

融合加性注意力机制的TCN-LSTM巷道变形预测模型

王龙飞, 陈俊智*

昆明理工大学国土资源工程学院, 云南 昆明 650093



摘要:为提高矿山巷道位移变形预测精度,提出融合加性注意力机制的时间卷积网络—长短期记忆网络(TCN-LSTM)预测模型。该模型通过TCN提取时序数据局部特征,采用LSTM捕捉长期依赖规律,并引入加性注意力机制动态优化特征权重分配。基于云南自走铁矿3个中段监测点的位移监测数据进行预测,试验结果表明:相较于单一模型和组合模型,融合加性注意力机制的TCN-LSTM模型平均绝对误差(MAE)降低了28.79%~60.63%,均方根误差(RMSE)减少了27.52%~64.08%,对称平均绝对百分比误差(SMAPE)下降了28.70%~68.48%,校正决定系数(R^2_{adj})稳定且高于0.97,预测误差均低于0.1 mm,证明该模型在预测精度和泛化能力方面具有显著优势,可为该矿山巷道安全监测提供技术支持。

关键词:时间卷积网络;长短期记忆网络;加性注意力机制;时间序列预测;矿山巷道;巷道变形预测;机器学习

中图分类号:TD76 文献标志码:A 文章编号:1005-2518(2025)05-1020-11 DOI:10.11872/j.issn.1005-2518.2025.05.103
引用格式:王龙飞,陈俊智.融合加性注意力机制的TCN-LSTM巷道变形预测模型[J].黄金科学技术,2025,33(5):1020-1030.
WANG Longfei, CHEN Junzhi. TCN-LSTM Roadway Deformation Prediction Model Based on Additive Attention Mechanism[J]. Gold Science and Technology, 2025, 33(5): 1020-1030.

随着矿山开采不断向深部发展,井巷作为连接地表与地下矿体的重要通道,其稳定性遇到诸多挑战。井巷围岩不仅承受高地应力的作用,而且受到围岩蠕变和采区扰动等长期因素,以及开挖卸荷和爆破振动等短期因素的共同影响(王旭锋等,2023),进而发生破裂和变形等破坏状况(康红普等,2015;谢和平等,2015;孙朝焱等,2024)。井巷失稳不仅影响矿山正常生产,还会威胁作业人员的安全。因此,对井巷变形进行准确预测并提前识别围岩破裂风险,对于矿山安全生产至关重要(芦楠楠等,2019)。

近年来,随着计算机和人工智能技术的发展,众多学者使用机器学习模型对井巷位移进行预测。

其中,一维卷积神经网络(1DCNN)、门控循环单元(GRU)、长短期记忆网络(LSTM)及其改进模型在井巷和隧道位移预测中得到广泛应用(赵楠等,2021; He et al, 2023; 崔靖奇等, 2024; 刘辉等, 2025)。此外,结合模拟退火算法的自适应神经模糊网络(SA-ANFIS)、基于鲸鱼优化算法的BP神经网络(WOA-BP)和采用粒子群优化算法的长短期记忆网络(PSO-LSTM)等方法,通过对神经网络模型的参数优化,在提升井巷位移的预测精度中也展现出优势(Xie et al, 2021; Du et al, 2022; 吴泽鑫等, 2024)。上述研究证实了使用神经网络模型对井巷位移进行预测能够取得较好的预测效果。然而,现有研究采用单一神经网络模型的方法存在一定局

收稿日期:2025-03-04;修订日期:2025-05-05

基金项目:国家自然科学基金项目“山区岩坡工程诱发灾害机理及其监测预警研究”(编号:U1602232)资助

作者简介:王龙飞(2000—),男,陕西宝鸡人,硕士研究生,从事矿山岩石力学方面的研究工作。E-mail:wanglongfei0322@foxmail.com

*通信作者:陈俊智(1974—),男,陕西榆林人,副教授,从事矿山安全和矿山信息自动化研究工作。E-mail:465424873@qq.com

© Editorial Department of Gold Science and Technology (CC BY-NC-ND)

限性。具体表现为BP神经网络易陷入局部极值,卷积神经网络在全局特征提取方面存在局限,循环神经网络易受梯度消失和爆炸问题困扰(郭风景等,2023;翟小伟等,2025)。此外,仅采用单一模型存在难以协同解析井巷位移中短期扰动响应与长期演化趋势的多时间尺度耦合特征的缺点。因此,构建具备多尺度特征解耦能力的预测模型,对于进一步提升井巷预测精度和增强模型泛化性能具有重要意义。

针对上述问题,本研究提出了一种融合加性注意力机制的时间卷积网络和长短期记忆网络(TCN-LSTM)模型。该模型结合时间卷积网络(TCN)与长短期记忆网络(LSTM)的优势,通过TCN提取时序数据的局部突变特征,利用LSTM捕捉长期依赖关系,引入加性注意力机制动态分配TCN和LSTM的输出贡献,通过多尺度特征解耦机制实现局部特征与全局记忆的优势互补,从而为复杂时序数据的精确预测提供新路径。

1 研究方法

1.1 时间卷积网络

时间卷积网络(TCN)是一种基于深度卷积神经网络的时序建模方法,能够通过层次化的时序特征提取机制和严格的因果约束,实现对时间序列数据的高效建模(周小燕等,2025)。TCN网络的核心

结构主要由扩张卷积、膨胀卷积和残差连接3个部分组成(图1)。因果卷积采用一维卷积核将数据沿时间轴滑动计算,并通过左填充的方式,使卷积核仅覆盖历史时间步,确保当前时刻的输出仅依赖于历史及当前时刻的输入,能够有效避免未来信息泄露。膨胀卷积通过指数级增长的膨胀因子控制采样间隔,能够在不增加参数量的前提下扩大其感受野,且通过浅层小膨胀因子能够有效捕捉局部突变特征。残差连接模块包含卷积层、激活函数、归一化层和跳跃连接,通过跳跃连接融合浅层细节特征和深层抽象表征,能够在缓解梯度消失和加速模型收敛的同时,增强模型对于复杂时序特征的表征能力。

膨胀因果卷积和残差连接的计算公式如下:

$$y_t = \sum_{i=0}^{k-1} w_i x_{t-di} \quad (1)$$

$$Output = Activation[F(x) + Conv_{1 \times 1}(x)] \quad (2)$$

式中: y_t 为时间步 t 的输出; w_i 为卷积核的第 i 个权重参数; x_{t-di} 为输出序列中时间步 $t-di$ 的值; d 为膨胀率,用来控制卷积核的间隔采样; k 为卷积核大小,决定覆盖时间步数量; $F(x)$ 为经过卷积层和激活函数后的特征; $Conv_{1 \times 1}$ 为 1×1 卷积调整通道数,使残差与主路径维度相匹配。

