DOI:10.20040/j.cnki.1000-7709.2025.20241202

基于注意力机制的CNN-BiLSTM过闸流量预测模型

何立新[a,b],沈正华[a,b],张 峥[a,b],雷晓辉[a,b]

(河北工程大学 a.水利水电学院;b. 河北省智慧水利重点实验室,河北 邯郸 056038)

摘要:在明渠调水工程中,精确掌握过闸流量对于提升渠道调控效率、保障输水系统安全等问题意义重大。为提高过闸流量预测精度,提出一种基于注意力机制,融合卷积神经网络(CNN)和双向长短期记忆网络(BiLSTM)的过闸流量预测模型。以洺河渡槽节制闸为例,选取其1年时间尺度的实际数据为模型输入,模型首先将输入数据标准化,再利用CNN提取特征信息,经过BiLSTM捕获序列数据中的前后向依赖关系,最后通过注意力机制评估信息的重要程度,对特征参数进行加权处理,实现对过闸流量的预测。结果表明,所建模型相比于传统的BP-NN、SVR、LSTM等预测模型具有更好的预测结果,模型的平均绝对误差、平均绝对百分比误差、均方根误差和决定系数分别为3.682、0.018、4.661、0.983,可为工程实践提供参考。

关键词:过闸流量预测;BiLSTM;注意力机制;神经网络

中图分类号:TV663 **文献标志码**:A **文章编号**:1000-7709(2025)05-0135-04

1 引言

明渠调水工程运行调度的核心在于渠道的水力控制,节制闸作为调水工程中的主要建筑物之一,可通过调节闸门开度达到准确调节水位和控制流量的效果[1]。精确测量过闸流量对于水利工程的安全运行、水资源的科学调度及工程运行效益等均有重要意义。目前,基于传统水力学法和量纲分析法建立闸门过流流量计算经验公式或建立流量预测模型的研究成果较丰富[2-4]。近年来,随着人工智能的发展,神经网络和深度学习法在水利工程中的应用也较广泛,如郭世圆等[5]对比了分别使用LSTM神经网络和圣维南方程组数值计算方法对明渠流量预测的结果,结果表明LSTM神经网络的准确性与效率均较优,验证了人工智能方法的可行性;王艺霖等[6]采用LSTM神经网络建立了弧形闸门的过流能力预测模型,结果表明LSTM神经网络预测效果整体优于量纲分析法;李珊珊等[7]采用支持向量机、广义回归神经网络、核函数极限学习机等多种机器学习方法预测弧形闸门淹没流态的流量,结果表明核函数极限学习机的性能最优,丰富了校核闸孔流量的方法;MUNDHER Y Z等[8]使用Bagging回归树(BGRT)、双向循环神经网络(Bi-RNN)等多种机器学习算法预测不同出流工况的闸门流量系数,总体上效果较好。鉴此,本文搭建了基于注意力机制的CNN-BiLSTM过闸流量预测模型,对收集到的原始闸渠水情数据进行预处理,消除异常数据的影响;再经过数据标准化等操作,使用CNN提取水情数据的空间特征,然后将提取的特征序列输入到BiLSTM层中,分别从前向和后向挖掘当前时刻数据同过去和未来时刻数据的内在联系,并融合注意力机制对各特征赋予不同的权重,以增强重要特征,实现对过闸流量的精准预测。

2 模型介绍

2.1 基于注意力机制的CNN-BiLSTM模型

基于注意力机制的CNN-BiLSTM的过闸流量组合预测模型预测流程见图1。

(1)卷积神经网络(CNN)模型。CNN的原理是利用其中卷积层中的卷积核对数据进行特征提取,利用池化层压缩提取的特征并生成更重要的特征信息,从而提高模型的泛化能力[9]。其结

收稿日期:2024-07-01,**修回日期**:2024-08-23

基金项目:国家重点研发计划(2023YFC3209400)

