

doi:10.6041/j.issn.1000-1298.2021.05.045

基于注意力机制的时空卷积数控机床热误差模型研究

杜柳青 李仁杰 余永维
(重庆理工大学机械工程学院, 重庆 400054)

摘要: 为了提高数控机床热误差模型的精度与泛化性,提出了基于注意力机制的长短时记忆卷积神经网络(Long short term memory convolutional neural network based on attention mechanism, AM-CNN-LSTM)热误差模型。利用卷积神经网络提取高维数据空间状态特征的能力和长短时记忆网络提取长时间序列状态特征的能力,构建具有2个支路的热误差模型,分别提取特征后输入到注意力机制中进行特征重要性重构,建立原始数据与热误差的特征映射,最后通过全连接层进行热误差预测。采用G460L型数控机床进行实验数据采集,将不同季节采集到的温度数据和热误差作为模型输入,采用循环学习率与正则化优化方法对模型进行训练。与LSTM、ConvLSTM和CNN-LSTM热误差模型对比,结果表明,AM-CNN-LSTM模型对特征还原能力最强,残差波动范围最小,其残差范围较最大值下降62.09%,模型预测精度在2.4 μm以内。

关键词: 注意力机制; 热误差模型; 数控机床; 长短时记忆卷积神经网络

中图分类号: TH115 文献标识码: A 文章编号: 1000-1298(2021)05-0404-08 OSID: 

Spatiotemporal Convolution Thermal Error Model of CNC Machine Tools Based on Attention Mechanism

DU Liuqing LI Renjie YU Yongwei
(College of Mechanical Engineering, Chongqing University of Technology, Chongqing 400054, China)

Abstract: In order to improve the accuracy and generalization of the thermal error model of CNC machine tools, a thermal error model of and long short term memory convolutional neural network based on attention mechanism (AM-CNN-LSTM) was proposed. A thermal error model with two branches was established by using the ability of convolutional neural networks to extract the space state features of high-dimensional data and the ability of long short term memory networks to extract long-term sequence state features, and the extracted features were input into the attention mechanism to reconstruct according to the importance, and then a feature map of original data and thermal error value was established. Finally, the thermal error prediction value was performed through the full connect layer. The G460L CNC lathe was used to collect experimental data, the temperature and thermal error collected in different seasons were used as the model input, and the model was trained using the cyclic learning rate and regularization optimization method. Compared with the thermal error model of LSTM, ConvLSTM and CNN-LSTM, the results showed that AM-CNN-LSTM model had the strongest ability to restore features and the smallest residual error range. It was decreased by 62.09%, and the prediction accuracy of the model was within 2.4 μm.

Key words: attention mechanism; thermal error model; CNC machine tool; long short term convolutional neural network

0 引言

高端装备制造业对机床精度要求越来越高。机

床误差大致可分为:几何和运动学误差、热误差、切削力引起的误差及其他误差。其中,热误差所导致的机床总加工几何误差为最主要因素^[1-2]。国内外

收稿日期: 2020-07-07 修回日期: 2020-08-12

基金项目: 国家自然科学基金项目(51775074)、重庆市重点产业共性关键技术创新重点研发项目(cstc2017zdcy-zdyfX0066、cstc2017zdcy-zdyfX0073)、重庆市技术创新与应用示范重点项项目(cstc2018jszx-cyzdX0144)、重庆市基础研究与前沿探索项目(cstc2018jcyjAX0352)和重庆市研究生科研创新项目(CYS19316)

作者简介: 杜柳青(1975—),女,教授,博士,主要从事机床精度设计和智能制造研究,E-mail: lqdu@cqut.edu.cn

学者通过研究热误差的产生机制,提出了理论热误差模型,JORGENSEN^[3]考虑不同润滑条件下利用集中质量法对某机床主轴系统进行了稳态求解;KIM 等^[4-5]采用质量集中法对机床丝杠系统建模,并与有限元模型进行对比。也有学者通过研究热误差的主要产生原因即温度与主轴热漂移的关系,提出了经验热误差模型,李彬等^[6]建立了遗传算法(Gentic optimization algorithm, GA)优化小波神经网络(Wavelet neural network, WNN)的数控机床热误差模型,在某型号五轴摆动卧式加工中心进行实验,验证该模型具有精度高、抗扰动能力和鲁棒性强等优点。夏军勇等^[7]通过对机床进给系统的热特性分析,得到随时间的延长,温升与热变形之间的关系会逐渐趋于稳定,但不可能达到绝对稳态的结论。此外,机床工作环境的变化也会导致机床温度敏感点的变化,在不同转速下的机床主轴热特性以及机床在空载与负载状态下的热特性都具有极大的不确定性。

近年来,深度学习模型在机床状态监测与误差预测领域得到了应用,与传统的数学模型、神经网络模型相比,深度学习模型能更好地对原始数据进行特征提取,缩短了对特征提取所需的时间,同时也减少了由于领域知识限制等而导致的特征信息丢失,这些均对模型的质量与性能产生影响。杜柳青等^[8]采用深度自编码器(Stacked automatic encoder, SAE)对筛选出的温度敏感点进行特征重构,利用遗传算法对 BP 神经网络进行参数寻优,建立了 SAE-GA-BP 热误差模型。余永维等^[9]利用相空间重构原理建立模型输入向量,提出并建立了基于长短时记忆(Long short term memory, LSTM)网络的时序深度学习数控机床运动精度预测模型,实验表明,预测相对误差不超过 7.96%。余永维等^[10]提出一种基于 Faster R-CNN 的数控机床运动误差溯因模型,实现了对数控机床圆运动轨迹的实时准确识别。杜柳青等^[11]提出无监督初始权值共享的并联深度网络训练方法,建立的并联深度信念网络热误差模型具有较高的预测精度与鲁棒补偿效果。陈英义等^[12]引入小波变换对原始数据进行降噪,利用 CNN 挖掘变量之间的关系,利用 LSTM 对溶解含氧量进行预测,建立的溶解含氧量预测模型决定系数达 0.954。高浩然^[13]建立了基于机床能耗数据的 LSTM 热误差模型,解决了卷积神经网络模型的误差累计问题,模型预测精度在 5 μm 以内。卷积神经网络通过卷积核的共享可以有效提取出高维数据的深层次特征。长短时记忆网络作为循环网络的变体,解决了传统循环神经网络(Recurrent neural