1.2 长短期记忆网络

长短期记忆网络(LSTM)作为循环神经网络的

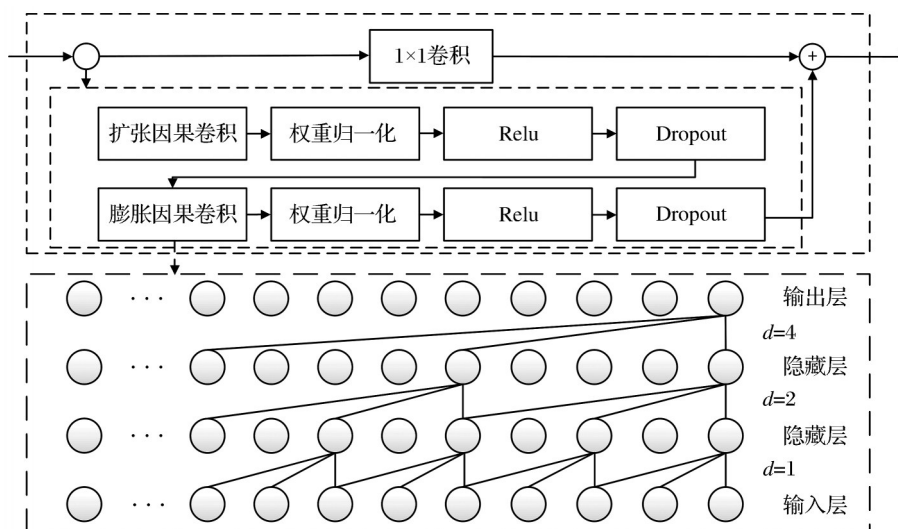


图1 TCN网络结构图

Fig. 1 Network structure diagram of TCN

重要改进架构,通过门控机制有效解决了传统RNN存在的梯度消失和梯度爆炸问题(荣光旭等,2023)。其核心优势体现在对时间序列数据长期依赖关系的建模能力,通过改进梯度反向传播过程中的信息传递效率,显著提升了对时间序列数据的处理性能。LSTM单元的核心在于构建了由遗忘门、输入门和输出门3种门控机制和具有长期记忆功能的细胞状态(图2)。其中,输入门依据当前输入向量与前一时刻状态,采用Sigmoid激活函数动态调节当前输入信息向细胞状态的传递强度;遗忘门通过Sigmoid函数计算遗忘概率,及时剔除失效信息以维持记忆系统的有效性;输出门通过Sigmoid函数和tanh双重激活机制,从更新后的细胞状态中提取与当前任务相关的特征表示。在各种机制的协同作用下,LSTM能够有效地记住对模型重要的信息,同时摒弃冗余干扰信息,在长序列数据处理任务中具有显著优势。

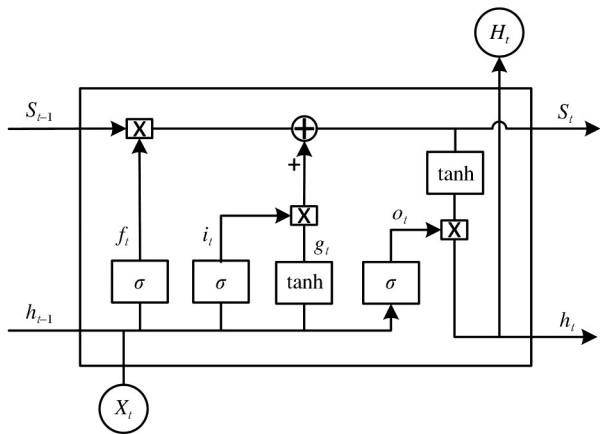


图2 LSTM网络结构图

Fig. 2 Network structure diagram of LSTM

LSTM单元计算过程如下

$$f_t = \sigma(W_{fx}x_t + W_{fh}h_{t-1} + b_f) \quad (3)$$

$$i_t = \sigma(W_{ix}x_t + W_{ih}h_{t-1} + b_i) \quad (4)$$

$$g_t = \varphi(W_{gx}x_t + W_{gh}h_{t-1} + b_g) \quad (5)$$

$$o_t = \sigma(W_{ox}x_t + W_{oh}h_{t-1} + b_o) \quad (6)$$

$$S_t = g_t \otimes i_t + S_{t-1} \otimes f_t \quad (7)$$

$$h_t = \varphi(S_t) \otimes o_t \quad (8)$$

式中: f_t, i_t, g_t, o_t, h_t 和 S_t 分别为 t 时刻遗忘门、输入门、输入节点、输出门、中间输出和状态单元的状态; h_{t-1} 为上一时刻的中间输出; $W_{fx}x_t, W_{ix}x_t, W_{gx}x_t,$

和 $W_{ox}x_t$ 分别为遗忘门、输入门、输入节点和输出门与当前时刻的输入 x 相乘的权重矩阵; $W_{fh}h_{t-1}, W_{ih}h_{t-1}, W_{gh}h_{t-1}$ 和 $W_{oh}h_{t-1}$ 分别为遗忘门、输入门、输入节点和输出门与上一时刻 $t-1$ 的中间输出 h_{t-1} 相乘的权重矩阵; b_f, b_i, b_g 和 b_o 分别为遗忘门、输入门、输入节点和输出门的偏置项; \otimes 表示向量中的元素按位相乘; σ 表示Sigmoid激活函数, φ 为tanh激活函数。

1.3 加性注意力机制

加性注意力是一种重要的注意力机制。在对时序数据进行分析时,引入加性注意力机制能够显著增强对时间序列内关键信息的捕捉能力,提高信息的利用率(Wu et al, 2021; 胡倩伟等, 2024)。在加性注意力运行过程中,将输入数据送入线性变换层通过权重矩阵转换为全局查询向量 $q \in R^d$,该过程能够高效地从输入中提取后续注意力计算需要的关键特征表示;再将查询向量和键向量 $k_j \in R^d$ 分别通过各自线性变换矩阵进行线性变换得到中间向量,随后通过可学习的参数向量,将中间特征映射为标量相似度评分,随后使用Softmax函数对评分进行归一化,生成注意力权重分布,最终依据权重对值向量 $v_j \in R^d$ 进行加权求和得到输出。

注意力权重计算公式为

$$e_i = v_a^T \tanh(W_a q + U_a k_i) \quad (9)$$

$$a_i = \frac{\exp(e_i)}{\sum_{j=1}^N \exp(e_j)} \quad (10)$$

式中: e_i 为第 i 个键向量 k_i 与查询向量 q 之间的相似度评分; $W_a \in R^{d \times d}$ 为可学习的线性变换矩阵,对查询向量 q 进行线性变换,映射至合适的特征空间; $U_a \in R^{d \times d}$ 为可学习的线性变换矩阵,对键向量 k_i 进行线性变换,使键向量 k_i 映射至与查询向量 q 变换后的相同特征空间; $v_a^T \in R^d$ 为可学习参数向量,将高维特征向量压缩为标量; N 为查询矩阵 Q 的序列长度;tanh为双曲正切函数; a_i 为第 i 个键向量 k_i 所对应的注意力权重。