作者简介:何立新(1966-),男,博士、教授、博导,研究方向为智慧水务与水文水资源,E-mail:helixin@hebeu.edu.cn

通讯作者:张峥(1981-),男,博士、副教授,研究方向为智慧水利,E-mail:zhangzheng@hebeu.edu.cn

中国知网 https://www.cnki.net

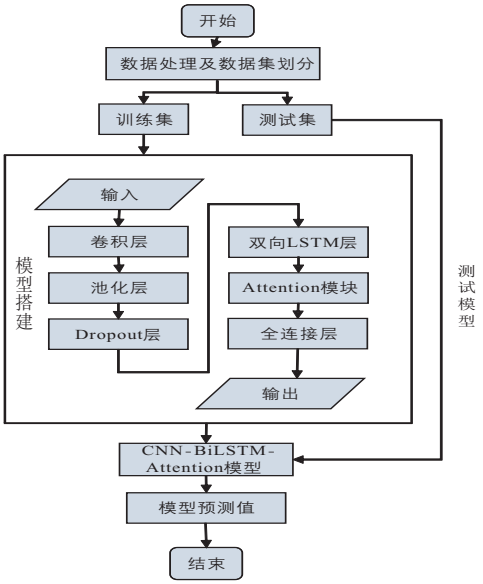

图 1 融合注意力机制的 CNN-BiLSTM 模型流程图

Fig. 1 Flow chart of CNN-BiLSTM model with fusion attention mechanism

构见图 2,卷积计算公式为:

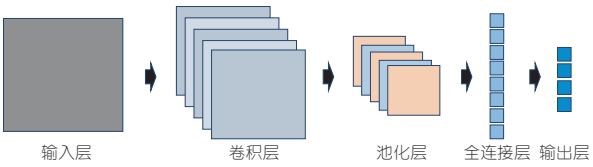

图 2 卷积神经网络结构图

Fig. 2 Convolutional neural network structure diagram

$$Y=f(X\boldsymbol{W}+b) \tag{1}$$

式中,Y、f、X、$\boldsymbol{W}$、b 分别为输出、激活函数、输入、权重矩阵、偏置。

(2)BiLSTM 神经网络模型。LSTM 模型主要引入了“门”的机制,包含遗忘门、输入门和输出门。通过这些门的组合,LSTM 能够在序列中有效地存储和检索信息,从而更好地捕捉长期依赖关系[10]。其原理为:

$$f_t=\sigma(\boldsymbol{W}_f\cdot[h_{t-1},x_t]+b_f) \tag{2}$$

$$i_t=\sigma(\boldsymbol{W}_i\cdot[h_{t-1},x_t]+b_i) \tag{3}$$

$$\bar{c}_t=\tanh(\boldsymbol{W}_c\cdot[h_{t-1},x_t]+b_c) \tag{4}$$

$$o_i=\sigma(W_0\times[h_{t-1},h_t]+b_0) \tag{5}$$

式中,f_t、i_t、o_i 分别为遗忘门输出(决定信息的去留)、输入门输出、输出门输出;σ 为 sigmoid 激活函数;$\boldsymbol{W}$、b 分别为各门的权重项、偏置项;$\boldsymbol{W}_0$ 为输出门权重矩阵;h_{t-1}、h_t、x_t 分别为 $t-1$ 时刻输出、t 时刻输出、t 时刻输入;$\bar{c}_t$ 用于更新信息状态,通过遗忘门舍弃部分信息,通过输入门添加新信息,得到新状态信息;b_0 为输出门的偏置项。

双向长短期记忆网络(BiLSTM)是一种改进的循环神经网络,相比于 LSTM,BiLSTM 在理解和表示序列数据方面更具有优势,能够通过前向和后向两个 LSTM 网络处理时间序列数据[11]。其表达式为:

$$y_t=\sigma(\boldsymbol{W}_y\cdot[\overrightarrow{h}_t,\overleftarrow{h}_t]+b_y) \tag{6}$$

式中,$\overrightarrow{h}_t$、$\overleftarrow{h}_t$、y_t 分别为 t 时刻前后向的隐藏层状态和输出;$\boldsymbol{W}_y$、b_y 分别为权重矩阵、偏置项。

BiLSTM 神经网络原理见图 3。

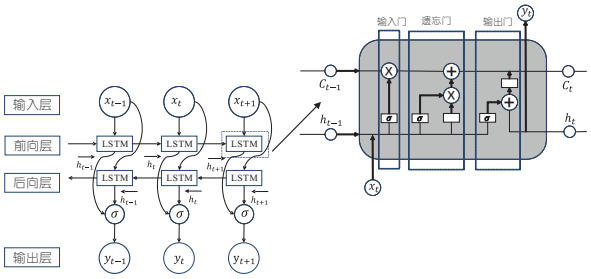

图 3 BiLSTM 神经网络原理图

Fig. 3 Schematic diagram of BiLSTM neural network

(3)注意力机制。注意力机制是深度学习领域中的一种重要技术,原理是对不同的输入特征赋予不同的权重,通常使用打分函数 score_i 增强重要特征[12],注意力向量是比较当前目标隐藏状态与每个源隐藏状态所得,其计算公式为:

$$\text{score}_i=\tanh(\boldsymbol{W}h_i+b) \tag{7}$$

$$a_i=\frac{\exp(\text{score}_i)}{\sum_{j=1}^{T}\exp(\text{score}_j)} \tag{8}$$

$$c_t=\sum_{i=1}^{T}\text{score}_i h_i \tag{9}$$

式中,$\boldsymbol{W}$、b 均为模型学习参数,随着训练进程更新迭代;a_i 为当前隐藏层的单元状态值 h_i 对应的注意力权重值;c_t 为隐藏层状态值。

