

基于双通道特征融合的热连轧厚度预测补偿

张晓东, 史靖文, 白广芝, 秦子轩

(中国石油大学(华东)青岛软件学院 计算机科学与技术学院, 山东 青岛 266580)

摘要: 针对热连轧轧制受油膜厚度、轧辊偏心等因素影响而造成精度降低的问题, 提出一种基于双通道特征融合的热连轧板带厚度误差预测模型。模型由时空特征提取和多尺度特征提取两部分组成: 时空特征提取部分是基于相邻机架板带厚度空间的相关性构建了前置时空矩阵 (FSTM), 通过卷积长短期记忆网络 (Conv-LSTM) 提取 FSTM 的时空关联特征; 多尺度特征提取部分是采用离散小波变换 (DWT) 对当前机架轧制厚度数据进行分解, 得到趋势项数据和细节项数据, 并采用差分自回归移动平均模型-长短期记忆网络 (ARIMA-LSTM) 和双向长短期记忆网络 (BiLSTM) 分别对趋势项数据和细节项数据进行特征提取。将上述特征融合后输入全连接层进行回归预测, 得到热连轧板带厚度误差预测值。实验结果表明: 双通道特征融合模型能有效提高厚度误差预测精度, 验证了模型的有效性。

关键词: 热连轧; 厚度补偿预测; 长短期记忆网络; 卷积神经网络; 特征提取

DOI: 10.13330/j.issn.1000-3940.2024.02.020

中图分类号: TG335

文献标志码: A

文章编号: 1000-3940 (2024) 02-0161-11

Thickness prediction and compensation for hot tandem rolling based on dual-channel feature fusion

Zhang Xiaodong, Shi Jingwen, Bai Guangzhi, Qin Zixuan

(Qingdao Institute of Software, College of Computer Science and Technology, China University of Petroleum (East China),
Qingdao 266580, China)

Abstract: For the problem of accuracy reduction in hot tandem rolling caused by factors such as oil film thickness and roll eccentricity, a thickness error prediction model for hot tandem rolling strips based on dual-channel feature fusion was proposed, which was composed of two parts, namely, spatiotemporal feature extraction and multi-scale feature extraction. For the spatiotemporal feature extraction part, a Frontal Spatio-Temporal Matrix (FSTM) was constructed based on the spatial correlation of adjacent rack plate thicknesses, and the spatiotemporal correlation features of the FSTM were extracted by the Convolutional Long Short-Term Memory Network (Conv-LSTM). For the multi-scale feature extraction part, the current rack rolling thickness data was decomposed by Discrete Wavelet Transform (DWT) to obtain trend item data and detail item data, and Autoregressive Integrated Moving Average Long Short-Term Memory Network (ARIMA-LSTM) and Bi-directional Long Short-Term Memory Network (BiLSTM) were used to extract features from the trend item data and detail item data, respectively. The above features were fused and input into the fully connected layer for regression prediction, the thickness error prediction value of the hot tandem rolling strip was obtained. Experimental results show that the dual-channel feature fusion model effectively improves the accuracy of thickness error prediction, validating the effectiveness of the model.

Key words: hot tandem rolling; thickness compensation prediction; long and short term memory network; convolutional neural network; feature extraction

带钢热连轧厚度控制系统通常由 6~7 台精轧机架组成。传统的带钢厚度预测方法主要是基于轧机弹跳方程建立的数学模型实现的, 研究人员在热连轧厚度控制方面做了大量的研究工作, 如模型预测

控制^[1-2]、自适应控制^[3-4]、模糊控制^[5-6]和神经网络控制^[7-9]等, 提高了厚度控制精度, 获得了较好的控制效果。然而, 传统的模型预测控制方法往往需要对系统进行较为复杂的数学建模, 需要考虑到各种影响因素, 建模过程较为繁琐, 且带钢厚度控制是一个复杂的动态过程, 其中包含了多种影响因素, 例如轧辊偏心、温度、张力以及弹性形变等。这些因素会相互作用、相互影响, 导致带钢厚度的变化过程具有高度的非线性和强耦合性。同时, 原

收稿日期: 2023-06-13; **修订日期:** 2023-09-15

基金项目: 国家自然科学基金资助项目 (61801517); 中央高校基本科研业务经费 (19CX02029A)

作者简介: 张晓东 (1979-), 男, 博士, 副教授

E-mail: zhangxiaodong@upc.edu.cn

始采集的数据往往包含了大量的噪声和干扰,为基于模型的厚度控制系统设计增加了难度。

越来越多的研究人员将数据驱动方法应用到热连轧生产控制过程中。Deng J F 等^[10]将人工神经网络 Artificial Neural Network, ANN) 应用到了热连轧带钢凸度预测,并证明了 ANN 与一些智能算法相结合可以提高热连轧参数的预测精度。Huang Y 等^[11]根据薄带钢轧制特性,提出了一种基于变分贝叶斯高斯混合聚类算法的径向基函数神经网络 (VBGM-RBF) 来预测冷轧薄带钢的厚度,可以获得更准确的建模和预测效果。Liu Y 等^[12]建立了基于轧制机理模型引导的机器学习的热轧带钢厚度、板形和板形预测模型,采用加权处理 (Weighted Processing, WP) 方法对轧制机理和过程数据进行融合,改善强相关性特征与模型之间的关系。杨静等^[13]针对热连轧实际的轧制过程因受多种因素影响导致传统模型的预测精度往往较低的问题,提出了并行异构极限学习机的轧制力预测模型,得到了较好的轧制力预测结果。Ding J G 等^[14]提出了一种将机理模型与机器学习相结合的钢板轧制后倾角预测模型,实现了对钢板轧制后倾角的精确控制。姬亚锋等^[15]提出了基于多输出支持向量回归 M-SVR 的板带宽度-厚度预测模型,并用精英策略的非支配排序遗传算法 NSGA-II 实现该模型的多目标参数优化。然而,传统的浅层机器学习方法存在一定的缺陷,如通常仅能对数据进行浅层次的学习,难以挖掘数据的深层次特征,容易受到噪声和局部最优解的干扰,导致模型的泛化性能差,容易出现过拟合的问题。

深度信念网络 (Deep Belief Network, DBN)、卷积神经网络 (Convolutional Neural Network, CNN) 和长短期记忆网络 (Long Short-Term Memory Network, LSTM), 已被许多学者用于连轧生产中来实现回归预测。Hu X 等^[16]采用多尺度卷积神经网络对热连轧板带的最终冷却温度进行预测,获得了较高的预测精度。Chen Y F 等^[17]通过考虑复杂的生成机制和空间维度相关性,提出了一种具有注意机制的长短期记忆模型并用于解决串联冷轧带钢平整度的预测问题。于加学等^[18]提出了一种基于深度学习的热轧带钢头部厚度的命中预测方法,实现了对带钢头部厚度命中的较为精准的预测。Huang S Y 等^[19]充分考虑了分位数回归的稳健性,结合了 DBN 可以从低层数据中学习更抽象的隐含层信息的优势,提出了一种基于深度信任网络和复合分位数回归模型的热轧带钢抗拉强度预测方法。张瑞成^[20]等针对冷连轧机振动数据

样本大、非线性强,且在时间上具有前后依赖关系的特点,提出了一种基于栈式稀疏自编码网络优化的长短时记忆网络 (SSAE-LSTM) 的轧机扭振预测方法。

本文提出了一种基于双通道特征融合的热连轧板带厚度误差预测方法。利用小波变换对目标机架单流原始厚度数据进行分解,得到不同尺度和频率上的信息。针对多机架对应的多流厚度数据在时空上紧密联系的特点,构建了前置时空矩阵 (Frontal Spatio-Temporal Matrix, FSTM)。然后,将得到的数据输入到构建的基于 Conv-LSTM 网络和 DWT-ARIMA-LSTM-BiLSTM 网络的热连轧轧制厚度预测模型中,进行热连轧板带厚度数据的预测研究,有效地补偿了热连轧厚度控制误差,提高了厚度控制精度。最后,利用某 1580 mm 热连轧生产线精轧机组末位两个机架的轧制厚度数据进行实验来评估所提出的模型的有效性,结果显示,该模型相较于其他模型具有较好的预测能力。

