

文章编号: 1671-251X(2024)02-0116-08

DOI: [10.13272/j.issn.1671-251x.2023110039](https://doi.org/10.13272/j.issn.1671-251x.2023110039)

# 基于改进门控循环神经网络的采煤机滚筒调高量预测

齐爱玲<sup>1</sup>, 王雨<sup>1</sup>, 马宏伟<sup>2</sup>

(1. 西安科技大学 计算机科学与技术学院, 陕西 西安 710054;  
2. 西安科技大学 机械工程学院, 陕西 西安 710054)

**摘要:** 采煤机自适应截割技术是实现综采工作面智能化开采的关键技术。针对采煤机在复杂煤层下自动截割精度较低的问题, 提出了一种基于改进门控循环神经网络(GRU)的采煤机滚筒调高量预测方法。鉴于截割轨迹纵向及横向相邻数据之间的相关性, 采用定长滑动时间窗法对获取的采煤机滚筒高度数据进行预处理, 将输入数据划分为连续、大小可调的子序列, 同时处理横向、纵向的特征信息。为提高模型预测效率, 满足循环截割的实时性要求, 提出了一种用因果卷积改进的门控循环神经网络(CC-GRU), 对输入数据进行双重特征提取和双重数据过滤。CC-GRU 利用因果卷积提前聚焦序列纵向的局部时间特征, 以减少计算成本, 提高运算速度; 利用门控机制对卷积得到的特征进行序列化建模, 以捕捉元素之间的长期依赖关系。实验结果表明, 采用 CC-GRU 模型对采煤机滚筒调高量进行预测, 平均绝对误差(MAE)为 43.80 mm, 平均绝对百分比误差(MAPE)为 1.90%, 均方根误差(RMSE)为 50.35 mm, 决定系数为 0.65, 预测时间仅为 0.17 s; 相比于长短时记忆(LSTM)神经网络、GRU、时域卷积网络(TCN), CC-GRU 模型的预测速度较快且预测精度较高, 能够更准确地对采煤机调高轨迹进行实时预测, 为工作面煤层模型的建立和采煤机调高轨迹的预测提供了依据。

**关键词:** 采煤机; 滚筒调高; 煤岩识别; 深度学习; 门控循环神经网络; 因果卷积

中图分类号: TD421.6 文献标志码: A

Prediction of height adjustment of shearer drum based on improved gated recurrent neural network

QI Ailing<sup>1</sup>, WANG Yu<sup>1</sup>, MA Hongwei<sup>2</sup>

(1. College of Computer Science and Technology, Xi'an University of Science and Technology, Xi'an 710054, China;  
2. College of Mechanical Engineering, Xi'an University of Science and Technology, Xi'an 710054, China)

**Abstract:** The adaptive cutting technology of the shearer is a key technology for achieving intelligent mining in fully mechanized working faces. In order to solve the problem of low automatic cutting precision of the shearer in complex coal seams, a prediction method for the height adjustment of shearer drum based on improved gated recurrent neural network (GRU) is proposed. Considering the correlation between adjacent data in the longitudinal and transverse directions of the cutting trajectory, the fixed length sliding time window method is used to obtain the height data of the shearer drum. The input data is divided into continuous and adjustable subsequences, while processing the feature information in the transverse and longitudinal directions. To improve the prediction efficiency of the model and meet the real-time requirements of cyclic cutting, causal convolution gated recurrent unit(CC-GRU) is proposed to perform dual feature extraction and dual data filtering on input data. CC-GRU utilizes causal convolution to focus on the local temporal features in the longitudinal direction of the sequence in advance, in order to reduce computational costs and improve computational speed. CC-GRU uses gating

收稿日期: 2023-11-13; 修回日期: 2024-02-19; 责任编辑: 胡娟。

基金项目: 国家自然科学基金资助项目(61674121)。

作者简介: 齐爱玲(1972—), 女, 陕西西安人, 副教授, 博士, 研究方向为煤岩识别、信号与图像处理, E-mail: [qalemail@126.com](mailto:qalemail@126.com)。

通信作者: 王雨(2000—), 女, 陕西榆林人, 硕士研究生, 研究方向为煤岩预测, E-mail: [21208223078@stu.xust.edu.cn](mailto:21208223078@stu.xust.edu.cn)。

引用格式: 齐爱玲, 王雨, 马宏伟. 基于改进门控循环神经网络的采煤机滚筒调高量预测[J]. 工矿自动化, 2024, 50(2): 116-123.

QI Ailing, WANG Yu, MA Hongwei. Prediction of height adjustment of shearer drum based on improved gated recurrent neural network[J]. Journal of Mine Automation, 2024, 50(2): 116-123.



扫码移动阅读

mechanism to serialize and model the features obtained from convolution to capture long-term dependencies between elements. The experimental results show that using the CC-GRU model to predict the height adjustment of the shearer drum, the MAE is 43.80 mm, MAPE is 1.90%, RMSE is 50.35 mm, the determination coefficient is 0.65, and the prediction time is only 0.17 seconds. Compared to long short term memory (LSTM) neural networks, GRU, and temporal convolutional network (TCN), the CC-GRU model has a faster prediction speed and higher prediction precision. It can more accurately predict the height adjustment trajectory of the shearer in real time. This provides a basis for the establishment of coal seam models in the working face and the prediction of the height adjustment trajectory of the shearer.