1.4 融合加性注意力机制的TCN-LSTM模型

综合上述方法,本研究构建了融合TCN、LSTM和加性注意力机制的巷道变形预测组合模型(图3)。该模型采用并联设计:时序数据同步输入TCN和

LSTM双通道,其中TCN通过膨胀因果卷积提取序列的局部突变特征,LSTM通过门控机制捕捉长期演化趋势;随后,加性注意力机制对双通道输出特

征进行动态权重分配,自适应调整TCN局部感知和LSTM全局记忆的贡献度;最后通过全连接层实现多尺度信息的整合。

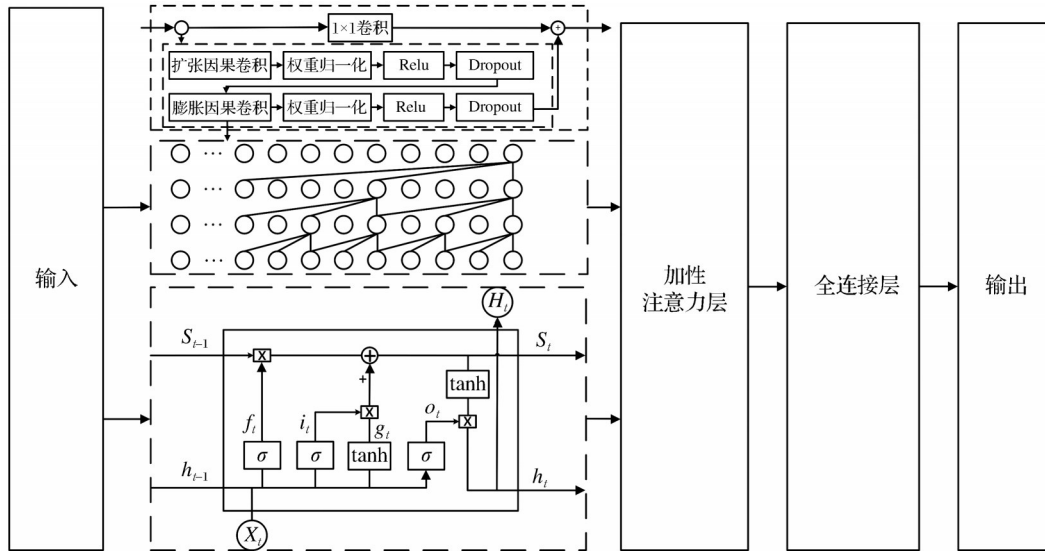


图3 融合加性注意力机制的TCN-LSTM模型网络结构图

Fig. 3 Network structure diagram of TCN-LSTM model integrated with additive attention mechanism

2 工程应用

2.1 工程概况

自走铁矿位于云南省玉溪市新平县城 120° 方向,直线距离为 29 km。矿区地处川滇 SN 向构造带南段,位于青杨断裂与通海山字型构造前缘石屏弧西翼的复合部位,处于断裂带区域,节理裂隙较为发育。区内岩体主要为沉积-变质的碳酸盐类岩石,揭露岩石以灰岩为主。矿床为以溶蚀裂隙为主的岩溶充水矿床,最低开采标高高于最低侵蚀基准面,地下水补给条件差,巷道围岩含水情况较好,对围岩稳定性无较大影响。矿山开拓方式为平硐开拓,目前 1 430 m 中段为主生产中段。为掌握次生应力随着采空区范围扩大和时间推移所发生的变化情况,从而判断围岩和间柱的稳定性,以便及时采取有效的处理与防范措施来确保生产安全,目前自走铁矿使用 NL4410 位移计等地压监测设备开展地压监测工作(图 4)。位移计量程为 100 mm,分辨率为 0.01 mm,数据可自动进行采集,采集到的数据通过各中段基站箱发送至服务器主机。



图4 监测设备现场图

Fig. 4 Field diagram of monitoring equipment

2.2 监测数据处理

本研究选取 W1430 中段的 W1430-10 和 W1430-11 监测点的位移数据进行分析(图 5)。W1430-10 监测点数据采集时段为 2022 年 10 月 20 日至 2024 年 8 月 9 日,共获取 3 310 组数据;W1430-11 监测点数据采集时间为 2022 年 11 月 7 日至 2023 年 9 月 13 日,共获取 1 555 组数据。采集频率均为每天 5 组。

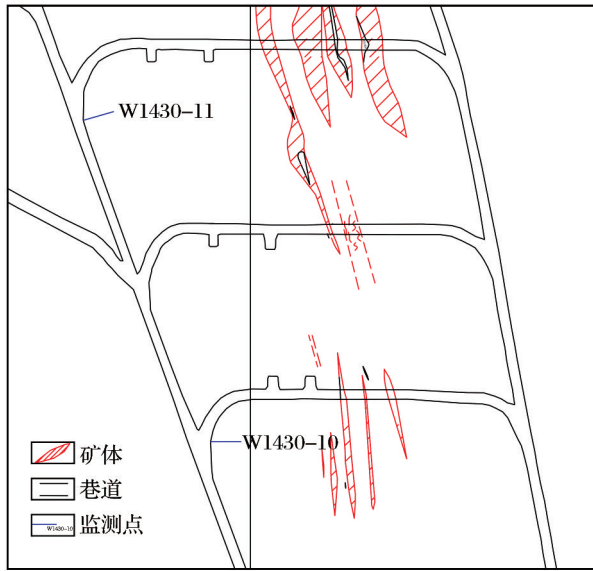


图 5 巷道监测点布置图

Fig. 5 Layout diagram of monitoring points in the roadway

由于受施工扰动、环境噪声和设备异常等因素的影响,监测数据存在缺失值和异常值现象。由于时间序列分析需要保证数据点的等时间间隔特性,本研究在分析前对原始数据进行预处理,具体处理内容包括异常值剔除、缺失值插补和数据归一化,以确保时序数据的完整性和可靠性。

