、还原、融化、脱碳反应后生成生铁, 含铁矿石中的杂质和一些溶剂发生反应生成炉渣, 最终下落到炉缸区与铁水混合, 铁水从铁口排出流经铁口, 最终在撇渣器处实现渣铁分离.

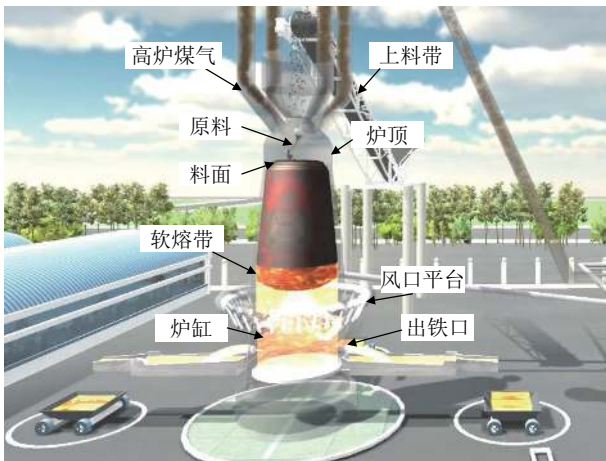


图 1 高炉三维仿真模拟图

Fig.1 Three-dimensional simulation diagram of the blast furnace cast field

1.2 问题描述

铁水硅含量是评价高炉冶炼过程的关键性能指标, 能在一定程度上表征炉内热状态和铁水质量. 炉温过高, 炉内透气性下降, 炉缸煤气流太过旺盛, 导致悬料、崩料等炉况故障及原材料的浪费; 炉温

过低, 铁水物理热不足, 炉内铁元素不能被充分还原, 冶炼效率降低. 为了保证高炉的顺行和铁水的质量, 操作者们需要对铁水硅含量进行实时在线的监控. 高炉系统可以由一个非线性模型描述, 其表达式为:

$$y^t = F(x^t) \quad (1)$$

式中, t 表示采样时间, $x^t \in \mathbf{R}^{d_x}$ 表示在采样时刻 t 的输入向量, y^t 表示在采样时刻 t 的输出, 即铁水硅含量的值. 因此, 铁水硅含量的在线检测等同于构建一个基于数据驱动的预测模型 F' , 使得模型 F' 能够尽量地逼近非线性的高炉系统 F 和尽可能准确地在线预测铁水硅含量. 为此, 本文提出了一种基于动态注意力机制的深度网络 (Attention deep network, ADNet) F' 来构建铁水硅含量预测模型, 通过引入动态注意力机制提高网络逐层抽取的特征的质量和增强黑箱模型的可解释性, 并且迁移红外视觉铁水测温系统检测的分钟级铁水温度数据训练的模型, 来进一步提高铁水硅含量预测模型 F' 的性能.

2 基于动态注意力机制模块的深度网络

本节主要介绍深度网络的一种基本结构单元——去噪自编码器, 考虑到入炉矿源品质的波动和冶炼条件的改变对预测目标铁水硅含量造成的动态影响, 提出一种动态注意力机制模块, 来描述动态的和非线性的高炉冶炼过程, 进而提高硅含量在线预测精度.

2.1 去噪自编码器

深度网络通过从低层到高层将数据逐层抽象, 进而自学习到原始数据的本质特征以适应于各种复杂的任务. 去噪自编码器是一种堆叠深度网络的基本单元, 是一种典型的 3 层无监督神经网络, 由输入层、隐含层和输出层组成^[25-26], 结构如图 2 所示. 训练目标是让输出尽可能地复现输入, 为了防止模型简单地用输出复制输入, 训练过程中, 隐含层神经元数量一般小于输入层神经元数量, 迫使隐含层神经元学习到输入的压缩抽象特征表示. 为了提取更鲁棒性的抽象特征表示, 进而更出色地完成去噪任务, 去噪自编码器在输入样本中, 加入了随机噪声进行干扰.

不失一般性, 假设模型输入为 $\mathbf{X} = [x^1, x^2, \dots, x^t]^T$, 其中 $x^t = [x_1^t, x_2^t, \dots, x_{d_x}^t]^T \in \mathbf{R}^{d_x}$, d_x 是样本的维度, 加入随机噪声污染后的第 t 个输入为 $\tilde{x}^t = [\tilde{x}_1^t, \tilde{x}_2^t, \dots, \tilde{x}_{d_x}^t]^T \in \mathbf{R}^{d_x}$. 对第 t 个污染后的输入向量 \tilde{x}^t 进行编码可以得到隐含层特征 $h^t = [h_1^t, h_2^t, \dots, h_{d_h}^t]^T \in \mathbf{R}^{d_h}$, 编码函数 f_θ 如下:

$$h^t = f_\theta(\tilde{x}^t) = f(\mathbf{W}\tilde{x}^t + \mathbf{b}) \quad (2)$$

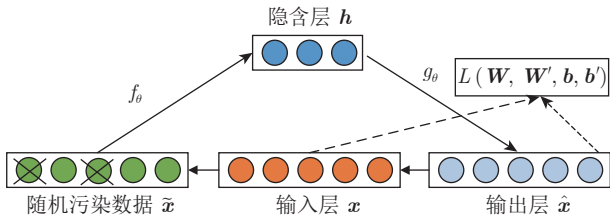


图 2 去噪自编码器基本结构

Fig.2 Architecture of a denoising autoencoder

式中, f 是隐含层的激活函数, \mathbf{W} 是大小为 $d_h \times d_x$ 的权值矩阵, $\mathbf{b} \in \mathbf{R}^{d_h}$ 是隐含层的偏置向量. 隐含层特征 \mathbf{h}^t 通过解码函数映射到重构特征 $\hat{\mathbf{x}}^t = [\hat{x}_1^t, \hat{x}_2^t, \dots, \hat{x}_{d_x}^t]^T \in \mathbf{R}^{d_x}$, 解码函数 g_θ 如下:

$$\hat{\mathbf{x}}^t = g_\theta(\mathbf{h}^t) = f(\mathbf{W}'\mathbf{h}^t + \mathbf{b}') \quad (3)$$

式中, \mathbf{W}' 和 \mathbf{W} 互为转置矩阵, $\mathbf{b}' \in \mathbf{R}^{d_x}$ 是输出层的偏置向量. 去噪自编码器通过梯度下降算法不断最小化目标函数来反复调整参数组 $\theta = (\mathbf{W}, \mathbf{W}', \mathbf{b}, \mathbf{b}')$, 其目标函数为:

$$L(\mathbf{W}, \mathbf{W}', \mathbf{b}, \mathbf{b}') = \frac{1}{2N} \sum_{t=1}^N \|\hat{\mathbf{x}}^t - \mathbf{x}^t\|^2 = \frac{1}{2N} \sum_{t=1}^N \sum_{d=1}^{d_x} (x_d^t - \hat{x}_d^t)^2 \quad (4)$$

考虑到单个去噪自编码器隐含层学习到的是原始数据的浅层特征, 为了得到更加抽象和鲁棒性的特征表示, 来提高硅含量在线预测精度, 可以通过堆叠多个去噪自编码器来搭建深度网络, 有多个非线性隐藏层的深度网络可以学习更复杂的输入样本和预测目标硅含量之间的关系, 从而在高层得到更具抽象和鲁棒的特征表示. 堆叠多个去噪自编码器过程如图 3 所示, 当第 1 个去噪自编码器 (DAE 1) 训练完成后, 其隐含层的特征表示 $\mathbf{H}_1 = [\mathbf{h}_1^1, \mathbf{h}_1^2, \dots,$

$\mathbf{h}_1^t]^T$, 其中 $\mathbf{h}_1^t = [h_1^t, h_2^t, \dots, h_{d_1}^t]^T \in \mathbf{R}^{d_1}$ 经过随机污染后作为第 2 个去噪自编码器 (DAE 2) 的输入, 通过编码和解码函数得到网络的输出矩阵 $\hat{\mathbf{H}}_1 = [\hat{\mathbf{h}}_1^1, \hat{\mathbf{h}}_1^2, \dots, \hat{\mathbf{h}}_1^t]^T$, 其中 $\hat{\mathbf{h}}_1^t = [\hat{h}_1^t, \hat{h}_2^t, \dots, \hat{h}_{d_1}^t]^T \in \mathbf{R}^{d_1}$, 则第 2 个去噪自编码器的误差函数为:

$$L^2(\mathbf{W}, \mathbf{b}) = \frac{1}{2N} \sum_{t=1}^N \sum_{d=1}^{d_1} (h_d^t - \hat{h}_d^t)^2 \quad (5)$$

通过梯度下降算法, 最小化误差函数训练模型的参数, 重复这个过程, 直到第 N 个去噪自编码器 (DAE N) 训练完成. 无监督的预训练过程完成后, 将 N 个训练好的去噪自编码器的隐含层权值和偏置矩阵取出来, 堆叠成一个深度网络, 在最后隐含层后面加上回归层并随机初始化网络回归层的权值和偏置矩阵, 使用带标签的硅含量数据集有监督地微调整个网络参数, 进而更高效地完成硅含量在线预测任务.

