

基于改进 GRU-TCN 的磨煤机故障诊断算法

马 记¹, 许伟强², 王荣昌¹, 徐良友¹, 陈世彪², 胡 勇³

(1. 华电莱州发电有限公司, 山东 莱州 261400;

2. 华电电力科学研究院有限公司, 杭州 310030;

3. 华北电力大学 控制与计算机工程学院, 北京 102206)

摘要: 磨煤机系统的监测与诊断对电厂的安全运行至关重要; 由于真实故障数据的稀缺性以及故障与正常数据之间的不平衡, 传统数据驱动的故障诊断方法在故障识别上表现不佳, 有时甚至会产生误判; 为了高效地识别磨煤机在不同工况下的典型故障, 设计了一种结合了卷积块注意力模块的 GRU-TCN 融合算法, 用于建立磨的故障识别模型, 新算法不仅能提升分类准确性, 还能实现故障的提前预警; 首先, 通过调整磨煤机故障生成模型的关键参数, 模拟断煤、堵煤和自燃 3 种典型故障, 获取大量不同工况下的故障样本数据; 然后, 采用新分类算法建立基于典型样本的故障预警模型, 旨在提高故障识别的准确性, 在故障初期提醒操作人员进行干预, 从而避免磨煤机故障进一步扩大。

关键词: 磨煤机故障诊断; 典型故障样本; 注意力机制; 门控循环单元 (GRU); 时间卷积网络 (TCN)

Fault Diagnosis Algorithm for Coal Mills Based on Improved GRU-TCN

MA Ji¹, XU Weiqiang², WANG Rongchang¹, XU Liangyou¹, CHEN Shibiao², HU Yong³

(1. Huadian Laizhou Power Generation Co., Ltd., Laizhou 261400, China;

2. Huadian Electric Power Research Institute Co., Ltd., Hangzhou 310030, China;

3. School of Control and Computer Engineering, North China Electric Power University, Beijing 102206, China)

Abstract: The monitoring and diagnosis of coal mill systems are critical for the safe operation of power plants. Due to the scarcity of real fault data and the imbalance between faulty and normal data, traditional data-driven fault diagnosis methods perform poorly in fault identification and sometimes even make misjudgments. In order to efficiently identify typical faults of coal mills under different operating conditions, this paper designs a gate recurrent unit and temporal convolutional network (GRU-TCN) fusion algorithm with convolutional block attention module, which is used to build a mill fault identification model. The new algorithm not only improves the classification accuracy, but also realizes the early warning of faults. Firstly, by adjusting the key parameters of the coal mill fault generation model, three typical faults of coal breakage, coal plugging and spontaneous combustion are simulated to obtain a large amount of fault sample data under different working conditions. Then, the new classification algorithm is used to establish the fault warning model based on typical samples, the purpose is to improve the accuracy of fault identification and reminder operators to intervene at the early stage of faults, thereby avoiding further expansion of coal mill faults.

Keywords: coal mill fault diagnosis; typical fault samples; attention mechanisms; GRU (gate recurrent unit); TCN (temporal convolutional network)

0 引言

在当今的中国电力产业中, 火力发电作为主导能源仍占据绝对优势。磨煤机在燃煤发电厂中起到关键的辅助作用, 其工作状态对机组的稳定与安全至关重要^[1]。当磨煤机发生故障, 无法确保供给锅炉的煤粉流量, 导

致锅炉蓄能不足, 无法满足汽轮机需求。从而导致主蒸汽压力快速下降, 这会引发机组的快速减负荷, 严重时会导致炉内火焰熄灭和机组紧急停机。从而给发电企业造成了巨大的经济损失, 并对电网产生影响^[2]。因此, 磨煤机典型故障的提前预警对燃煤发电厂的安全运行具有重要意义。

收稿日期:2024-01-03; 修回日期:2024-03-06。

作者简介:马 记(1983-),男,硕士,高级工程师。

引用格式:马 记,许伟强,王荣昌,等. 基于改进 GRU-TCN 的磨煤机故障诊断算法[J]. 计算机测量与控制, 2025, 33(4): 17-23, 122.

磨煤机的复杂性为故障诊断带来了困扰。由于磨煤机系统本身参数繁多、耦合性强,单个故障往往会引起多个相关参数异常,难以直接判断故障,且其有些关键参数如磨内煤粉量信号难以准确捕获^[3]。因此,研究磨煤机典型故障的快速准确报警,提醒运行人员及时响应,是有必要的。