network, RNN)的梯度消失、梯度爆炸等问题,能更好地解决长时间序列问题,从而被广泛应用在语音识别、机器翻译等领域。程成^[14]利用注意力机制(Attention mechanism, AM)和 ConvLSTM 实现对传感器数据的时空特征提取,建立发动机剩余寿命预测模型,引入注意力机制 ConvLSTM 模型的 RMSE 与原模型相比降低了 14.07%。机床误差信息也属于连续的长时间序列,利用 LSTM 的时间序列特征提取能力,能有效弥补 CNN 在特定时空场景下特征提取的不足^[15-16]。

本文提出一种基于注意力机制的长短时记忆卷积神经网络热误差模型。利用 CNN 提取原始数据中的深层次空间特征,利用 LSTM 提取原始数据中的长时间序列特征,引入注意力机制使模型着重于某时刻局部的重要特征,以高的重要特征权重,使得模型具有更高的预测精度。在模型中引入正则化与循环学习率的训练方法,在提高训练速度的同时避免模型过拟合。通过实验验证该模型的有效性与可行性。

1 基于 AM-CNN-LSTM 热误差模型原理

1.1 热误差模型基本框架

提出了一种基于注意力机制的时空卷积热误差模型(图 1),该模型具有两个支路:支路 1 为卷积神经网络支路,输入传感器采集的图像化温度数据,通过 3 个卷积层和 2 个池化层,提取原始高维温度数据中的多维度空间特征输出到全连接层中;支路 2 为长短时记忆网络,输入数据为时间序列温度数据,经过 LSTM 层后提取出长时间序列温度数据中的时序特征输出到全连接层中。支路 1 与支路 2 融合在全连接层后输入到注意力层中进行重要特征重构,最后将重构得到的特征输入回归层得到热误差预测值。

1.2 LSTM 网络支路设计

LSTM 神经网络是一种特殊类型的 RNN,具有利用反馈连接来学习长期依赖性的能力。RNN 存在梯度消失问题的困扰,该问题限制了模型学习远程依赖关系。因此,LSTM 通过在存储单元上存储有用的信息并消除不必要的信息来解决该问题,从而获得比传统 RNN 更好的性能^[17-18]。

每个 LSTM 单元都由 1 个存储单元和 3 个主门组成:输入门、输出门和遗忘门。通过这种结构,LSTM 通过确定一些信息必须“忘记”而另一些信息必须“记住”来创建受控信息流,从而设法学习长期依赖关系。遗忘门 f_t 控制在时刻 $t-1$ 必须消失或必须保留在存储单元上的过去信息,即

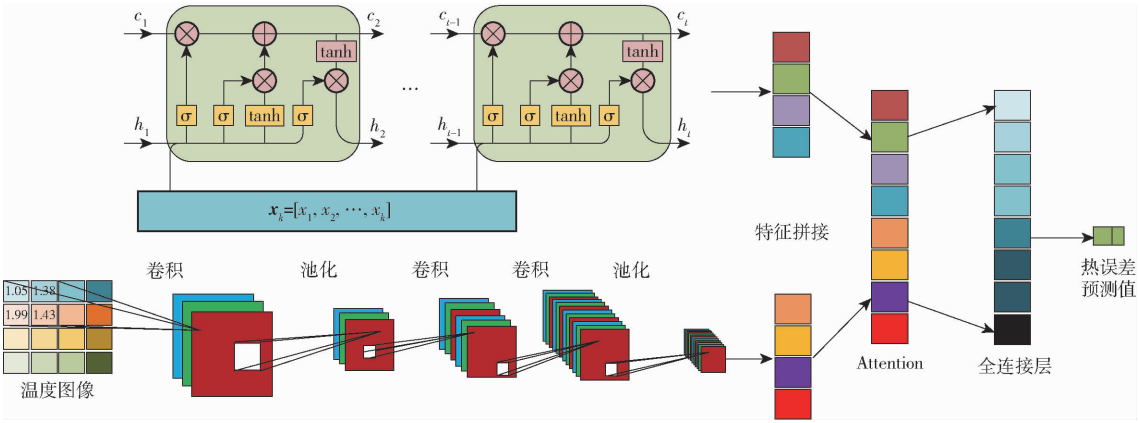


图1 AM-CNN-LSTM 热误差模型结构

Fig.1 Structure of thermal error model of AM-CNN-LSTM

$$f_i = \sigma(U_g x_i + W_g h_{i-1} + b_g)$$

式中 x_i ——温度序列数据

W_g 、 U_g ——权重矩阵

b_g ——偏置

σ ——sigmoid 激活函数

输入门 i_t 与单元状态 c_t 同时控制在时刻 t 存储到存储状态 c_t^* 中的新信息。在获得遗忘门和输入门的温度特征序列信息后,即可得到当前更新状态的状态单元 c_t ,计算方法为