**Key words:** shearer; drum height adjustment; coal rock recognition; deep learning; gated recurrent neural network; causal convolution

## 0 引言

无人化智能开采是煤矿智能化建设的中心目标<sup>[1]</sup>。当前煤矿智能化建设过程中,综采工作面的智能化水平制约着煤矿智能化向更高水平发展<sup>[2]</sup>。作为综采工作面的核心设备之一,采煤机的截割精度直接关系到采煤的质量和效率。随着开采工作向着少人化、无人化的目标迈进,采煤机滚筒自动调高技术成为亟需攻克的技术难题。采煤机滚筒自动调高是指根据截割作业中工作面工况的变化自动对截割滚筒高度进行调整,以避免截割顶底板岩石。

截割轨迹与煤岩边界线的拟合程度对于截割精度具有关键性影响<sup>[3]</sup>,因此煤岩界面的精准识别可作为采煤机滚筒自动调高的依据。针对煤岩界面识别问题,学者们提出了基于反射光谱<sup>[4]</sup>、红外成像<sup>[5]</sup>等的直接识别方法,但由于井下环境复杂恶劣,这些方法无法实现常态化精准应用<sup>[6-8]</sup>。为解决这些问题,间接感知煤岩界面的记忆截割技术被提出。传统的记忆截割技术将重复再现目标轨迹作为控制目标,不考虑轨迹本身的合理程度,虽然提高了采煤机的自动化程度,但是对于顶板起伏差异较大的煤层,仍需要人工调整截割路径<sup>[9]</sup>。为减少人工干预频率,田立勇等<sup>[10]</sup>提出了一种基于灰色预测理论的采煤机截割路径规划方法,提高了采煤机滚筒适应煤层的能力。王廷栋<sup>[11]</sup>提出一种基于遗传算法的采煤机截割路径优化方法,使截割过程更加平稳。基于修正策略的截割轨迹预测方法虽然在一定程度上提高了截割精度,但是当前状态参数传递及计算需要时间,并且特征参数容易冗余,容易导致采煤机无法及时调整姿态并快速适应煤层变化,不可避免会截割到岩石。这些缺陷促使采煤机自适应截割技术从识别走向预测,最好在当前位置获知下一段煤层边界状况。

近年来人工智能等前沿技术被应用到采煤机截

割轨迹预测中,支持向量机、神经网络等智能技术可自动从具有多特征参数的轨迹数据中学习到复杂的特征表示,捕捉轨迹数据中的时序信息、空间关系及其他相关特征,从而提高截割轨迹预测的准确性和鲁棒性<sup>[12-14]</sup>。煤层作为空间分布的一种层状地质实体,有较强的全局规律性。综采工作面动态生产时的采煤机历史截割轨迹数据可反映煤层分布形态随煤层空间几何形态变化而不断变化的趋势,这些数据实质上是时间序列数据<sup>[15]</sup>,并且具有较强的波动性和非线性。为进一步提高预测效率及精度,研究人员尝试使用时间序列预测技术提取截割轨迹中所隐含的特征信息,采用长短时记忆(Long Short-term Memory, LSTM)神经网络构建采煤机截割轨迹预测模型,取得了较好的效果,但仍存在输入参数过多、计算成本增加等问题<sup>[16-18]</sup>。

在实际采煤过程中,煤层起伏不定,需要进行迭代式预测,为保证误差不会逐级累计,对预测模型的实时性提出了较高要求。因此,本文提出一种用因果卷积改进的门控循环神经网络(Causal Convolution Gated Recurrent Unit, CC-GRU),并基于CC-GRU实现采煤机滚筒调高量预测。首先,采用滑动窗口方法将输入数据划分为连续、大小可调的子序列,同时处理横向、纵向的特征信息;其次,构造CC-GRU预测模型,对输入数据进行双重特征提取和双重数据过滤,其中因果卷积模块使用卷积核获取输入特征图的局部特征表示,以减少计算冗余,门控循环神经网络(Gated Recurrent Unit, GRU)处理迭代序列数据,以发挥其捕捉长期依赖关系的能力;最后,通过全连接层输出最终预测结果。

## 1 采煤机滚筒调高量预测方法

### 1.1 模型选择

采煤机滚筒调高量预测可定义为一个多元时间序列预测问题。多元时间序列数据具有高维度和复

杂性,单一模型无法捕捉到数据中的所有复杂结构和关系,会导致模型在预测未来时间序列数据时精度不足,所以需将多个模型融合或多个结构层叠加来克服这一缺陷。因此,本文采用因果卷积叠加递归神经网络结构的方式进行特征提取,尽可能地让模型挖掘到数据的多形态、多状态特征。

传统预测方法忽略了采煤机截割轨迹数据的时序性,不能根据数据的差异性为不同时刻的数据赋予相应权重。递归神经网络由于本身结构与时间序列数据特性的契合性,将其用于搭建预测模型可大幅提高预测精度,但所需参数较多,预测效率低。因果卷积捕捉序列数据的依赖关系,采用并行计算方式同时处理多个时间步的信息,解决了递归神经网络逐个时间步顺序处理的低效问题。

本文融合时域卷积网络(Temporal Convolutional Network, TCN)及 GRU 构建 CC-GRU 预测模型。通过一维因果卷积对历史数据提取局部时序特征,GRU 利用门控机制对卷积得到的特征进行序列化建模,捕捉元素之间的长期依赖关系。与单一模型相比,CC-GRU 预测模型减少了过拟合的风险,提高了模型的泛化能力,使预测结果更加稳定可靠。

## 1.2 预测流程

基于 CC-GRU 的采煤机滚筒调高量预测流程如图 1 所示。首先,获取原始采高数据,对数据进行归一化处理,然后采用滑动窗口方法划分数据,构建输入数据样本。然后,将样本划分为训练集和测试集,训练集数据用于训练 CC-GRU 预测模型学习时序特征,测试集数据用于测试最终训练完成的模型性能。最后,输出模型预测结果,并进行反归一化处理,得到真实的预测结果。