数据驱动的故障检测方法通过历史数据提取特征,建立数据特征与设备状态的关系,从而判断设备的故障类型。为了提高磨煤机的故障预警准确度,众多学者进行了深入研究。

基于数据驱动的故障诊断方法,如支持向量机^[4]、极限学习机^[5]和神经网络^[6],能有效提取故障数据中的关键特征。这些方法首先训练历史数据建立故障识别的映射,再用训练好的模型对磨煤机运行数据进行故障判断。但这些方法可能受到数据量和数据均衡性的制约,导致分类效果受限^[7]。解决这个问题,有两种主要方法:1)小样本故障诊断算法。从少量故障数据中提取特征,建立特征与设备状态的关系。文献[8]采用灰色关联度的方式,针对磨煤机典型故障,实现早期故障预测并有效判别故障类型。文献[9]提出了基于模糊决策融合的制粉系统故障诊断方法。考虑磨断煤、磨堵煤、磨自燃3种故障,用单层神经网络进行故障识别,用灰色关联度指标识别故障的严重程度。文献[10]提出了基于核主元分析和最小二乘支持向量机的磨煤机故障诊断新方法,用于判断磨煤机早期故障并判断故障类型。文献[11]采用FINCH分类器,对小样本故障集进行分类,实现对设备的故障诊断。小样本故障预警不能遍历机组的所有工况,当磨煤机运行工况变化后,可能不再准确;2)数据与机理融合的故障诊断方法。采用机理模型产生大量故障样本,解决数据不平衡问题,再通过神经网络算法建立故障类别映射模型用于故障预警,有效利用海量的磨煤机故障数据,提升故障判别准确性。袁存波等人^[12]建立反映磨煤机动态过程的MPS中速磨煤机模型,模拟磨煤机故障过程,并采用D-S证据理论进行多证据融合故障诊断。文献[7]采用堆叠式自动编码器对机理模型生成的故障样本进行分类,实现了对磨煤机故障识别。然而生成故障样本的磨煤机模型是用于控制的,在生成故障样本的方面可能不太适合,因此文献[13]提出了适合生成故障样本的磨煤机机理模型,本文基于该模型生成不同工况下的故障样本,并采用融合算法建立分类模型。

人工智能算法的性能影响分类效果,从而影响故障判断的准确性,且时序算法能实现故障的提前预警。在深度学习的早期阶段,数据的分类训练主要依赖于反向传播(BP, backpropagation)算法,该算法通过有效地计算损失函数的梯度来迭代地优化网络权重。随着技术

的进步和可用数据的增加,卷积神经网络(CNN)^[14]逐渐成为图像分类任务的主流方法,因为它能够自动学习并层次化地表示空间特征,从而显著提高了分类准确性。然而,对于时间序列或文本等序列数据的分类,传统的CNN结构往往难以捕获长期的依赖关系。为了解决这一问题,长短时记忆网络(LSTM)^[15]和门控循环单元(GRU)^[16]被提出。这两种结构都是循环神经网络(RNN)的变体,它们通过引入特定的门控机制来有效地捕获和记忆长期的序列信息。其中GRU是LSTM的简化版本,具有一个结构更简洁、参数更少的结构,能达到LSTM相似或甚至更好的性能。为了进一步提高序列数据的处理效率,时间卷积网络(TCN)^[17]被提出。TCN通过使用扩张卷积和因果卷积结构,成功地将卷积的思想扩展到了时间序列数据,从而为序列分类任务提供了一个高效且强大的工具。因此,本文将GRU与TCN结合,用于时间序列样本分类,提升分类的快速性和准确性。注意力机制的引入为深度学习模型提供了更大的灵活性和能力,使其能够更加关注数据中的重要部分。卷积块注意力模块^[18]包括空间注意力模块和通道注意力模块,可以集成到各种框架中。引入到融合模型中确保网络能在数据样本分类时关注到序列模型的关键特征。

通过建立磨煤机的机理模型来模拟生成故障数据,旨在解决真实故障数据稀缺的问题^[19-20]。这种思路虽然为故障诊断带来了新的可能,但由于现有的分类模型尚不够精确需进一步优化。因此,本文采用磨煤机机理模型来生成大量的故障数据。同时设计一种创新的融合算法,即结合了GRU和TCN的深度学习模型,并融入了卷积注意力机制(CBAM),用于对故障数据进行分类。新模型不仅能快速预警,还能提高分类准确度减少误报。

1 磨煤机模型及典型故障模拟

1.1 用于模拟故障的磨煤机模型

磨辊式中速磨煤机因其低能耗、稳定输出和长维护周期在火电机组中常用。如图1所示,它由原煤仓、给

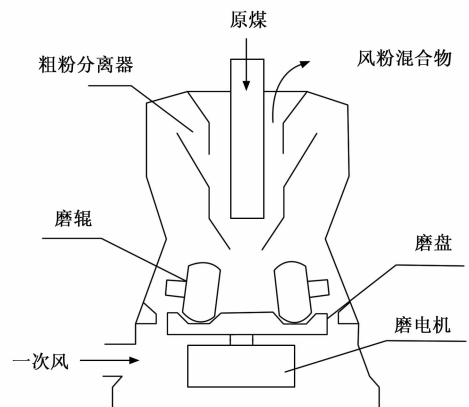


图1 中速磨煤机的结构

煤机、磨煤机及风机等部分组成。原煤通过给煤机送入磨机研磨。空气被分为两路: 一路预热后进入热一次风管, 另一路进入冷一次风管。混合后的空气进入磨煤机, 完成煤粉的干燥和输送。磨煤机内部主要由槽式磨板和磨辊组成, 将原煤研磨成煤粉。研磨后的煤粉随一次风进入粗粉分离器, 大颗粒煤粉回落继续研磨, 细颗粒与一次风一同进入炉膛燃烧。