$$i_t = \sigma(U_i x_t + W_i h_{t-1} + b_i)$$

$$c_t^* = \tanh(U_c x_t + W_c h_t + b_c)$$

$$c_t = f_t \odot c_{t-1} + i_t \odot c_t^*$$

式中 $\odot$ ——逐元数乘积

输出门 o_t 控制可以将一些信息用于存储单元的输出,并将 LSTM 网络提取的长时序温度数据特征向量 f_{LSTM} 输出到特征融合层。

$$o_t = \sigma(U_o x_t + W_o h_{t-1} + b_o)$$

最后,构成存储单元输出的隐藏状态 h_t 为

$$h_t = o_t \odot \tanh(c_t)$$

如多个 LSTM 层堆叠在一起,则每个 LSTM 层的存储状态 c_t 和 h_t 隐藏状态都作为输入转发到下一个 LSTM 层。

1.3 CNN 网络支路设计

卷积神经网络基本结构由输入层、卷积层 (Convolution layer)、池化层 (Pooling layer)、全连接层 (Full connect layer) 以及输出层组成,其网络结构可根据目标对象不同而调整,卷积层与池化层可交替连接存在一个或多个。卷积层是实现特征提取的关键,其建立依赖于卷积核的设计,卷积神经网络的特点之一是共享卷积核,优势是处理高维数据快。每个卷积核提取对应输入图像的一个特征,网络输入经过卷积层后生成的新图像称特征图,每个下层的特征图是由上层的几个特征图形成映射。特征图

的每个神经元与其输入进行局部连接,并通过对应的连接权值与局部输入进行加权求和再加上偏置值,得到该神经元输入值,该过程等同于卷积过程,卷积层的映射公式为

$$x_j^i = f\left(\sum_{i \in M_i} x_i^{t-1} * k_{ij}^l + b_i^l\right)$$

式中 x ——特征图

k ——卷积核

M_i ——感受野

$*$ ——卷积运算符号

$f(\cdot)$ ——激活函数

采用 ReLU 型激活函数加速网络的收敛,计算公式为

$$f(z) = \begin{cases} 0 & (z < 0) \\ z & (z \geq 0) \end{cases}$$

池化层也称为下采样层,能够减小温度特征图的尺寸,在尽可能保留较多原始特征信息的同时降低特征的维度,减少了运算量,并且具有平移不变性的特点。池化层映射公式为

$$x_j^l = f(\beta_j^l \text{down}(x_j^{l-1}) + b_j^l)$$

式中 $\text{down}(\cdot)$ ——池化函数

β ——权重 b ——偏置

全连接层的功能是将多维的特征图展成单一维度的向量,向下传递进入全连接层网络,其公式为

$$x^l = f(w^l x^{l-1} + b^l)$$

采用 sigmoid 激活函数得到 CNN 网络提取的空间温度数据特征 f_{CNN} ,并输入到特征融合层。

1.4 注意力机制

注意力机制^[19]是为了网络在每一步挑选信息时都能从更大的信息集里挑出有用信息。对输入模型的特征信息进行权重的重新分配,对重要的特征分配更多的注意力,突出典型关键特征对结果的影响。在时间序列的数据中引入注意力机制能更好地使模型关注到更有用的信息。Attention 机制主要由 3 个阶段组成,第 1 阶段计算输入特征相似度,输入特征向量为 CNN 与 LSTM 两条支路提取的特征融

合 $\mathbf{x}_k = (f_{\text{LSTM}}, f_{\text{CNN}}) = [x_1, x_2, \dots, x_k]$, 计算公式为

$$s_k = g(\mathbf{x}_k) = \text{activation}(W\mathbf{x}_k + \mathbf{b}) \tag{11}$$

为了更好地使模型注意到每个时刻的时间与空间特征, 激活函数 activation 设置为 sigmoid 。

第 2 阶段根据第 1 阶段产生的不同特征值相似分数进行数值转换, 通过 softmax 归一化函数的内在机制使重要特征元素的权重更加突出, 计算方法为

$$\alpha_k = \text{softmax}(s_k) = \frac{e^{s_i}}{\sum_{i=1}^k e^{s_i}} \tag{12}$$

第 3 阶段对第 2 阶段得到的特征权重进行加权求和计算即可得到 Attention 数值, 即经由 Attention 层进行重要性特征重构后的特征值。

$$x^* = \sum_{i=1}^k \alpha_i x_i \tag{13}$$

2 实验数据采集

提出的 $\text{AM} - \text{CNN} - \text{LSTM}$ 模型具有良好的泛化性, 适用于变化状态下的热误差预测, 在此通过 G460L 型数控车床进行实验数据采集。实验所建立的数据采集系统如图 2 所示。温度传感器采用 Fluke 公司的 PT100, 测量精度为 $\pm 0.2^\circ\text{C}$; 位移传感器型号为 ML33-B-B2-3, 测量精度为 $0.5\text{ }\mu\text{m}$, 采集到的模拟数据信号经过 ART 公司 USB3120 数据采集卡后转换为数字信号, 由 LabVIEW 数据采集软件存储与显示。