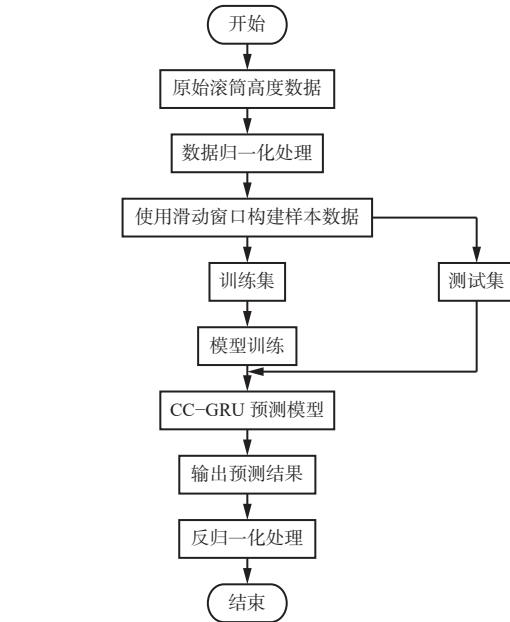


图 1 基于 CC-GRU 的采煤机滚筒调高量预测流程

Fig. 1 Prediction process of height adjustment of shearer drum based on causal convolution gated recurrent unit (CC-GRU)

## 2 CC-GRU 预测模型

CC-GRU 预测模型结构如图 2 所示。通过一维因果卷积层有效提取序列数据中的局部关联特征,并将特征映射到多维度特征向量。通过多层次门控机制对特征序列进行更精细的特征提取,捕获全局长期依赖关系。经过因果卷积层第 1 次特征过滤后,数据杂质显著减少,为后续特征提取和数据过滤提供了有利条件。使用 GRU 层对剩余特征进行选择性保留,完成第 2 次特征过滤。CC-GRU 的双层结构实现了多次特征提取和多次数据杂质过滤,有效提升了网络性能。

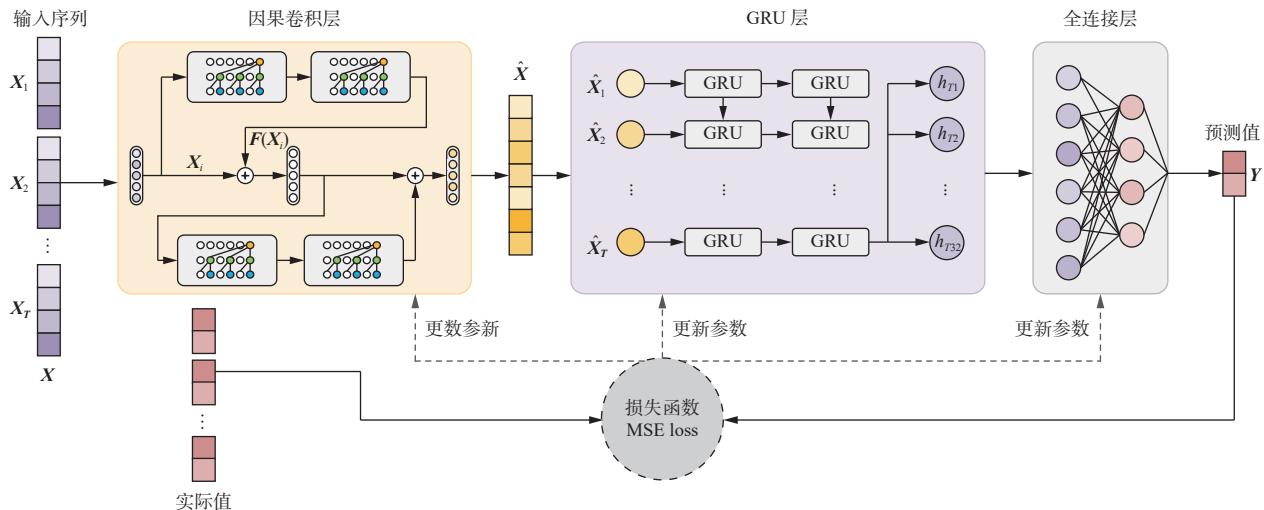


图 2 CC-GRU 预测模型结构

Fig. 2 Structure of CC-GRU prediction model

## 2.1 输入层

影响当前采样点采高的因素除不同刀具位置采样点的高度数据外,还有邻近采样点的高度数据。如果将所有数据无差别地输入到单个模型中进行训练,不能有效提取出截割轨迹纵向的时间特征及横向相邻数据之间的相关特征,因此,采用滑动窗口方法进行预处理。使用固定大小的滑动窗口沿着时间序列滑动,截取历史数据。对原始数据进行预处理后,得到输入样本:

$$\mathbf{X} = \begin{bmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \\ \vdots \\ \mathbf{X}_T \end{bmatrix} = \begin{bmatrix} x_1^1 & x_1^2 & \cdots & x_1^s \\ x_2^1 & x_2^2 & \cdots & x_2^s \\ \vdots & \vdots & & \vdots \\ x_T^1 & x_T^2 & \cdots & x_T^s \end{bmatrix} \quad (1)$$

式中:  $\mathbf{X}_i (i=1, 2, \dots, T)$  为采煤机进行第  $i$  刀截割时的高度序列;  $T$  为输入序列中的截割总刀数;  $x_i^j$  为第  $i$  刀的第  $j$  个采样点的高度;  $s$  为 1 刀的总采样点数。