1 问题描述

热连轧精轧厚度控制系统示意图如图 1 所示。该系统由多个机架协同压下实现钢板的轧制,通过末端机架的厚度检测装置测量带钢的实际厚度,其中, F_i 表示第 i 个机架的轧制力, N 表示机架总数。

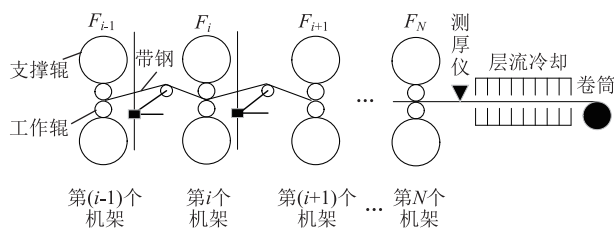


图 1 热连轧系统示意图

Fig. 1 Schematic diagram of hot tandem rolling system

从图 1 中可以观察到,上游机架的出口厚度值等于下游机架的入口厚度值。在时刻 k , 第 i 个机架能够获取第 $(i-1)$ 个机架的轧制出口厚度信息,同时能够获取当前机架的轧制出口厚度信息,即在轧制力 F_i 的作用下产生厚度增量。因此,上游机架轧制厚度的变化会对下游机架轧制精度产生影响,这体现为机架间存在较强的耦合效应。由于轧制环境的恶劣性,无法实时获取钢板的实际轧制厚度,只能依赖于末位机架后方的测厚仪获取,但由于存在较大的时滞,且基于弹跳方程计算得到的轧制厚度精度不够,尤其是受到轧辊偏心、油膜厚度等因素的影响,使得厚度偏差难以消除。因此,基于数据分析的厚度补偿成为提高热连轧厚度精度的关键。

2 混合预测模型

2.1 基于 Conv-LSTM & DWT-ARIMA-LSTM-BiLSTM 混合模型的热连轧厚度误差预测

本文提出了一种基于双通道特征融合的混合深

度学习模型用于热连轧板带厚度预测，如图2所示。其中，Conv-LSTM通道能够捕获 FSTM 的时空特征，DWT-ARIMA-LSTM-BiLSTM 通道能够通过将待预测机架的厚度数据分解，捕获其高频细节特征和低频近似特征。

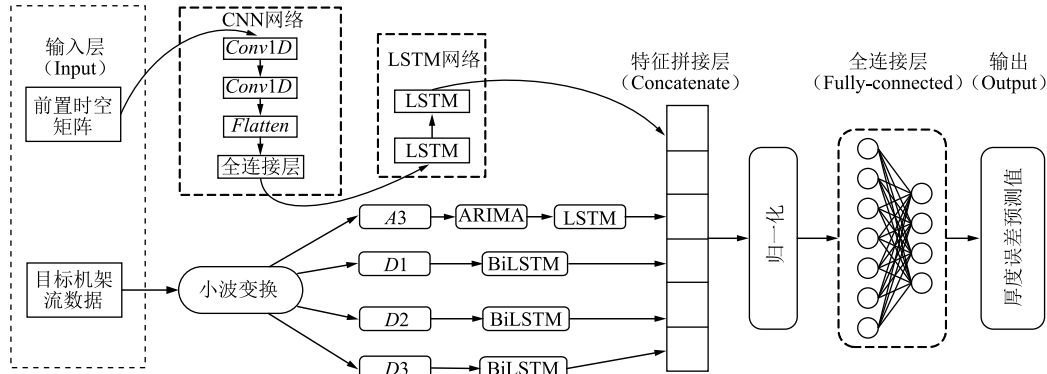


图2 热连轧厚度预测的混合深度学习模型

Fig. 2 Hybrid depth learning model of thickness prediction for hot tandem rolling

首先，利用当前机架及上游机架厚度时序数据构建 FSTM。同时，将当前机架轧制厚度数据经过离散小波变换分解为低频近似数据 A3 和高频细节数据 D1、D2 和 D3。其中，低频近似数据反映了数据在稳态条件下的变化趋势，而高频细节数据反映了数据在低频稳态基础上的波动和突变。

然后，构建 Conv-LSTM 网络，FSTM 囊括了热连轧厚度数据的空间依赖信息和时间依赖信息。其中，CNN 网络用于提取厚度时序数据的空间特征，LSTM 网络用于提取厚度时序数据的时间特征。同时，构建 ARIMA-LSTM 网络，提取低频近似数据的趋势特征，采用 BiLSTM 时序预测模型提取高频细节数据中的上下文信息和依赖关系。

最后，将两个通道提取的时空特征、低频近似趋势特征和高频细节特征进行特征融合。通过将特征融合后的序列输入全连接网络，得出热连轧板带厚度的最终预测结果。双通道融合策略充分利用了不同通道提取的特征信息，提高了厚度预测的准确性和稳定性。

2.2 构建前置时空矩阵（FSTM）

考虑每台机架不同的轧制特性及轧制数据在时空上的关联性，建立前置时空矩阵，设 f_l^i 表示板带第 l 个点经第 i 个机架轧制后的厚度，选取 k 个点位的历史数据。则该机架从第 $(l-k)$ 个点位到第 l 个点位的历史板带厚度数据可以表示为 $[f_{l-k}^i, f_{l-k+1}^i, \dots, f_l^i]^T$ ，根据板带成形顺序，减少预测的时滞，仅考虑将相邻上游机架的历史板带厚度时间序

列组合成如式（1）所示的时空矩阵 X_l ：

$$X_l = \begin{bmatrix} X_{l-k} \\ X_{l-k+1} \\ \vdots \\ X_l \end{bmatrix}^T = \begin{bmatrix} f_{l-k}^1 & f_{l-k+1}^1 & \dots & f_l^1 \\ f_{l-k}^2 & f_{l-k+1}^2 & \dots & f_l^2 \\ \vdots & \vdots & \vdots & \vdots \\ f_{l-k}^i & f_{l-k+1}^i & \dots & f_l^i \end{bmatrix} \quad (1)$$

式中： X_l 为带钢第 l 点位经过不同机架轧制后的厚度值， $X_l = [f_l^1, f_l^2, \dots, f_l^i]^T$ 。

2.3 卷积长短期记忆网络（Conv-LSTM）

该网络用于时空序列预测任务，输入为时空矩阵。网络由两层一维卷积层 Conv1D、一层 Flatten 层、一层全连接层和两层 LSTM 层构成。网络的前 3 层用于提取时空特征，而后两层 LSTM 层则用于建模序列的时序关系。具体地，第 1 层一维卷积层提取局部特征；第 2 层一维卷积层在第 1 层的基础上提取更高级别的特征；Flatten 层将卷积层输出的三维张量展平成二维张量；全连接层进一步提取特征并将其映射到 LSTM 层的输入维度上；在 LSTM 层中，第 1 层 LSTM 层用于提取时序信息，第 2 层 LSTM 层则用于进一步捕获更复杂的时序关系。