传感器布置如图 3 所示, 在主轴径向与轴向方向分别布置 2 个电涡流位移传感器用于测量主轴径

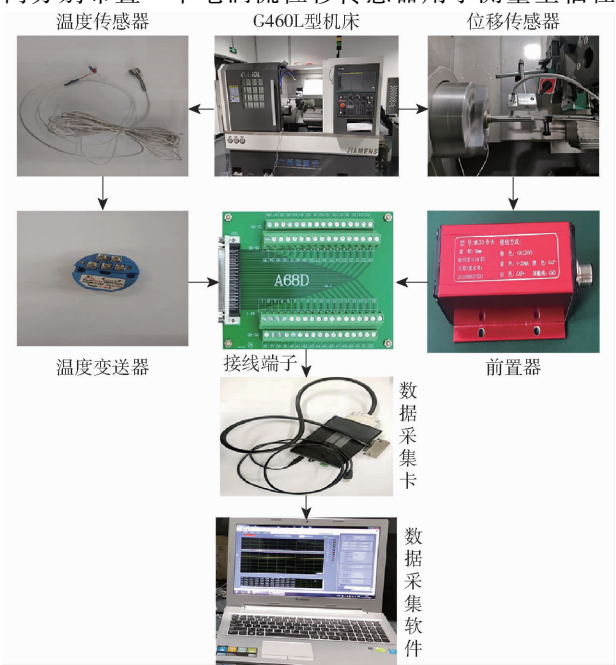


图 2 数据采集系统

Fig. 2 Data collection system

向与轴向热误差。温度传感器的布置贴近所有的激励热源, 对同一激励热源处测点间隔一定距离均匀分布布置, 难以测量的激励热源位置采用相近温度测点布置方法, 温度传感器 T1 ~ T16 布置如表 1 所示。

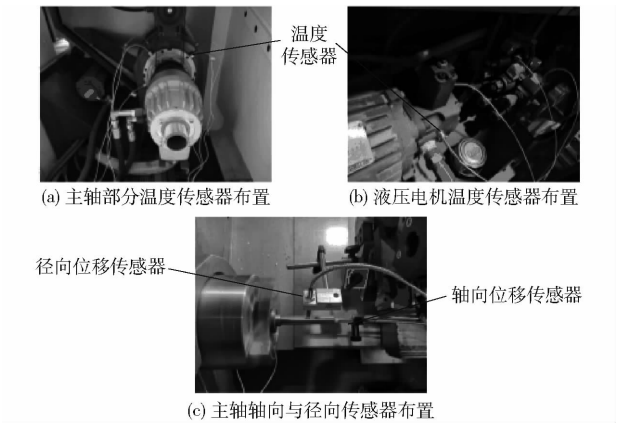


图 3 传感器布置实物图

Fig. 3 Picture of sensor arrangement

表 1 传感器布置位置

Tab. 1 Sensor arrangement positions

传感器	布置位置
T1、T2、T3、T4	主轴前轴承
T5、T6、T7、T8	主轴后轴承
T9、T10	主轴外壳体
T11、T12、T13	液压泵电机
T14、T15	液压换向阀
T16	机床床身

采集数据时, 机床主轴以恒定的转速(1 500、3 000 r/min)空转, 数据采样间隔为 30 s, 总测量时间为 240 min。根据上述实验方法, 在不同的季节内共测得 12 批次原始数据, 如表 2 所示。

表 2 数据采集实验参数

Tab. 2 Data acquisition experiment parameters

数据批次	实验时间	主轴转速/(r·min ⁻¹)	环境温度/℃
K1	2019-12-31	1 500	15.0
K2	2020-01-01	1 500	15.8
K3	2020-01-02	3 000	15.0
K4	2020-01-03	3 000	15.7
K5	2020-01-04	3 000	15.4
K6	2020-01-05	3 000	15.4
K7	2020-06-01	1 500	26.0
K8	2020-06-02	1 500	26.0
K9	2020-06-03	3 000	24.6
K10	2020-06-04	3 000	25.0
K11	2020-06-05	3 000	25.8
K12	2020-06-06	3 000	26.0

将数据采集系统的数据转换为增量形式后保存, 得到 12 批次的热误差如图 4 所示。从图 4 中可以看出, 在不同季节下机床主轴达到相对热平

衡状态所需时间在 150 min 左右,达到平衡状态下的热误差随季节温度的增加而增加。第 7 批次的温度如图 5 所示。从图 5 可以看出,机床温度具有较大的耦合性且邻近主轴位置温度测点的温度变化最大。

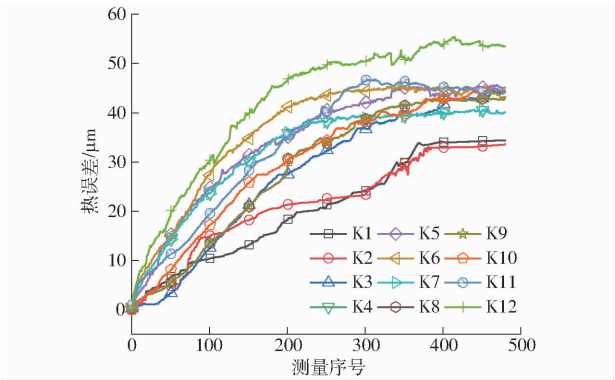


图 4 K1 ~ K12 热误差变化曲线
Fig. 4 Thermal error data of K1 ~ K12

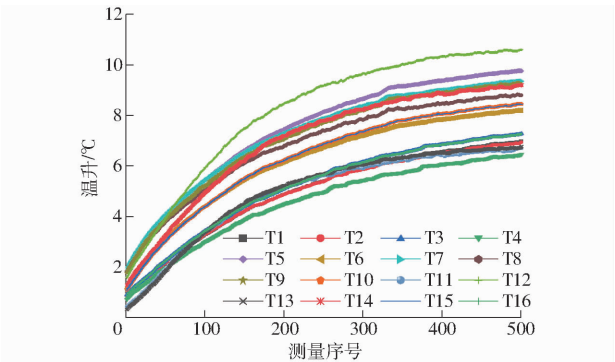


图 5 第 7 批次温度变化曲线
Fig. 5 Temperature data of K7

3 模型训练及热误差预测

3.1 模型训练

所建立 AM - CNN - LSTM 热误差模型基于 TensorFlow 框架,集成开发环境为 Pycharm,实验硬件配置为 Intel Core i3 - 7100 CPU,8 GB 内存。