## 2.2 因果卷积层

2018 年, Bai Shaojie 等<sup>[19]</sup>提出一种用于分析时间序列数据的神经网络——TCN。TCN 通过扩展因果卷积和残差模块并行处理输入序列,不仅避免了梯度爆炸和梯度消失等现象,提高了模型训练速度,而且有效抑制了循环神经网络中的误差累计问题,在时间序列数据预测任务中表现出良好性能。TCN 由多个残差模块组成,残差模块中最为重要的扩张因果卷积结构如图 3 所示。

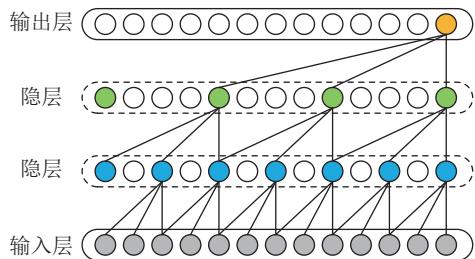


图 3 扩张因果卷积结构

Fig. 3 Causal dilated convolutional structure

对于输入时间序列  $\mathbf{X}$ ,通过因果卷积进行特征提取,即滑动卷积核  $\mathbf{F}=(f_1, f_2, \dots, f_K)$ ( $f_k$  为位置  $k$  处的权重,  $k=1, 2, \dots, K$ ,  $K$  为卷积核尺寸)分别与输入序列的不同部分进行卷积运算,得到的卷积值为

$$\mathbf{F}(\mathbf{X}_i) = \sum_{k=1}^K f_k X_{i-(K-k)d} \quad (2)$$

式中:  $d$  为扩张因子,当  $d=1$  时,对输入数据进行常规因果卷积运算,当  $d \neq 1$  时,对输入数据进行扩张因果卷积运算。

为更好地捕捉序列中的长期依赖关系,在卷积层后添加残差连接:

$$\mathbf{o}_i = \ell(\mathbf{F}(\mathbf{X}_i) + \mathbf{X}_i) \quad (3)$$

式中  $\ell(\cdot)$  为 ReLU 激活函数。

残差连接将输入序列的信息直接传递到卷积层的输出中,能帮助网络更好地学习并保留输入序列的重要特征,同时,使用非线性激活函数可增强模型的非线性建模能力。

## 2.3 GRU 层

经过因果卷积层卷积计算后的特征序列为  $\hat{\mathbf{X}} = (\hat{\mathbf{X}}_1, \hat{\mathbf{X}}_2, \dots, \hat{\mathbf{X}}_T)$ , 采用有效处理时间序列的循环神经网络对采煤机滚筒调高量进行预测。

2014 年, K. Cho 等<sup>[20]</sup>提出更易于计算的 GRU。与 LSTM 相比,GRU 需要的训练参数更少,收敛速度更快,并且可在时间序列数据有限的情况下降低模型过度拟合的风险<sup>[21]</sup>。GRU 单元主要由重置门和更新门组成,结构如图 4 所示。重置门决定前一时刻与当前时刻输入的结合程度,更新门决定前一时刻状态信息在当前时刻中的保留程度。将  $\hat{\mathbf{X}}$  作为 GRU 神经网络的输入,GRU 单元第  $t$  个时间步的输入为  $\hat{\mathbf{X}}_t$ ,输出为  $\mathbf{h}_t$ 。

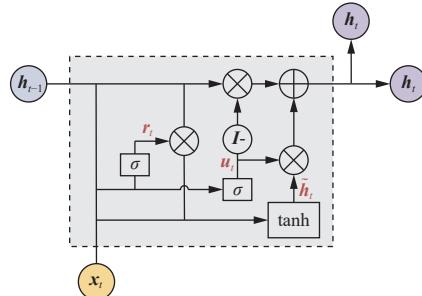


图 4 GRU 结构

Fig. 4 Structure of gated recurrent unit

首先,将当前时刻输入  $\hat{\mathbf{X}}_t$  与前一时刻隐藏状态  $\mathbf{h}_{t-1}$  的拼接矩阵输入到 GRU 单元中,得到  $t$  时刻重置门门控状态  $\mathbf{r}_t$  及更新门门控状态  $\mathbf{u}_t$ :

$$\mathbf{r}_t = \sigma(\mathbf{W}_r [\mathbf{h}_{t-1}, \hat{\mathbf{X}}_t] + \mathbf{b}_r) \quad (4)$$

$$\mathbf{u}_t = \sigma(\mathbf{W}_u [\mathbf{h}_{t-1}, \hat{\mathbf{X}}_t] + \mathbf{b}_u) \quad (5)$$

$$\sigma(p) = \frac{1}{1 + \exp(-p)} \quad (6)$$

式中:  $\sigma(\cdot)$  为 Sigmoid 激活函数;  $\mathbf{W}_r$ ,  $\mathbf{W}_u$  分别为重置门与更新门的权值矩阵;  $[\cdot]$  为向量拼接操作符号;  $\mathbf{b}_r$ ,  $\mathbf{b}_u$  分别为重置门与更新门的偏置向量;  $p$  为激活函数自变量。