将文献 [13] 中的磨煤机模型整理为状态方程的形式, 如下所示:

$$\begin{cases} \dot{x}_1 = 0.0655 (1.2001 \cdot u_4 + 0.3045 \cdot u_5)^{1-(x_5+x_6)/75} \cdot u_1 - x_1 \\ \dot{x}_2 = 0.00475 (k_1 \cdot u_1 \cdot u_3 + k_2 \cdot u_5 \cdot u_2 - k_3 \cdot k_4 \cdot x_2) \\ \dot{x}_3 = 0.123 \left[0.4704 \times \left(1 - \frac{x_5 + x_6}{75} \right) \cdot x_1 - x_3 \right] \\ \dot{x}_4 = \frac{k_5 - k_6}{19.039(3481.6 + x_5 + x_6)} \\ \dot{x}_5 = u_6 - 0.0095k_7 \cdot x_5 \\ \dot{x}_6 = 0.0095k_7 \times x_5 - k_9 \\ \dot{x}_7 = \frac{0.1333u_6 - k_8 - k_9 \cdot x_7}{x_5 + x_6} \end{cases} \quad (1)$$

$$\begin{cases} (k_1, k_2, k_3, k_{11}) = 0.0001 \times (u_3, u_4, x_2, x_4) + 1.0099 \\ k_4 = u_4 \sqrt{575.32 \times (u_1 - x_1)} + u_5 \sqrt{110.75 \times (u_1 - x_1)} \\ k_5 = k_3 \cdot x_2 \cdot k_4 + 43.126u_6 + 0.6k_7 + 0.1k_1u_3k_4 \\ k_6 = 1.1 \times k_{11}x_4k_4 + 4.18x_4k_8 + k_10x_4k_9 + 0.005 \times k_5 \\ k_7 = 0.4332 \times x_5 + 0.3746 \times x_6 + 22.139 \\ k_8 = 0.1333 \cdot u_6 \cdot (1 - e^{-k_1/28.155}) \cdot (0.0563 \times x_4 - 3.3664) \\ k_9 = 0.0075 \cdot x_3 \cdot x_6 \cdot k_4 \\ k_{10} = 1.844 \times (1 - 0.01 \times x_7) + 0.04187 \times x_7 \end{cases} \quad (2)$$

式中, 模型的输入 u , 包含 u_1 为一次风机出口压力 (kPa)、 u_2 为热风温度 ($^{\circ}\text{C}$)、 u_3 为冷风温度 ($^{\circ}\text{C}$)、 u_4 为热风阀开度 (%)、 u_5 为冷风阀开度 (%)、 u_6 为给煤量 (kg/s)。模型的状态变量为 x , 包含 x_1 为磨入口风压 (kPa)、 x_2 为磨入口风温 ($^{\circ}\text{C}$)、 x_3 为磨出口风压 (kPa)、 x_4 为磨出口风温 ($^{\circ}\text{C}$)、 x_5 为原煤储存量 (kg)、 x_6 为煤粉储存量 (kg)、 x_7 为煤粉水分 (%)。模型的辅助变量 k , 包含 k_1, k_2, k_3 , 分别为冷空气、热空气和混合空气的比热容 ($\text{kJ}/(\text{kg} \cdot ^{\circ}\text{C})$)、 k_4 为磨入口一次风量 (kg/s)、 k_5 为 Q_m 磨煤机系统的总输入能量 (kJ), 由一次风、原煤、磨煤机粉磨和密封风带来的热量组成、 k_6 为磨煤机系统总输出能量 (kJ), 由一次风带热量、蒸发水带热量、煤粉带热量和热损失组成、 k_7 为 I_b 磨的研磨电流 (A)、 k_8 为原煤蒸发水量 (kg/s)、 k_9 为出口煤粉流量 (kg/s)、 k_{10} 为煤粉比热容 ($\text{kJ}/(\text{kg} \cdot ^{\circ}\text{C})$)、 k_{11} 为磨出口蒸汽温度的比热容 ($\text{kJ}/$

($\text{kg} \cdot ^{\circ}\text{C})$)。

生成故障样本的磨煤机机理模型分为一次风模型、磨内研磨模型、磨出口温度模型和煤粉水分模型 4 部分。其中, 一次风模型用于干燥煤粉和将煤粉送出磨煤机。磨内研磨模型, 表示原煤进入磨煤机后, 研磨成煤粉的过程。磨出口温度模型表示一次风将研磨好的煤粉吹出磨煤机的过程。煤粉水分模型为研磨过程中对煤粉水分的估计。从状态方程中可以看出, 用于生成故障样本的模型涉及到磨煤机工作过程, 建模更加精细, 生成的故障样本匹配模型真实的故障。

1.2 磨煤机典型故障特征模拟

由于磨煤机系统故障率较低, 故障样本数据非常有限, 难以覆盖系统的所有运行工况。因此, 如何有效地获取大量典型故障样本是解决基于数据的故障诊断的关键。在模拟闭环控制的情况下, 改变关键参数或模型参数的设定值来进行磨煤机系统的断煤、堵煤和自燃故障模拟, 产生接近实际故障样本的故障数据。