为了提高网络学习速度,避免原始数据中存在的异常数据和不同数据维度导致网络在训练过程中无法收敛,将 12 个批次的数据进行最大最小归一化处理,计算公式为

$$X_{\text{norm}} = \frac{X - X_{\min}}{X_{\max} - X_{\min}} \tag{14}$$

式中 $X_{\max}$ 、 $X_{\min}$ ——原始数据中的最大值与最小值
将归一化后的数据 80% 用于训练,20% 用于验证。根据本文提出的基于注意力机制的时空卷积热误差模型(图 1),支路 1 为 CNN 模型,模型输入为 (4,4) 图像化温度,通过数据增强方法将原始图像尺寸增大到 (12,12)。将输入模型的数据进行随机打乱,以提取原始数据的多维特征。采用 Adam 优

化方法对网络的超参数进行更新,采用循环学习率^[20] (Cyclical learning rates, CLR)用以加快模型的训练。CLR 方法是使学习率在合理的边界值之间循环变化,采用周期性学习率可以提高模型的精度,而无需进行手动调整且迭代的次数少。

支路 2 为 LSTM 模型,模型的输入为 (1,16) 的温度向量,步长设置为 8,共设置 2 个 LSTM 层深度不断加深,激活函数设置为 tanh。为防止过拟合,在训练过程中加入 Dropout 操作和 BatchNormalization 批正则化方法,保留概率参数为 0.7,批尺寸为 24,训练次数为 120,CLR 中最小学习率为 0.001,最大学习率为 0.008,半个学习率循环步长为 20,在整个训练周期中共有 3 个学习率调整周期,且保证最后一个周期中对应的学习率达到最小,从而保证模型收敛于一个局部最优解,学习率循环模式为三角形减半下降模式。

2 个支路各自提取出 4 个空间与时间上的特征后输入到 Attention 层中进行特征权重系数分配。最后在全连接层中采用 sigmoid 激活函数得到具有时间连续性的热误差。

为了进行验证对比,建立了 4 种热误差模型,如表 3 所示。LSTM 模型为传统的 3 层长短期记忆神经网络;ConvLSTM 模型为时空长短期记忆网络,将权重与输入之间的点乘操作变换为卷积操作,模型的输入变为温度图像。CNN - LSTM 模型为长短期记忆卷积神经网络,输入数据经过卷积神经网络提取特征后输入到长短期记忆神经网络中进行热误差的预测。这 3 种模型都采用循环学习率、Dropout 操作和 BatchNormalization 批正则化方法,训练数据集设置、各项参数设置均与 AM - CNN - LSTM 相同。图 6 为模型训练损失值对比。由图 6 可知,经过 120 轮训练后所有模型都达到了收敛状态。

图 7 为 CNN 与 LSTM 两条支路在所提取的特征在融合层的特征值热力图,热力图中颜色越深代表特征值越大,对模型的影响也越大。图 8 为融合层特征经过 Attention 层后对特征进行重要性分配后的特征热力图,随着测量序列节点的变化,热力图中颜色表明不同的特征在不同的序列节点被赋予不同的权重,更好地挖掘了空间层次的信息^[21-22]。

3.2 预测结果与分析

利用热误差模型预测得到热误差曲线如图 9 所示。从图 9 可以看出,热误差模型都能对长时间序列下热误差进行较为准确的预测,可以提取出原始数据中点时间序列特征,由于获取了原始数据在不同时间空间状态下的特征,AM-CNN-

表 3 热误差模型信息	
模型	模型信息
LSTM	LSTM1 神经元数:16
	LSTM2 神经元数:32
	LSTM3 神经元数:64
ConvLSTM	ConvLSTM1,卷积核大小 3×3,卷积层数 16
	ConvLSTM2,卷积核大小 3×3,卷积层数 32
	ConvLSTM3,卷积核大小 3×3,卷积层数 64
	Attention 层,神经元数 1 024
	全连接层,神经元数 64-32-8-1
CNN-LSTM	卷积层 1,卷积核大小 3×3,卷积层数 5
	池化层 1,最大池化大小 2×2
	卷积层 2,卷积核大小 3×3,卷积层数 10
	池化层 2,最大池化大小 2×2
	卷积层 3,卷积核大小 3×3,卷积层数 32
	卷积层 4,卷积核大小 3×3,卷积层数 64
	LSTM1 神经元数:64
	LSTM2 神经元数:128
	Attention 层,神经元数:128
	全连接层,神经元数:32-1
AM-CNN-LSTM	支路 1
	卷积层 1,卷积核大小 3×3,卷积层数 5
	池化层 1,最大池化大小 2×2
	卷积层 2,卷积核大小 3×3,卷积层数 10
	池化层 2,最大池化大小 2×2
	卷积层 3,卷积核大小 3×3,卷积层数 32
	卷积层 4,卷积核大小 3×3,卷积层数 64
	全连接层,神经元数:64-16-4
	支路 2
	LSTM1 神经元数:32
	LSTM2 神经元数:64
	全连接层神经元数:32-4
	融合层
	Attention 层,神经元数:8
	全连接层,神经元数:8-1