2.2 模型评价指标

选用平均绝对误差(M_{MAE})、均方根误差(R_{RMSE})、平均绝对百分比误差(M_{MAPE})及决定系数(R^2)作为评价指标。

3 实例分析

3.1 数据收集和处理

洺河渡槽位于邯郸市永年区境内,全长 930 m,设计流量 230 m^3/s,是南水北调中线总干渠上最大的一座河渠交叉建筑物,其进口段设置了三孔弧形闸门,单孔净宽 7 m。取 2020 年 10 月 1 日至 2021 年 10 月 1 日洺河渡槽进口节制闸的实时运行数据共 17 494 条,采集频率为 30 min,洺河节制闸的实测水情数据见表 1。

在数据收集过程中,由于多种主客观因素,包括传感器故障、数据传输错误及人为操作失误等,可能会导致收集到的数据中存在空缺、异常值,故

中国知网 https://www.cnki.net

表 1　洺河节制闸实测水情数据

Tab. 1　Actual water situation data of Minghe control gate

时间	闸 1 开度/mm	闸 2 开度/mm	闸 3 开度/mm	闸后水位/m	闸前水位/m	过闸流量/($m^3 \cdot s^{-1}$)
2020-10-01 00:00:00	2 300	2 300	2 302	87.36	88.05	178.42
2020-10-01 00:00:30	2 300	2 300	2 302	87.36	88.05	177.22
2020-10-01 00:01:00	2 300	2 300	2 302	87.36	88.05	177.11
2020-10-01 00:01:30	2 300	2 300	2 302	87.36	88.06	178.50
⋮	⋮	⋮	⋮	⋮	⋮	⋮

需清洗原始数据。对于部分特征缺失且其余特征未变化的水情数据,可沿用上一时刻的水情数据。若全部特征缺失即出现空白值,则根据相邻的上下条数据采用线性插值法补充。采用皮尔逊相关系数法对处理后的数据集各项特征进行相关性分析,见图 4。可以看出,过闸流量和闸后水位相关性最大,高达 0.99,闸前水位和过闸流量相关性却为−0.13,这是因为闸门启闭时需要保持闸前水位稳定,故闸前水位变幅较小,且过闸水流状态时常处于淹没出流状态[13],闸后水位和淹没系数之间相关程度较高,各闸门相对开度和流量之间相关程度均较低。综合考虑,选用三孔闸门各孔开度、闸前水位、闸后水位及过闸流量作为输入特征。

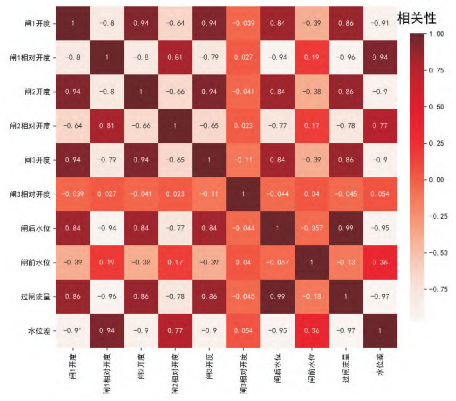

图 4　各特征相关性热力图

Fig. 4　Heat map of correlation between various features

3.2　模型验证

选取数据集前 10 个月的样本作为训练集,后 2 个月的样本作为测试集。各模型的测试参数均默认为最大迭代次数为 500 次,初始学习率为 0.01,且每经过 100 次迭代训练降低学习率,下降因子为 0.5。BiLSTM 的维度均设置为 64,激活函数为 Relu,且在每个 BiLSTM 隐藏层后面附加一个随机失活(Dropout)层防止过拟合,Attention 层激活函数采用 Softmax。此外,模型使用 Adam 优化器和均方误差作为损失函数进行编译,批量大小为 32。模型经多次运行后,对基于 LSTM 改进的各模型性能指标进行 20 次随机抽样分析。