在数据处理阶段,首先对原始数据的异常值进行剔除。本试验使用 Hampel 滤波器对异常值进行筛选(王文强等, 2024),其核心原理是通过滑动窗口内的中位数和中位数绝对偏差(MAD)识别异常点。当时间窗口为 k 时,对于时间序列中的每个点 x_i ,选取其前后共 $2k+1$ 个点组成窗口序列 $W_i = \{x_{i-k}, x_{i-k+1}, \dots, x_i, \dots, x_{i+k}\}$,每个窗口序列的中值和 MAD 的计算公式为

$$M_i = \text{median}(W_i) \tag{11}$$

$$MAD_i = 1.4826 \times \text{median} \{(W_i - M_i)\} \tag{12}$$

式中: M_i 为窗口序列 W_i 的中值; MAD_i 为每个 x_i 对应窗口 M_i 的中位数绝对偏差;常量 1.4826 为该偏移量在数据呈正态分布时等于标准偏移量的情况;本次时间窗口大小为 5。

异常值判定公式为

$$|x_i - M_i| \geq \lambda \times MAD_i \tag{13}$$

式中: λ 通常取 3。如式(13)成立,则认为 x_i 为异常点并将其剔除。

对于剔除异常值后的时间序列,使用 3 次样条插值法对缺失处进行插值,3 次样条插值法通过构造分段 3 次多项式 $S(x)$ 来拟合给定数据点 $\{(x_i, y_i)\}_{i=0}^n$,其中 $(x_0 < x_1 < \dots < x_n)$,且保证函数 $S(x)$ 及其一阶和二阶导数在节点数连续,其数学表达式为

$$S_i(x) = y_i + \left[\frac{y_{i+1} - y_i}{h_i} - \frac{h_i}{6} (2M_i + M_{i+1}) \right] (x - x_i) + \frac{M_i}{2} (x - x_i)^2 + \frac{M_{i+1} - M_i}{6h_i} (x - x_i)^3 \tag{14}$$

$$h_{i-1}M_{i-1} + 2(h_{i-1} + h_i)M_i + h_iM_{i+1} = 6 \left(\frac{y_{i+1} - y_i}{h_i} - \frac{y_i - y_{i-1}}{h_{i-1}} \right) \tag{15}$$

式中: $h_i = x_{i+1} - x_i$ 为区间步长; $M_i = S''(x_i)$ 为节点处的二阶导数,对于自然样条,边界条件为 $M_0 = M_n = 0$ 。若已知端点一阶导数 f'_0 和 f'_n ,采用固定边界条件,固定边界形式可表示为

$$2h_0M_0 + h_0M_1 = 6 \left(\frac{y_1 - y_0}{h_0} - f'_0 \right) \tag{16}$$

$$h_{n-1}M_{n-1} + 2h_{n-1}M_n = 6 \left(f'_n - \frac{y_n - y_{n-1}}{h_{n-1}} \right) \tag{17}$$

通过追赶法求解上述方程组后,将 M_i 代入可确定多项式系数。使用 3 次样条插值法可确保插值曲线更光滑,从而规避高阶多项式插值的龙格现象(Song et al, 2023)。处理后的位移监测数据如图 6 所示。

随后对处理后的数据进行归一化操作。归一化处理公式为

$$X'_i = \frac{2(X_i - X_{\min})}{X_{\max} - X_{\min}} - 1 \tag{18}$$

式中: X_i 为原始输入数据; X'_i 为经过归一化处理后的数据; X_{\max} 和 X_{\min} 分别为训练样本的最大值和最小值。

经过归一化处理,将数据按照 8:1:1 的比例划分为训练集、验证集和测试集,输入到融合加性注意力机制的 TCN-LSTM 模型进行训练(模型巷道

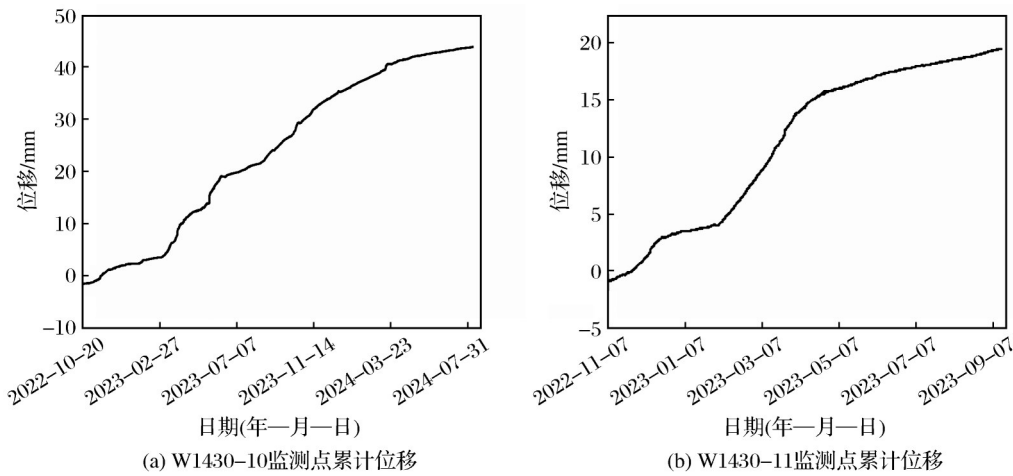


图6 W1430-10和W1430-11监测点累计位移
Fig. 6 Cumulative displacement of W1430-10 and W1430-11 monitoring points

变形预测流程图见图7)。

2.3 模型参数设置

模型具体参数配置如下:输入时间窗口步长设置为35,TCN模块由单个残差单元构成,采用大小为3的一维因果卷积核,并在层间设置0.1的Dropout率防止过拟合;LSTM模块采用双层堆叠;加性注意力机制通过32维隐藏层动态融合TCN提取的局部特征和LSTM生成的全局时序状态,强化关键信息的表征能力。训练过程采用Adam优化器进行200轮迭代,以均方误差(MSE)作为损失函数,并应用阈值为0.5梯度裁剪以稳定训练过程。为提升模型泛化能力,设置早停机制,在验证集损失连续30轮未下降时终止训练。通过网格搜索对模型学习率、LSTM隐藏层维度、TCN通道数和批量大小等超参数进行优化,其中学习率的取值范围为 $[0.0001, 0.0005, 0.001, 0.005, 0.01]$,LSTM隐藏层大小取值范围为 $[32, 64, 128]$,TCN卷积通道数取值范围为 $[16, 32, 64]$,批量大小取值范围为 $[16, 32, 64]$ 。确定最优参数组合如下:学习率为0.001,LSTM隐藏层维度为64,TCN卷积通道数为32,批量大小为32。

2.4 模型权重可视化

融合加性注意力机制的TCN-LSTM模型权重贡献分析结果如图8和图9所示。在W1430-10监测点采集的时间序列预测分析中,LSTM模块在整个监测周期内均持续维持高权重贡献(权重 >0.8),这一结果表明,该时间序列具有显著的长周期趋势