首先，输入的时空矩阵经过两层一维卷积层进行特征提取。每层卷积层后接一个 ReLU 激活函数以增强模型的非线性表达能力，具体卷积操作如下：

$$Y_t = \sum_{m=1}^M W_m X_{(t-m+1)} \quad (2)$$

式中： Y_t 为卷积结果输出序列； M 为卷积核大小； W_m 为权向量； $X_{(t-m+1)}$ 为输入信号序列。

然后, 经过卷积层得到的特征被展平成一维向量, 并输入到全连接层中, 全连接层对展平后的特征进行线性变换, 并再次使用 ReLU 激活函数, 得到输出特征 FC , 其表达式为:

$$FC(\mathbf{x}_f) = \text{ReLU}(\mathbf{W}\mathbf{x}_f + b) \quad (3)$$

式中: \mathbf{W} 为权重; \mathbf{x}_f 为展平后的特征向量。

输出的特征序列被输入到两层 LSTM 网络中进行序列建模, LSTM 通过长短期记忆单元实现对序列特征的抽取和记忆。LSTM 中的遗忘门和输入门分别决定丢弃或保留哪些信息, 而细胞态表示长期记忆, 记忆体表示短期记忆。LSTM 的结构如图 3 所示, 具体的表达式如下:

$$f_t = \sigma(\mathbf{W}_f[x_t, h_{t-1}] + b_f) \quad (4)$$

$$i_t = \sigma(\mathbf{W}_i[x_t, h_{t-1}] + b_i) \quad (5)$$

$$\tilde{C}_t = \tanh(\mathbf{W}_c[x_t, h_{t-1}] + b_c) \quad (6)$$

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t \quad (7)$$

$$O_t = \sigma(\mathbf{W}_o[x_t, h_{t-1}] + b_o) \quad (8)$$

$$h_t = O_t \cdot \tanh(C_t) \quad (9)$$

式中: σ 和 \tanh 表示模型中特定的激活函数; \mathbf{W}_f 、 \mathbf{W}_i 、 \mathbf{W}_c 、 \mathbf{W}_o 分别为遗忘门、输入门、细胞态和输出门的权重矩阵; x_t 为当前时刻 t 的时序数据输入; h_{t-1} 为上一个 LSTM 单元的输出; C_{t-1} 为前一时刻的细胞态; b_f 、 b_i 、 b_c 、 b_o 分别为遗忘门、输入门、细胞态和输出门的偏置; f_t 、 i_t 、 \tilde{C}_t 、 C_t 、 O_t 、 h_t 分别为当前时刻 t 的遗忘门的输出、输入门的输出、新的候选细胞状态、更新后的细胞态、输出门的输出、当前时刻的隐藏状态。

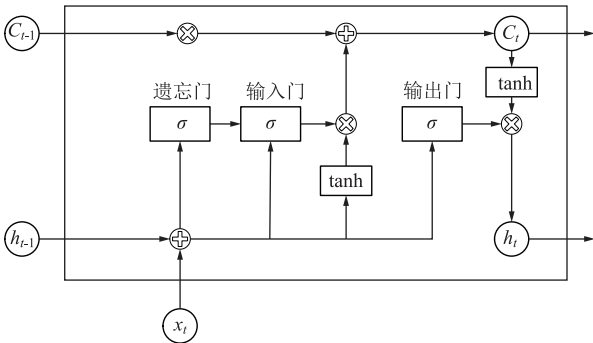


图 3 LSTM 节点原理图

Fig. 3 Principle diagram of LSTM node

2.4 DWT-ARIMA-LSTM-BiLSTM 网络

2.4.1 离散小波变换 (DWT)

板带厚度数据呈现非线性、非稳定性和波动性的特点。直接使用原始数据会使模型无法准确把握数据的短期突变特征。本文采用离散小波变换

(Discrete Wavelet Transform, DWT) 对原始数据进行多尺度特征提取, 将波动分解为低频近似序列和不同频率的高频细节序列, 能更为有效且快速地获得频域特征, 降低计算的复杂度。采用带 Daubechies 小波积函数的离散小波变换将原始数据分解为多个不同的子序列, 可以表征去噪后的低频分量和去趋势化的高频分量。最大分解层数为 s 的离散小波变换公式如下:

$$DWT_f(u, v) = \alpha_0^{-\frac{u}{2}} \int_{-\infty}^{+\infty} f(t) \varphi(\alpha_0^{-u}t - v\tau_0) dt \quad (10)$$

式中: u 和 v 分别为伸缩因子和平移因子; α_0 和 τ_0 分别为尺度参数和平移参数; $\varphi(t)$ 为复共轭函数; $f(t)$ 为时间序列。

将每一部分得到的低频子信号进一步分解为高频子信号和低频子信号。考虑数据的维度和平滑度, 将分解层数设置为 3, 实验中所采用的小波变换过程如图 4 所示。将原始数据 $Y(t)$ 分解为 $A3$ 、 $D1$ 、 $D2$ 和 $D3$ 共 4 个分量, 其中 $A3$ 为第 3 层分解后的低频分量, $D1$ 、 $D2$ 和 $D3$ 分别为第 1、2 和 3 层分解后的高频分量。

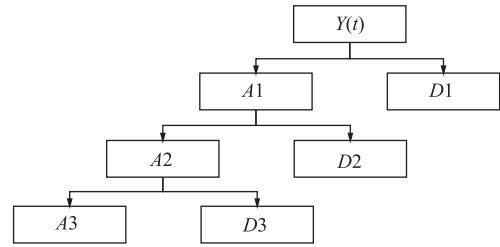


图 4 小波变换结构

Fig. 4 Structure of wavelet transform

2.4.2 差分自回归移动平均模型 (ARIMA)

差分自回归移动平均 (Autoregressive Integrated Moving Average Model, ARIMA) 模型是一种常用的线性回归模型, 用于处理时间序列数据。ARIMA 模型结合了自回归 (Autoregression, AR) 和移动平均 (Moving Average, MA) 的概念, 并通过差分操作将非平稳的时间序列数据转化为平稳序列, 从而使其更适合用于建模和预测。

ARIMA 模型表示为 $ARIMA(p, d, q)$, 其中 p 为自回归阶数, d 为差分阶数, q 为移动平均阶数。ARIMA 模型的基本原理是根据时间序列数据的自相关性 (AR) 和滞后误差 (MA) 来建立模型, 并通过差分操作 (d 次) 来使数据平稳化。ARIMA 模型的数学表示如式 (11) 所示:

$$Z_t = c + \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \cdots + \phi_p Z_{t-p} - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \cdots - \theta_q \varepsilon_{t-q} + \varepsilon_t \quad (11)$$

式中: Z_t 为 t 时刻的时间序列值; c 为常数; ϕ_1 、 ϕ_2 、 \dots 、 ϕ_p 为 AR 模型的系数; θ_1 、 θ_2 、 \dots 、 θ_q 为 MA 模型的系数; ε_t 为在时刻 t 的误差项。

在热连轧板带厚度预测中, 首先利用 ARIMA 模型拟合经过小波分解得到的时序近似数据的线性趋势, 并计算得到残差序列。这些残差序列包含了原始数据中无法被线性趋势描述的非线性特征。将这些残差序列作为输入传递给 LSTM 模型, 以提取其时序的非线性特征。LSTM 模型能够捕捉长期依赖关系和时序模式, 进一步提升预测的准确性和稳定性。通过将 ARIMA 处理后的残差与 LSTM 模型相结合, 可以综合利用线性和非线性建模能力, 该组合作为网络组成部分融入到建模中, 可以更好地预测当前热连轧板带厚度的变化趋势。

2.4.3 双向长短期记忆网络 (BiLSTM)