磨断煤表现为磨煤机没有原煤输入, 通常由原煤仓的堵塞或给煤机故障引起。因此将给煤机给煤流量信号降为 0, 即可实现给磨断煤模拟, 如图 2 所示。

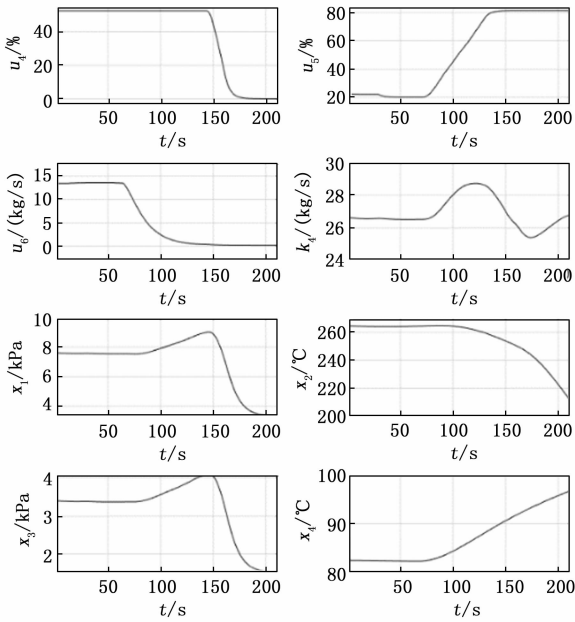


图 2 磨断煤故障模拟

磨堵煤表示煤粉堵在磨煤机出口处, 难以吹到炉膛中, 通过降低磨煤机的研磨能力来模拟堵煤故障。研磨系数表示磨煤机将原煤磨成煤粉的能力。在状态方程中, 研磨系数, 在磨电流不变的情况下研磨能力下降, 因此将系数 0.009 5 减少, 就能模拟堵煤故障。随着粉磨能力的降低, 磨内原煤量迅速增加, 导致流动阻力增大, 出口风压降低, 磨堵煤故障模拟如图 3 所示。

磨自燃表示为磨煤机内温度过高引燃煤粉。如图 4

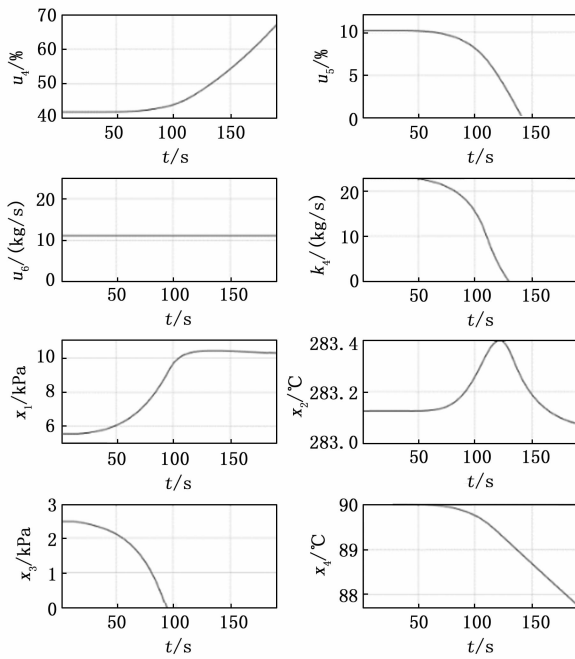


图 3 磨堵煤故障模拟

所示,在煤粉自燃模拟中,煤粉分为两部分:一部分由一次风带入炉内燃烧;另一部分在磨煤机内自燃。因此,在模拟煤粉自燃故障时,将自燃产生的热量加入总输入能量,将 k_5 的公式修改如下:

$$k_5 = k_3 x_2 k_4 + 2.1563 T_{rc} u_6 + 0.6 k_7 + 0.1 k_1 u_3 k_4 + \varphi \cdot k_9 \cdot q_{\text{net},\text{ar}} \quad (3)$$

其中: Φ 为自燃比例系数, $q_{\text{net},\text{ar}}$ 为原煤净热值 (kJ/kg)。增大自燃比例系数,模拟煤粉故障。随着自燃故障的发生,即使热风门关死,冷风门开满,磨出口温度依旧逐渐上升,符合磨自燃的情况。

通过建立磨煤机的机理模型来模拟生成故障数据,旨在解决真实故障数据稀缺的问题^[19-20]。这种思路虽然为故障诊断带来了新的可能,但由于现有的分类模型尚不够精确,使得其实际效果尚待进一步优化。因此,本文采用磨煤机机理模型来生成大量的故障数据。同时设计一种创新的融合算法,即结合了 GRU 和 TCN 的深度学习模型,并融入了卷积注意力机制 (CBAM),用于对故障数据进行分类。新模型不仅能快速预警,还能提高分类准确度减少误报。

2 改进的 GRU-TCN 算法

2.1 GRU 算法原理

GRU 与 LSTM 相似性在于均是 RNN 的变体,能够有效地捕捉长序列之间的关联,同时能够改善 RNN 中的梯度消失和梯度爆炸问题。因此这些算法被广泛应用于语音识别、文本翻译和时间序列预测等领域,其性能优于传统的 RNN,而且相对于 LSTM 来说,计算复