2.2 动态注意力机制模块

预训练好的深度网络使用带标签的铁水硅含量样本最小化损失函数, 微调整个网络结构来提高硅含量在线预测的性能, 有监督微调的损失函数如下:

$$L_{loss} = \frac{1}{2N_{Si}} \sum_{t=1}^{N_{Si}} (y_{Si}^t - \hat{y}_{Si}^t)^2 = \frac{1}{2N_{Si}} \sum_{t=1}^{N_{Si}} \left(y_{Si}^t - f^{(N+1)}(\mathbf{W}^{(N+1)}(\dots f^{(2)}(\mathbf{W}^{(2)}(f^{(1)}(\mathbf{W}^{(1)}\mathbf{x}_{Si}^t + \mathbf{b}^{(1)} + \mathbf{b}^{(2)})) \dots + \mathbf{b}^{(N+1)})) \right)^2 \quad (6)$$

式中, y_{Si}^t 为预测目标真实值, \hat{y}_{Si}^t 为模型的预测值, $f^{(N+1)}$ 为堆叠去噪自编码器 (Stacked DAE, S-DAE)

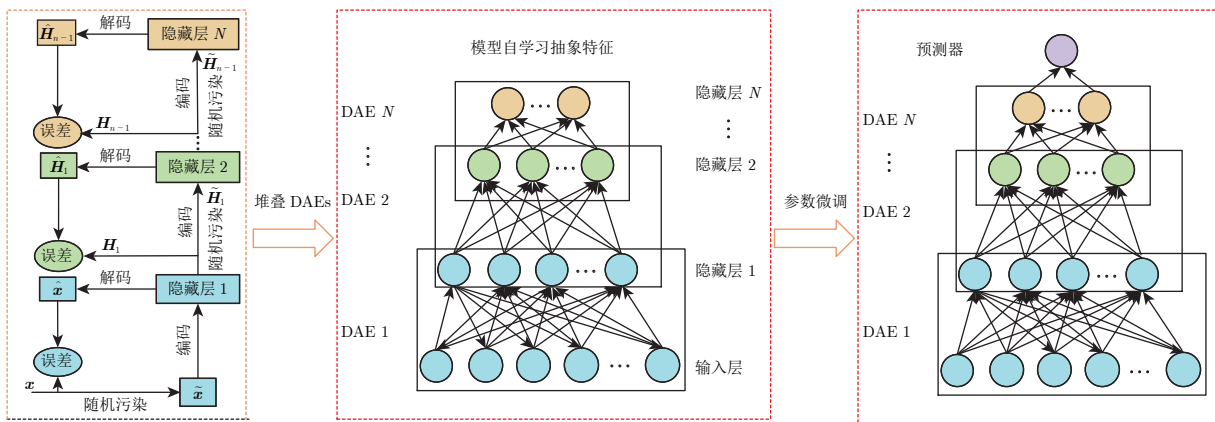


图 3 堆叠去噪自编码器训练过程

Fig.3 The training process of stacking denoising autoencoders

第 $(N+1)$ 层神经元的非线性激活函数, $\mathbf{W}^{(N+1)}$ 和 $\mathbf{b}^{(N+1)}$ 分别为第 $(N+1)$ 层神经元与前一层神经元之间的权值矩阵与偏置矩阵. 由式 (6) 可知, 堆叠去噪自编码器形成的深度网络在训练时, 对输入样本的每一维过程变量都给予了相同的关注度, 使得模型前向传播时, 无差别地提取抽象特征表示. 但是在高炉冶炼过程中, 输入样本的过程变量与铁水硅含量相关系数是不一样的, 并且随着入炉矿源品质的波动和冶炼条件的改变, 影响铁水硅含量的过程变量的重要性, 随时间呈现出一种动态变化的趋势. 在工程应用技术中, 一般是采取先降维后建模的思路, 通过变量相关性分析确定主要过程变量, 忽略次要过程变量对铁水硅含量的影响, 且主要过程变量的重要性在建模的过程中, 并不会得到体现. 显然, 传统的深度的静态建模思路无法准确地描述高炉冶炼过程的动态特性, 且无差别地提取抽象特征无法准确地表征输入变量与铁水硅含量之间的非线性关系. 因此, 本文提出了一种动态的注意力机制模块, 能实时地为每个输入样本的过程变量计算动态的注意力分数, 使得模型能动态地为每个样本中有效的和有价值的过程变量分配更多的注意力, 进而更高效地完成铁水硅含量的预测任务.

为了准确地描述完整的动态高炉冶炼过程, 提出的动态注意力机制模块必须满足两个基本准则: 1) 能够学习样本的过程变量和铁水硅含量之间的非线性关系; 2) 能够描述样本的过程变量和铁水硅含量之间的动态关系. 基于此设计的动态注意力机制模块, 由注意力得分模块和注意力聚焦模块两部分构成, 其目标是通过挖掘样本过程变量与铁水硅含量之间的动态关系, 从而提高深度网络自学习的抽象特征质量, 基本结构如图 4 所示.

注意力得分模块由多个全连接层 (Fully connected layers, FC) 构成, 隐含层和输出层神经元中非线性激活函数的存在, 使得网络能模拟高炉冶炼过程的复杂非线性特点. 注意力聚焦模块由注意力得分模块计算的输入样本各过程变量的注意力分数, 与之对应的输入样本各过程变量点乘, 进而描述样本各过程变量与铁水硅含量之间的动态关系. 动态注意力机制模块的数学表达过程描述如下, 具体地, 假设模块的输入向量为 $\mathbf{X}_{Si} = [\mathbf{x}_{Si}^1, \mathbf{x}_{Si}^2, \dots, \mathbf{x}_{Si}^t]^T$, 经过注意力得分模块计算后输出记为 ω_{Si} :

$$\omega_{Si} = f^{(M)} \left(\mathbf{W}^{(M)} \left(\dots f^{(2)} \left(\mathbf{W}^{(2)} \left(f^{(1)} \left(\mathbf{W}^{(1)} \mathbf{X}_{Si} + \mathbf{b}^{(1)} \right) + \mathbf{b}^{(2)} \right) \dots + \mathbf{b}^{(M)} \right) \right) \right) \quad (7)$$

式中, $f^{(M)}$ 为第 M 层神经元的非线性激活函数, $\mathbf{W}^{(M)}$ 和 $\mathbf{b}^{(M)}$ 分别为第 M 层神经元与前一层神经

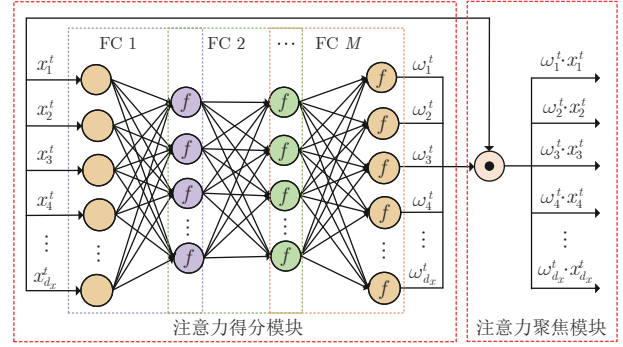


图 4 动态注意力机制模块

Fig. 4 The dynamic attention mechanism module

元之间的权值矩阵与偏置矩阵. 将 ω_{Si} 在每个样本的各维度上的过程变量上展开可得:

$$\omega_{Si} = [\omega_{Si}^1, \omega_{Si}^2, \dots, \omega_{Si}^t]^T = \begin{bmatrix} \omega_1^1 & \omega_2^1 & \dots & \omega_{d_x}^1 \\ \omega_1^2 & \omega_2^2 & \dots & \omega_{d_x}^2 \\ \vdots & \vdots & \ddots & \vdots \\ \omega_1^t & \omega_2^t & \dots & \omega_{d_x}^t \end{bmatrix} \quad (8)$$

为了描述样本的过程变量在不同时刻与铁水硅含量之间的动态关系, 将注意力得分模块计算得分矩阵 ω_{Si} 与输入变量 \mathbf{X}_{Si} 做哈达玛积操作:

$$\omega_{Si} \odot \mathbf{X}_{Si} = [\omega_{Si}^1, \omega_{Si}^2, \dots, \omega_{Si}^t]^T \odot \begin{bmatrix} \mathbf{x}_{Si}^1 & \mathbf{x}_{Si}^2 & \dots & \mathbf{x}_{Si}^t \end{bmatrix}^T = \begin{bmatrix} \omega_1^1 \cdot x_1^1 & \omega_2^1 \cdot x_2^1 & \dots & \omega_{d_x}^1 \cdot x_{d_x}^1 \\ \omega_1^2 \cdot x_1^2 & \omega_2^2 \cdot x_2^2 & \dots & \omega_{d_x}^2 \cdot x_{d_x}^2 \\ \vdots & \vdots & \ddots & \vdots \\ \omega_1^t \cdot x_1^t & \omega_2^t \cdot x_2^t & \dots & \omega_{d_x}^t \cdot x_{d_x}^t \end{bmatrix} \quad (9)$$

输入样本 \mathbf{X}_{Si} 与得分矩阵 ω_{Si} 对应位置元素点乘的矢量矩阵, 作为预训练好的堆叠去噪自编码器的输入, 基于动态注意力机制模块的深度去噪自编码器网络如图 5 所示, 整个网络输出可表示为:

$$\hat{\mathbf{Y}}_{Si} = f^{(N+1)} \left(\mathbf{W}^{(N+1)} \left(\dots f^{(2)} \left(\mathbf{W}^{(2)} \left(f^{(1)} \left(\mathbf{W}^{(1)} \left(\omega_{Si} \odot \mathbf{X}_{Si} \right) + \mathbf{b}^{(1)} \right) + \mathbf{b}^{(2)} \right) \dots + \mathbf{b}^{(N+1)} \right) \right) \right) \quad (10)$$

因此, 使用带标签的硅含量样本有监督地微调基于动态注意力机制的深度网络时的损失函数为:

$$L_{loss} = \frac{1}{2N_{Si}} \sum_{t=1}^{N_{Si}} (y_{Si}^t - \hat{y}_{Si}^t)^2 = \frac{1}{2N_{Si}} \sum_{t=1}^{N_{Si}} (y_{Si}^t - f^{(N+1)}(\mathbf{W}^{(N+1)}(\dots f^{(2)}(\mathbf{W}^{(2)}(f^{(1)}(\mathbf{W}^{(1)}(\omega_{Si}^t \odot \mathbf{x}_{Si}^t) + \mathbf{b}^{(1)}) + \mathbf{b}^{(2)})) \dots + \mathbf{b}^{(N+1)})))^2 \quad (11)$$

由式 (11) 可以看出, 网络在训练和预测的过程中, 考虑了每个输入样本的过程变量与预测目标之间的动态关系, 能有区别地为每个样本提取目标相关的抽象特征表示, 来描述复杂的高炉冶炼过程, 进而提高硅含量在线预报模型的性能. 基于动态注意力机制模块的深度去噪自编码器网络如图 5 所示. 注意力得分模块目的在于捕获输入样本的过程变量与铁水硅含量之间的相关关系, 并强调样本的各个过程变量对预测的硅含量值的动态贡献度. 注意力聚焦模块目的在于增强与铁水硅含量相关性较大的过程变量的影响, 抑制与铁水硅含量相关性较小的过程变量的影响. 设计的注意力机制模块是一个轻量化的结构单元, 能够在网络任意层之间嵌入, 从而提高网络自学习特征的质量. 且注意力得分模块能实时动态地给出每个样本各过程变量对硅含量的贡献度, 因此对黑箱模型具有一定的可解释性.