带钢轧制厚度变化信息具有很强的前后关联性, 同一机架轧制的带钢厚度不仅受历史轧制厚度的影响, 同时也会影响之后机架轧制的带钢厚度。采用双向长短期记忆网络 (Bi-directional Long Short-Term Memory Network, BiLSTM) 可以综合利用数据上下文信息。BiLSTM 可以同时从前向和后向对输入序列进行建模和学习。在该模型中, 输入序列被分别输入到前向和后向 LSTM 网络中, 每个 LSTM 单元结构包含输入门、遗忘门和输出门, 通过长短期记忆单元实现对序列特征的抽取和记忆。BiLSTM 的输出是前向和后向 LSTM 网络的隐藏状态的拼接, 具体计算公式如式 (12) 所示:

$$h_B = [\vec{h}_t \oplus \overleftarrow{h}_t] \quad (12)$$

式中: h_B 为 BiLSTM 在时刻 t 的输出; \vec{h}_t 和 \overleftarrow{h}_t 分别为前向和后向 LSTM 网络在时刻 t 的隐状态; \oplus 表示拼接操作。

3 实验过程及模型配置

3.1 数据预处理

实验数据来源于国内某钢铁厂的 1580 mm 带钢生产线精轧机组, 涵盖了第 3 机架~第 6 机架 (记为 F3~F6) 的轧制厚度时序数据, 如图 5 所示。板带钢的轧制时间约为 15~20 s, 采样时间为 10 ms。据轧钢自动化生产验收要求, 剔除了带钢首卷和起始卷的数据。最终获得 1080 组样本数据。每组样本点包括 4 个参数, 分别为该热连轧生产线精轧机组的 F3~F6 的轧辊出口厚度数据。首先对数据进行预处理, 将预处理后数据的前 80% 作为训练集, 用于训练网络模

型, 20% 的数据用作测试集, 将测试集中得到预测值与真实值进行对比, 从而评估测试结果的准确性。

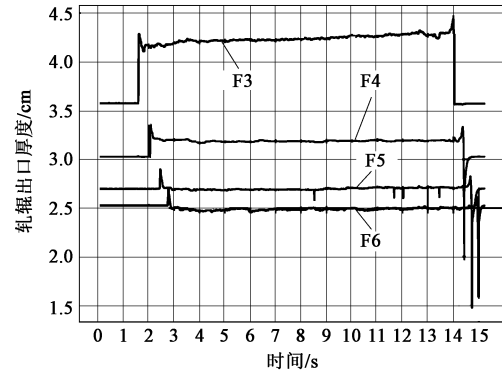


图5 精轧机组出口厚度

Fig. 5 Outlet thickness of finishing mill group

3.1.1 数据空间对齐

由于生产线上的轧机位置不同, 导致测量数据存在延迟, 因此, 需要将数据与生产线上的设备分布和轧制速度进行同步。通过轧机编号进行测量数据同步。以带钢上的一个点 E 通过 F4 机架的时刻作为参考点, 计算之后轧机上测得的厚度数据相对于该参考点的延迟量如式 (13) 所示:

$$n_{S+1} = \sum_{i=3}^S \frac{b}{V_{i+1}g} \quad (13)$$

式中: n_{S+1} 为第 $(S+1)$ 个机架的延迟点位数个 (2 ≤ S+1 ≤ 6); c 为机架总数; b 为机架间距; V_{i+1} 为第 $(i+1)$ 个机架的滚动速度; g 为传感器采样的时间间隔。

3.1.2 数据清洗

仪器故障、数据库存储和外部干扰等不可控因素会导致采集到的数据集存在离群点这种数据异常的情况。因此, 需要在输入到网络模型前进行相应的处理操作以提高数据的质量。

本文进行了数据清洗以保证数据集的准确性和可靠性。针对异常值, 根据数理统计方法中的三倍标准差原则进行原始数据的筛选。由于生产型数据服从正态分布, 故有:

$$P[(\bar{K} - 3\tau) < K_\gamma < (\bar{K} + 3\tau)] = 99.7\% \quad (14)$$

$$\bar{K} = \frac{1}{\alpha} \sum_{\gamma=1}^{\alpha} K_\gamma \quad (15)$$

$$\tau = \sqrt{\frac{1}{\alpha-1} \sum_{\gamma=1}^{\alpha} (K_\gamma - \bar{K})^2} \quad (16)$$

式中: P 为概率函数; \bar{K} 为样本均值; τ 为样本标准差; K_γ 为样本数据; γ 为样品编号, $\gamma=1, 2, \dots, \alpha$;

α 为观测样本数量。

3.1.3 数据归一化

为消除量纲的影响并提升数据的收敛速度,采用 Min-Max 归一化方法对数据进行归一化处理,计算公式如式 (17) 所示,数据被映射到 $[0, 1]$ 之间。

$$y' = \frac{y - y_{\min}}{y_{\max} - y_{\min}} \quad (17)$$

式中: y 为原始数据; y_{\min} 和 y_{\max} 分别为数据的最小值和最大值; y' 为归一化后的数据。

3.1.4 原始数据分解

实验中采用 F5 号和 F6 号机架的带钢厚度数据进行研究,利用带 Daubechies 小波积函数离散小波分解,分解后得到一个低频近似部分 $A3$ 和 3 个高频细节部分 $D1$ 、 $D2$ 和 $D3$,如图 6 和图 7 所示。

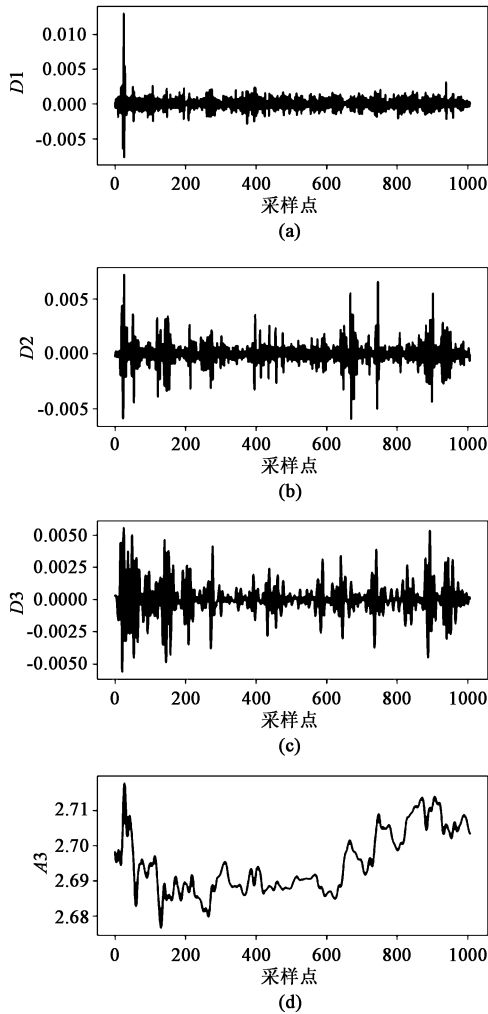


图 6 F5 号机架板厚小波分解

(a) $D1$ 高频分量 (b) $D2$ 高频分量 (c) $D3$ 高频分量
(d) $A3$ 低频分量

Fig. 6 Wavelet decomposition of plate thickness for F5

(a) $D1$ high frequency component (b) $D2$ high frequency component
(c) $D3$ high frequency component (d) $A3$ low frequency component