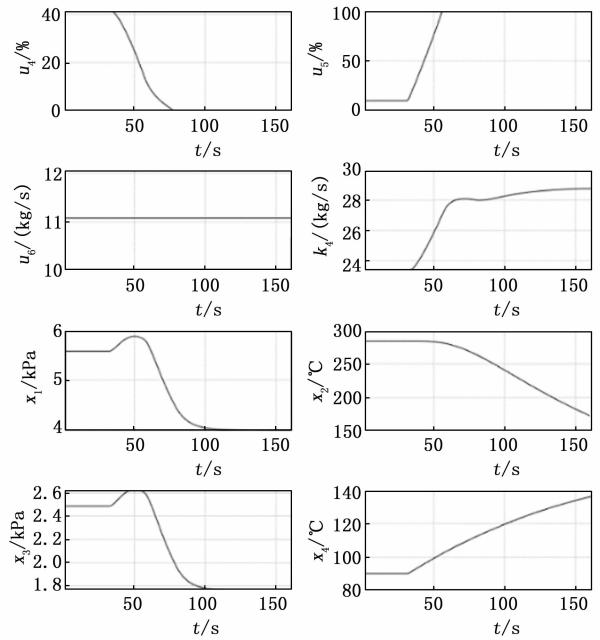


图 4 磨自燃故障模拟

杂度较低。门控循环单元的结构如图 5 所示。

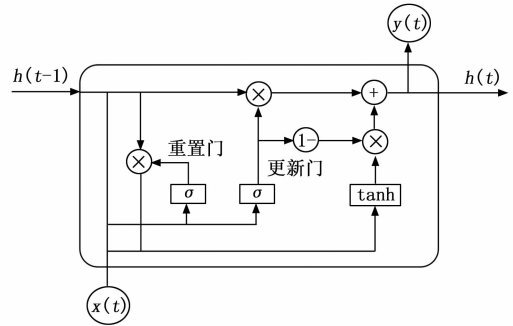


图 5 门控循环单元的结构

GRU 与 LSTM 的主要区别在于其内部结构的复杂度较低,并引入了重置门 (Reset Gate) 和更新门 (Update Gate) 的概念。这两个门控制了 GRU 单元的关键部分,其中更新门类似于 LSTM 的输入门和遗忘门,用于控制前一时刻的隐藏状态在当前时刻的更新程度,而重置门用于控制前一时刻的隐藏状态在当前候选隐藏状态中的更新程度。重置门和更新门的输入均来自当前时刻的输入向量 $\mathbf{x}_t$ 和上一时刻的隐藏状态 $\mathbf{h}_{t-1}$,模块输出通过全连接层的激活函数计算得出:

$$r_t = \sigma[\mathbf{W}_r \cdot (\mathbf{h}_{t-1}, \mathbf{x}_t)] \quad (4)$$

$$z_t = \sigma[\mathbf{W}_z \cdot (\mathbf{h}_{t-1}, \mathbf{x}_t)] \quad (5)$$

$$\mathbf{h}'_t = \tanh[\mathbf{W}_h \cdot (r_t \times \mathbf{h}_{t-1}, \mathbf{x}_t)] \quad (6)$$

$$\mathbf{h}_t = (\mathbf{I} - z_t) \times \mathbf{h}_{t-1} + z_t \times \mathbf{h}'_t \quad (7)$$

式 (4) 表示重置门,式 (5) 表示更新门。其中 $\mathbf{x}_t$ 表示当前时刻的输入向量, $\mathbf{h}_{t-1}$ 是上一时刻的状态记忆变量,通过这两者的组合,可以计算出重置门 r_t 和更新门 z_t 。其中, r_t 的取值范围在 $[0, 1]$ 之间,当它接近

1 时, 表示过去的信息更重要。而 Z_t 的取值范围也在 $[0, 1]$ 之间, 信号接近 1 表示过去的信息被更多地保留。在式 (7) 中, h_t 代表候选隐藏状态, 式 (7) 前半部分表示对原本的隐藏状态的选择性丢弃, 而式 (7) 后半部分则表示对当前节点的候选隐藏状态进行选择保留。

W_r 、 W_z 和 W_h 均表示各自的权值矩阵, I 表示单位矩阵, “ $\cdot$ ” 表示矩阵点乘, “[]” 表示向量连接, “ $\times$ ” 表示矩阵乘积。此外, 激活函数采用了 sigmoid 和 tanh 函数, 其具体表达式如下:

$$\text{sigmoid}(x) = \frac{1}{1 + e^{-x}}$$

(8)

$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$$

(9)

2.2 TCN 算法原理

TCN 是一种由膨胀因果卷积和残差网络结合而成的网络结构。该网络包括多层一维卷积结构, 其中膨胀因果卷积的输入和输出长度相同, 并使用残差网络连接, 以避免像 RNN 结构一样在处理时序数据时出现梯度爆炸的问题。TCN 的卷积层既包括因果卷积结构和膨胀卷积结构。时间卷积网络的结构如图 6 所示。

