

• 机械工程 •

DOI:10.15961/j.jsuese.202100340



本刊网刊

# 基于注意力机制的数控机床热误差深度学习预测方法

杜柳青, 李仁杰, 李宝钏

(重庆理工大学 机械工程学院, 重庆 400054)

**摘 要:**为提高热误差预测精度和鲁棒性, 提出一种基于注意力机制和深度学习网络的数控机床热误差预测模型。采用数据转化策略, 将数控机床原始温度数据转化为温度图像, 直接作为深度学习网络的输入; 提出一种基于注意力机制的温度敏感点识别网络, 根据温度测点与热误差关联程度, 赋予各温度测点不同的权值, 避免了温度测点人为选取的弊端; 建立 12 层深度 CNN 学习预测网络, 利用其强大的图像特征学习能力, 挖掘温度图像与热误差的非线性映射关系, 无需对温度关键点进行预选择, 保留了更多的热误差与机床温度特征关系, 显著提高了模型预测精度。引入 Dropout 正则化方法和 Adam 优化算法, 对深度卷积神经网络的结构与参数进行了优化, 以提高热误差模型的精度与泛化能力。该方法在对 G460L 型数控车床的热误差验证中表现出较高的预测精度。与基于 BP 神经网络、多元回归、CNN 网络等热误差模型比较, 本文方法在泛化性性能上表现更优。

**关键词:**温度测点; 卷积神经网络; 机床热误差; 深度学习

中图分类号: TH151

文献标志码: A

文章编号: 2096-3246(2021)06-0194-10

## Deep Learning Prediction for Thermal Error of CNC Machine Tools Based on Attention Mechanism

DU Liuqing, LI Renjie, LI Baochuan

(College of Mechanical Eng., Chongqing Univ. of Technol., Chongqing 400054, China)

**Abstract:** Thermal error prediction and compensation of CNC machine tools is an important technology to improve the machining accuracy and reliability of CNC machine tools. The thermal error of machine tool is time-varying and nonlinear. To improve the accuracy and robustness of thermal error prediction, a numerical control machine tool thermal error prediction model based on attention mechanism and deep learning network was proposed. Using the data conversion strategy, the original temperature data of CNC machine tool was transformed into temperature image, which could be directly used as the input of deep learning network. The complete information of the temperature field of the machine tool was retained by converting the temperature field data into the temperature image points. At the same time, the nonlinear and coupling problems between the temperature measuring points were avoided by using the deep learning modeling method. A recognition network of temperature sensitive points based on attention mechanism was proposed. According to the correlation degree between temperature measuring points and thermal error, different weights were given to each temperature measuring point to avoid the disadvantages of artificial selection of temperature measuring points. A 12-layer deep CNN learning prediction network was established to mine the nonlinear mapping relationship between temperature image and thermal error by using its powerful image feature learning ability. This method does not need to preselect the key temperature points, retained more relationship between thermal error and machine temperature characteristics, and can significantly improve the prediction accuracy of the model. In order to improve the accuracy and generalization ability of thermal error model, dropout regularization method and Adam optimization algorithm were introduced to optimize the structure and parameters of deep convolution neural network. The method shows high prediction accuracy in the thermal error verification of G460L CNC lathe. Compared with the thermal error models based on BP neural network, mul-

收稿日期: 2021-04-18

基金项目: 国家自然科学基金项目(51775074); 重庆市自然科学基金项目(cstc2021jcyj-msxmX0372); 重庆市基础研究与前沿探索项目(cstc2018jcyjAX0352); 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。研究方向: 机床精度; 机器视觉。E-mail: lqdu@cqut.edu.cn

网络出版时间: 2021-11-08 09:28:12 网络出版地址: <https://kns.cnki.net/kcms/detail/51.1773.TB.20211105.1125.003.html>

multiple regression and CNN network, the proposed method performs better in generalization performance.