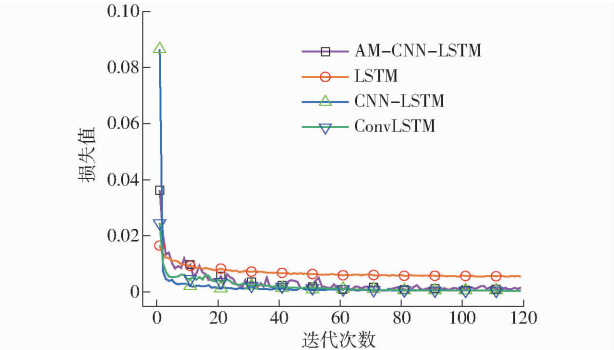


图 6 各模型训练损失值对比

Fig.6 Comparison graph of each model training loss values

LSTM 模型预测效果最好,预测出的热误差曲线更接近真实的测量曲线。LSTM 模型仅提取出原始数据中的时序特征的预测值波动幅度较大。CNN-LSTM

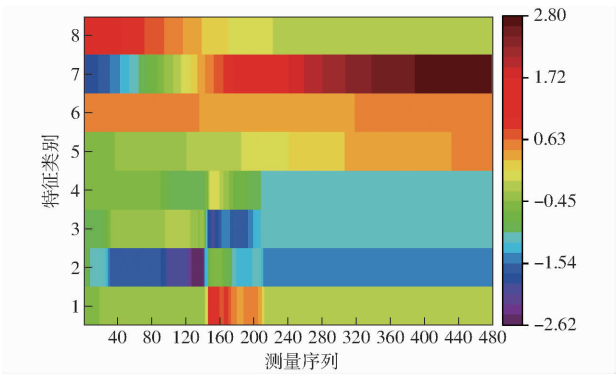


图 7 融合层特征热力图

Fig.7 Heat map of merge layer feature

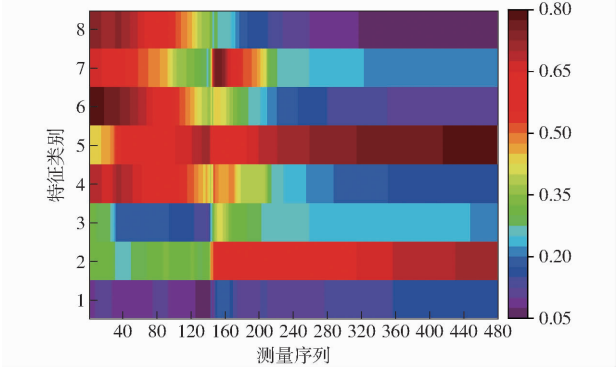


图 8 Attention 层特征热力图

Fig.8 Heat map of Attention layer feature

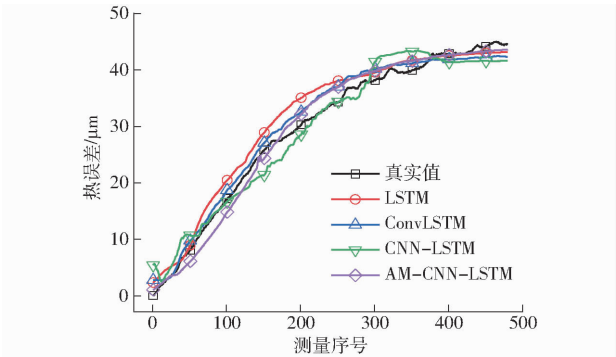


图 9 4 种模型热误差预测曲线

Fig.9 Thermal error prediction curves of four models

LSTM 模型由于在 CNN 输出阶段将特征图转换为特征向量,丢失了一部分特征信息,且模型的输入数据是随机打乱进入模型训练的,也丢失了一部分时序特征,使得模型的预测效果低于 AM-CNN-LSTM 模型。

为了进一步评估模型的性能,采用平均绝对误差 (Mean absolute error, MAE)、均方根误差 (Root mean squared error, RMSE)、均方差 (Mean squared error, MSE) 和纳什系数 (Nash-sutcliffe efficiency coefficient, NSE)对模型的性能进行量化的评定,计算公式为

$$MAE = \frac{1}{n} \sum_{i=1}^k |y_i - y_{pre}|$$

(15)

$$MSE = \frac{1}{n} \sum_{i=1}^k (y_i - y_{pre})^2$$

(16)

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^k (y_i - y_{pre})^2}$$

(17)

$$NSE = 1 - \frac{\sum_{i=1}^k (y_i - y_{pre})^2}{\sum_{i=1}^k (y_i - \bar{y})^2}$$

(18)

$$f(v_i) = \frac{|v - v_i|}{v}$$

(19)

式中 y_i ——实际值 y_{pre} ——预测值
 k ——预测值个数
 $\bar{y}$ —— k 个实际值的平均数

NSE 越接近 1 表示模型的预测效果越好,精度越高。式(19)用于描述数据相对增长量,反映 2 组数据在不同情况下的变化情况。

4 种模型的残差对比如图 10 所示。从图 10 可以看出,当机床接近达到相对热误差平衡状态时,4 种热误差模型都提取到了相同的特征,但 AM-CNN-LSTM 模型对特征的还原能力最强。CNN-LSTM 模型残差波动范围最大,为 $-5.447\ 3 \sim 4.500\ 8\ \mu\text{m}$ 。AM-CNN-LSTM 模型的残差波动范围最小,为 $-2.416\ 1 \sim 1.355\ 3\ \mu\text{m}$,AM-CNN-LSTM 模型的残差相较于 CNN-LSTM 模型下降 62.09%。证明模型预测精度在 $2.4\ \mu\text{m}$ 以内^[23-24]。