为进一步验证本文所提模型 CNN-BiLSTM-Attention 的预测性能,选择 4 种浅层机器学习模型(高斯过程回归(GPR)、反向传播神经网络(BP-NN)、支持向量机(SVR)和极限学习机(ELM))进行对比,各模型评价指标见表 2,测试集预测结果的绝对误差分布区间见图 5。

表 2　各模型评价指标

Tab. 2　Evaluation indicators for each model

模型	训练集				测试集			
	M_{MAE}	M_{MAPE}	R_{RMSE}	R^2	M_{MAE}	M_{MAPE}	R_{RMSE}	R^2
GPR	0.940	0.006	1.243	0.998	17.486	0.078	19.679	0.667
ELM	3.368	0.019	4.516	0.992	11.154	0.050	15.085	0.828
BP-NN	1.394	0.008	1.987	0.998	8.347	0.038	12.352	0.883
SVR	2.329	0.013	3.515	0.995	6.740	0.032	10.110	0.923
LSTM	2.415	0.014	3.274	0.996	6.415	0.030	8.599	0.944
BiLSTM	2.985	0.018	3.910	0.994	5.460	0.027	6.908	0.963
CNN-BiLSTM	3.760	0.022	4.804	0.991	4.755	0.024	6.600	0.967
CNN-BiLSTM-Attention	3.860	0.022	4.904	0.990	3.682	0.018	4.661	0.983

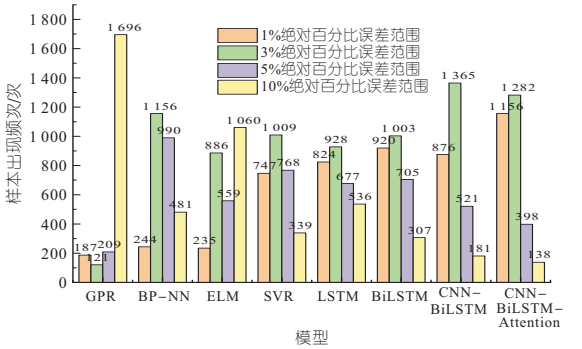

图 5　绝对误差分布区间柱状图

Fig. 5　Bar chart of absolute error distribution interval

由表 2、图 5 可知:①对比 LSTM、BiLSTM、CNN-BiLSTM 和 CNN-BiLSTM-Attention 模型的各项指标,可看出 CNN-BiLSTM-Attention 的模型预测精度最高,模型测试集的决定系数为 0.983,M_{MAE}、M_{MAPE}、R_{RMSE} 分别为 3.682、0.018、4.661,按照模型预测性能从大到小排序为 CNN-BiLSTM-Attention ＞ CNN-BiLSTM＞ BiLSTM ＞ LSTM。从整体上看,在 LSTM 模型基础上相继引入了前后向学习神经网络、CNN 及注意力机制,模型的各项指标随着模型的改进而有所提升。其中,BiLSTM 模型性能较 LSTM 模型明显提升,这是因为 BiLSTM 能够同时从前向和后向两个维度捕捉数据的连续性和关联性,从而更有效地挖掘出数据内部的深层次联系。当引入 CNN 后,模型性能提升较小,但同时减少了各评价指标的振荡。融合注意力机制后的模型的各指标波动较为平稳,注意力机制对各特征赋予了不同的权重,提高了预测的稳定性。相较于未引入注意力机制的 CNN-BiLSTM,其 M_{MAE}、M_{MAPE}、R_{RMSE} 分别降低了 22.56%、23.78%、29.38%,CNN-BiLSTM-Attention 的模型预测值整体上较其他模型的预测值更贴近真实值。②GPR、BP-NN、

中国知网　https://www.cnki.net

ELM、SVR 预测模型在训练集中各指标均表现较好，但在测试集中各指标表现较差，说明这些模型均产生了不同程度的过拟合现象。本文提出的 CNN-BiLSTM-Attention 模型在训练集和测试集上的各项指标均表现稳定，说明该模型不存在过拟合现象，且模型的预测能力均优于 GPR、BP-NN、ELM、SVR 预测模型。由图 5 可知，本文提出的模型在 1%、3%、5%的绝对百分比误差范围内的样本出现总频率高于其他对比模型，高达 95%的测试集样本处于该范围内，说明本文模型具备较高的预测精度。

4 结论

本文建立了基于注意力机制的 CNN-BiLSTM 过闸流量预测模型，该模型结合了 CNN 特征提取、BiLSTM 双向学习及注意力机制增强特征权重的优点。实例验证结果表明，相比于其他预测模型，该组合预测模型可有效提高过闸流量预测精度。

参考文献：

[1] 龙岩，雷晓辉，杨艺琳，等. 南水北调工程突发水污染事件分级体系研究 [J]. 水力发电学报，2019，38(3)：12-22.