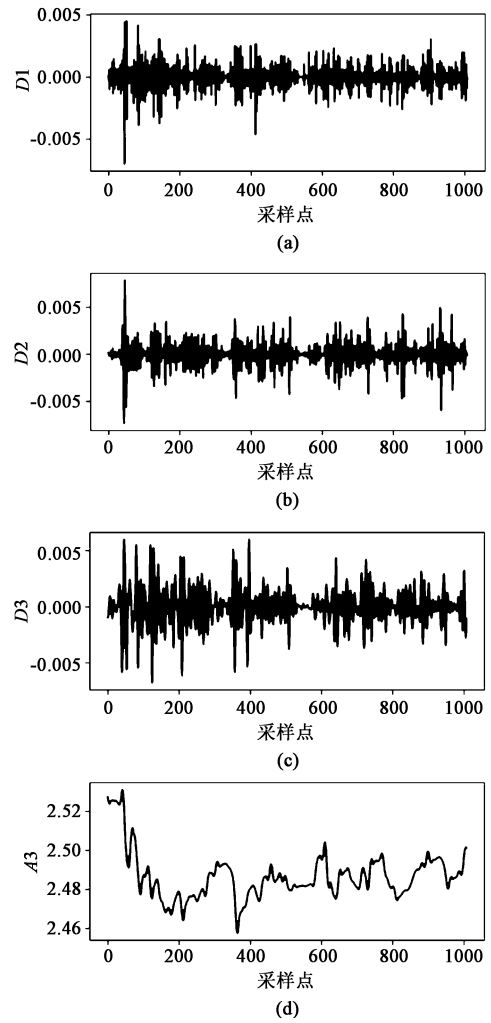


图 7 F6 号机架板厚小波分解

(a) $D1$ 高频分量 (b) $D2$ 高频分量 (c) $D3$ 高频分量
(d) $A3$ 低频分量

Fig. 7 Wavelet decomposition of plate thickness for F6

(a) $D1$ high frequency component (b) $D2$ high frequency component
(c) $D3$ high frequency component (d) $A3$ low frequency component

3.2 预测结果评估指标

为评估预测结果的有效性和预测性能,使用平均绝对误差 MAE 、均方根误差 $RMSE$ 以及决定系数 R^2 对实验预测结果进行处理来评估模型的性能。 $RMSE$ 和 MAE 值越小, R^2 值越大,模型性能越好,定义如下:

$$RMSE = \sqrt{\frac{1}{\mu} \sum_{j=1}^{\mu} [y_i - \hat{y}_i]^2} \quad (18)$$

$$MAE = \frac{1}{\mu} \sum_{j=1}^{\mu} |y_i - \hat{y}_i| \quad (19)$$

$$R^2 = 1 - \frac{\sum_{j=1}^{\mu} [y_i - \hat{y}_i]^2}{\sum_{j=1}^{\mu} [y_i - \bar{y}]^2} \quad (20)$$

式中: μ 为数据集的总数; y_i 为第 j 个时间点厚度误差; \hat{y}_j 为第 j 个时间点厚度误差预测值; \bar{y} 为实际值的均值。

3.3 对比实验与实验参数

进行了与传统机器学习算法和消融模型的对比实验, 以验证所提出的模型的有效性。其中, 采用的传统机器学习算法包括支持向量机 (SVR)、多层感知机 (MLP)、时间卷积网络 (TCN)、LSTM 和门控循环单元 (GRU), 而消融实验对比则包括不

带小波分解的 dual Conv-LSTM & ARIMA-LSTM、DWT-ARIMA-LSTM-BiLSTM、DWT-ARIMA-LSTM、DWT-LSTM-BiLSTM、顺序模型的 Conv-LSTM、顺序模型的 ARIMA-LSTM。通过这些对比实验, 能够全面评估不同模型的预测性能表现。同时, 通过消融对比实验, 还可以研究模型中不同组成部分对厚度误差预测效果的影响。这些实验结果不仅有助于验证所提出的模型的有效性, 还可以为进一步优化模型设计和参数设置提供指导。实验结果分别如图 8 和图 9 所示。

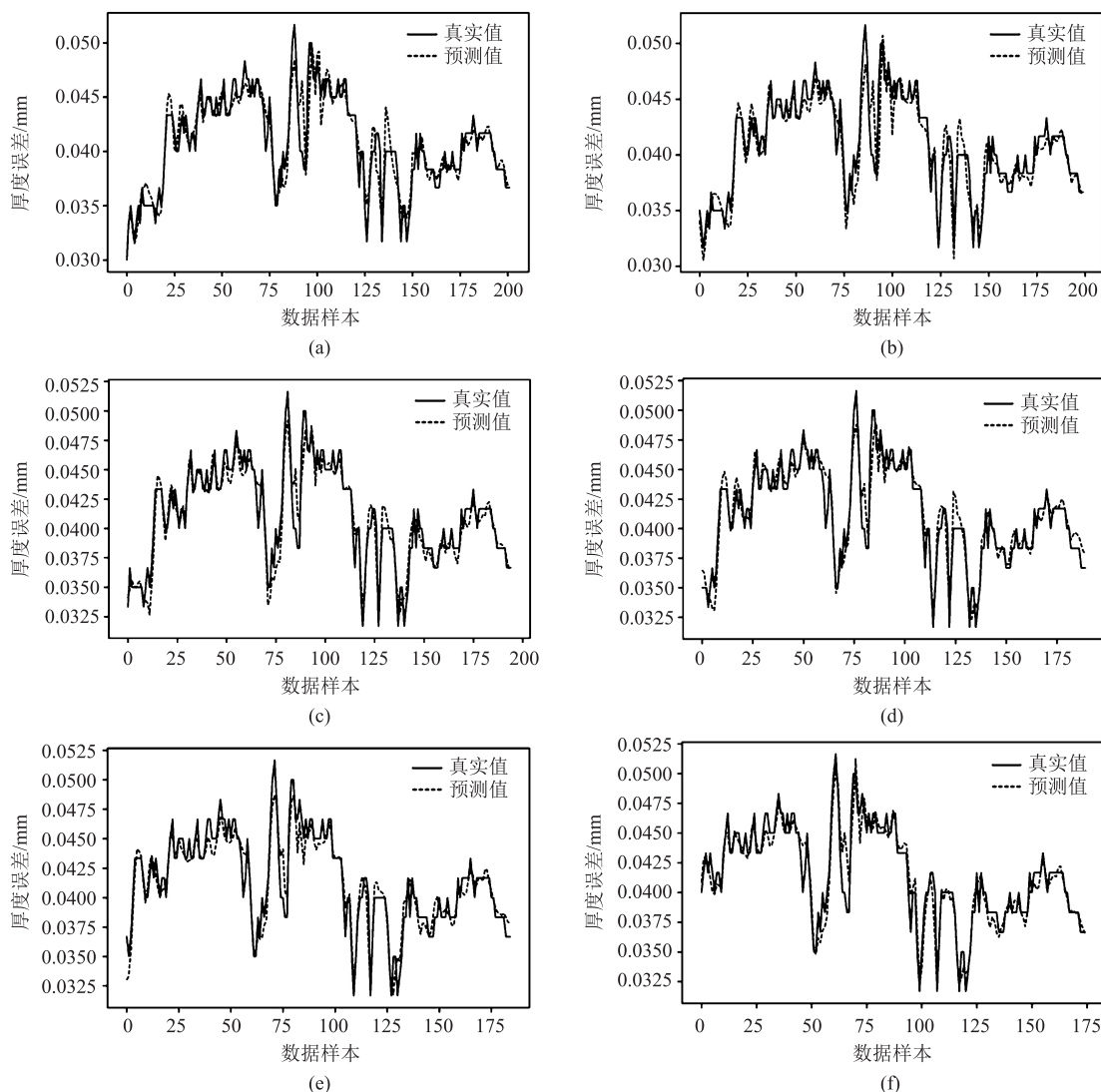


图 8 5 号机架不同步长时的预测结果

(a) 第 3 步 (b) 第 5 步 (c) 第 10 步 (d) 第 15 步 (e) 第 20 步 (f) 第 30 步

Fig. 8 Predicted results of rack 5 in different steps

(a) Step 3 (b) Step 5 (c) Step 10 (d) Step 15 (e) Step 20 (f) Step 30

本文所提出的 dual Conv-LSTM & DWT-ARIMA-LSTM-BiLSTM 模型及对比算法模型均在 Python 3.7.11 和 Keras 2.3.1 环境下运行。实验中所用的 LSTM 模块的神经元数均为 15。Conv-LSTM 网络剩