3 基于动态注意力机制模块的深度迁移网络

本节主要介绍基于动态注意力深度迁移网络的铁水硅含量在线预测模型, 为了降低硅含量预测模型训练过程中对标签数据的依赖, 分析了铁水温度

与硅含量之间的相关关系, 并介绍和引入了前期工作中研发的高炉铁口铁水温度红外视觉检测系统, 迁移了基于分钟级的铁水温度数据训练好的深度模型结构, 并使用少量的硅含量标签数据, 进一步微调网络参数进而提高硅含量在线预测的精度.

3.1 基于红外视觉的铁水温度检测系统

通过在线同步采集中国西南某炼铁厂 2650 m³ 的 2 号高炉撇渣器处热电偶检测铁水温度和硅含量化验数据, 绘制了如图 6 所示的铁水温度与铁水硅含量之间的散点分布图. 通过分析可知, 铁水温度与硅含量之间存在一定的正相关性, 建模时加入铁水温度数据, 对提升硅含量在线预测模型的准确率有一定的帮助. 而高炉现场主要采用快速热电偶来检测高炉撇渣器处的铁水温度, 一次出铁周期只能获取有限个数据, 导致铁水温度数据的质量和数量得不到保证. 在前期工作中, 本文原创研发了铁水测温仪和高炉出铁口铁水温度红外视觉检测系统, 能实时在线检测高炉出铁口铁水温度, 这为建立硅含量在线预报模型提供了独有的数据来源^[23-24].

研制的铁水测温仪结构如图 7(a) 所示, 在中国西南某炼铁厂 2650 m³ 的 2 号高炉的 1 号、2 号和 3 号出铁口进行了安装, 并能长期稳定工作. 基于硬件设备, 搭建了由铁水测温仪、防护装置、专用光缆、三维云台、控制柜和计算机等构成的红外视觉铁水温度检测系统, 开发了大型高炉铁水质量监控平台, 实现了高炉出铁口铁水温度的在线精确检测. 实时的铁水温度数据能为硅含量的在线检测模型提供更关键和全面的信息来源.

3.2 基于动态注意力深度迁移网络的铁水硅含量在线预测模型

前期工作研发的铁水温度红外视觉检测系统,

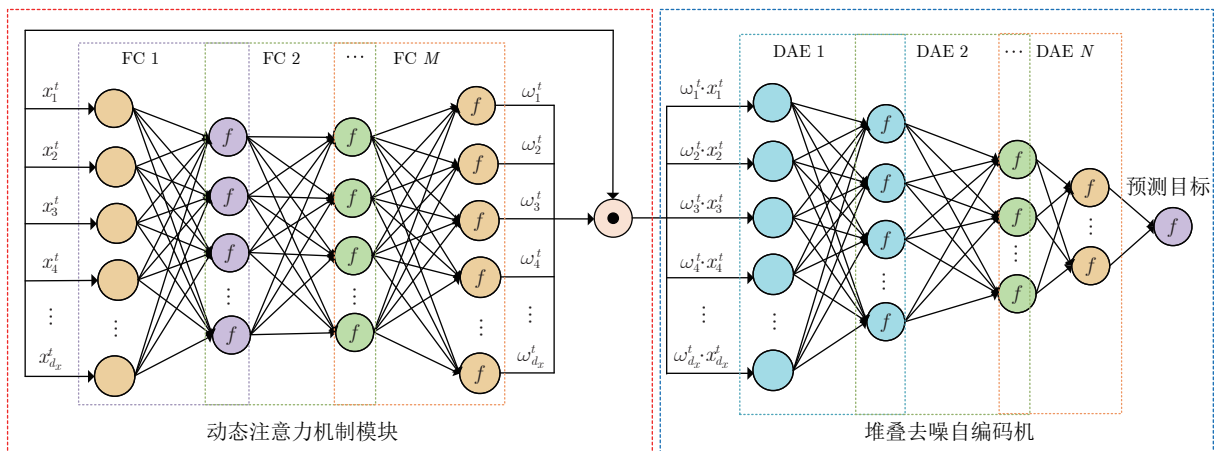


图 5 基于动态注意力机制模块的深度去噪自编码器网络

Fig. 5 Deep denoising autoencoders network based on dynamic attention mechanism module

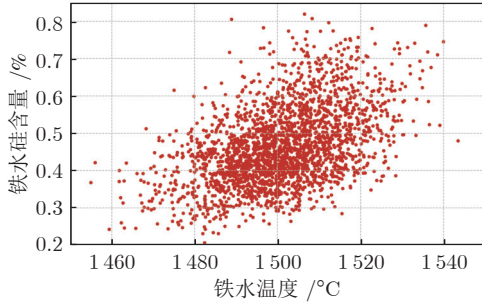


图 6 铁水温度与铁水硅含量的散点图

Fig.6 The scatter plot of temperature and silicon content of molten iron

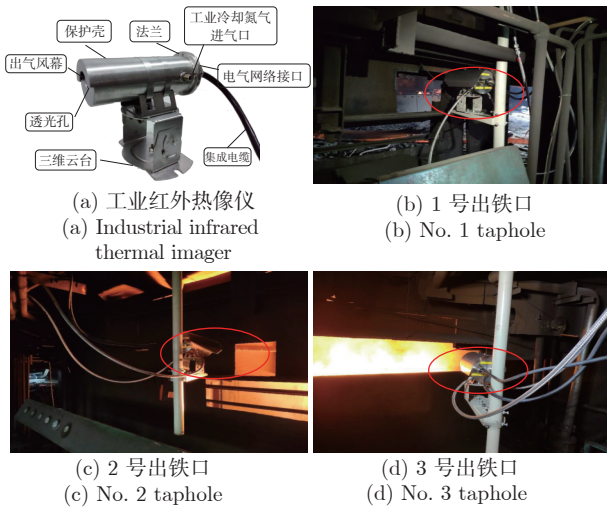


图 7 高炉铁水测温系统

Fig.7 Molten iron temperature measuring system in a blast furnace

能实现秒级出铁口温度数据的检测,但考虑到图像数据的存储空间的限制和铁水温度渐近式变化特点,现场数据库中保存了分钟级铁水温度的数据.高炉现场安装的过程变量检测传感器的采样频率约为 10 s 一次.而现场对铁水硅含量的检测主要是通过人工定期抽样,离线化验分析,在一个班次(8 h)内约有 12~16 组数据.因此,数据库中有大量能反映铁水硅含量的过程变量数据,而对应的需要高成本标注的硅含量数据是明显不足的.由图 6 可知,铁水温度与铁水硅含量之间存在一定的正相关性,因此在建模过程中,铁水温度数据作为一维主要过程变量的加入,会在一定程度上提高模型的预测性能.但为了数据集时间轴上的匹配,分钟级采样的铁水温度数据需要经过相关处理,来匹配小时级化验的铁水硅含量数据,这样会导致大量的铁水温度数据被压缩或丢失.

考虑到红外视觉铁水测温系统和过程变量采样

频率的相对快速性,为了充分利用铁水温度数据中的信息,通过构建分钟级优质数据集训练铁水温度在线预测模型,再将模型上学习到的知识迁移到铁水硅含量在线预测任务中,降低硅含量模型训练时对标签数据的依赖.由于铁水温度与铁水硅含量之间的正相关性和反映铁水温度和硅含量的过程变量的相对一致性,通过迁移从已训练的铁水温度在线预测任务中学习的知识,来提升铁水硅含量在线预测任务是可行的,且建模过程中不存在很大的源领域和目标域漂移的问题.基于深度迁移学习的铁水硅含量在线预报模型结构如图 8 所示,需要注意的是,为了让铁水温度预测模型学习到的参数能更好地适配铁水硅含量在线预测任务,动态注意力机制模块中输入是能反映铁水温度和硅含量的共同过程变量.唯一的区别是,铁水温度模型数据是分钟级别采样的,而铁水硅含量模型的数据是小时级别采样的.为进一步提高模型的性能,在基于堆叠去噪自编码机构建的铁水硅含量深度预测模型的输入层,拼接了对硅含量值有影响的特有过程变量.基于动态注意力深度迁移网络的铁水硅含量在线预报过程具体步骤如下:

步骤 1. 根据专家经验和高炉冶炼过程机理分析,从高炉历史数据库中挑选出相关的数据,用于分别构建铁水温度和铁水硅含量在线预报模型.

步骤 2. 对采集的数据进行相关数据预处理,包括输入输出样本时间配准、异常值剔除、缺失值处理、归一化处理、变量相关性分析.预处理后的铁水温度数据集记为 $\{\mathbf{X}_{Fe}, \mathbf{Y}_{Fe}\} = \{(\mathbf{x}_{Fe}^1, \mathbf{y}_{Fe}^1), (\mathbf{x}_{Fe}^2, \mathbf{y}_{Fe}^2), \dots, (\mathbf{x}_{Fe}^t, \mathbf{y}_{Fe}^t), \dots, (\mathbf{x}_{Fe}^{N_{Fe}}, \mathbf{y}_{Fe}^{N_{Fe}})\}$, 铁水硅含量数据集表示为 $\{\mathbf{X}_{Si}, \mathbf{Y}_{Si}\} = \{(\mathbf{x}_{Si}^1, \mathbf{y}_{Si}^1), (\mathbf{x}_{Si}^2, \mathbf{y}_{Si}^2), \dots, (\mathbf{x}_{Si}^t, \mathbf{y}_{Si}^t), \dots, (\mathbf{x}_{Si}^{N_{Si}}, \mathbf{y}_{Si}^{N_{Si}})\}$, 其中 N_{Fe} 和 N_{Si} 分别为铁水温度数据集和铁水硅含量数据集中样本的个数.