然后,重置门可对  $\mathbf{h}_{t-1}$  选择性地重置,决定上一时刻信息的保留程度。将  $\mathbf{h}_{t-1}$  经过重置门处理后与  $\hat{\mathbf{X}}_t$  相加,并应用 tanh 激活函数得到当前时刻候选状态:

$$\tilde{\mathbf{h}}_t = \phi(\mathbf{W}_h [\mathbf{r}_t \odot \mathbf{h}_{t-1}, \hat{\mathbf{x}}_t] + \mathbf{b}_h) \quad (7)$$

$$\phi(p) = \frac{\exp(p) - \exp(-p)}{\exp(p) + \exp(-p)} \quad (8)$$

式中:  $\phi(\cdot)$  为  $\tanh$  激活函数;  $\mathbf{W}_h$  为候选状态的权值矩阵;  $\odot$  为逐元素乘法符号;  $\mathbf{b}_h$  为对应的偏置向量。

最后, GRU 单元结合  $\mathbf{u}_t$  和  $\tilde{\mathbf{h}}_t$  处理当前时刻信息, 结合  $\mathbf{I} - \mathbf{u}_t$  ( $\mathbf{I}$  为单位矩阵) 和  $\mathbf{h}_{t-1}$  处理历史时刻信息, 两者相加得到当前时刻 GRU 单元的输出:

$$\mathbf{h}_t = (\mathbf{I} - \mathbf{u}_t) \odot \mathbf{h}_{t-1} + \mathbf{u}_t \odot \tilde{\mathbf{h}}_t \quad (9)$$

GRU 能够有效捕获长期序列中的关键信息。在每个时刻, GRU 都能够动态保留和更新相关信息, 并传递到网络的下一个时间步, 直到模型处理完整个序列。最后一个时间步的隐藏状态  $\mathbf{h}_T$  会作为输入传递给全连接层, 通过线性映射产生最终预测结果。

### 3 实验准备与评价指标

#### 3.1 实验环境

实验使用的硬件配置为 Intel(R) Core(TM) i5-8265U CPU @ 1.80 GHz, 运行内存为 8 GiB。实验过程基于深度学习 Pytorch 框架, 使用的编程语言为 Python。

#### 3.2 实验数据及预处理

为验证 CC-GRU 预测模型的有效性, 选用实际截割数据进行实验。本文所使用的数据是某矿采煤机实际截割过程中记录的数据, 即采煤机多次截割过程中在各采样点的前滚筒高度。采样间隔为 3 m, 采煤机截深为 0.8 m, 每一次走刀包含 40 个采样点。由于已知数据较少, 采用克里金插值算法对实际数据进行插值, 得到的插值曲面如图 5 所示。

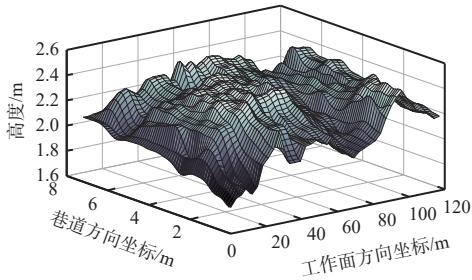


图 5 插值曲面

Fig. 5 Interpolation surface

由于原始采高数据中存在噪声, 且不同截割循环采高数据没有统一标准, 所以在进行采高预测前需要对原始数据进行数据预处理, 以提高数据的质量和可靠性。数据的预处理主要包括数据标准化、基于滑动窗口构造样本数据 2 个部分。

#### 3.2.1 数据标准化

为消除特征间单元差异和尺度差异的影响, 需对采煤机采高数据进行归一化处理, 使采高数据限定在 [0, 1] 区间。本文采用 min-max 归一法进行采高数据归一化, 归一化公式为

$$z^* = \frac{z - z_{\min}}{z_{\max} - z_{\min}} \quad (10)$$

式中:  $z^*$  为归一化后的数据;  $z$  为归一化前某刀某采样点的截割滚筒高度;  $z_{\min}$  为训练集中最小采高;  $z_{\max}$  为训练集中最大采高。

为评估模型性能, 需对模型输出的预测结果进行反归一化处理, 即将预测值还原到原始数据的尺度上, 以计算真实误差。反归一化公式为

$$y^* = y(z_{\max} - z_{\min}) + z_{\min} \quad (11)$$

式中  $y^*$  为反归一化后的数据;  $y$  为模型直接输出的预测值。

#### 3.2.2 滑动窗口

在实际生产中, 采煤机需要在工作面进行循环截割, 且每刀截割的相隔时间一般较短。为了保证采煤机工作的实时性, 预测模型需对截割轨迹进行连续预测。因此, 选择在预测前根据滚动方式构造多元样本数据。

在原始高度序列数据  $\mathbf{H} = \{\mathbf{H}_1, \mathbf{H}_2, \dots, \mathbf{H}_T\}$  中,  $\mathbf{H}_i$  为第  $i$  刀 40 个采样点的截割滚筒高度数据。设置滑动窗口大小为  $m$ , 将  $\mathbf{H}_i$  作为目标输出, 前  $m$  刀数据  $\{\mathbf{H}_{i-m}, \mathbf{H}_{i-m+1}, \dots, \mathbf{H}_{i-1}\}$  作为输入, 以此构建 1 个样本, 而后将窗口后移, 即将样本输入中的第 1 个向量  $\mathbf{H}_{i-m}$  剔除, 加入第  $i$  刀数据  $\mathbf{H}_i$ , 则输入为  $\{\mathbf{H}_{i-m+1}, \mathbf{H}_{i-m+2}, \dots, \mathbf{H}_i\}$ , 输出为第  $i+1$  刀数据  $\mathbf{H}_{i+1}$ , 构建 1 个新的样本, 依此类推, 构建所需样本集合。每组样本中包含  $m$  刀输入数据及对应的 1 刀输出数据。所构建的样本形式为

$$\left\{ \begin{array}{l} (\mathbf{H}_1, \mathbf{H}_2, \dots, \mathbf{H}_m), \mathbf{H}_{m+1} \\ (\mathbf{H}_2, \mathbf{H}_3, \dots, \mathbf{H}_{m+1}), \mathbf{H}_{m+2} \\ \vdots \\ (\mathbf{H}_{T-m+1}, \mathbf{H}_{T-m+2}, \dots, \mathbf{H}_T), \mathbf{H}_{T+1} \end{array} \right\} \quad (12)$$