余的初始设置参数设置分为以下两个部分: (1) CNN 特征提取模块: 两层一维卷积层 Conv1D, 卷积核数均为 15, 大小为 3; (2) LSTM 模块: 两层 LSTM 层, 其后连接全连接层 Fully-connected, Fully-

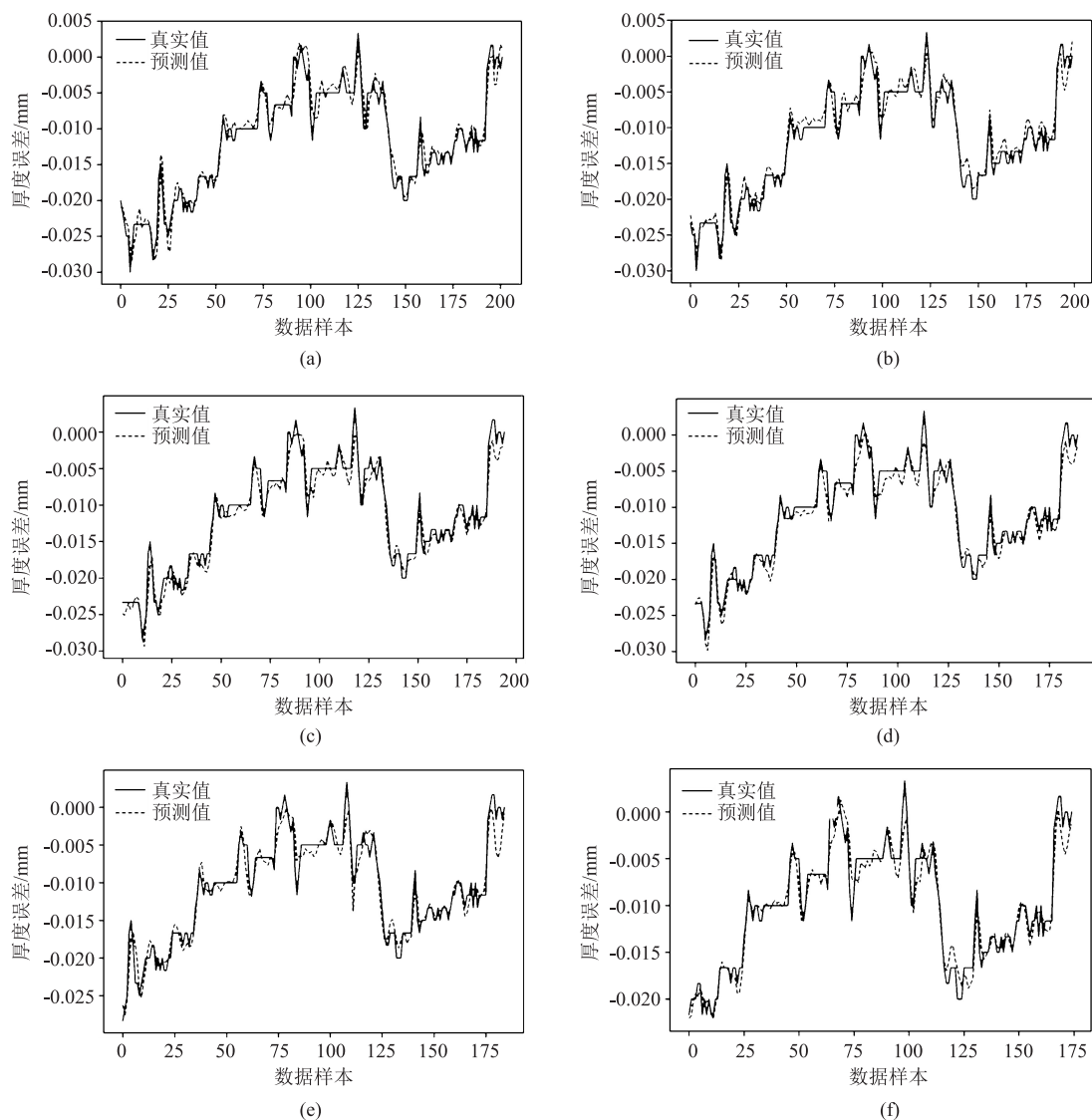


图 9 6号机架不同时步长的预测结果

(a) 第3步 (b) 第5步 (c) 第10步 (d) 第15步 (e) 第20步 (f) 第30步

Fig. 9 Predicted results of rack 6 in different steps

(a) Step 3 (b) Step 5 (c) Step 10 (d) Step 15 (e) Step 20 (f) Step 30

connected 起到了强化数据特征学习的作用。

DWT-ARIMA-LSTM-BiLSTM 网络剩余的初始设置参数为: (1) DWT 算法采用 db 系列小波, 分解层数为 3; (2) ARIMA-LSTM 网络中, 基于 AIC 准则确定在不同机架预测趋势所对应的 ARIMA (p, d, q) 模型的参数; (3) 批次大小设置为 18, 迭代次数设置为 200, 学习率设置为 0.001, 并且通过提前终止策略 (Early Stopping) 防止过拟合, 优化器采用 Adam 优化器, 损失函数采用 mse。

3.4 设置时间步长

为探究采取不同的时间步长对模型预测准确性的影响, 将滑动数据划分为不同的时间序列, 并在

提出的热连轧厚度误差预测模型中, 设置 6 个时间步长, 分别为 3、5、10、15、20 和 30 进行实验比较, 得到了如图 8 和图 9 所示的实验结果。

实验结果表明, 不同时间步长的设置对预测结果的准确性有较大的影响。步长越长, 训练时间越长。当时间步长为 3 时, 由于时间步长较短且未考虑全局因素的影响, 预测结果的偏差较大, 数据具有一定的波动性。当时间步长设置为 30 时, 考虑的时间范围太大, 容易在短时间内忽略突变因素等的影响, 预测结果不准确。而当步长设置为 10 时, 误差最小, 精度最高。因此, 最终将时间步长设置为 10, 以获得更准确的预测结果。

3.5 对比实验与实验参数

以 F5 号和 F6 号精轧机架上的带钢厚度数据集进行实验, 得到的各种方法在评价指标上的对比结果如表 1 所示。对比图 10 和图 11 可知, 相较其他算法的输出曲线, 本文所提算法的预测曲线与实际

输出的拟合效果最好, 获得了较好的预测效果。对比实验中用到的 11 个模型, 从表 1 中能够明显看出, 本文所提出的模型在各个指标上均有最佳的预测效果, 并且在带钢经过两个机架后产生的厚度误差预测拟合度均优于其他模型。