步骤 3. 用数据 $\mathbf{X}_{Fe} = [\mathbf{x}_{Fe}^1, \mathbf{x}_{Fe}^2, \dots, \mathbf{x}_{Fe}^t, \dots, \mathbf{x}_{Fe}^{N_{Fe}}]^T$ 无监督地预训练第 1 个去噪自编码器网络,采用误差反向传播算法最小化误差函数 $\sum_{t=1}^{N_{Fe}} \times \sum_{d=1}^{d_x} ((x_d^t - \hat{x}_d^t)^2 / (2N_{Fe}))$ 并保存训练好的隐含层的权值和偏置矩阵 $[\mathbf{W}^1, \mathbf{b}^1]$. 将第 1 个去噪自编码器的隐含层输出 $\mathbf{H}_1 = [h_1^1, h_1^2, \dots, h_1^t, \dots, h_1^{N_{Fe}}]^T$, 其中 $\mathbf{h}_1^t = [h_1^t, h_2^t, \dots, h_{d_1}^t]^T \in \mathbf{R}^{d_1}$ 作为第 2 个去噪自动编码器的输入,利用误差反向传播算法最小化误差函数 $\sum_{t=1}^{N_{Fe}} \sum_{d=1}^{d_1} ((h_d^t - \hat{h}_d^t)^2 / (2N_{Fe}))$ 并保存训练好的隐含层的权值和偏置矩阵 $[\mathbf{W}^2, \mathbf{b}^2]$. 重复上述步骤,直到第 N 个去噪自编码器训练完成,保存权值矩阵 $\mathbf{W}^{deep} = [\mathbf{W}^1, \mathbf{W}^2, \dots, \mathbf{W}^N]^T$ 和偏置矩阵 $\mathbf{b}^{deep} = [\mathbf{b}^1, \mathbf{b}^2, \dots, \mathbf{b}^N]^T$.

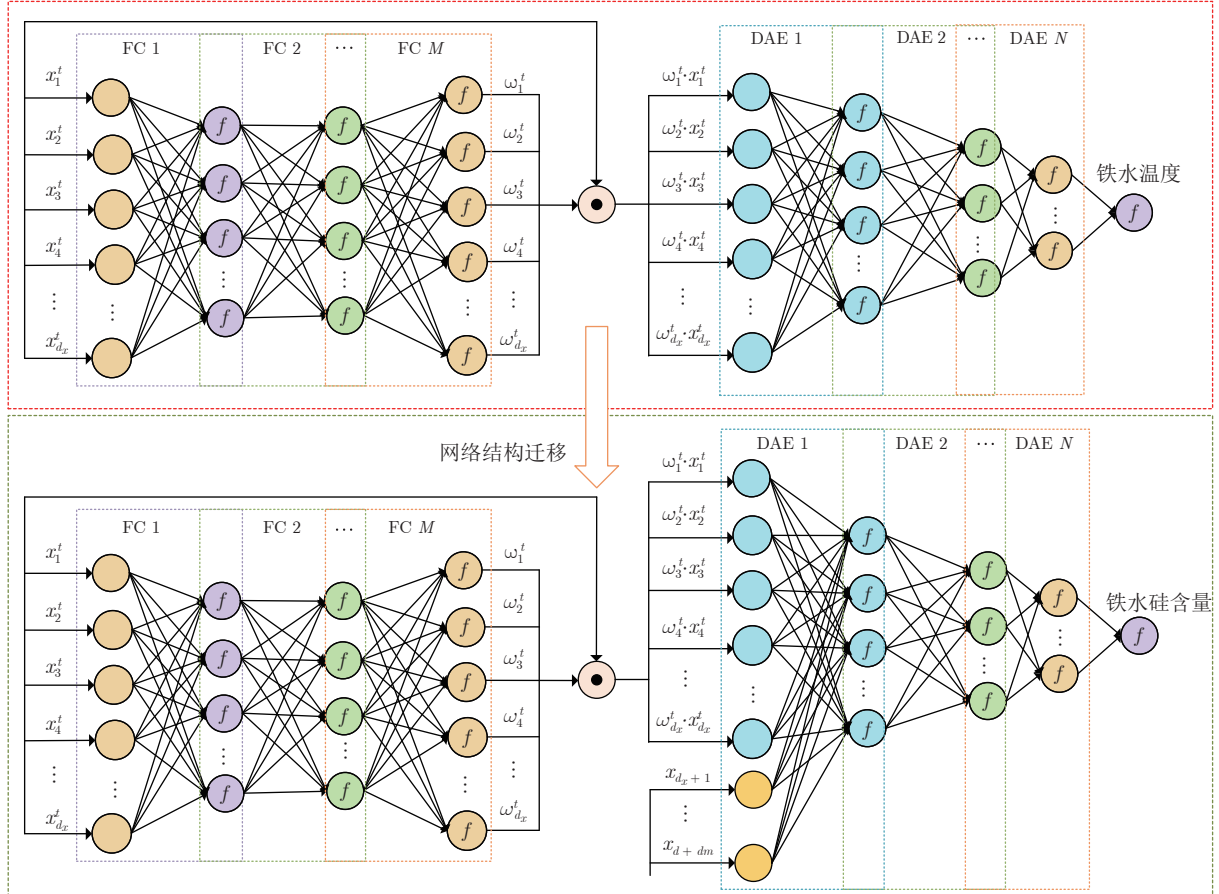


图 8 基于深度迁移网络的铁水硅含量在线预报模型

Fig.8 Online prediction model of silicon content in molten iron based on deep transfer network

步骤 4. 将 N 个训练好的去噪自编码器的隐含层权值和偏置矩阵取出来, 堆叠成一个深度网络, 在预训练好的深度网络的前端嵌入动态注意力机制模块, 并随机初始化模块权值矩阵 $\mathbf{W}^{attention} = [\mathbf{W}^1, \mathbf{W}^2, \dots, \mathbf{W}^M]^T$ 和偏置矩阵 $\mathbf{b}^{attention} = [\mathbf{b}^1, \mathbf{b}^2, \dots, \mathbf{b}^M]^T$, 把铁水温度数据 $\mathbf{X}_{Fe} = [\mathbf{x}_{Fe}^1, \mathbf{x}_{Fe}^2, \dots, \mathbf{x}_{Fe}^{t_{Fe}}, \dots, \mathbf{x}_{Fe}^{N_{Fe}}]^T$ 输入到动态注意力机制模块中得到各样本过程变量的注意力得分矩阵 ω_{Fe} , 将 \mathbf{X}_{Fe} 与得分矩阵 ω_{Fe} 做哈达玛积操作, 得到预训练好的深度网络的输入为 $\mathbf{X}_{Fe} = [\omega_{Fe}^1 \odot \mathbf{x}_{Fe}^1, \omega_{Fe}^2 \odot \mathbf{x}_{Fe}^2, \dots, \omega_{Fe}^t \odot \mathbf{x}_{Fe}^t, \dots, \omega_{Fe}^{N_{Fe}} \odot \mathbf{x}_{Fe}^{N_{Fe}}]^T$, 其中 $\omega_{Fe}^t \odot \mathbf{x}_{Fe}^t = [\omega_{Fe}^t x_{Fe}^t, \omega_{Fe}^t x_{Fe}^t, \dots, \omega_{Fe}^t x_{Fe}^t]^T$.

步骤 5. 在嵌入动态注意力机制模块的深度网络的后端, 即最后一个隐含层上再加一层输出层并随机初始化参数 $[\mathbf{W}^{N+1}, \mathbf{b}^{N+1}]$, 输入铁水温度数据 $\{\mathbf{X}_{Fe}, \mathbf{Y}_{Fe}\} = \{(\mathbf{x}_{Fe}^1, \mathbf{y}_{Fe}^1), (\mathbf{x}_{Fe}^2, \mathbf{y}_{Fe}^2), \dots, (\mathbf{x}_{Fe}^{t_{Fe}}, \mathbf{y}_{Fe}^{t_{Fe}}), \dots, (\mathbf{x}_{Fe}^{N_{Fe}}, \mathbf{y}_{Fe}^{N_{Fe}})\}$ 利用误差反向传播算法最小化误差函数 $\sum_{t=1}^{N_{Fe}} ((y_{Fe}^t - \hat{y}_{Fe}^t)^2 / (2N_{Fe}))$ 并调整整个网络结构参数 $\mathbf{W} = [\mathbf{W}^{attention}, \mathbf{W}^{deep}]$ 和 $\mathbf{b} = [\mathbf{b}^{attention}, \mathbf{b}^{deep}]$.

步骤 6. 将训练好的铁水温度模型迁移到铁水硅含量在线预报模型上, 把 $\mathbf{X}_{Si} = [\mathbf{x}_{Si}^1, \mathbf{x}_{Si}^2, \dots, \mathbf{x}_{Si}^t, \dots, \mathbf{x}_{Si}^{N_{Si}}]^T$ 输入到动态注意力机制模块中, 得到深度网络输入为 $\mathbf{X}_{Si} = [\omega_{Si}^1 \odot \mathbf{x}_{Si}^1, \omega_{Si}^2 \odot \mathbf{x}_{Si}^2, \dots, \omega_{Si}^t \odot \mathbf{x}_{Si}^t, \dots, \omega_{Si}^{N_{Si}} \odot \mathbf{x}_{Si}^{N_{Si}}]^T$, 并在深度网络的输入层拼接了对铁水硅含量有影响的特有的数据 $\hat{\mathbf{X}}_{Si} = [\hat{\mathbf{x}}_{Si}^1, \hat{\mathbf{x}}_{Si}^2, \dots, \hat{\mathbf{x}}_{Si}^t, \dots, \hat{\mathbf{x}}_{Si}^{N_{Si}}]^T$, 其中 $\hat{\mathbf{x}}_{Si}^t = [x_{dx+1}^t, x_{dx+2}^t, \dots, x_{dx+dm}^t]^T \in \mathbf{R}^{d_m}$, 利用误差反向传播算法最小化误差函数 $\sum_{t=1}^{N_{Si}} ((y_{Si}^t - \hat{y}_{Si}^t)^2 / (2N_{Si}))$ 并微调网络结构参数, 完成对基于深度迁移网络的铁水硅含量在线预报模型的训练.