**Key words:** temperature measuring points; convolutional neural network; thermal error of machine tool; deep learning

数控机床热误差预测与补偿是提高数控机床加工精度和可靠性的重要技术<sup>[1-2]</sup>。有效补偿的关键是热误差准确建模,近年来,国内外学者对机床热误差预测模型的研究分为理论热误差模型与经验热误差模型<sup>[3-5]</sup>。理论热误差模型依靠对机床部件的物理机理进行实验研究,进而进行物理建模<sup>[6]</sup>。Kim<sup>[7]</sup>、Yun<sup>[8]</sup>等利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。谢飞等<sup>[9]</sup>发现了机床伪滞后现象,建立了基于优化最小二乘支持向量机的热误差伪滞后预测模型。谭峰等<sup>[10]</sup>利用二进制鲸鱼优化算法对最小二乘支持向量机的超参数进行优化,其损失函数设计考虑最大化预测精度和最小化最优温度关键点,实验结果表明热误差模型精度提高约62.8%,温度测点从20个减至3个。上述研究在机床热态性能方面均取得较好进展,但现有热误差模型受温度测点的影响很大,模型的建立和精度依赖于机床温度敏感点进行布局并优化选取。

热误差具有时变、非线性和耦合的特点,工况不同,机床的温度场分布模式和热误差均会发生变化。而温度传感器的优化布置一直是热误差补偿研究和实际应用中的一个瓶颈。魏弦等<sup>[11]</sup>利用模糊聚类与相关性分析对温度测点进行优化后,采用特征提取法得到预测模型的自变量,从而建立了自适应模糊聚类热关键点优化方法,消除了复共线性对模型精度与鲁棒性的影响。张伟等<sup>[12]</sup>采用灰色关联法计算温度与热误差之间的相关系数,并选取最优温度变量;再用模糊聚类对所选变量进行聚类,确定关键温度点;最后,建立线性回归模型,将机床热误差由 $41.3\text{ }\mu\text{m}$ 减小到 $7.6\text{ }\mu\text{m}$ 。凡志磊等<sup>[13]</sup>提出了一种基于偏相关分析的数控机床温度测点优化方法,实现了温度测点的优化,并提高了所建立的多元回归热误差模型的精确性与鲁棒性。Chen等<sup>[14]</sup>通过分别计算9个温度传感器与3个方向的位移传感器之间的关系,利用判定系数,确定进给量的5次多项式与6个温度敏感点作为模型的输入。以上温度优化研究均需面对一个难以克服的问题,即数控机床的温度分布具有极强的耦合性并且存在热弹性效应<sup>[15]</sup>,数控机床的热特性不可能达到绝对的稳定状态,且随着机床的工况不同,以及季节与气候等影响因素的作用,机床的温度敏感点是可能发生变化的<sup>[16]</sup>。使用传统热误差建模方法对机床热误差进行预测就会产生较大的预测误差,导致数控机床加工精度降低,满足不了不同工况运行下的精度要求。

深度学习算法具有强大的数据处理与特征自学习能力。目前,针对深度学习的研究主要集中在图像处理、字符识别、语义分割等方面,并取得了一些成果<sup>[17-19]</sup>。为避免温度测点变化与敏感点人工选择对机床精度的影响,提高机床热误差预测精度和鲁棒性,提出一种基于注意力机制的温度敏感点自动识别网络,根据计算出的温度测点与热误差关联程度,赋予各温度测点不同的权值;基于卷积神经网络(convolutional neural network, CNN),建立数控机床热误差深度学习预测模型;利用深度CNN的强大特征提取能力,自动提取机床温度数据图像的特征,准确建立温度图像与热误差的非线性映射关系,在实验中获得较高的热误差预测精度。

## 1 基于深度卷积神经网络的热误差模型构建

### 1.1 热误差模型

作者建立了一种基于注意力机制与卷积神经网络的深度学习热误差模型,结构如图1所示,主要包括数据转换模块、敏感点注意力网络、深度CNN预测网络等。该模型的输入为图像化的温度数据,将数控机床原始温度数据转化为温度图像,建立的数控机床温度分布图与深度卷积神经网络系统特性具有高度契合性,也即全部原始温度数据直接作为深度学习网络的输入,完整保留了数控机床温度特征。敏感点注意力网络自动挖掘温度敏感点,避免了温度测点人工优化,有利于提高效率。深度CNN预测网络进行特征自动提取,从多维带有耦合性的数据中自动提取出机床热误差的高层次特征,通过训练自动找到合适的卷积核,避免了传统热误差建模方法中需要人为优化温度敏感点而存在的种种弊端,从而提升模型的精度与鲁棒性。