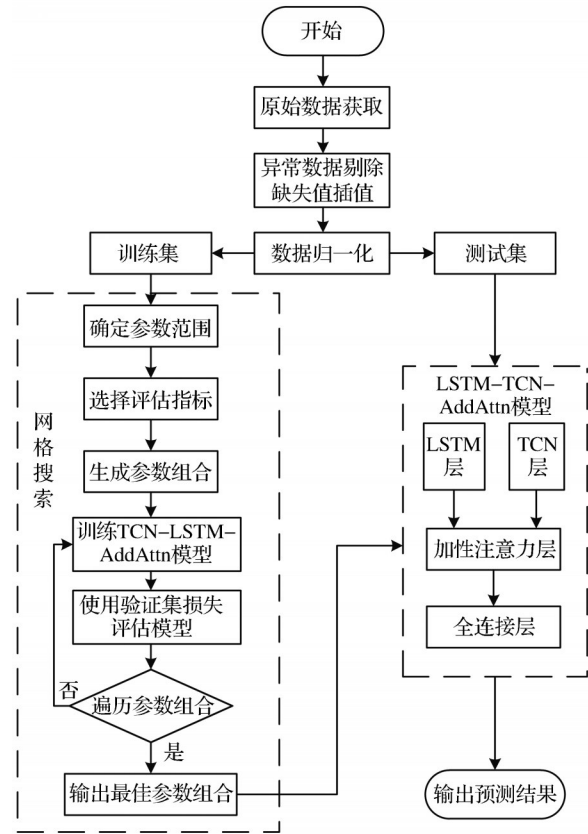


图7 融合加性注意力机制的TCN-LSTM模型
巷道变形预测流程图

Fig. 7 Flow chart of TCN-LSTM model integrated with additive attention mechanism for roadway deformation prediction

特征。由此可知,在整个周期内巷道变形受岩体蠕变机制主导,其力学响应表现出明显的长期依赖

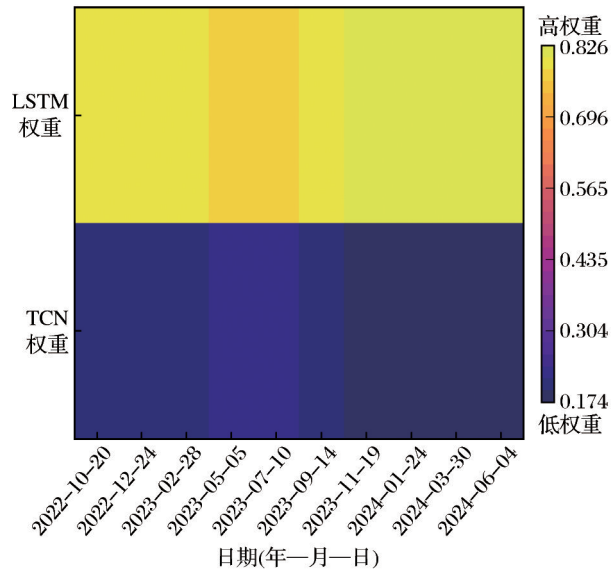


图8 W1430-10 监测点模型权重贡献热力图

Fig. 8 Heatmap of model weight contributions at monitoring points W1430-10

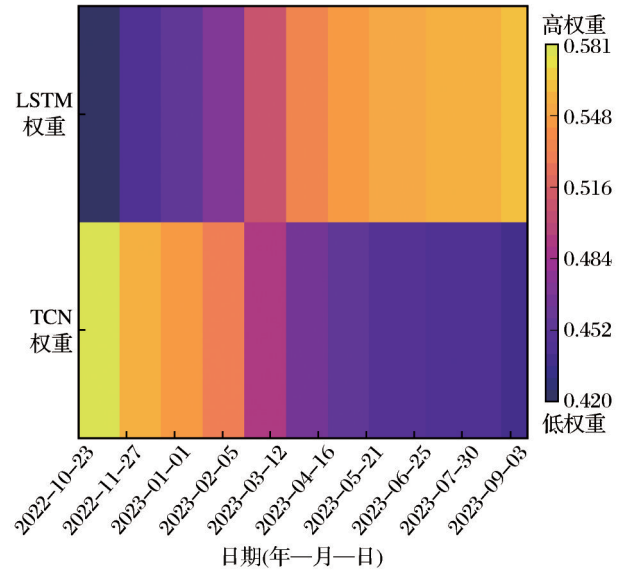


图9 W1430-11 监测点模型权重贡献热力图

Fig. 9 Heatmap of model weight contributions at monitoring points W1430-11

性,短期扰动因素并未对巷道变形产生显著影响。

在对 W1430-11 监测点进行预测的过程中,TCN 和 LSTM 模块的权重贡献呈现出动态博弈的差异化特征。在监测序列初始阶段,TCN 网络凭借其膨胀因果卷积结构的优势,精准捕捉到位移数据中的局部突变特征,在模型输出中占据主导;随着时间的推移,LSTM 模块的权重贡献开始逐步攀升,表明巷道变形后期逐渐回归到以岩体蠕变为主导,系统的长期记忆效应在变形预测中的控制作用显著提升。这种动态权重分配机制验证了融合模型架构在协同捕捉短期事件响应和长期趋势演化方面具有优越性。

3 预测结果分析

为验证融合加性注意力机制的 TCN-LSTM 模型的预测稳定性和泛化性能,基于不同中段的 W1430-10、W1430-11、W1480-7 和 W1530-11 监测点数据,将测试集的预测结果与 TCN、LSTM 和 TCN-LSTM 模型的预测误差和评价指标进行对比分析。

3.1 预测误差对比

不同模型的预测误差分析结果如图 10 所示。通过预测误差对比,得出如下结论:在不同中段 4 个监测点的预测任务中,单一 LSTM 和 TCN 模型

的预测误差中位数较高且箱线波动范围较大,说明二者对于复杂时序特征的捕捉能力有限;TCN-LSTM 模型通过结合 TCN 模型的局部特征提取能力和 LSTM 模型的长期依赖能力,降低了预测的误差中位数,同时减小了波动范围;融合加性注意力机制的 TCN-LSTM 模型 (TCN-LSTM-AddAttn) 的预测误差中位数最低,且箱体波动范围更小,说明加性注意力机制通过动态分配特征权重,增强了模型对于关键时序特征的敏感性,且强化了模型的泛化能力和稳定性。

3.2 预测评价指标

本研究采用平均绝对误差 (MAE)、均方根误差 (RMSE)、校正决定系数 (R^2_{adj}) 和对称平均绝对百分比误差 (SMAPE) 4 项指标对模型性能进行评价。

R^2_{adj} 通过引入自由度惩罚项,能够在一定程度上缓解模型因复杂度增加导致的传统 R^2 虚高的问题,从而更合理地平衡模型复杂度与拟合能力的关系 (Akossou et al, 2013; Chicco et al, 2021)。