步骤 7. 将与特有变量拼接好的测试样本 $\mathbf{X}_{Test} = [\mathbf{x}_{Test}^1, \mathbf{x}_{Test}^2, \dots, \mathbf{x}_{Test}^t, \dots, \mathbf{x}_{Test}^{N_{Test}}]^T$ 输入到训练好的铁水硅含量预测模型中, 其中 $\mathbf{x}_{Test}^t = [x_1^t, x_2^t, \dots, x_{dx}^t, x_{dx+1}^t, x_{dx+2}^t, \dots, x_{dx+dm}^t]^T \in \mathbf{R}^{d_x+d_m}$, N_{Test} 是测试集样本的数量. 把测试样本 $1 \sim d_x$ 维度的过程变量输入到动态注意力机制模块, 把 $(d_x + 1) \sim (d_x + d_m)$ 维度的过程变量作为硅含量独有的特征并拼接深度网络的输入层, 训练好的网络最后一层输出为铁水硅含量的预测结果, 记为

$$\hat{Y}_{Test} = [\hat{y}_{Test}^1, \hat{y}_{Test}^2, \dots, \hat{y}_{Test}^t, \dots, \hat{y}_{Test}^{N_{Test}}]^T.$$

4 工业验证

为了验证基于动态注意力深度迁移网络的铁水硅含量在线预报模型的有效性, 将本文提出算法在中国西南某炼铁厂 2 号高炉上采集的数据进行工业试验. 实验结果验证了本文方法的有效性和可行性, 模型的预测结果和样本过程变量对硅含量的贡献度, 能为现场操作者提供实时的冶炼状态信息和更为直观的操作指导.

4.1 过程数据分析和预处理

高炉历史数据库中记录了大量能反映铁水温度和硅含量的传感器检测数据, 根据高炉的冶炼工艺机理、已安装的传感器检测设备和现场专家经验, 确定了对铁水温度和硅含量有影响的过程变量, 详细描述见表 1. 提取 2020 年 8 月 1 日至 2020 年 12 月 17 日数据用于建模分析, 其中过程变量有 1160141 组, 铁水温度数据有 172352 组, 硅含量数据有 7282 组. 考虑到数据库中的数据会因为设备故障或者人工录入错误等原因出现错误, 或高炉休风等特殊情造成数据缺失, 也会因为现场冶炼过程的干扰使得测量数据出现大量的噪声. 因此在建模前, 需要对数据进行相关预处理, 得到标准的、干净的和连续的数据提供给后续模型.

表 1 过程变量最大互信息系数

Table 1 Maximal information coefficient of process variables

过程变量	MIC _{MIT}	MIC _{Si}	过程变量	MIC _{MIT}	MIC _{Si}
富氧率	0.104	0.115	冷风压力	0.104	0.094
透气性指数	0.104	0.111	全压差	0.104	0.130
CO	0.103	0.104	热风压力	0.103	0.116
CO ₂	0.111	0.145	实际风速	0.100	0.113
标准风速	0.117	0.111	冷风温度	0.102	0.109
富氧流量	0.120	0.129	热风温度	0.101	0.115
冷风流量	0.117	0.111	顶温	0.120	0.162
鼓风动能	0.101	0.107	顶温下降管	0.111	0.155
炉腹煤气量	0.108	0.127	阻力系数	0.103	0.110
炉腹煤气指数	0.109	0.128	鼓风湿度	0.135	0.140
顶压	0.128	0.156	富氧压力	0.103	0.096
本小时实际喷煤量	0.100	0.136	上一小时实际喷煤量	0.110	0.165

考虑到撇渣器处人工取样的铁水, 需要花费 1 h 左右时间化验铁水硅含量百分比值, 因此将数据库中记录的铁水硅含量数据往前推一个小时, 与传感

器检测的过程变量在时间维度上进行配准. 对于设备故障或者人工录入错误而导致的异常数据, 通过箱线图直接剔除. 直接删除休风和设备故障等原因造成的缺失数据. 为了构建完备的输入、输出样本集, 把经过数据预处理后的铁水温度数据集按每分钟取均值处理, 铁水硅含量的数据按每小时取均值处理. 经过处理后的铁水温度数据集共有 111041 组数据, 铁水硅含量数据集共有 3117 组数据. 针对数据集中样本的不同过程变量量纲存在较大的差异, 因此在建模前需要对数据进行归一化处理, 进而消除不同量纲对模型的影响, 本文采用式 (12) 对数据归一化处理:

$$\bar{x}_{d_x}^t = \frac{x_{d_x}^t - x_{d_x}^{\min}}{x_{d_x}^{\max} - x_{d_x}^{\min}} \quad (12)$$

式中, $\bar{x}_{d_x}^t$ 是第 t 个数据样本中的第 d_x 个过程变量归一化处理后的结果, $x_{d_x}^t$ 为第 t 个数据样本中第 d_x 个过程变量的值, $x_{d_x}^{\max}$ 和 $x_{d_x}^{\min}$ 分别为第 d_x 个过程变量在所有的数据样本中的最大值和最小值.

为了定量描述表 1 中通过专家经验挑选的过程变量与铁水温度和铁水硅含量之间的相关性, 本文采用最大互信息系数 (Maximal information coefficient, MIC)^[27] 来计算两个变量之间的相关性, MIC 定义如下:

$$\text{MIC} = \max_{|x|, |y| < B} \frac{\sum_{x, y} p(x, y) \log_2 \frac{p(x, y)}{p(x)p(y)}}{\log_2(\min(|x|, |y|))} \quad (13)$$

式中, $p(x)$ 为数据点落在 x 列的概率, $p(y)$ 为数据点落在 y 行的概率, $p(x, y)$ 为变量 x 与变量 y 联合概率, B 为变量最优推荐值, 即数据量的 0.6 次方. 各过程变量与铁水温度和硅含量的最大信息系数分别记为 MIC_{MIT} 和 MIC_{Si}, 如表 1 所示.

由表 1 可知, 通过专家经验和工艺机理分析确定的过程变量与铁水温度和铁水硅含量, 都有较强的相关性, 因此这 24 个过程变量作为铁水温度和铁水硅含量预测模型动态注意力机制模块的输入. 需要注意的是, 在铁水硅含量预测模型的深度网络输入层拼接了对铁水硅含量有影响的特有过程变量, 即: $q^{-1}x_{\text{硅含量}}^t$ 、 $q^{-2}x_{\text{硅含量}}^t$ 、 $q^{-3}x_{\text{硅含量}}^t$ 、 $x_{\text{红外铁水温度}}^t$ 、 $q^{-1}x_{\text{红外铁水温度}}^t$ 、 $q^{-2}x_{\text{红外铁水温度}}^t$ 、 $q^{-1}x_{\text{CO}_2}^t$ 、 $q^{-1}x_{\text{富氧流量}}^t$ 、 $q^{-1}x_{\text{顶压}}^t$ 、 $q^{-1}x_{\text{全压差}}^t$ 、 $q^{-2}x_{\text{顶温}}^t$ 、 $q^{-1}x_{\text{鼓风湿度}}^t$, 其中 $q^{-1}x^t = x^{t-1}$ 、 $q^{-2}x^t = x^{t-2}$ 和 $q^{-3}x^t = x^{t-3}$ 分别表示在第 $t-1$ 、 $t-2$ 、 $t-3$ 时刻过程变量 x 的取值, 这部分特有的过程变量主要考虑的是高炉冶炼过程中大惯性和时序性对铁水硅含量的影响.

需要注意的是,用归一化后的数据训练的模型输出的仍是归一化后的结果,因此需要对输出结果进行反归一化处理才能得到最终的实际预测结果.铁水温度和铁水硅含量的反归一化,分别如式(14)和式(15)所示:

$$\hat{y}_{Fe}^t = \bar{y}_{Fe}^t (y_{Fe}^{\max} - y_{Fe}^{\min}) + y_{Fe}^{\min} \quad (14)$$

$$\hat{y}_{Si}^t = \bar{y}_{Si}^t (y_{Si}^{\max} - y_{Si}^{\min}) + y_{Si}^{\min} \quad (15)$$

式中, \hat{y}_{Fe}^t 和 \hat{y}_{Si}^t 是铁水温度模型和铁水硅含量模型的反归一化后的真实输出结果, \bar{y}_{Fe}^t 和 \bar{y}_{Si}^t 分别是铁水温度模型和铁水硅含量模型的输出结果, y_{Fe}^{\max} 和 y_{Fe}^{\min} 分别是铁水温度真实数据中的最大值和最小值, y_{Si}^{\max} 和 y_{Si}^{\min} 分别是铁水硅含量真实数据中的最大值和最小值.

4.2 模型性能评价指标

为了评价模型的性能,通过引入均方根误差(Root mean squared error, RMSE)和平均绝对误差(Mean absolute error, MAE)来衡量模型的预测值与实际化验值之间的差异程度, RMSE 和 MAE 定义如下:

$$\text{RMSE} = \sqrt{\frac{1}{N_{Si}} \sum_{t=1}^{N_{Si}} (y_{Si}^t - \hat{y}_{Si}^t)^2} \quad (16)$$

$$\text{MAE} = \frac{1}{N_{Si}} \sum_{t=1}^{N_{Si}} |y_{Si}^t - \hat{y}_{Si}^t| \quad (17)$$

统计指标 RMSE 和 MAE 越小,表示模型的性能越好.根据现场专家经验,预测值与实际化验值的误差绝对值在 0.1% 范围内为可接受的结果,为了更直观地展示模型的预测效果,模型的预测命中率定义为 HR:

$$\text{HR} = \frac{1}{N_{Si}} \sum_{t=1}^{N_{Si}} (H(t)) \times 100\% \quad (18)$$

式中, $H(t)$ 是第 t 个样本的 Heaviside 函数,定义为:

$$H(t) = \begin{cases} 1, & |y_{Si}^t - \hat{y}_{Si}^t| \leq 0.1 \\ 0, & |y_{Si}^t - \hat{y}_{Si}^t| > 0.1 \end{cases} \quad (19)$$

4.3 实验结果和分析

经过数据预处理后,某钢铁厂 2 号高炉上采集的 111041 组铁水温度数据,用来建立基于动态注意力机制的深度网络预测模型,其中 10 万组数据用于训练模型,11041 组数据用于测试模型.3117 组铁水硅含量数据,用于建立基于动态注意力深度迁移网络的预测模型,其中 2837 组数据用于模型训练,280 组数据用于模型测试.通过实验过程中对

模型结构的反复训练和调整,预测模型中的动态注意力机制模块的全连接层层数 M 取 3,各层神经元个数分别设置为 256、128、24,第 1~2 个隐含层激活函数设置为 ReLU (Rectified linear unit) 函数^[28],第 3 个隐含层激活函数设置为 Sigmoid 函数.4 个去噪自编码器堆叠形成一个结构为 24-160-120-80-40-1 的深度网络,其中第 1 个、第 2 个、第 3 个和第 4 个去噪自编码器的结构(输入层-隐含层-输出层)分别设置为 24-160-2、160-120-160、120-80-120 和 80-40-80.激活函数都设置为 ReLU 函数,加性高斯噪声的系数设为 0.01.铁水温度模型的训练批次设置为 1024,迭代次数设置为 100,铁水硅含量模型的训练批次设置为 32,迭代次数设置为 300.网络权值使用 Xavier 初始化为均值为 0,方差为 0.1 的均匀分布,偏置初始化为 0,学习率设置为 0.001.