[2] 徐苏，窦明，米庆彬，等. 五种过闸流量计算方法在槐店闸的适用性及改进的杜屿公式精度分析 [J]. 水电能源科学，2016，34(3)：36-39，10.

[3] 李琳，赵帅杰. 垂直梯形闸门自由出流水力特性试验研究 [J]. 水利学报，2023，54(6)：677-685，695.

[4] 陈晓楠，乔雨，许新勇，等. 基于双调和样条插值和量纲分析的过闸流量计算 [J]. 华北水利水电大学学报(自然科学版)，2024，45(3)：26-32.

[5] 郭世圆，马为之，卢瑞麟，等. 基于 LSTM 神经网络的复杂工况下明渠流量预测 [J]. 清华大学学报(自然科学版)，2023，63(12)：1924-1934.

[6] 王艺霖，靳燕国，陈晓楠，等. LSTM 神经网络和量纲分析法在弧形闸门过流计算中的对比 [J]. 南水北调与水利科技(中英文)，2022，20(3)：590-599.

[7] 李珊珊，曹顶业，沈桂莹，等. 基于机器学习和全局敏感性的弧形闸门淹没特性 [J]. 农业工程学报，2023，39(15)：25-33.

[8] MUNDHER Y Z，ALAWI O A，MOHAMMED A A，et al. Development of advanced data-intelligence models for radial gate discharge coefficient prediction：Modeling different flow scenarios [J]. Water resources management，2023，37(14)：5677-5705.

[9] 岳明哲，陈旭东，李俊杰. 基于 CNN-LSTM 的混凝土坝渗流预测 [J]. 水电能源科学，2020，38(9)：75-78.

[10] ZHANG M，YUAN Z M，DAI S S，et al. LSTM RNN-based excitation force prediction for the real-time control of wave energy converters [J]. Ocean engineering，2024，306：118023.

[11] 任建吉，位慧慧，邹卓霖，等. 基于 CNN-BiLSTM-Attention 的超短期电力负荷预测 [J]. 电力系统保护与控制，2022，50(8)：108-116.

[12] 马良玉，胡景琛，段晓冲，等. 具有注意力机制的 CNN-GRU 模型在风电机组异常状态预警中的应用 [J/OL]. 南京信息工程大学学报. https://doi.org/10.13878/j.cnki.jnuist.20240509001.

[13] 侯冬梅，王才欢，刘毅. 调水工程输水渠道堰闸流量计算方法探讨 [J]. 长江科学院院报，2013，30(8)：79-83.

CNN-BiLSTM Overpass Flow Prediction Model Based on Attention Mechanism

HE Li-xin[a,b]，SHEN Zheng-hua[a,b]，ZHANG Zheng[a,b]，LEI Xiao-hui[a,b]

(a. School of Water Conservancy and Hydroelectric Power；b. Hebei Key Laboratory of Intelligent Water Conservancy，Hebei University of Engineering，Handan 056038，China)

Abstract：In the open channel water transfer project，it is of great significance to accurately grasp the flow through the gate for improving the efficiency of channel regulation and ensuring the safety of the water transfer system. In order to improve the prediction accuracy of gate flow，this paper proposes a model of gate flow prediction based on attention mechanism，combining convolutional neural network (CNN) and BiLSTM. Taking the control gate of the Minghe aqueduct as an example，the actual data of the 1-year time scale is selected as the model input. The model first standardizes the input data，and the feature information is extracted using CNN. BiLSTM is used to capture the forward and backward dependencies in sequence data. Finally，the attention mechanism is used to evaluate the importance of information，and the characteristic parameters are weighted to achieve the prediction of gate flow. Compared with the traditional prediction models such as BP-NN，SVR and LSTM，the results show that the model has better prediction results. The mean absolute error (MAE)，mean absolute percentage error (MAPE)，root mean square error (RMSE) and determination coefficient (R^2) of the model are 3.682，0.018，4.661 and 0.983，respectively，which can provide reference for engineering practice.

Key words：gate flow prediction；BiLSTM；attention mechanism；neural network

中国知网 https://www.cnki.net