SMAPE 通过修正传统平均绝对百分比误差 (MAPE) 的分母不对称性,解决了后者因预测值小于真实值时导致误差被过度放大造成的评估偏差问题,确保误差评估的公平性和方向无关性 (Ribeiro et al, 2023)。

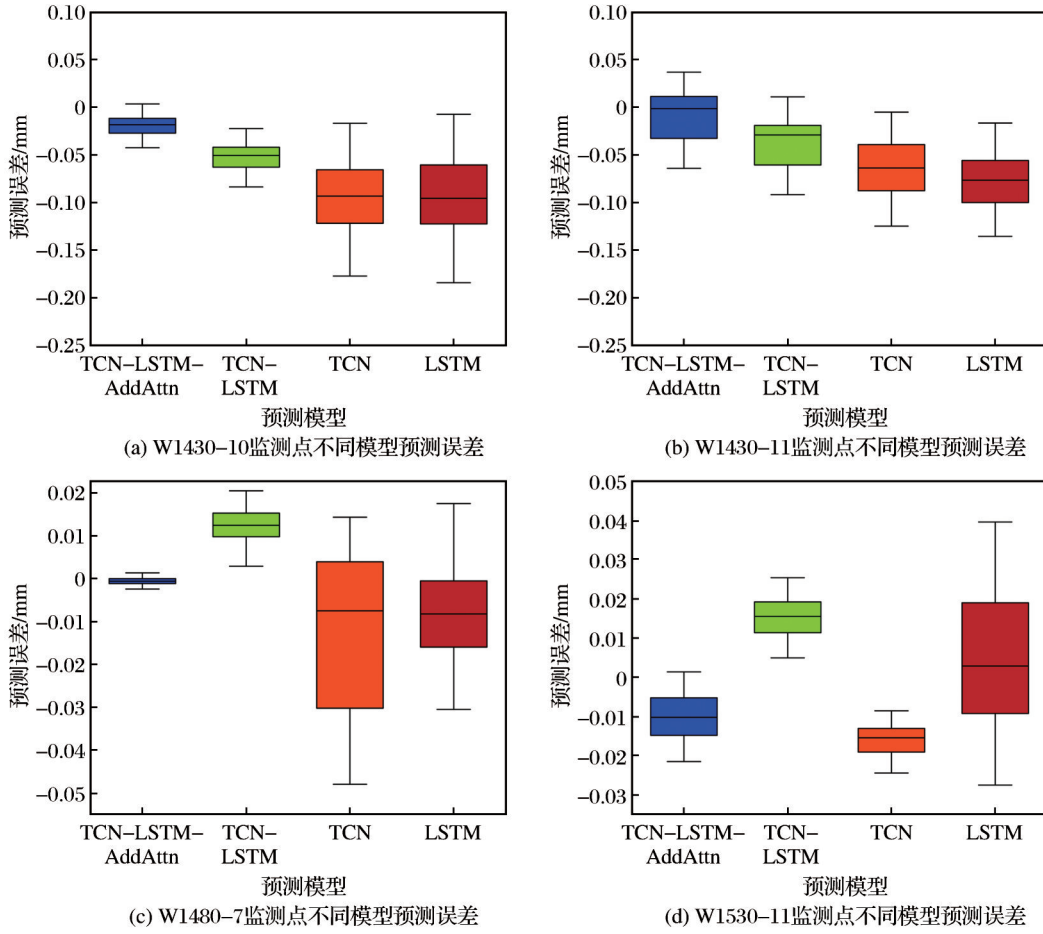


图 10 不同模型预测误差对比

Fig. 10 Prediction error comparison of different models

MAE 、 $RMSE$ 、 R^2_{adj} 和 $SMAPE$ 指标的计算公式分别表示为

$$MAE = \frac{1}{N} \sum_{i=1}^N |Y_i - \hat{Y}_i| \quad (19)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N [Y_i - \hat{Y}_i]^2} \quad (20)$$

$$R^2 = 1 - \frac{\sum_{i=1}^N [\hat{Y}_i - Y_i]^2}{\sum_{i=1}^N [\bar{Y}_i - Y_i]^2} \quad (21)$$

$$R^2_{adj} = 1 - \frac{(1 - R^2)(N - 1)}{N - P - 1} \quad (22)$$

$$SMAPE = \frac{100\%}{N} \sum_{i=1}^N \frac{|\hat{Y}_i - Y_i|}{(|\hat{Y}_i| + |Y_i|)/2} \quad (23)$$

式中： Y_i 为模型预测值； \hat{Y}_i 为实际值； N 为样本数量； P 为样本特征数量。

性能评价指标见表 1。 MAE 、 $RMSE$ 和 $SMAPE$ 指标越小,表明建立的神经网络模型对数据的预测精度越高; R^2_{adj} 越接近 1,说明模型对数据的拟合效果越好。

通过对表 1 中的 4 项指标 (MAE 、 $RMSE$ 、 $SMAPE$ 和 R^2_{adj}) 进行分析可知:融合加性注意力机制的 TCN-LSTM 模型,即 TCN-LSTM-AddAttn 模型在各项指标上均优于其他模型。

在 W1430-10 监测点数据集上,相较于 TCN-LSTM、TCN 和 LSTM 模型,TCN-LSTM-AddAttn 模型的 MAE 分别降低了 47.95%、68.09% 和 68.50%, $RMSE$ 分别降低了 37.34%、60.00% 和 61.21%, $SMAPE$ 分别降低了 47.91%、68.12% 和 68.48%;在 W1430-11 监测点数据集上,TCN-LSTM-AddAttn 模型的 MAE 分别降低了 28.79%、55.02% 和 61.71%, $RMSE$ 分别降低 27.52%、50.27% 和 56.27%, $SMAPE$

表1 各模型性能评价指标对比

Table 1 Comparison of performance evaluation indexes of each model

监测点	预测模型	MAE	RMSE	SMAPE	R^2_{adj}
W1430-10	TCN-LSTM-AddAttn	0.0292	0.0396	0.0337	0.9861
	TCN-LSTM	0.0561	0.0632	0.0647	0.9646
	TCN	0.0915	0.0990	0.1057	0.9131
	LSTM	0.0927	0.1021	0.1069	0.9077
W1430-11	TCN-LSTM-AddAttn	0.0282	0.0366	0.0738	0.9799
	TCN-LSTM	0.0396	0.0505	0.1035	0.9616
	TCN	0.0627	0.0736	0.1636	0.9186
	LSTM	0.0737	0.0837	0.1924	0.8947
W1480-7	TCN-LSTM-AddAttn	0.0057	0.0082	0.0049	0.9805
	TCN-LSTM	0.0122	0.0133	0.0105	0.9487
	TCN	0.0151	0.0212	0.0130	0.8689
	LSTM	0.0134	0.0166	0.0115	0.9201
W1530-11	TCN-LSTM-AddAttn	0.0100	0.0111	0.0340	0.9808
	TCN-LSTM	0.0154	0.0159	0.0523	0.9604
	TCN	0.0186	0.0220	0.0628	0.9246
	LSTM	0.0254	0.0309	0.0862	0.8518