#### 3.3 实验评价指标

为准确评估模型的预测效果, 采用平均绝对误差 (Mean Absolute Error, MAE)、平均绝对百分比误差 (Mean Absolute Percentage Error, MAPE)、均方根误差 (Root Mean Squared Error, RMSE)、决定系数  $R^2$  作为评价指标, 其计算公式分别为

$$MAE = \frac{1}{n} \sum_{j=1}^n |y_j - \hat{y}_j| \quad (13)$$

$$\text{MAPE} = \frac{1}{n} \sum_{j=1}^n \left| \frac{y_j - \hat{y}_j}{y_j} \right| \times 100\% \quad (14)$$

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{j=1}^n (y_j - \hat{y}_j)^2} \quad (15)$$

$$R^2 = 1 - \frac{\sum_{j=1}^n (y_j - \hat{y}_j)^2}{\sum_{j=1}^n (y_j - \bar{y})^2} \quad (16)$$

式中:  $n$  为测试样本数;  $y_j$  为第  $j$  个样本的实际值;  $\hat{y}_j$  为第  $j$  个样本的预测值;  $\bar{y}$  为样本实际值的平均值。

## 4 实验结果及分析

### 4.1 参数对比实验

CC-GRU 涉及的超参数主要包括网络层数、隐

层节点数等,不同的超参数对于训练效果有较大影响。依据经验将隐层节点数的搜索范围限制为 [16, 64]。首先通过对比实验确定 CC-GRU 的网络层数及隐层节点数。设定模型迭代次数均为 50, 在相同训练数据集下选取 3 个指标值为对照标准。不同参数下 CC-GRU 模型的预测结果见表 1。

由表 1 可知,在网络层数为 2 且每个隐层节点数为 32 时,模型各项评价指标值均相对较低,MAE 为 43.80 mm, MAPE 为 1.90%, RMSE 为 50.35 mm,  $R^2$  为 0.65,表明该模型预测效果最佳。从整体趋势来看,随着各隐层节点数的增加,预测误差呈现先增后减的趋势。这是由于网络结构过于简单时可能导致欠拟合,而过于复杂时又会出现过拟合问题。并且,如果各隐层节点数相同,网络层数增加反而会降低模型性能。

表 1 不同参数下 CC-GRU 模型的预测结果

Table 1 Prediction results of CC-GRU model under different parameters

隐层层数	隐层节点数			MAE/mm	MAPE/%	RMSE/mm	$R^2$
	第1层	第2层	第3层				
2	16	16	—	46.54	2.02	53.41	0.61
2	32	16	—	51.88	2.25	64.73	0.35
<b>2</b>	<b>32</b>	<b>32</b>	—	<b>43.80</b>	<b>1.90</b>	<b>50.35</b>	<b>0.65</b>
2	64	32	—	47.44	2.06	54.38	0.55
2	64	64	—	50.58	2.19	57.40	0.45
3	16	16	16	47.84	2.07	54.15	0.54
3	32	16	16	50.98	2.21	61.70	0.41
3	32	32	16	48.89	2.12	55.86	0.56
3	32	32	32	53.71	2.32	67.33	0.29

不同参数下 CC-GRU 模型的预测结果对比如图 6 所示。可看出,隐层节点数为(32, 32)(第 1 层、第 2 层节点数均为 32)时,全局差异波动较小,预测值更符合实际滚筒调高量变化趋势。

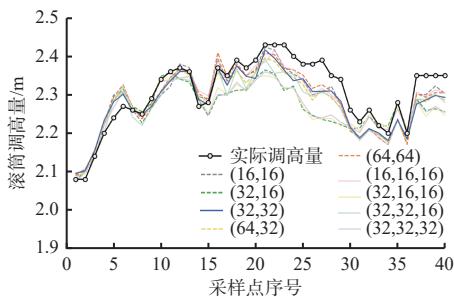


图 6 不同参数下 CC-GRU 模型的预测结果对比

Fig. 6 Comparison of prediction results of CC-GRU model under different parameters

### 4.2 模型对比实验

为对比 CC-GRU 模型与其他模型的性能差别,设置以下实验:在相同实验条件下,设置滑动窗口大

小为 4,采用 LSTM, GRU, TCN, CC-GRU 在同一数据集下对滚筒调高量进行预测。不同模型预测结果对比如图 7 所示。由图 7 可知:TCN 模型的预测效果最差,所得预测值均低于实际值;GRU 及 LSTM 模型效果次之,在采样点为[20, 30]的折线中可看出明显偏差;CC-GRU 模型预测结果最接近实际值,更符合变化趋势。