为了比较本文提出的基于动态注意力深度迁移网络的铁水硅含量预测模型的性能,将基于动态注意力机制的深度网络用于建模讨论,模型参数与深度迁移网络中铁水硅含量预测模块中的保持一致,目的是为了分析迁移训练好的铁水温度网络对铁水硅含量预测性能的影响.为了验证本文提出的动态注意力机制模块对模型性能的贡献,基于堆叠去噪自编码器的深度网络被用来建模分析,其网络结构为 36-160-120-80-40-1,模型的输入考虑了影响铁水温度和硅含量的 24 个共有过程变量和影响硅含量的 12 个特有过程变量.考虑到支持向量回归机在小样本建模上具有很好的优势,因此内核为径向基函数的支持向量回归机也被用来建模比较.所有模型都是在深度学习框架 PyTorch 中运行,计算机 CPU 为英特尔酷睿 i7-9700,运行频率为 3.0 GHz,内存为 32 GB.详细的实验对比结果如表 2 所示.

由表 2 可以看出,相比于浅层的支持向量回归机模型,深度网络模型 S-DAE、ADNet、ADTNet 的性能都优于支持向量回归机,可能原因是深度网络自学习到的抽象的和鲁棒性的特征表示,有助于挖掘数据中隐含的关系,进而提高模型的性能.相比于堆叠去噪自编码器搭建而成的深度网络,带有动态注意力机制的深度网络和深度迁移网络性能更好,这说明本文设计的动态注意力模块,实时地考虑每个样本不同的过程变量与硅含量的动态关系,模型能为每个样本学习到目标相关的抽象特征表示,来提高硅含量在线预报模型的性能.对比基于动态注意力机制的深度网络,基于动态注意力深度迁移网络的均方根误差和平均绝对误差分别为 0.0649 和 0.0509,命中率能达到 90%,表明前期工作开发的红外视觉铁水测温系统提供的实时出铁口处铁水温度信息能为硅含量的在线预测提供有用的

表 2 基于不同模型的预测性能
Table 2 Prediction performance based on different models

模型	RMSE	MAE	HR (%)
SVR	0.0832	0.0635	77.5
S-DAE	0.0794	0.0616	84.6
ADNet	0.0772	0.0583	86.4
ADTNet	0.0649	0.0509	90.0

信息,且使用铁水温度大数据训练后的迁移深度网络模型,能充分挖掘过程变量与铁水硅含量之间的关系,进而提高模型的性能,并在一定程度上降低硅含量预测模型对带标签数据样本的依赖.需要说明的是,本文提出的深度模型的训练时间是要远远大于浅层模型的,这主要是因为深度模型的训练有无监督的预训练和有监督的微调两个阶段.但在将模型部署到工业应用现场时,无监督地预训练完全可以离线完成,只需使用带标签的样本有监督地在线微调预训练好的网络结构.因此,本文提出的方法在工业现场也具有一定的应用价值.

为了进一步直观、清晰地对比模型的性能和铁水硅含量预测的细节信息,绘制了基于支持向量回归机模型、堆叠去噪自编码器网络、带有动态注意力机制的深度网络和动态注意力深度迁移网络的模型预测值曲线图和实际化验值的曲线图,如图 9~12 所示.通过对比发现,基于动态注意力深度迁移网络预测的硅含量和实际化验硅含量曲线在数值与趋势上吻合较好,进一步说明了本文提出方法的有效性和优越性.由图 9 可以看出,基于浅层的支持向量回归机预测值显著偏离了实际化验值,尽管根据定义模型的命中率能达到 77.5%,但实际模型预测值波动范围较窄,无法正确地跟踪实际的硅含量数值与趋势.由图 10 可以看出,基于堆叠去噪自编码器预测性能显著优于支持向量回归机,模型能以较高的准确率命中硅含量实际化验值,且样本预测的误差进一步缩小,说明了深度网络对复杂的高炉冶炼过程非线性描述能力强于浅层网络.由图 11 可以看出,基于动态注意力机制的深度网络预测结果,能基本一致地跟踪硅含量实际化验值,且预测误差基本稳定在绝对值正负 0.1% 范围内,这也进一步说明了本文提出动态注意力模块对模型性能的提升.由图 12 可以看出,基于深度迁移网络的预测性能是最佳的,且在入炉矿源频繁波动情况下有较强的稳定性,相比于基于动态注意力机制的深度网络预测结果,基于深度迁移网络的数值和趋势跟踪得更好,样本预测误差更小,这也进一步表明了独有的铁水温度数据的加入对模型性能的提升.模型

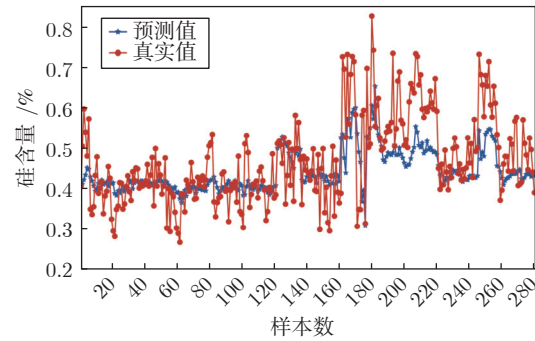


图 9 基于支持向量回归机的铁水硅含量预测结果
Fig. 9 Prediction results of silicon content based on SVR

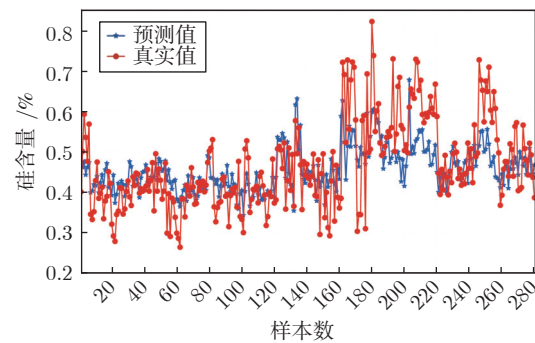


图 10 基于堆叠去噪自编码器的铁水硅含量预测结果
Fig. 10 Prediction results of silicon content based on S-DAE

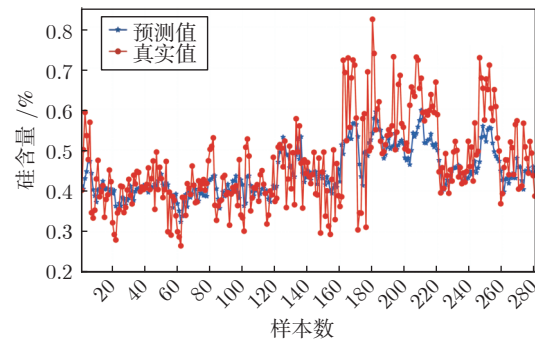


图 11 基于动态注意力机制深度网络的铁水硅含量预测结果
Fig. 11 Prediction results of silicon content based on ADNet

对 160~170 组硅含量预测性能欠佳,这是因为该数据区间内硅含量化验值超过 0.6%,通过分析训练集样本分布发现,超过 0.6% 硅含量样本数量在总样本中占比较少,导致模型对该部分少量样本拟合能力欠佳.需要说明的是,图 9~12 的预测结果表明,当实测硅含量值偏小时,预测硅含量数值普遍偏大;当实测硅含量数值偏大时,预测硅含量数值

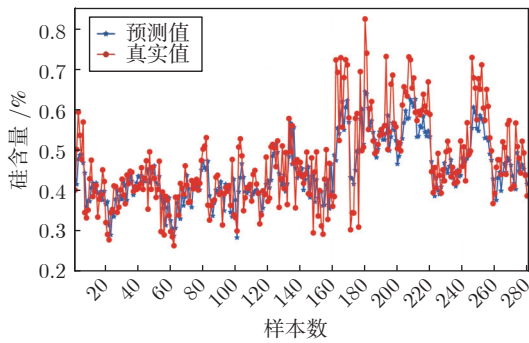


图 12 基于动态注意力深度迁移网络的铁水硅含量预测结果

Fig. 12 Prediction results of silicon content based on ADTNet

普遍偏小. 这是因为数据预处理操作中通过箱线图删除了异常的数据样本, 使得模型的预测结果不会出现极端预测值.

为了进一步更加直观地展示基于不同模型的铁水硅含量预测误差分布情况, 计算测试样本与对应训练样本的差值, 绘制了如图 13 所示的误差分布曲线. 从基于支持向量回归机预测结果绘制的误差曲线来看, 该模型误差波动范围比较大, 无法保证预测结果的可靠性. 相比浅层网络, 基于堆叠去噪自编码机和基于动态注意力机制的深度网络计算的误差大部分更加接近现场可接受的范围, 且动态注意力机制的引入, 使得基于动态注意力机制的深度模型相比于基于堆叠去噪自编码器模型对 160 ~ 280 组之间波动较大的样本的预测效果更好, 误差更小. 而本文的基于动态注意力深度迁移网络的误差普遍分布在工业现场可接受的范围 $[-0.1, 0.1]$ 内, 甚至大部分样本的误差能保持在 $[-0.05, 0.05]$ 内, 超出可接受范围的样本基本都是硅含量化验值超过 0.6% 或低于 0.3%. 根据中国西南某炼铁厂 2 号高炉冶炼的要求, 铁水硅含量需要严格控制在 0.3% ~ 0.6%, 这会导致超出控制范围的样本数量在总样本中占比较少, 在模型的训练过程中难以挖掘到小样本数据中的隐藏信息, 进而使得模型对该类样本预测性能欠佳. 从整体上来看, 本文模型能以较高的准确率预测铁水硅含量并具有较强的稳定性. 为了进一步阐述模型的稳定性, 将不同模型预测的硅含量值和硅含量实际化验值分别作为横、纵坐标, 绘制如图 14 所示的散点分布图. 由图 14 可以看出, 基于支持向量回归机模型的散点图有大量的点分布在 $y_2 = x - 0.1$ 的下方, 这说明模型对这部分样本的预测性能欠佳. 而基于堆叠去噪自编码机和基于动态注意力机制的深度模型的散点图有部分分布在 $y_1 = x + 0.1$ 和 $y_2 = x - 0.1$ 两条临界线之外, 且分