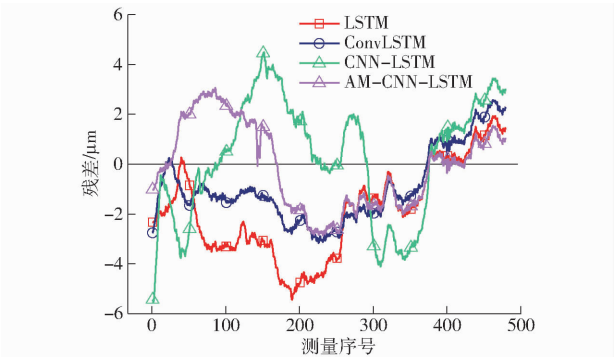


图 10 4 种模型残差对比
Fig. 10 Comparison chart of residual values of four models

表 4 为模型性能评价。由表 4 可以看出,CNN-LSTM 模型的 MSE、RMSE、MAE 均为最大,模型的各方面性能表现较差。ConvLSTM 模型的预测精度较好,由于将原始输入时间序列转换为图像序列,通过卷积操作的方式也能提取部分时间与空间特征,模型的 MSE 和 MAE 也相较 LSTM 和 CNN-LSTM 模型有一定的下降。ConvLSTM 模型的预测精度与残差波动范围与 AM-CNN-LSTM 模型都较为接近,在局部预测效果上优于 AM-CNN-LSTM 模型,但是考虑不同时间的适用性等综合因素,AM-CNN-LSTM 模型的综合性能更优。

表 4 4 种模型性能评价结果

Tab. 4 Performance evaluation of four models

模型	MSE/ μm^2	RMSE/ μm	MAE/ μm	NSE
LSTM	6. 653 5	2. 579 4	2. 200 4	0. 955 4
ConvLSTM	3. 839 8	1. 959 5	1. 775 7	0. 975 1
CNN-LSTM	6. 902 8	2. 627 3	2. 259 7	0. 956 1
AM-CNN-LSTM	2. 751 1	1. 658 6	1. 453 8	0. 985 5

AM-CNN-LSTM 模型的纳什系数最高,达到了 0.985 5,表明模型对原始数据的解释能力最好,模型精度高。相较于 LSTM、ConvLSTM 和 CNN-LSTM 模型,由式(19)计算可知纳什系数分别提高了 3.05%、1.06% 和 2.98%。原始温度数据来自不同季节不同转速下,也使得模型具有良好的泛化性。与 LSTM 模型相对比,融合模型在热误差预测中性能有了较大提升,融合深度学习模型能从原始数据中挖掘更深层空间状态特征和更全面的时间状态特征,证明了融合深度学习模型在热误差建模领域的应用可行性和有效性。

4 结论

- (1)相较于 LSTM 模型、ConvLSTM 模型和 CNN-LSTM 模型,AM-CNN-LSTM 模型相对预测精度分别提高了 3.05%、1.06% 和 2.98%。
- (2)采用不同深度学习模型来提取高维温度数据中的不同特征,将模型融合后得到的特征映射效果优于原始单一深度学习模型。

参 考 文 献

[1] MAYR J, JEDREJEWSKI J, UHLMANN E, et al. Thermal issues in machine tools[J]. CIRP Annals, 2012,61(2):771-791.

[2] LI Yang, ZHAO Wanua, LAN Shuhuai, et al. A review on spindle thermal error compensation in machine tools [J]. International Journal of Machine Tools and Manufacture, 2015,95:20-38.

[3] JORGENSEN B R. Robust modeling of high-speed spindle-bearing dynamics under operating conditions [D]. West Lafayette: Purdue University,1996.

[4] KIM S K, CHAO D W. Real-time estimation of temperature distution in a ball-screw system [J]. International Journal of Machine Tools and Manufacture, 1997, 37(4):451-464.

[5] YUN W S, KIM S K, CHAO D W. Thermal error analysis dynamic thermo-mechanical spindle model[J]. International Journal of Machine Tools and Manufacture, 2004, 44(4):347-364.

[6] 李彬,张云,王立平,等. 基于遗传算法优化小波神经网络数控机床热误差建模[J]. 机械工程学报,2019,55(21):215-220.

LI Bin, ZHANG Yun, WANG Liping, et al. Modeling for CNC machine tool thermal error based on genetic algorithm