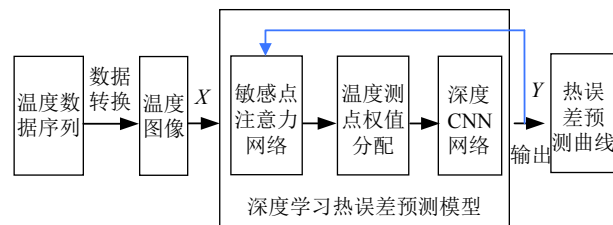


图1 基于注意力机制的深度CNN热误差模型

Fig. 1 Thermal error modeling of CNN based on attention mechanism

### 1.2 机床温度场转化

数控机床温度场是在某一瞬间机床内各点温度分布点的总称,是时间与空间位置点分布函数,其温度分布属于非稳态温度场。采用温度场数据转换为

温度图像点的方式保留了机床温度场的完整信息;同时,利用深度学习建模方法也避免了温度测点之间的非线性与耦合问题。转化思路基于图像の数値属性,即图像由0~256数值组成,转换原理如图2所示。

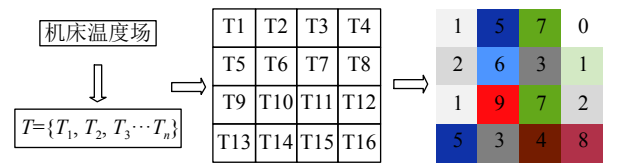


图 2 温度场转换示意图

Fig. 2 Schematic diagram of temperature field conversion

将同一时刻温度测点数据转换为温度矩阵,即组成同一时刻具有机床完整温度信息的温度图像。

1.3 深度CNN预测网络构建

构建的深度学习热误差模型一共有12层,除输入层和输出层以外,还含有10个隐藏层,分别为4个卷积层、3个池化层、1个全连接层和2个回归层,具体网络结构如图3所示,结构参数如表1所示。

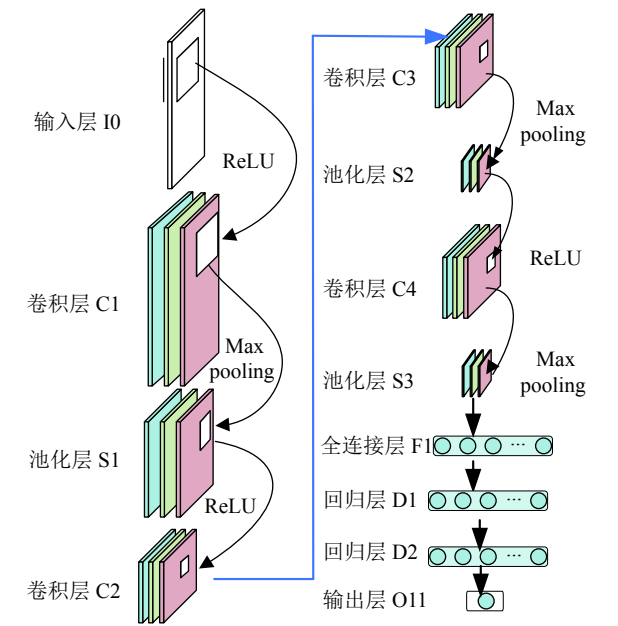


图 3 深度CNN预测网络

Fig. 3 Deep learning prediction net

卷积层与池化层交替连接。卷积层实现特征自动提取,其映射公式为:

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

式中: $x$ 为热误差模型的层次特征图; $k$ 为提取温度抽象特征而构造的卷积核; $M$ 为温度特征学习的感受野; $b$ 为层次特征图的偏置; $l$ 为热误差模型的卷积层序号; $j$ 为当前层特征图序号; $i$ 为前一层特征图序号; $*$ 为卷积运算; $f(\cdot)$ 为激活函数,采用ReLU型激活函数,其计算速度与收敛速度快于Sigmoid和tanh,同时

表 1 CNN具体结构

Tab. 1 CNN structure

| 层数 | 类型     | 特征图个数 | 特征图尺寸 | 卷积核 | 参数数量   |
|----|--------|-------|-------|-----|--------|
| 0  | 输入层I0  | 1     | 12×12 | —   | —      |
| 1  | 卷积层C1  | 5     | 12×12 | 3×3 | 50     |
| 2  | 池化层S1  | 5     | 6×6   | 2×2 | —      |
| 3  | 卷积层C2  | 10    | 6×6   | 3×3 | 460    |
| 4  | 卷积层C3  | 32    | 6×6   | 3×3 | 2912   |
| 5  | 池化层S2  | 32    | 3×3   | 2×2 | —      |
| 6  | 卷积层C4  | 64    | 3×3   | 3×3 | 18 496 |
| 7  | 池化层S3  | 64    | 1×1   | 2×2 | —      |
| 8  | 全连接层F1 | 1     | 1×64  | —   | 1 950  |
| 9  | 回归层D1  | 1     | 30    | —   | —      |
| 10 | 回归层D2  | 1     | 10    | —   | 310    |
| 11 | 输出层O11 | 1     | 1     | —   | 11     |