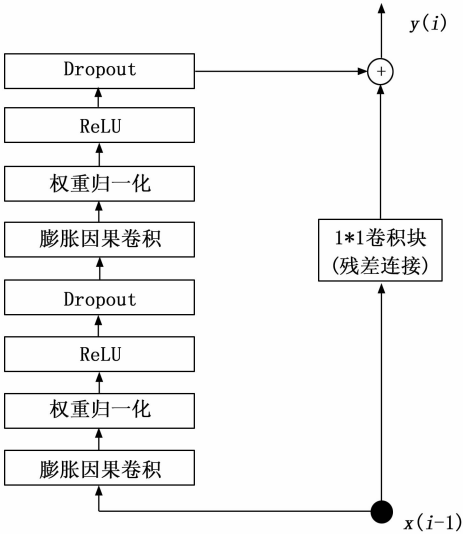


图 6 时间卷积网络的结构

如图 7 所示, 因果卷积是一种单向时间流动的模型, 只能按一个时间方向进行计算, 具有严格的时间条件。因此, 它要求模型的输出长度与输入长度相同, 并且在处理当前时间步时不能获取未来时间步的信息, 以

保持未来信息的未知性。

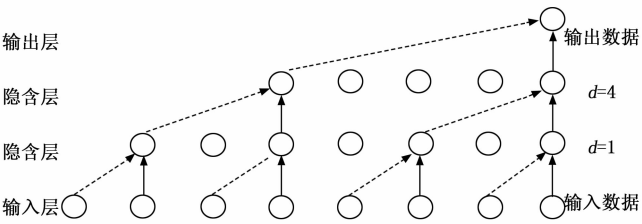


图 7 因果卷积结构

然而, 因果卷积在处理长时间序列时, 因为时间跨度太大会导致训练复杂性增加, 输出拟合效果不佳等问题。为解决这些问题, 研究人员引入了膨胀卷积 (Dilated Convolution) 的概念。膨胀卷积的核心思想是允许在卷积操作中对输入数据进行间隔采样, 与普通卷积不同, 它同时进行卷积和采样操作。

膨胀卷积计算公式:

$$F(t) = \sum_{i=0}^{k-1} f(i) \cdot x_{t-d \times i}$$

(10)

2.3 卷积注意力机制原理

CBAM 由通道注意力 (CAM, channel attention module) 和空间注意力 (SAM, spatial attention module) 结合而成。CBAM 能够同时考虑通道和空间两方面的注意力, 更合理地分配数据的权重, 从而取得更好的效果。CBAM 注意力机制是一种轻量级的注意力模块, 能够在通道和空间维度上将重要数据赋予高权重, 从而优先保留这些数据, 同时对次要和不重要的数据赋予低权重, 减少计算和输出的概率。这类似于将注意力集中在不同的输入部分。CBAM 的结构由 CAM 和 SAM 两种模块并联而成, 分别对输入数据进行通道和空间上的注意力机制计算。这种操作既节省了参数和计算时间, 还具有较高的灵活性, 可以轻松集成到现有网络架构中使用。

2.3.1 通道注意力模块

通道注意力结构如图 8 所示。

通过对输入特征 (尺寸为 $H \times W \times C$) 进行基于宽度和高度的全局最大池化和全局平均池化操作, 分别得到两个尺寸为 $1 \times 1 \times C$ 的特征矩阵。随后, 这两个特征矩阵分别传递给一个深度为两层的 BP 神经网络。第一层的神经元个数为 C/r (其中 r 为神经元减少率), 激活函数为 ReLU, 而第二层的神经元个数为 C 。此外, 这两

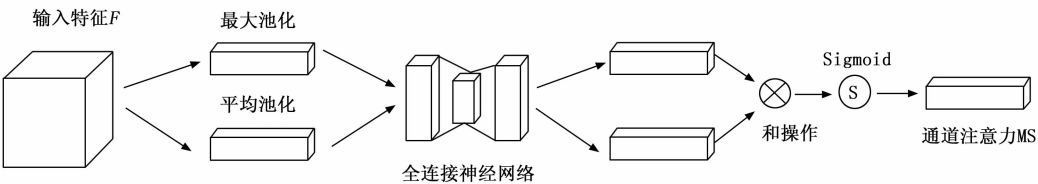


图 8 通道注意力结构

层神经网络的权值可以共享。最后,神经网络的输出特征矩阵与之前计算的矩阵相加,并经过 Sigmoid 激活函数,得到通道注意力特征矩阵,计算公式如下:

$$M_C(F) = \sigma\{MLP[AvgPool(F)] + \sigma\{MLP[MaxPool(F)] = \sigma\{W_1[W_0(F_{Avg}^C)] + W_1[W_0(F_{Max}^C)]\} \quad (11)$$