- optimization wavelet neural networks[J]. Journal of Mechanical Engineering, 2019, 55(21):215–220. (in Chinese)
- [7] 夏军勇,胡友民,吴波,等. 热弹性效应分析与机床进给系统热动态特性建模[J]. 机械工程学报,2010,46(15):191–198. XIA Junyong, HU Youmin, WU Bo, et al. Analysis on thermoelastic domino effect and modeling on thermal dynamic characteristic of machine tools feed system[J]. Journal of Mechanical Engineering, 2010,46(15):191–198. (in Chinese)
- [8] 杜柳青,王承辉,余永维,等. 基于深度自编码器的大型龙门加工中心热误差建模方法[J/OL]. 农业机械学报,2019, 50(10):395–400. DU Liuqing, WANG Chenghui, YU Yongwei, et al. Thermal error modeling method based on stacked auto-encoder for large gantry five-sided machining center[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2019, 50(10): 395–400. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?flag=1&file_no=20191046&journal_id=jcsam. DOI:10.6041/j.issn.1000-1298.2019.10.046. (in Chinese)
- [9] 余永维,杜柳青,易小波,等. 基于时序深度学习的数控机床运动精度预测方法[J/OL]. 农业机械学报,2019,50(1): 421–426. YU Yongwei, DU Liuqing, YI Xiaobo, et al. Prediction method of NC machine tools' motion precision based on sequential deep learning[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2019, 50(1):421–426. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?flag=1&file_no=20190149&journal_id=jcsam. DOI:10.6041/j.issn.1000-1298.2019.01.049. (in Chinese)
- [10] 余永维,杜柳青. 深度学习框架下数控机床运动误差溯源方法[J]. 仪器仪表学报,2019,40(1):28–34. YU Yongwei, DU Liuqing. Motion error tracing of NC machine tools based on deep learning framework[J]. Chinese Journal of Scientific Instrument, 2019, 40(1):28–34. (in Chinese)
- [11] 杜柳青,余永维. 基于并联深度信念网络的数控机床热误差预测方法[J/OL]. 农业机械学报,2020,51(8):414–419. DU Liuqing, YU Yongwei. Thermal error prediction method of CNC machine tools based on parallel depth belief network[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2020,51(8):414–419. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?flag=1&file_no=20200847&journal_id=jcsam. DOI:10.6041/j.issn.1000-1298.2020.08.047. (in Chinese)
- [12] 陈英义,方晓敏,梅思远,等. 基于 WT-CNN-LSTM 的溶解氧含量预测模型[J/OL]. 农业机械学报,2020,51(10):284–291. CHEN Yingyi, FANG Xiaomin, MEI Siyuan, et al. Dissolved oxygen prediction model based on WT-CNN-LSTM[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2020,51(10):284–291. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?flag=1&file_no=20201032&journal_id=jcsam. DOI:10.6041/j.issn.1000-1298.2020.10.032. (in Chinese)
- [13] 高浩然. 基于能耗数据和环境温度的机床主轴热变形模型研究[D]. 武汉:华中科技大学,2019. GAO Haoran. Research on thermal deformation model of machine tool spindle based on energy dissipation data and ambient temperature[D]. Wuhan: Huazhong University of Science & Technology, 2019. (in Chinese)
- [14] 程成. 基于深度学习的设备剩余寿命预测算法研究及其应用[D]. 北京:北京化工大学,2020. CHENG Cheng. Research and application of equipment remaining life prediction algorithm based on deep learning[D]. Beijing: Beijing University of Chemical Technology, 2020. (in Chinese)
- [15] WU Chengyang, XIANG Sitong, XIANG Wansheng. Thermal error modeling of rotary axis based on convolutional neural network[J]. Journal of Manufacturing Science and Engineering, 2021, 143(5):1–13.
- [16] ZHANG Xiaomei, YANG Liu, LOU Ping, et al. Thermal error modeling for heavy duty cnc machine tool based on convolution neural network[C] // 2019 IEEE 3rd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC). IEEE, 2019: 665–669.
- [17] SUN Chun, LI Le, WEN Zejun. Remaining useful life prediction via a vibrational autoencoder and a time-window-based sequence neural network[J]. Quality and Reliability Engineering International, 2020, 36(5):1639–1656.
- [18] XU Xingwei, TAO Zhengrui, MING Weiwei, et al. Intelligent monitoring and diagnostics using a novel integrated model based on deep learning and multi-sensor feature fusion[J]. Measurement, 2020, 165:108086.
- [19] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[C] // Advances in Neural Information Processing Systems, 2017: 5998–6008.
- [20] SMITH L N. Cyclical learning rates for training neural networks[C] // Proc. of 2017 IEEE Conference on Application of Computer Vision, 2015:464–472.
- [21] AN Qinglong, TAO Zhengrui, XU Xingwei, et al. A data-driven model for milling tool remaining useful life prediction with convolutional and stacked LSTM network[J]. Measurement, 2019, 154:107461.
- [22] 杜圣东,李天瑞,杨燕,等. 一种基于序列到序列时空注意力学习的交通流预测模型[J]. 计算机研究与发展,2020, 57(8):1715–1728. DU Shengdong, LI Tianrui, YANG Yan, et al. A sequence-to-sequence spatial-temporal attention learning model for urban traffic flow prediction[J]. Journal of Computer Research and Development, 2020,57(8):1715–1728. (in Chinese)
- [23] 李兵,苏文超,魏翔,等. 基于环境温度模型库分段式加权的数控机床热误差建模[J/OL]. 农业机械学报,2020, 51(7):413–419. LI Bing, SU Wenchao, WEI Xiang, et al. Piecewise weighted thermal error modeling of CNC machine tools based on model library of ambient temperature[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2020,51(7):413–419. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?flag=1&file_no=20200747&journal_id=jcsam. DOI: 10.6041/j.issn.1000-1298.2020.07.047. (in Chinese)
- [24] 余文利,姚鑫骅,孙磊,等. 基于 PLS 和改进 CVR 的数控机床热误差建模[J/OL]. 农业机械学报, 2015,46(2):357–364. YU Wenli, YAO Xinhua, SUN Lei, et al. Thermal error modeling of CNC machine tool based on partial least squares and improved core vector regression[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2015, 46(2): 357–364. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?flag=1&file_no=20150253&journal_id=jcsam. DOI: 10.6041/j.issn.1000-1298.2015.02.053. (in Chinese)