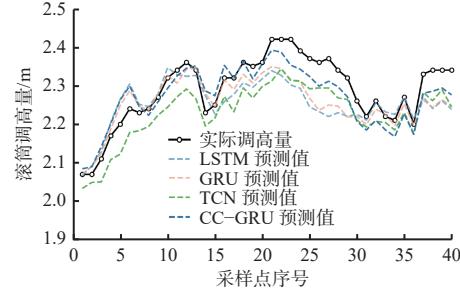


图 7 不同模型预测结果对比

Fig. 7 Comparison of prediction results of different models

不同模型的评价指标见表 2。分析表 2 可知,与

其他模型相比,CC-GRU 模型预测效果最佳,误差较小且决定系数较高。CC-GRU 模型相较于 LSTM 模型,MAE 降低了 14.1%,MAPE 降低了 13.9%,RMSE 降低了 21.1%,决定系数提高了 78.9%;相较于 GRU 模型,MAE 降低了 9.2%,MAPE 降低了 9.4%,RMSE 降低了 15.5%,决定系数提高了 69.3%;相较于 TCN 模型,MAE 降低了 28.5%,MAPE 降低了 28%,RMSE 降低了 24.2%,决定系数提高了 99.6%。这表明 CC-GRU 模型具有较强的鲁棒性,泛化性能较好。

表 2 不同模型评价指标

Table 2 Evaluation indicators of different models

模型	MAE/mm	MAPE/%	RMSE/mm	$R^2$
LSTM	51.01	2.21	63.85	0.36
GRU	48.25	2.10	59.58	0.38
TCN	61.27	2.64	66.41	0.32
<b>CC-GRU</b>	<b>43.80</b>	<b>1.90</b>	<b>50.35</b>	<b>0.65</b>

评价指标对比如图 8 所示。可看出,TCN 模型的 MAE 较 LSTM,GRU 模型明显增大,说明循环神经网络更能关注到序列的关键特征信息;CC-GRU 模型的决定系数最高,说明 CC-GRU 模型的拟合能力较强,能够很好地解释变化趋势。

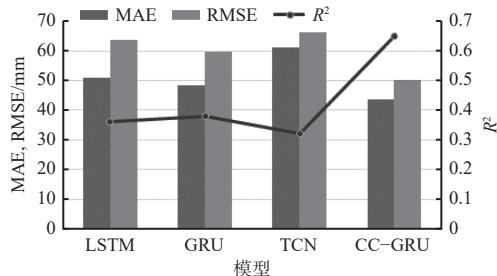


图 8 评价指标对比

Fig. 8 Comparison of evaluation indicators

不同模型预测时间对比见表 3。由表 3 可知,与其他模型相比,CC-GRU 模型的预测速度较快,每次预测时间仅为 0.17 s,可满足采煤机实际工作需求。相较于 LSTM,GRU 模型,其训练时间分别减少了 82.2%,80.5%;相较于 TCN 模型,其训练时间增加了 103.3%,但预测精度取得了较大提升。

表 3 不同模型预测时间对比

Table 3 Comparison of prediction time of different models

模型	预测时间/s
LSTM	0.96
GRU	0.88
TCN	0.08
<b>CC-GRU</b>	<b>0.17</b>

TCN 模型可快速处理多个时间步的信息,但在每个时间步上只在局部区域内提取特征,对于具有复杂长期依赖关系的序列,无法捕捉到所有的相关信息;循环神经网络相比 TCN 网络结构更为复杂,更适于处理长序列依赖关系,其中 GRU 比 LSTM 所需要的参数更少,但训练阶段计算复杂度过高,时间成本增加;具有双层结构的 CC-GRU 模型既关注序列中的局部特征,又可捕捉序列的长期依赖关系,在计算效率和性能之间取得了平衡。

实验结果表明,CC-GRU 模型可更好地提取学习滚筒高度的非线性变化特征,能够精准、快速进行预测,并取得良好的预测效果。

## 5 结论

1) 提出了一种基于 CC-GRU 的采煤机滚筒调高量预测方法。通过因果卷积提前聚焦序列的局部时序特征,并行处理输入序列,输出特征向量,并将其输入到多层 GRU 中对非线性和非平稳数据进行拟合。通过 CC-GRU 的双层结构实现了双重特征提取和双重数据杂质过滤,在保证精度的同时,满足了实时性要求。

2) 使用真实数据对采煤机滚筒调高量进行预测,CC-GRU 预测模型的 MAE 为 43.80 mm,MAPE 为 1.90%,RMSE 为 50.35 mm,决定系数为 0.65,预测时间仅为 0.17 s。实验结果表明,该模型可快速、有效预测滚筒调高量。

3) 与 LSTM 模型、GRU 模型、TCN 模型进行对比实验,结果表明,CC-GRU 预测模型的 MAE,MAPE,RMSE 均低于其他模型,决定系数更高,预测时间更短,表明 CC-GRU 预测模型具有较高的可靠性,能够在采煤机滚筒调高量预测中提供更为准确的结果。

## 参考文献(References):

- [1] 王国法,张良,李首滨,等.煤矿无人化智能开采系统理论与技术研发进展[J].煤炭学报,2023,48(1):34-53.  
WANG Guofa, ZHANG Liang, LI Shoubin, et al. Progresses in theory and technological development of unmanned smart mining system[J]. Journal of China Coal Society, 2023, 48(1): 34-53.
- [2] 王国法,富佳兴,孟令宇.煤矿智能化创新团队建设与关键技术研发进展[J].工矿自动化,2022,48(12):1-15.  
WANG Guofa, FU Jiaxing, MENG Lingyu. Development of innovation team construction and key technology research in coal mine intelligence[J]. Journal of Mine Automation, 2022, 48(12): 1-15.
- [3] HOLM M, BOHN C. Modelling of the cutting process