能消除负信号的干扰,解决了梯度消失的问题,表达式为:

$$f(z) = \begin{cases} z, & z < 0; \\ z, & z \geq 0 \end{cases} \tag{2}$$

池化层尽可能保留较多原始特征信息,同时降低特征维度,减少运算量,并且具有平移不变性的特点。池化层的映射公式为:

$$x_j^l = f\left(\beta_j^l \text{down}\left(x_j^{l-1}\right) + b_j^l\right) \tag{3}$$

式中,down(·)为池化函数, $\beta$ 为权重, $b$ 为池化层偏置。

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,传递式为:

$$x^l = f\left(w^l x^{l-1} + b^l\right) \tag{4}$$

式中, $b$ 为全连接层偏置, $w$ 为权值。

回归层用于将卷积神经网络输出的传统分类模式变成回归预测模式,计算热误差预测值,计算式为:

$$x^l = h\left(g^l x^{l-1} + s^l\right) \tag{5}$$

式中, $h$ 为回归层的偏置, $s$ 为权值, $h(\cdot)$ 为线性变换函数。

1.4 基于注意力网络的敏感点权值分配

注意力机制(attention mechanism, AM)<sup>[20]</sup>能从较大的信息集里面挑选出有用信息。设计一种温度敏感点注意力网络,自动挖掘温度测点与机床热误差间的关联关系,根据关联程度,对温度测点赋予不同的重要性,以提高机床热误差预测的效率和准确性。温度敏感点注意力网络如图4所示。



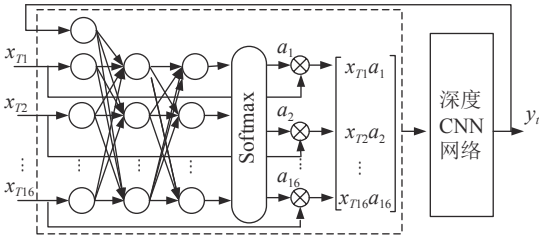


图4 敏感点注意力网络

Fig. 4 Sensitive attention network

敏感点注意力网络由3部分组成:特征相似度计算层、softmax权重重要性分配层与特征重构层。状态信号除注意力网络的输入温度测点数据 $x_i$ 外,还包含深度CNN预测网络的输出,即: $x_k = [x_1, x_2, x_3, \dots, x_r, y]$ 。第一层网络挖掘温度测点与热误差的相关性,其输出为相关性系数 $r$ :

$$r = \text{activation}(Wx_k + b) \quad (6)$$

注意力网络采用softmax函数实现注意力评分操作,softmax将上一层网络输出的相关性系数转化为权值,即状态信号权重矩阵 $a$ ,得到各温度测点对机床热误差的重要程度。

$$a_k = \frac{\exp(r)}{\sum_k \exp(r_k)} \quad (7)$$

然后,对输入的温度数据对应加权,对不同测点数据采取不同的权重,加权后的温度数据再作为深度CNN预测网络的输入。

$$\tilde{x} = [(x_{T1}a_1, x_{T2}a_2, \dots, x_{T16}a_{16})]^T \quad (8)$$

## 2 深度学习热误差模型训练及优化

深度CNN热误差模型网络结构较复杂,导致模型在训练过程中提取数据特征时,可能相应对随机噪声特征进行了提取,从而导致模型的过拟合。因此,在模型中引入Dropout<sup>[21]</sup>正则化方法防止模型过拟合,以一定概率随机放弃一些被激活的神经元,让模型不会太依赖某些局部的特征,使网络具有稀疏特性,模型更健壮,泛化能力更强。