2.3.2 空间注意力模块

空间注意力模块如图 9 所示。

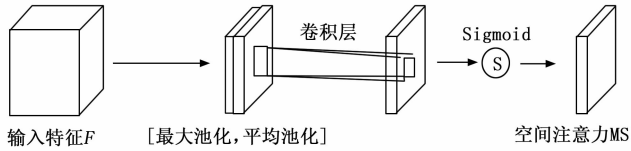


图 9 空间注意力结构

SAM 模块将其输出特征 F 与输入特征进行逐元素相乘,得到本模块的输入特征矩阵。首先,进行基于通道的全局最大池化和全局平均池化操作,得到两个尺寸为 $H \times W \times 1$ 的特征矩阵,然后将这两个特征矩阵进行通道拼接操作。如果特征参数数量为 n ,随后,这样处理后的数据经过一个 $n \times n$ 卷积操作,将其降维为一个通道特征,即尺寸为 $H \times W \times 1$ 的特征矩阵。最后,经过 Sigmoid 激活函数,得到最终的空间注意力特征矩阵,计算公式见式 (12):

$$M_S(F) = \sigma\{1^{6 \times 6}[(AvgPool(F);(MaxPool(F)))] = \sigma\{1^{6 \times 6}[(F_{Avg}^S;F_{Max}^S)]\} \quad (12)$$

2.4 引入注意力机制的 GRU-TCN 融合算法

GRU 网络更适合捕捉平稳信号的时间序列特征和变化规律,适用于数据波动较小且趋势相对稳定的短期预测任务。TCN 的膨胀卷积通过控制采样率,能够涵盖不同的感受野范围,从而适应长序列数据预测任务。此外,膨胀卷积的并行性能有助于有效跟踪长时间波动数据的时间序列变化。这使得模型能够充分利用短期和长期的故障特征参数,增强了故障特征的可信度,弥补了只考虑短期特征的不足。

模型引入 CBAM 模块后 CBAM-GRU-TCN 模型结构如图 10 所示。CBAM 是一种轻量级通用单元,可以方便地嵌入到类似 TCN 的架构中,并应用于每个卷积块的输出。考虑到不同特征在辨识任务中的重要性不同,CBAM 引入后可以根据每个特征向量对任务的重要性分配权重,从而提高模型的辨识准确度和鲁棒性。同时为了防止 TCN 网络在长距离时间特征提取中丢失短期信息,在网络架构中加入残差块,将数据直接与 TCN 提取的特征相加,之后送到 GRU 网络中。由 GRU 网络预测输出特征参数,输出经过 Dropout 正则化处理后,使用 SoftMax 激活函数作为网络的最后一层,将神经网络提取到的预测值转化分概率分布,并确

保多分类的概率总和为 1。使用交叉熵损失函数衡量模型的预测概率分布与真实标签之间的差异。交叉熵损失函数如下:

$$Loss = - \sum_{i=1}^{output} y_i \cdot \log \hat{y}_i \quad (13)$$

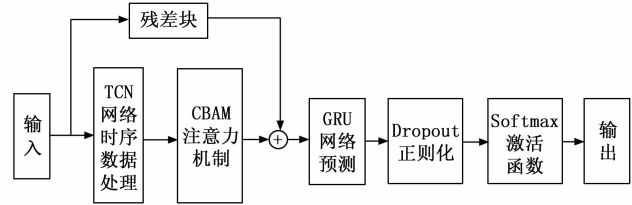


图 10 CBAM-GRU-TCN 模型结构

2.5 改进 GRU-TCN 算法故障预测流程

图 11 展示了磨煤机故障诊断的整体流程,具体如图 11 所示。

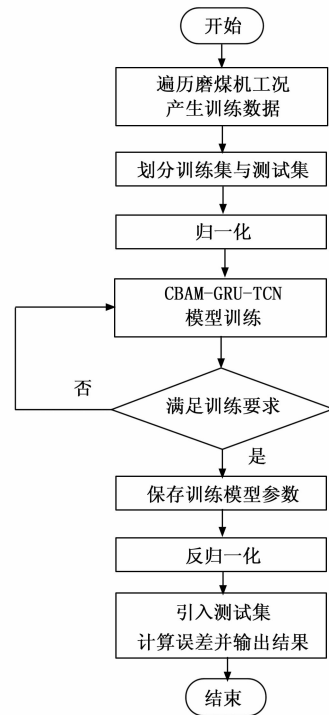


图 11 基于 CBAM-GRU-TCN 的磨煤机故障诊断流程

1) 遍历磨煤机模型的不同工况,分别持续采集断煤、堵煤、自燃和正常 4 种类型的数据样本。按照一定比例划分训练集和测试集样本。

2) 对样本数据进行均值归一化处理,公式如下:

$$x^* = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \quad (14)$$

3) 初始化设置。确定好网络模型,进行模型超参数的设置,如表 1 所示。

4) 模型训练部分。将预处理后的训练数据集输入到已经设置好的算法模型中进行。在迭代中找到最优参数,保留最优模型。

表 1 融合模型的超参数

输入变量	最大训练代数	学习率	批量大小	CBAM 缩减率	CBAM 卷积核	TCN 核大小	TCN 扩张率	TCN 滤波器数量	TCN 堆叠层数	GRU 隐含单元数	GRU 层数
8	200	0.001	32	4	7	3	2	32	2	32	2