分别降低 28.70%、54.89% 和 61.64%；在 W1480-7 监测点数据集上，TCN-LSTM-AddAttn 模型的 MAE 分别降低 53.28%、62.25% 和 57.46%，RMSE 分别降低 38.35%、61.32% 和 50.60%，SMAPE 分别降低 53.33%、62.31% 和 57.39%；在 W1530-11 监测点数据集上，相较于 TCN-LSTM、TCN 和 LSTM 模型，TCN-LSTM-AddAttn 模型的 MAE 分别降低了 35.06%、46.24% 和 60.63%，RMSE 分别降低了 30.19%、49.55% 和 64.08%，SMAPE 分别降低了 34.99%、45.86% 和 60.56%，并且 TCN-LSTM-AddAttn 模型在 4 个数据集上的 R^2_{adj} 最接近 1，充分表明融合加性注意力机制的 TCN-LSTM 模型预测精度最高、误差分布最集中，同时具有最强的数据解释能力和泛化能力。

4 结论

本研究构建了融合加性注意力机制的 TCN-LSTM 模型，使用自走铁矿不同中段的监测点位移监测数据进行预测，通过预测误差分析和 4 项评价指标对比，得出如下结论：

(1) 融合加性注意力机制的 TCN-LSTM 模型通过并联架构集成 TCN 的局部特征提取能力和 LSTM

的长期依赖建模优势，并结合加性注意力机制动态优化特征权重分配，显著提升了复杂时序数据的预测精度。在不同中段 4 个监测点的预测任务中，该模型的 MAE、RMSE 和 SMAPE 较 TCN-LSTM 模型分别降低了 28.79%~53.28%、27.52%~38.35% 和 28.70%~53.33%，且 R^2_{adj} 均达到 0.97 以上，验证了模型具有优异的非线性拟合能力和泛化能力。

(2) 针对 W1430-10、W1430-11、W1480-7 和 W1530-11 监测点预测任务中，融合加性注意力机制的 TCN-LSTM 模型预测结果与巷道位移监测值高度同步，预测误差波动范围在 0.1 mm 以内，证明模型能够精准捕捉围岩变形的局部突变特征和长期演化趋势。

(3) 通过使用融合加性注意力机制的 TCN-LSTM 模型对未来变形量进行精准预测，结合变形速率分析，有望提前识别潜在高风险变形区域，引导施工人员进行超前支护，从而为矿山巷道稳定性维护和安全预警提供技术支持。

参考文献 (References) :

Akossou A Y J, Palm R, 2013. Impact of data structure on the estimators R-square and adjusted R-square in linear regression [J]. International Journal of Mathematics and Computation, 20: 84-93.

Chicco D, Warrens M J, Jurman G, 2021. The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation [J]. PeerJ Computer Science, 7: e623.

Du B, Yi H H, Yang F, 2022. Deformation prediction and analysis of soft rock roadway with high altitude and large buried depth based on particle swarm optimization LSTM model [J]. Mobile Information Systems, (1): 5907051.

He Y C, Chen Q N, 2023. Construction and application of LSTM-based prediction model for tunnel surrounding rock deformation [J]. Sustainability, 15(8): 6877.

Ribeiro M H D M, da Silva R G, Ribeiro G T, et al, 2023. Cooperative ensemble learning model improves electric short-term load forecasting [J]. Chaos, Solitons and Fractals, 166: 112982.

Song H, Yuan R, Lü Y, et al, 2023. Cubic spline interpolation-based refined composite multiscale dispersion entropy and its application to bearing fault identification [J]. Structural