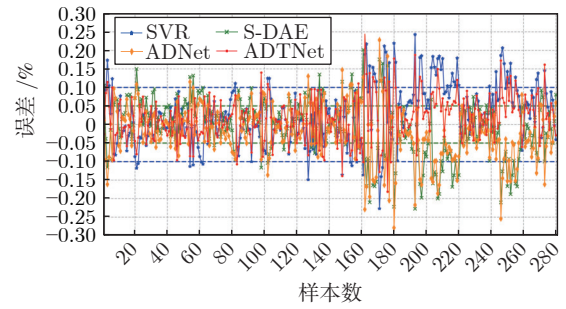


图 13 基于不同模型的铁水硅含量误差分布图

Fig. 13 Prediction errors of silicon content in molten iron based on different models

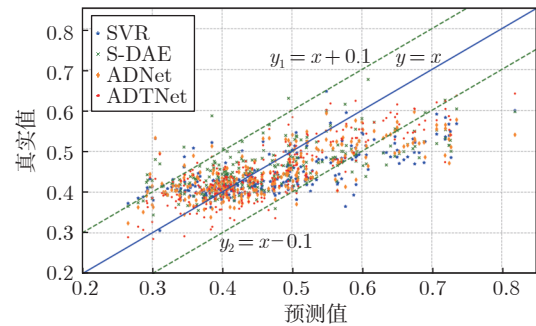


图 14 基于不同模型的预测和实际铁水硅含量分布散点图

Fig. 14 The scatter plot of predictive and observed silicon content based on different models

布在临界线内部的部分散点分布较为分散, 说明这两个模型预测的结果虽然满足了现场可接受范围, 但预测的精度还有进一步提升的空间. 而基于动态注意力深度迁移网络模型的散点普遍集中在 $y = x$ 附近, 只有少量的点分布在两条临界线之外, 且分布在临界线之外的点大多是硅含量化验值超过 0.6% 或者低于 0.3% 的样本, 这表明了本文模型对硅含量分布范围在 $[0.3\%, 0.6\%]$ 的数据具有较高的预测准确率, 而对于少数波动范围大的样本预测值则不能及时准确地跟踪检测值, 这也进一步说明了模型对部分超过 0.6% 或者低于 0.3% 数据的欠拟合能力. 考虑到现场入炉矿源的品质和铁水质量的要求, 铁水硅含量化验值大部分时间稳定在 0.3% ~ 0.6% 范围内, 因此对分布在 $[0.3\%, 0.6\%]$ 范围内硅含量样本的高预测准确率对现场操作也有较高的指导意义.

本文提出动态注意力深度迁移模型的另一个优势是, 设计的动态注意力机制模块能够实时地给出每个样本各个维度的过程变量对硅含量的动态注意力得分, 能实时地反映各过程变量对当前硅含量值的贡献程度. 本文给出测试集中前 24 个样本的过程变量对硅含量的贡献度, 绘制了如图 15 所示的

样本过程变量得分热力图, 横坐标代表过程变量, 纵坐标代表样本编号, 横纵坐标确定的矩阵中的数字代表的是该样本的过程变量对硅含量的贡献程度. 本文的注意力分数由 Sigmoid 函数计算出来, Sigmoid 函数值分布在 (0, 1) 范围内, 得分热力图给出了各过程变量的后 3 位小数. 尽管每个样本各维度过程变量注意力分数变化幅度较小, 但小幅度的变动是归一化后的数据和注意力模块计算累积的结果, 小幅度的变动在一定程度上能反映过程变量对硅含量贡献度的变化情况. 由图 15 可以看出, 各过程变量对第 1 个硅含量化验样本 (0.4015%) 与第 2 个硅含量化验样本 (0.5003%) 之间的贡献度的变化情况. 根据现场专家经验, 铁水硅含量的升高, 主要原因有 5 个: 1) 软熔带中滴落带的高度升高; 2) 炉料分布与煤气流分布不合理, 中心料柱透气性

差; 3) 焦炭强度不好且焦炭灰分高; 4) 高炉渣的二元碱度降低; 5) 鼓风动能不足或者透气性指数小难以消除炉缸中心死料柱, 确保炉缸中心活跃. 原因 1) 和原因 2) 是多种过程变量和现场操作共同作用的结果, 原因 3) 主要受原料成分的影响, 原因 4) 主要受高炉渣中氧化镁的含量的影响. 由图 15 可以看出, 当硅含量化验值上升时, 透气性指数和鼓风动能的注意力分数分别减小了 0.005 和 0.003, 其他过程变量的注意力分数发生了动态的变化, 这也能说明本文提出的注意力机制模块给出的过程变量对铁水硅含量值影响的合理性. 因此, 注意力机制模块不仅能捕获输入样本的过程变量与硅含量之间的相关关系并强调样本的各个过程变量在预测过程中的作用, 也能动态地给出各过程变量的注意力分数前后的差异, 可以为后续高炉冶炼过程的调节提供

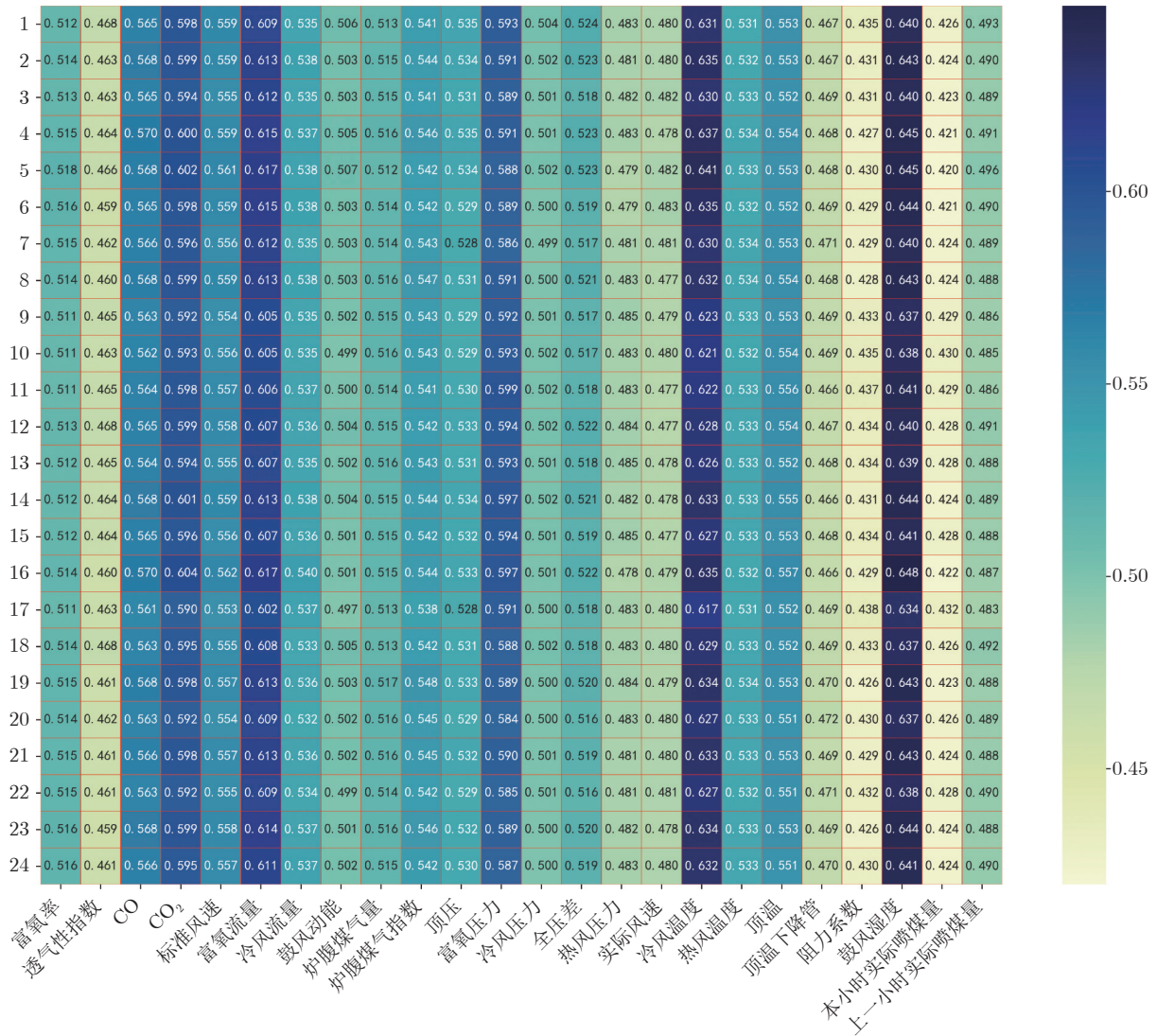


图 15 过程变量注意力得分热力图

Fig. 15 The heat map of process variables attention scores

重要的参考依据并在一定程度上增强黑箱模型的可解释性。

5 结束语

高炉铁水硅含量是高炉炼铁过程中表征铁水质量和炉缸内部热状态的重要和关键技术指标, 本文针对硅含量难以在线检测的实际工程难题, 提出了一种基于动态注意力深度迁移网络的铁水硅含量在线预报方法. 该方法提出了一种动态的注意力机制模块, 来捕获样本的各过程变量与硅含量之间的关系, 使得后端的深度网络在训练过程中逐层加强对预测性能有提升的主要过程变量的特征提取和抑制次要过程变量的特征提取. 考虑到铁水温度与铁水硅含量之间的正相关关系, 结合前期工作中研发的红外视觉铁水测温系统, 迁移了基于铁水温度数据训练的深度模型结构, 利用少量硅含量标签数据微调模型参数, 提高了铁水硅含量预测模型的性能并同时降低了模型对硅含量标签数据的依赖. 该模型不仅能实时地给出铁水硅含量的预测值, 还能给出对应的过程变量对预测值的贡献度, 不仅在一定程度上增强黑箱模型的可解释性, 也能为现场工人的操作和炉况调控提供重要的参考信息.