- of underground coal mining with shearer loaders for adaptive horizon control[C]. 19th SGEM International Multidisciplinary Scientific GeoConference, 2019: 423-430. DOI: [10.5593/sgem2019/1.3/S03.054](https://doi.org/10.5593/sgem2019/1.3/S03.054).
- [4] YANG Jianjian, CHANG Boshen, ZHANG Yuchen, et al. CNN coal and rock recognition method based on hyperspectral data[J]. International Journal of Coal Science & Technology, 2022, 9(1). DOI: [10.1007/s40789-022-00516-x](https://doi.org/10.1007/s40789-022-00516-x).
- [5] WANG Haijian, LIU Lili, ZHAO Xuemei, et al. Pre-perception and accurate recognition of coal–rock interface based on active excitation infrared characterization[J]. Journal of Computational Design and Engineering, 2022, 9(5): 2040-2054.
- [6] SUN Chuanmeng, LI Xinyu, CHEN Jiaxin, et al. Coal-rock image recognition method for complex and harsh environment in coal mine using deep learning models[J]. IEEE Access, 2023, 11: 80794-80805.
- [7] 张强, 张润鑫, 刘峻铭, 等. 煤矿智能化开采煤岩识别技术综述[J]. 煤炭科学技术, 2022, 50(2): 1-26.  
ZHANG Qiang, ZHANG Runxin, LIU Junming, et al. Review on coal and rock identification technology for intelligent mining in coal mines[J]. Coal Science and Technology, 2022, 50(2): 1-26.
- [8] 毛明仓, 张孝斌, 张玉良. 基于透明地质大数据智能精准开采技术研究[J]. 煤炭科学技术, 2021, 49(1): 286-293.  
MAO Mingcang, ZHANG Xiaobin, ZHANG Yuliang. Research on intelligent and precision mining technology based on transparent geological big data[J]. Coal Science and Technology, 2021, 49(1): 286-293.
- [9] WANG Chen, ZHOU Jie. New advances in automatic shearer cutting technology for thin seams in Chinese underground coal mines[J]. Energy Exploration & Exploitation, 2022, 40(1): 3-16.
- [10] 田立勇, 毛君, 王启铭. 基于采煤机摇臂惰轮轴受力分析的综合煤岩识别方法[J]. 煤炭学报, 2016, 41(3): 782-787.  
TIAN Liyong, MAO Jun, WANG Qiming. Coal and rock identification method based on the force of idler shaft in shearer's ranging arm[J]. Journal of China Coal Society, 2016, 41(3): 782-787.
- [11] 王廷栋. 基于遗传算法的采煤机截割控制系统研究[J]. 煤炭技术, 2021, 40(10): 206-208.  
WANG Tingdong. Research on cutting control system of coal shearer based on genetic algorithm[J]. Coal Technology, 2021, 40(10): 206-208.
- [12] 司垒, 王忠宾, 刘新华, 等. 基于煤层分布预测的采煤机截割路径规划[J]. 中国矿业大学学报, 2014, 43(3): 464-471.  
SI Lei, WANG Zhongbin, LIU Xinhua, et al. Cutting path planning of coal mining machine based on prediction of coal seam distribution[J]. Journal of China University of Mining & Technology, 2014, 43(3) : 464-471.
- [13] XIE Jiacheng, YAN Zewen, WANG Xuewen, et al. A memory cutting method of virtual shearer based on shape track prediction of AFC[J]. Mining, Metallurgy & Exploration, 2021, 38(5): 2005-2019.
- [14] WANG Weibing, JING Zelin, ZHAO Shuanfeng, et al. Intelligent height adjustment method of shearer drum based on rough set significance reduction and fuzzy rough radial basis function neural network[J]. Applied Sciences, 2023, 13(5): 2877-2898.
- [15] 梁要, 王世博, 葛世荣, 等. 综采工作面煤层三维模型动态修正方法研究[J]. 工矿自动化, 2022, 48(7): 58-65, 72.  
LIANG Shua, WANG Shibo, GE Shirong, et al. Study on dynamic modification method of 3D model of coal seam in fully mechanized working face[J]. Journal of Mine Automation, 2022, 48(7): 58-65, 72.
- [16] 崔涛. 基于真实开采数据的采运装备虚拟规划方法研究[D]. 太原: 太原理工大学, 2022.
- CUI Tao. Research on virtual planning method of mining and transportation equipment based on real mining data[D]. Taiyuan: Taiyuan University of Technology, 2022.
- [17] 陈伟华, 南鹏飞, 闫孝姮, 等. 基于深度学习的采煤机截割轨迹预测及模型优化[J]. 煤炭学报, 2020, 45(12): 4209-4215.  
CHEN Weihua, NAN Pengfei, YAN Xiaoheng, et al. Prediction and model optimization of shearer memory cutting trajectory based on deep learning[J]. Journal of China Coal Society, 2020, 45(12): 4209-4215.
- [18] 唐志章. 综采工作面集中控制关键技术研究[D]. 徐州: 中国矿业大学, 2021.  
TANG Zhizhang. Research on key technologies of centralized control in fully mechanized mining face[D]. Xuzhou: China University of Mining and Technology, 2021.
- [19] BAI Shaojie, KOLTER J Z, KOLTUN V. Trellis networks for sequence modeling[C]. The 7th International Conference on Learning Representations, New Orleans, 2018: 1-18.
- [20] CHO K, MERRIENBOER B, GULCEHRE C, et al. Learning phrase representations using RNN encoder-decoder for statistical machine translation[C]. Conference on Empirical Methods in Natural Language Processing, Doha, 2014: 1724-1734.
- [21] RUÍZ-GUIROLA D E, LÓPEZ O L A, MONTEJO-SÁNCHEZ S, et al. Performance analysis of ML-based MTC traffic pattern predictors[J]. IEEE Wireless Communications Letters, 2023, 12(7): 1144-1148.