- Health Monitoring, 22(4):2651-2673.
- Wu C H, Wu F Z, Qi T, et al, 2021. Fastformer: additive attention can be all you need[J]. Arxiv Preprint: 2108.09084.
- Xie C Y, Nguyen H, Bui X N, et al, 2021. Predicting roof displacement of roadways in underground coal mines using adaptive neuro-fuzzy inference system optimized by various physics-based optimization algorithms [J]. Journal of Rock Mechanics and Geotechnical Engineering, 13 (6) : 1452-1465.
- 崔靖奇, 吴顺川, 程海勇, 等, 2024. 滇中引水软岩隧洞围岩位移时序预测[J]. 清华大学学报(自然科学版), 64(7): 1215-1225. Cui Jingqi, Wu Shunchuan, Cheng Haiyong, et al, 2024. Time series prediction of the surrounding rock displacement of a soft rock tunnel in the Central Yunnan Water Diversion Project[J]. Journal of Tsinghua University (Science and Technology), 64(7): 1215-1225.
- 郭风景, 王斌, 贾澎涛, 等, 2023. 基于梯度提升决策树的瓦斯浓度在线预测[J]. 中国煤炭, 49(10): 61-67. Guo Fengjing, Wang Bin, Jia Pengtao, et al, 2023. Online prediction of gas concentration based on gradient boosting decision tree[J]. China Coal, 49(10): 61-67.
- 胡倩伟, 王秀青, 安阳, 等, 2024. A2 former 模型在时间序列预测中的应用研究[J]. 人工智能科学与工程, 49(1): 41-50. Hu Qianwei, Wang Xiuqing, An Yang, et al, 2024. A study on the application of A² former model in time series forecasting [J]. Artificial Intelligence Science and Engineering, 49(1): 41-50.
- 康红普, 范明建, 高富强, 等, 2015. 超千米深井巷道围岩变形特征与支护技术[J]. 岩石力学与工程学报, 34(11): 2227-2241. Kang Hongpu, Fan Mingjian, Gao Fuqiang, et al, 2015. Deformation and support of rock roadway at depth more than 1 000 meters [J]. Chinese Journal of Rock Mechanics and Engineering, 34(11): 2227-2241.
- 刘辉, 李国强, 朱晓峻, 等, 2025. 基于深度学习的井筒变形预测模型与应用[J]. 煤炭学报, 50(2): 732-747. Liu Hui, Li Guoqiang, Zhu Xiaojun, et al, 2025. Exploration and application of deep learning based wellbore deformation forecasting model [J]. Journal of China Coal Society, 50(2): 732-747.
- 芦楠楠, 陈永珩, 马占国, 等, 2019. 深部地下巷道围岩结构监测信息化技术研究[J]. 工矿自动化, 45(6): 1-9. Lu Nannan, Chen Yongheng, Ma Zhanguo, et al, 2019. Research on informatization technology for roadway surrounding rock structure monitoring in deep underground [J]. Industry and Mine Automation, 45(6): 1-9.
- 荣光旭, 李宗洋, 2023. CNN-LSTM模型在边坡可靠度分析中的应用[J]. 黄金科学技术, 31(4): 613-623. Rong Guangxu, Li Zongyang, 2023. Application of CNN-LSTM model in slope reliability analysis [J]. Gold Science and Technology, 31(4): 613-623.
- 孙朝焱, 陈从新, 夏开宗, 等, 2024. 金属矿山复杂采空区围岩破坏及岩层移动机制研究[J]. 岩石力学与工程学报, 43(3): 670-682. Sun Chaoyi, Chen Congxin, Xia Kaizong, et al, 2024. Mechanism of surrounding rock failure and strata movement in complex goaf of metal mines [J]. Chinese Journal of Rock Mechanics and Engineering, 43(3): 670-682.
- 王文强, 燕波, 张俊儒, 2024. 基于 Hampel-DeepAR 模型的浅埋隧道洞口塌方处治变形预测研究[J]. 自动化与仪器仪表, (4): 40-46. Wang Wenqiang, Yan Bo, Zhang Junru, 2024. Prediction model for deformation of shallow buried tunnel portal collapse treatment based on Hampel-DeepAR model [J]. Automation and Instrumentation, (4): 40-46.
- 王旭锋, 陈旭阳, 王纪尧, 等, 2023. 平顶山矿区深部软岩巷道围岩蠕变破坏机制及控制[J]. 采矿与安全工程学报, 40(6): 1139-1150. Wang Xufeng, Chen Xuyang, Wang Jiyao, et al, 2023. Creep failure mechanism and control of the deep soft rock roadway in Pingdingshan mining area [J]. Journal of Mining and Safety Engineering, 40(6): 1139-1150.
- 吴泽鑫, 张成良, 张华超, 等, 2024. 基于 WOA 优化神经网络的斜坡道拱顶沉降预测研究[J]. 有色金属工程, 14(4): 150-160, 174. Wu Zexin, Zhang Chengliang, Zhang Huachao, et al, 2024. Study on optimization of neural network based on whale algorithm for prediction of vault subsidence in slope [J]. Nonferrous Metals Engineering, 14(4): 150-160, 174.
- 谢和平, 高峰, 鞠杨, 2015. 深部岩体力学研究探索[J]. 岩石力学与工程学报, 34(11): 2161-2178. Xie Heping, Gao Feng, Ju Yang, 2015. Research and development of rock mechanics in deep ground engineering [J]. Chinese Journal of Rock Mechanics and Engineering, 34(11): 2161-2178.
- 翟小伟, 王博, 蒋学明, 等, 2025. 基于 DBO-BP 神经网络的采空区煤温预测方法及应用[J]. 煤矿安全, 56(4): 75-81. Zhai Xiaowei, Wang Bo, Jiang Xueming, et al, 2025. Prediction method and application of coal temperature in goaf based on DBO-BP neural network [J]. Safety in Coal Mines, 56(4): 75-81.
- 赵楠, 李洁, 2021. 基于 LSTM-SVM 的隧道围岩位移预测[J]. 公路, 66(6): 404-407. Zhao Nan, Li Jie, 2021. Displacement prediction of tunnel surrounding rock based on

LSTM-SVM[J].Highway,66(6):404-407.
周小燕,李双平,冉鲁光,等,2025.基于深度学习的大坝边坡深部变形时空预测模型研究[J].中国农村水利水电,(7):182-187,195. Zhou Xiaoyan, Li Shuangping, Ran

Luguang, et al, 2025. Research on spatiotemporal prediction model for deep deformation of dam slope based on deep learning [J]. China Rural Water and Hydropower, (7):182-187,195.

TCN-LSTM Roadway Deformation Prediction Model Based on Additive Attention Mechanism

WANG Longfei, CHEN Junzhi

Faculty of Land and Resources Engineering, Kunming University of Science and Technology, Kunming 650093, Yunnan, China

Abstract: As mining activities continue to intensify, underground roadways encounter unprecedented stability challenges due to increasingly complex stress environments, dynamic geological variations, and anthropogenic disturbances. Accurate prediction of roadway deformation is essential for ensuring mining safety and optimizing operational layouts. However, conventional single neural network models often struggle to effectively capture both abrupt local features and long-term evolutionary trends in nonlinear displacement time series. To address these limitations, this study introduces an innovative hybrid model combining a Temporal Convolutional Network and Long Short-Term Memory (TCN-LSTM-AddAttn) architecture, enhanced with an additive attention mechanism, to achieve high-precision predictions of roadway deformation. The proposed architecture employs a parallel framework to leverage the strengths of both TCN and LSTM. An additive attention mechanism is incorporated to dynamically prioritize critical patterns from both networks. The model employs learnable parameters to calculate feature similarity and utilizes the Softmax function to generate normalized weights, facilitating the adaptive fusion of multi-scale representations. Validation of the model is conducted using displacement data from four monitoring points(W1430-10, W1430-11, W1480-7, W1530-11) across various roadways in the Yunnan Zizou iron mine. Data preprocessing involves the removal of outliers using a Hampel filter, interpolation of missing values via cubic spline, and min-max normalization to standardize input scales. The processed data are divided into training, validation, and test sets in an 8:1:1 ratio. Hyperparameters, including TCN channels (32), LSTM hidden dimensions (64), batch size (32), and learning rate (0.001), are optimized through grid search to ensuring generalization across diverse mining scenarios. Experimental results indicate that the TCN-LSTM-AddAttn model outperforms standalone TCN, LSTM, and the TCN-LSTM hybrid models. In the case of W1430-10, the TCN-LSTM-AddAttn model demonstrates a Mean Absolute Error (MAE) of 0.0292 mm, representing a 47.95% reduction compared to the TCN-LSTM model. Additionally, it achieves a Root Mean Square Error (RMSE) of 0.0396 mm, marking a 37.34% reduction, and a Symmetric Mean Absolute Percentage Error (SMAPE) of 0.0337, indicating a 47.91% reduction. The Adjusted R^2 (R_{adj}) value of 0.9861 suggests near-perfect prediction accuracy. For W1430-11, the model records an MAE of 0.0282 mm (28.79% reduction), an RMSE of 0.0366 mm (27.52% reduction), a SMAPE of 0.0738 (28.70% reduction), and an R_{adj} of 0.9799. Comparable improvements are noted for W1480-7 and W1530-11, with prediction errors consistently remaining below 0.1 mm. By incorporating multi-scale feature decoupling and dynamic weighting, the proposed model offers robust technical support for assessing mine roadway stability, identifying risk zones, and providing early safety warnings.

Key words: temporal convolutional network; long short-term memory network; additive attention mechanism; time series prediction; mine roadway; prediction of roadway deformation; machine learning