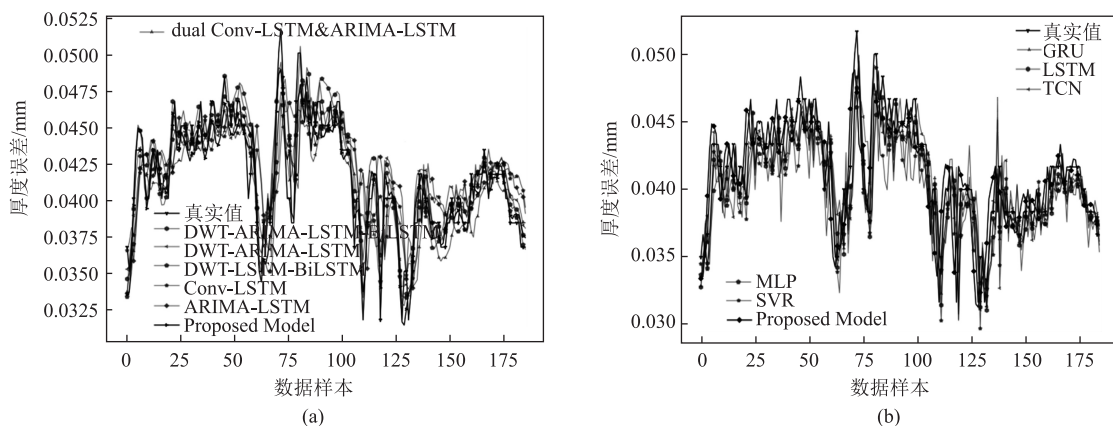


图 10 5 号机架带钢厚度 180 个点位的预测结果

(a) 消融实验 (b) 机器学习算法对比实验

Fig. 10 Predicted results of strip steel thickness for frame 5 at 180 points

(a) Ablation experiments (b) Comparison experiment of machine learning algorithms

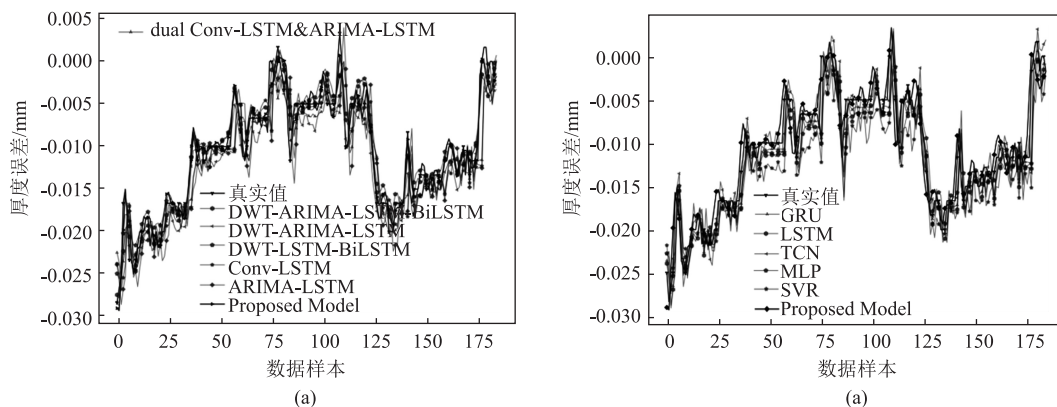


图 11 6 号机架带钢厚度 180 个点位的预测结果

(a) 消融实验 (b) 机器学习算法对比实验

Fig. 11 Predicted results of strip steel thickness for frame 6 at 180 points

(a) Ablation experiments (b) Comparison experiment of machine learning algorithms

在本文模型构建上, 将小波分解与其他模型融合, 对分解得到的子序列特征分别进行处理, 使得模型能够结合数据的整体走势和细节特征达到精准的预测效果, 相比 Conv-LSTM 和 ARIMA-LSTM 模型, 获得了较好的预测效果。

在前置时空矩阵构建上, 本文提出的模型在两组机架板厚数据集上效果均上优于 DWT-ARIMA-LSTM-BiLSTM 模型, 说明利用 Conv-LSTM 学习所构建的带钢厚度前置时空矩阵的时空特征能够有助

于提高模型的预测精度。ARIMA-LSTM 模型在两个数据集上的预测曲线与真实曲线的贴合程度优于单一的 ARIMA 模型和 LSTM 模型, 预测效果更好, 证明了组合传统预测模型和深度学习模型的可行性。

差分自回归移动平均模型能够捕获低频近似数据中的线性趋势。对于某些具有一定趋势性的时间序列数据, 可提高模型对线性趋势的预测能力, 从而提高整体预测的准确率。DWT-ARIMA-LSTM-BiLSTM 相对于 DWT-LSTM-BiLSTM 具有更好的预测效果。

表 1 不同板带厚度误差预测模型的性能指标对比

Table 1 Performance index comparison of different error prediction models for strip thickness

模型	F5 号机架			F6 号机架		
	$MAE/(\times 10^{-3})$	$RMSE/(\times 10^{-3})$	R^2	$MAE/(\times 10^{-3})$	$RMSE/(\times 10^{-3})$	R^2
SVR	2.657	3.460	0.218	3.378	4.236	0.594
MLP	2.362	3.147	0.353	2.512	3.464	0.728
TCN	2.452	3.249	0.311	2.494	3.541	0.716
LSTM	2.341	3.147	0.353	2.700	3.656	0.698
GRU	2.211	3.027	0.402	2.445	3.465	0.728
dual Conv-LSTM & ARIMA-LSTM	1.628	2.130	0.7036	2.234	3.128	0.780
DWT-ARIMA-LSTM-BiLSTM	1.149	1.457	0.861	1.264	1.684	0.936
DWT-ARIMA-LSTM	1.386	1.774	0.795	1.771	2.200	0.890
DWT-LSTM-BiLSTM	1.521	2.082	0.717	1.431	1.870	0.921
Conv-LSTM	2.030	2.787	0.493	2.237	3.164	0.774
ARIMA-LSTM	2.119	2.933	0.438	2.530	3.630	0.702
本文模型	0.997	1.322	0.887	0.980	1.301	0.962

对比 DWT-ARIMA-LSTM 和 DWT-LSTM-BiLSTM 的实验预测结果可知, 将原本用 BiLSTM 处理的地方换为 LSTM 处理后, 模型在两个测试集上的 MAE 分别下降了 2.34×10^{-4} 和 5.07×10^{-4} 。在本实验中, 使用 BiLSTM 对周期性数据进行建模, 能够更好地捕捉到数据的时序特征, 相比于 LSTM 更具优势。

4 结语

本文提出了一种时间序列混合预测模型用于热连轧厚度的预测。该模型由 Conv-LSTM 通道和 DWT-ARIMA-LSTM-BiLSTM 通道构成, 以多机架带钢厚度数据组成的前置时空矩阵和待预测带钢厚度数据作为输入数据, 充分考虑到机架组轧制的带钢厚度数据在时间特征和空间特征上的联系, 并基于小波变换将待预测机架带钢的厚度数据分解为低频近似子序列和高频细节子序列, 进而利用 ARIMA-LSTM 和 BiLSTM 在不同通路上分别提取出不同频率范围内的特征信息, 消除了各个分量对模型学习的干扰。通过对实际生产数据的分析和学习, 证明该模型可以动态对热连轧板带的厚度误差进行预测。

参考文献:

- [1] Wang Y, Luo X C, Zhang F, et al. GPU-based model predictive control for continuous casting spray cooling control system using particle swarm optimization [J]. Control Engineering Practice, 2019, 84: 349-364.
- [2] Sun J, Hou F, Hu Y J, et al. Application of distributed model predictive control based on neighborhood optimization in gauge-looper integrated system of tandem hot rolling [J]. Journal of Iron and Steel Research International, 2022, 30 (2): 277-292.
- [3] Li B Q, Guo J R, Zhao Z Y, et al. Application on neural PID control of MN-AGC in continuous hot strip rolling [A]. 22nd Chinese Control and Decision Conference [C]. Xuzhou: IEEE, 2010.
- [4] Meng F W, Hu Y B, Ma P Y, et al. Practical control of a cold milling machine using an adaptive PID controller [J]. Applied Sciences, 2020, 10 (7): 2516.
- [5] Mei Y, Wang Y D, Shen H, et al. Adaptive event-triggered fuzzy control for hot strip finishing mill via a disturbance observer-based method [J]. Applied Mathematical Modelling, 2023, 117: 417-430.
- [6] Li B Q, Wang L, Zhang S L, et al. Research on loop decoupling control based on fuzzy RBF neural network [A]. 33rd Chinese Control and Decision Conference [C]. Kunming: IEEE, 2021.
- [7] Zhang Z X. PLC realization and algorithm research of cold rolling mill thickness control system based on neural network [A]. 6th International Conference on Mechatronics and Intelligent Robotics [C]. Kunming: SPIE, 2022, 12301: 848-854.
- [8] Qian C, Hua C C, Zhang L L, et al. Adaptive neural torsional vibration suppression of the rolling mill main drive system subject to state and input constraints with sensor errors [J]. Journal of the Franklin Institute, 2020, 357 (17): 12886-12903.
- [9] Sedykh I, Istomin V. Management of the strip cooling process using neural networks based on clustered data [A]. 2nd International Conference on Control Systems, Mathematical Modeling, Automation and Energy Efficiency [C]. Russia: IEEE, 2020.
- [10] Deng J F, Sun J, Peng W, et al. Application of neural networks for predicting hot-rolled strip crown [J]. Applied Soft Computing, 2019, 78: 119-131.