传统网络训练采用随机梯度下降法(stochastic gradient descent, SGD)求解局部最优值,由于方向选择的问题,其得到的结果不一定是全局最优。因此,在深度CNN热误差模型的构建中引入Adam(adaptive momentum, Adam)<sup>[22]</sup>自适应动量的随机优化方法,通过控制超参数和调整学习率改善模型的训练过程。

温度图像数据直接输入网络的第一层,即输入层;经过中间各隐层,逐层变换,逐层映射,直到输出层。

前向传播过程包含数据前向传播与Dropout正则化过程,设置热误差深度学习网络的丢弃率 $p$ (一般为0.3~0.6),生成一个0~1之间均匀分布的随机数组 $U_1$ , $U_1$ 中小于 $p$ 的被置0,大于 $p$ 的被置1。

设置神经元 $H_1=U_1$ , $U_1$ 中置0位对应的神经元被放弃,即随机选择神经层中的一些单元,并将其临时隐藏。

使用Dropout正则化后神经元的连接方式变化如图5所示。

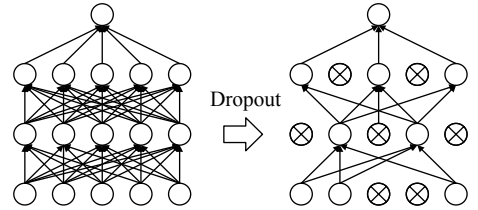


图5 Dropout正则化

Fig. 5 Dropout regularization

Adam基于训练数据迭代更新热误差深度学习网络参数,热误差模型参数的更新不受梯度的伸缩变换影响,计算效率高,对内存需求少。该算法从梯度均值及梯度平方2个角度进行更新,步长自适应地调节,而不是直接取决于前梯度值。

## 3 热误差模型应用与结果

### 3.1 温度数据时变性分析

在G460L高速线轨刀塔式数控车床上进行了验证应用。采用PT100温度传感器采集机床温度信息,采用ML33-B-V2-3型电涡流位移传感器测量主轴径向与轴向热误差,并分别经过SBWZ-2460温度变送器、前置器,以及USB-3120数据采集卡将测得模拟信号转变为数字信号,输出到数据采集软件进行处理。部分传感器布置如图6所示。

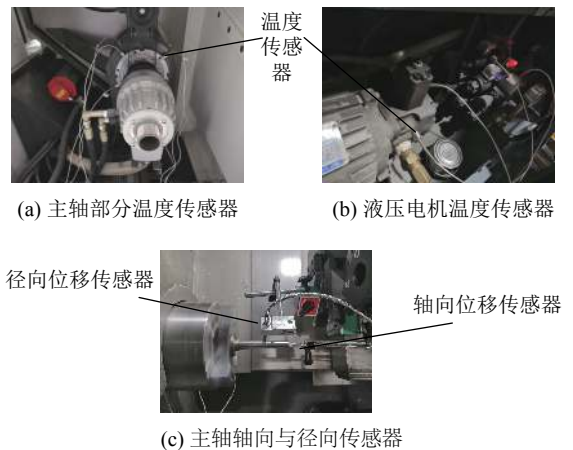


图6 部分传感器布置

Fig. 6 Partial sensors arrangement

温度传感器在机床各部件的布置如表2所示。

表 2 温度传感器布置

Tab. 2 Arrangement of temperature sensors

| 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  |

机床工况发生变化,以及不同季节条件下,实际的机床温度场将发生不确定的变化。因此,在不同季节、不同时间、不同工况下,采集G460L机床热误差与温度数据,如表3所示。噪声问题在空载时更为突出,空载曲线最富有意义<sup>[23-24]</sup>,因此,设定机床主轴为空转模式,机床主轴以1 500和3 000 r/min 2种转速模拟实际加工的空载转速;为了提高模型的泛化性,也设置了6 000 r/min的高转速空载。数据采集系统的取样间隔为30 s,为了降低高采样频率带来的误差影响,对测量数据进行移动均值滤波处理。