5) 当模型分类误差小于设定值或循环代数超过设定值时,跳出循环,结束训练,保存模型参数。否则返回步骤 4)。

6) 记录测试结果评价指标的分类准确程度,并给出测试样本中时间序列上的故障可能性。

3 仿真实验分析

正常运行数据样本来自实际运行数据,故障数据样本来自模拟故障数据。从模拟故障数据中,选择了 10 000 组数据作为训练集,3 000 组数据作为测试集。这些数据集中涵盖了热风门开度 u_4 、冷风门开度 u_5 、给煤量 u_6 、磨入口一次风流量 k_4 、磨煤机入口压力 x_1 、磨煤机入口温度 x_2 、磨煤机出口压力 x_3 和磨煤机出口温度 x_4 共 8 个关键特征作为模型输入特征。设置两组消融实验,分别证明带 CBAM 的 GRU-TCN 算法在时序样本分类和磨故障识别的效果。

第一组仿真实验中,将 CBAM-GRU-TCN 算法与 GRU, TCN 和 GRU-TCN 算法比较。首先在 10 000 组训练集数据上进行模型训练,而后在测试集上比较网络的分类效果。测试集中共 3 000 组数据,包含磨断煤、磨堵煤、磨自燃和正常运行的数据每种 750 组数据。算法分类效果对比如表 2 所示。

表 2 算法分类效果对比 %

算法	磨断煤	磨堵煤	磨满煤
TCN	91.73	89.73	92.40
GRU	93.73	94.93	95.60
GRU-TCN	96.40	97.47	97.07
CBAM-GRU-TCN	98.93	99.07	98.00

从表 2 中可以看出,磨断煤、磨堵煤及磨自燃的判断正确率分别从 TCN 的 91.73%、89.73% 和 92.40% 上升到 GRU 的 93.73%、94.93% 和 95.60%,TCN 主要在于对长序列数据的特征捕捉,不在于对数据的预测,因此预测效果不如 GRU。而 GRU 与 TCN 融合后使 3 种故障的算法分类指标进一步上升到 96.93%、97.47% 和 97.07%,在 GRU-TCN 中引入 CBAM 后,3 种故障的分类准确度进一步上升到 98.93%、99.07% 和 98.00%,说明 CBAM-GRU-TCN 的预测效果相对于基准模型表现出更强的分类能力,注意力机制的引入增强了网络对故障特征提取的能力,从而分类的准确率有所提升。

在第二组仿真实验中,将 CBAM-GRU-TCN 算法

与 GRU-TCN, GRU 和 TCN 共 3 种算法对比在时序样本上的磨煤机故障预警效果。使用上一组仿真中的训练模型,对现场实际磨断煤数据正向传播测试。测试效果如图 12 所示。

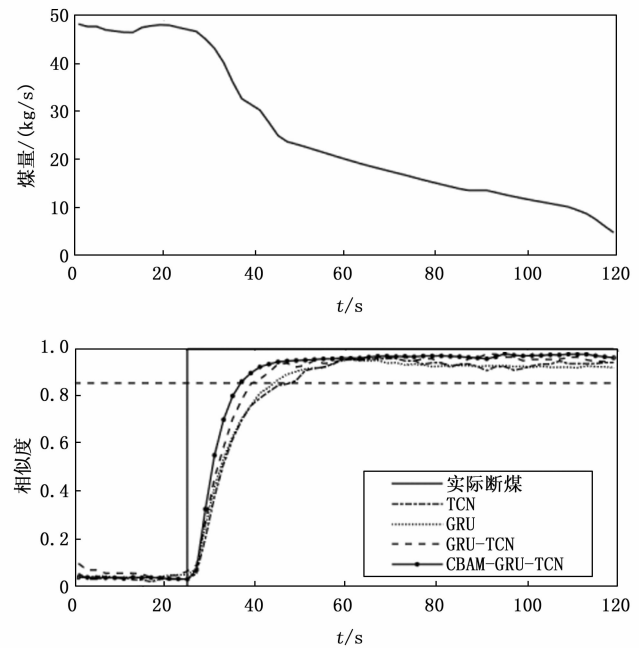


图 12 磨断煤故障判断的效果对比

磨断煤故障的置信度指标为 0.85,该值人为设定,使网络的分类效果最好,当相似度超过 0.85 时,发生断煤预警。图 12 中显示了磨煤机因发生断煤故障而引发跳闸这个过程中,4 种网络的报警情况,从图中可以看出,在 25 s 时发生断煤故障,CBAM-GRU-TCN 网络断煤发生后 12 s 时给出了断煤故障预警,GRU-TCN, GRU 和 TCN 分别在断煤发生后 14、19 和 23 s 发出预警信号,均实现了磨煤机故障的提前预警。CBAM-GRU-TCN 网络能更好地给出预警信号,这是由于新算法网络有效结合了时序结构和注意力机制,能比其他网络更早地识别出磨断煤故障,说明新网络的响应时间更加灵敏。提前发出的断煤信号为运行人员争取到响应时间裕度,运行人员收到断煤信号后,快速启动另一台磨煤机,从而减少总给煤量波动,维持机组主蒸汽压力的稳定,避免机组运行进一步恶化。

4 结束语

本文提出了一种结合 CBAM 注意力机制的 GRU-TCN 磨煤机故障这一诊断方法。GRU-TCN 作为一种时 (下转第 122 页)