- [11] Huang Y, Zhou X M, Gao Z Y. Thickness prediction of thin strip cold rolling based on VBGm-RBF [J]. The International Journal of Advanced Manufacturing Technology, 2022, 120 (9-10): 5865-5884.
- [12] Liu Y, Wang X J, Sun J, et al. Strip thickness and profile-flatness prediction in tandem hot rolling process using mechanism Model-Guided machine learning [J]. Steel Research International, 2023, 94 (1): 1-14.
- [13] 杨静, 任彦, 高晓文, 等. 基于 GA-PELM 的板材热连轧轧制力预测 [J]. 锻压技术, 2022, 47 (1): 43-48.
Yang J, Ren Y, Gao X W, et al. Rolling force prediction of hot strip rolling based on GA-PELM [J]. Forging & Stamping Technology, 2022, 47 (1): 43-48.
- [14] Ding J G, He Y H C, Kong L P, et al. Camber prediction based on fusion method with mechanism model and machine learning in plate rolling [J]. ISIJ International, 2021, 61 (10): 2540-2551.
- [15] 姬亚锋, 刘瑜, 宋乐宝, 等. 基于 M-SVR 的热连轧板带宽度-厚度预测 [J]. 塑性工程学报, 2022, 29 (4): 58-64.
Ji Y F, Liu Y, Song L B, et al. Width-thickness prediction of hot-rolled strip based on M-SVR [J]. Journal of Plasticity Engineering, 2022, 29 (4): 58-64.
- [16] Hu X, Zhang D H, Tan R J, et al. Controlled cooling temperature prediction of hot-rolled steel plate based on multi-scale convolutional neural network [J]. Metals, 2022, 12 (9): 1-14.
- [17] Chen Y F, Peng L G, Wang Y, et al. Prediction of tandem cold-rolled strip flatness based on Attention-LSTM model [J]. Journal of Manufacturing Processes, 2023, 91: 110-121.
- [18] 于加学, 孙杰, 张殿华. 基于深度学习的热轧带钢头部厚度的命中预测 [J]. 钢铁, 2021, 56 (9): 19-25.
Yu J X, Sun J, Zhang D H. Prediction of head thickness accuracy for hot-rolled strip steel based on deep learning [J]. Iron and Steel, 2021, 56 (9): 19-25.
- [19] Huang S Y, Tian T. Prediction of mechanical properties of hot rolled strip based on DBN and composite quantile regression [A]. 4th International Conference on Algorithms, Computing and Artificial Intelligence [C]. Sanya, 2021.
- [20] 张瑞成, 刘力菲, 梁卫征. 基于 SSAE-LSTM 模型的冷连轧机扭转预测 [J]. 锻压技术, 2023, 48 (4): 193-198.
Zhang R C, Liu L F, Liang W Z. Torsional vibration prediction on tandem cold rolling mill based on SSAE-LSTM [J]. Forging & Stamping Technology, 2023, 48 (4): 193-198.

(上接第 106 页)

- Zhong H J, Liao D, Xiong S, et al. Control of pock marks on surface of the outer panel [J]. Forging & Metalforming, 2022, (22): 52-55.
- [3] 赵子海, 孟宇轩, 刘英堂, 等. 汽车覆盖件铝板冲压工艺研究 [J]. 模具制造, 2020, (2): 29-32.
Zhao Z H, Meng Y X, Liu Y T, et al. Research on stamping process of aluminum sheet for automobile cover parts [J]. Die Mould Manufacture, 2020, (2): 29-32.
- [4] 曹晨华, 才佳, 张一闯. 汽车白车身轻量化-铝板技术的应用 [J]. 汽车实用技术, 2019, (21): 130-132.
Cao C H, Cai J, Zhang Y C. Car body lightweight-application of aluminum plate technology [J]. Automobile Applied Technology, 2019, (21): 130-132.
- [5] 佟琳. 汽车轻量化—汽车铝板在白车身和覆盖件减重中的应用 [J]. 世界有色金属, 2014, (2): 54-56.
Tong L. Automobile light weighting-The application of automotive aluminum sheets in body-in-white and component weight reduction [J]. World Nonferrous Metals, 2014, (2): 54-56.
- [6] 陈垒, 姚海军, 王同, 等. 关于汽车冲压板料油膜控制的研究 [J]. 锻造与冲压, 2021, (2): 61-65.
Chen L, Yao H J, Wang T, et al. Research on oil film control of the sheet stamping for car [J]. Forging & Metalforming, 2021, (2): 61-65.
- [7] 谢晖, 刘翔, 王诗恩, 等. 时效硬化对 6021-T4 铝板成形性影响及冲压工艺稳健设计 [J]. 塑性工程学报, 2016, 23 (1): 21-26.
Xie H, Liu X, Wang S E, et al. Age hardening effect on formability of 6021-T4 aluminum and robust stamping process design [J]. Journal of Plasticity Engineering, 2016, 23 (1): 21-26.
- [8] 曾渝, 王颖婧, 张晓蕾, 等. 6016 铝板性能解析及挽救工艺研究 [J]. 汽车工艺与材料, 2018, (8): 1-5.
Zeng Y, Wang Y J, Zhang X L, et al. Performance analysis of 6016 aluminum sheet and remedy process research [J]. Automobile Technology & Material, 2018, (8): 1-5.
- [9] 于燕, 赵洪运, 宋起峰, 等. 表面形貌和润滑条件对轿车用国产冷轧薄板的成形性影响 [J]. 精密成形工程, 2000, 18 (1): 42-43.
Yu Y, Zhao H Y, Song Q F, et al. Influence of surface and lubrication on forming ability of cold-rolled thin plate for home made cars [J]. Journal of Netshape Forming Engineering, 2000, 18 (1): 42-43.
- [10] 陈明达. 基于 AutoForm 软件的铝合金汽车冲压件成形模拟 [J]. 汽车制造业, 2018, (6): 54-56.
Chen M D. Simulation of forming aluminum alloy automotive stamping parts based on autoForm software [J]. Automobil Industrie, 2018, (6): 54-56.
- [11] 陈道钟, 李建华, 周洁, 等. 汽车覆盖件板成形的数值模拟技术及其在车身模具开发中的应用 [J]. 模具制造, 2003, (8): 11-14.
Chen D Z, Li J H, Zhou J, et al. Simulation technology of the sheet forming for automotive panel and its application on the development of body die [J]. Die Mould Manufacture, 2003, (8): 11-14.
- [12] 王祝堂. 未来 13 年是乘用车汽车铝板增长最快时期 [J]. 轻金属, 2014, (11): 49.
Wang Z T. The next 13 years will be the fastest period of growth for automotive aluminum sheets in passenger vehicles [J]. Light Metals, 2014, (11): 49.