在今后的研究中, 将以预测的实时铁水硅含量数据和测温系统检测的铁水温度数据为基础, 黑箱模型的部分可解释性为辅助, 融合工艺条件、操作参数之间的耦合关系, 开展基于深度学习的高炉炼铁过程操作参数优化方法的研究.

References

- Zhou Ping, Zhang Li, Li Wen-Peng, Dai Peng, Chai Tian-You. Modeling of blast furnace multi-element molten iron quality with random weight neural network based on self-encoding and PCA. *Acta Automatica Sinica*, 2018, **44**(10): 1799–1811 (周平, 张丽, 李温鹏, 戴鹏, 柴天佑. 集成自编码与 PCA 的高炉多元铁水质量随机神经网络建模. *自动化学报*, 2018, **44**(10): 1799–1811)
- Zhou H, Zhang H F, Yang C J. Hybrid model based intelligent optimization of ironmaking process. *IEEE Transaction on Industrial Electronics*, 2020, **67**(3): 2469–2479
- Jiang K, Jiang Z H, Xie Y F, Pan D, Gui W H. Abnormality monitoring in the blast furnace ironmaking process based on stacked dynamic target-driven denoising autoencoders. *IEEE Transactions on Industrial Informatics*, 2022, **18**(3): 1854–1863
- Gao Chuan-Hou, Jian Ling, Chen Ji-Ming, Sun You-Xian. Data-driven modeling and prediction algorithm for complex blast furnace ironmaking process. *Acta Automatica Sinica*, 2009, **35**(6): 725–730 (郇传厚, 靳令, 陈积明, 孙优贤. 复杂高炉炼铁过程的数据驱动建模及预测算法. *自动化学报*, 2009, **35**(6): 725–730)
- Chen S H, Gao C H. Linear priors mined and integrated for transparency of blast furnace black-box SVM model. *IEEE Transactions on Industrial Informatics*, 2020, **16**(6): 3862–3870
- Zhou P, Lv Y B, Wang H, Chai T Y. Data-driven robust RVFLNs modeling of a blast furnace ironmaking process using Cauchy distribution weighted M-estimation. *IEEE Transaction on Industrial Electronics*, 2017, **64**(9): 7141–7151
- Song He-Da, Zhou Ping, Wang Hong, Chai Tian-You. Nonlinear subspace modeling of multi-variate molten iron quality in blast furnace ironmaking and its application. *Acta Automatica Sinica*, 2016, **42**(11): 1664–1679 (宋贺达, 周平, 王宏, 柴天佑. 高炉炼铁过程多元铁水质量非线性子空间建模及应用. *自动化学报*, 2016, **42**(11): 1664–1679)
- Spirin N A, Onorin O P, Istomin A S. Study of transition processes of blast-furnace smelting by the mathematical model method. In: Proceedings of the IOP Conference Series, Materials Science and Engineering. Suzhou, China: Institute of Physics Publishing, 2018. 12–73
- Spirin N, Onorin O, Alexander I. Prediction of blast furnace thermal state in real-time operation. *Solid State Phenomena*, 2020, **299**: 518–523
- Spirin N A, Polinov A A, Gurin I A, Pishnograev S N. Information system for real-time prediction of the silicon content of iron in a blast furnace. *Metallurgist*, 2020, **63**(9): 898–905
- Saxen H, Gao C H, Gao Z W. Data-driven time discrete models for dynamic prediction of the hot metal silicon content in the blast furnace: A review. *IEEE Transactions on Industrial Informatics*, 2013, **9**(4): 2213–2225
- Li Wen-Peng, Zhou Ping. Blast furnace hot metal quality robust regularization random weight neural network modeling. *Acta Automatica Sinica*, 2020, **46**(4): 721–733 (李温鹏, 周平. 高炉铁水质量鲁棒正则化随机神经网络建模. *自动化学报*, 2020, **46**(4): 721–733)
- Jiang Zhao-Hui, Xu Chuang, Gui Wei-Hua, Jiang Ke. Prediction method of hot metal silicon content in blast furnace based on optimal smelting condition migration. *Acta Automatica Sinica*, 2022, **48**(1): 194–206 (蒋朝辉, 许川, 桂卫华, 蒋珂. 基于最优工况迁移的高炉铁水硅含量预测方法. *自动化学报*, 2022, **48**(1): 194–206)
- Zhou P, Guo D W, Wang H, Chai T Y. Data-driven robust M-LS-SVR-based NARX modeling for estimation and control of molten iron quality indices in blast furnace ironmaking. *IEEE Transactions on Neural Networks and Learning Systems*, 2018, **29**(9): 4007–4021
- Jiang Zhao-Hui, Dong Meng-Lin, Gui Wei-Hua, Yang Chun-Hua, Xie Yong-Fang. Two-dimensional prediction for silicon content of hot metal of blast furnace based on Bootstrap. *Acta Automatica Sinica*, 2016, **42**(5): 715–723 (蒋朝辉, 董梦林, 桂卫华, 阳春华, 谢永芳. 基于 Bootstrap 的高炉铁水硅含量二维预报. *自动化学报*, 2016, **42**(5): 715–723)
- Li J P, Hua C C, Yang Y N, Guan X P. Bayesian block structure sparse based T-S fuzzy modeling for dynamic prediction of hot metal silicon content in the blast furnace. *IEEE Transactions on Industrial Electronics*, 2017, **65**(6): 4933–4942
- Hinton G E, Osindero S, Teh Y W. A fast learning algorithm for deep belief nets. *Neural Computing*, 2006, **18**(7): 1527–1554
- Hinton G E, Deng L, Yu D, Ahl G E, Mohamed A, Jaitly N, et al. Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups. *IEEE Signal Process*, 2012, **29**(6): 82–97
- Ma J L, Wu F, Zhu J, Xu D, Kong D X. A pre-trained convolutional neural network based method for thyroid nodule diagnosis. *Ultrasonics*, 2017, **73**: 221–230
- Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks. In *Process Advance Neural Information Process System*, 2012: 1097–1105
- Jiang K, Jiang Z H, Xie Y F, Chen Z P, Pan D, Gui W H. Classification of silicon content variation trend based on fusion of multilevel features in blast furnace ironmaking. *Information Sciences*, 2020, **521**: 32–45
- Wang Y L, Pan Z F, Yuan X F, Yang C H, Gui W H. A novel deep learning based fault diagnosis approach for chemical pro-

- cess with extended deep belief network. *ISA Transactions*, 2020, **96**: 457–467
- 23 Pan D, Jiang Z H, Chen Z P, Jiang K, Gui W H. Compensation method for molten iron temperature measurement based on heterogeneous features of infrared thermal images. *IEEE Transactions on Industrial Informatics*, 2020, **16**(11): 7056–7066
- 24 Pan D, Jiang Z H, Chen Z P, Gui W H, Xie Y F, Yang C H. Temperature measurement and compensation method of blast furnace molten iron based on infrared computer vision. *IEEE Transactions on Instrumentation and Measurement*, 2018, **68**(10): 3576–3588
- 25 Vincent P, Larochelle H, Lajoie I, Bengio Y, Manzagol P. Stacked denoising autoencoders: Learning useful representations in a deep network with a local denoising criterion. *Journal of Machine Learning Research*, 2010, **11**: 3371–3408
- 26 Vincent P, Larochelle H, Bengio Y, Manzagol P A. Extracting and composing robust features with denoising autoencoders. In: Proceedings of the 25th International Conference Machine Learning. Helsinki, Finland: 2008. 1096–1103
- 27 Reshef D N, Reshef Y A. Detecting novel associations in large data sets. *Science*, 2011, **334**(6062): 1518–1524
- 28 Agarap A F. Deep learning using rectified linear units (ReLU). arXiv preprint arXiv: 1803.08375, 2018.



蒋珂 中南大学博士研究生. 2019 年获得中南大学硕士学位. 主要研究方向为数据驱动的工业过程建模与控制, 过程数据分析和机器学习.

E-mail: jiangke@csu.edu.cn

(JIANG Ke Ph.D. candidate at the School of Automation, Central South University. She received her master degree from Central South University in 2019. Her research interest covers data-based modeling and control of industrial process, process data analysis, and machine learning.)



蒋朝辉 中南大学自动化学院教授. 2011 年获得中南大学博士学位. 主要研究方向为智能传感与检测技术, 图像处理与智能识别, 人工智能和机器学习. 本文通信作者.

E-mail: jzh0903@csu.edu.cn

(JINAG Zhao-Hui Professor at the School of Automation, Central South University. He received his Ph.D. degree from Central South University in 2011. His research interest covers intelligent sensing and detection technology, image processing and intelligent recognition, artificial intelligence, and machine learning. Corresponding author of this paper.)



谢永芳 中南大学自动化学院教授. 1993 年获得中南工业大学学士学位. 主要研究方向为分散控制, 鲁棒控制, 过程控制, 工业大数据和知识自动化. E-mail: yfxie@csu.edu.cn

(XIE Yong-Fang Professor at the School of Automation, Central South University. He received his bachelor degree from Central South University of Technology in 1993. His research interest covers decentralized control, robust control, process control, industrial big data, and knowledge automation.)



潘冬 中南大学自动化学院讲师. 分别于 2015 年和 2021 年获得中南大学学士和博士学位. 2019 年至 2021 年, 在加拿大拉瓦尔大学电子与计算工程系联合培养. 主要研究方向为红外热成像, 视觉检测, 图像处理和深度学习. E-mail: pandong@csu.edu.cn

(PAN Dong Lecturer at the School of Automation, Central South University. He received his bachelor and Ph.D. degrees from Central South University in 2015 and 2021, respectively. He was a joint training Ph.D. candidate in the Department of Electrical and Computing Engineering from Université Laval, Canada, from 2019 to 2021. His research interest covers infrared thermal imaging, vision-based measurement, image processing, and deep learning.)



桂卫华 中国工程院院士, 中南大学自动化学院教授. 1981 年获得中南矿冶学院硕士学位. 主要研究方向为复杂工业过程建模, 优化与控制应用, 故障诊断与分布式鲁棒控制. E-mail: gwh@csu.edu.cn

(GUI Wei-Hua Academician of Chinese Academy of Engineering, professor at the School of Automation, Central South University. He received his master degree from Central South Institute of Mining and Metallurgy in 1981. His research interest covers complex industrial process modeling, optimization and control applications, fault diagnosis, and distributed robust control.)