表 3 G460L温度数据采集

| 数据批次            | 采集时间       | 主轴转速/(r·min <sup>-1</sup> ) | 环境温度/℃ |
|-----------------|------------|-----------------------------|--------|
| K <sub>1</sub>  | 2019-12-31 | 1 500                       | 5~10   |
| K <sub>2</sub>  | 2020-01-01 | 1 500                       | 5~10   |
| K <sub>3</sub>  | 2020-01-02 | 3 000                       | 4~10   |
| K <sub>4</sub>  | 2020-01-03 | 3 000                       | 5~7    |
| K <sub>5</sub>  | 2020-01-04 | 3 000                       | 5~9    |
| K <sub>6</sub>  | 2020-01-05 | 3 000                       | 5~8    |
| K <sub>7</sub>  | 2020-06-01 | 1 500                       | 20~26  |
| K <sub>8</sub>  | 2020-06-02 | 1 500                       | 21~27  |
| K <sub>9</sub>  | 2020-06-03 | 3 000                       | 21~29  |
| K <sub>10</sub> | 2020-06-04 | 3 000                       | 22~28  |
| K <sub>11</sub> | 2020-06-05 | 3 000                       | 23~29  |
| K <sub>12</sub> | 2020-06-06 | 3 000                       | 21~27  |
| K <sub>13</sub> | 2020-07-27 | 1 500                       | 28~37  |
| K <sub>14</sub> | 2020-07-29 | 1 500                       | 28~39  |
| K <sub>15</sub> | 2020-07-30 | 3 000                       | 28~39  |
| K <sub>16</sub> | 2020-07-31 | 3 000                       | 29~37  |
| K <sub>17</sub> | 2020-08-01 | 3 000                       | 26~38  |
| K <sub>18</sub> | 2020-08-02 | 3 000                       | 25~31  |
| K <sub>19</sub> | 2020-08-03 | 3 000                       | 27~36  |
| K <sub>20</sub> | 2020-08-04 | 3 000                       | 26~33  |
| K <sub>21</sub> | 2020-09-02 | 3 000                       | 22~29  |
| K <sub>22</sub> | 2020-09-03 | 6 000                       | 23~29  |
| K <sub>23</sub> | 2020-09-04 | 6 000                       | 23~30  |

随机选取K<sub>8</sub>与K<sub>11</sub>这2组热误差和温升的数据进行对比,如图7所示。从图7可以看出:在开始测量的

前100 min内,机床X方向上热误差内小于机床Z方向热误差。随着测试时间的增加,X向热误差逐渐增大并超过Z向热误差;X向热误差值总体上呈现指数形式增长,在测量时长达到150 min左右时,机床达到了准热特性平衡状态,热误差值变化较小。随着测量时间的增加,机床温度也在不断地增加。部分测点温度变化趋势十分相近,且呈现出一定的耦合性与相关性;在不同转速下,同一温度测点的温度值变化也可能发生较大变化,这与环境温度及车间的空气流通有一定的关系。

因此,机床的温度变化呈现出较大耦合性与不确定性,这对热误差模型的建立产生了极大影响,使模型的鲁棒性及预测精度在一定程度上受到影响。相关研究指出,温度敏感点在不同工况下会发生变化,故根据现有热误差模型选择温度敏感点作为模型的输入,在机床工况发生变化时会产生模型预测精度下降和泛化性降低等问题。

3.2 模型应用效果分析

为了减少长时间测量过程中其他因素对数据的影响,采用增量形式记录数据;同时,也对数据进行归一化处理,减小因数据维度差距过大所带来的影响。将训练数据集的顺序进行随机打乱,避免每次训练都以相同的数据集顺序传入网络模型中。这种措施可以在一定程度上缓解过度拟合。

在模型训练过程中,采用了水平或垂直翻转图像、随机角度旋转图像、比例缩放图像、随机剪裁图像、位移图像、添加高斯噪声、颜色增强等“数据增强”方法提高热误差模型的准确性与泛化性。对卷积神经网络在每个图像的多个不同副本上进行训练,有利于针对同一对象的变化进行泛化,从而提高模型的泛化能力与鲁棒性。

采用平均绝对误差(mean absolute error, MAE)、均方差(mean squared error, MSE)和修正决定系数 $R^2_{\text{adj}}$ 对模型的性能进行量化评定,弥补 $R^2$ 随样本量增加而增大的缺陷。决定系数 $R^2$ 越接近1,模型预测效果越好,精度越高。模型训练后的决定系数 $R^2$ 值最大的同时,MAE与MSE越小,模型预测误差越小。 $f(\text{value})$ 为数据的相对变化值,上述指标的计算公式如下:

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i| \tag{9}$$

$$\text{MSE} = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2 \tag{10}$$

$$R^2 = 1 - \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N (y_i - \bar{y})^2} \tag{11}$$

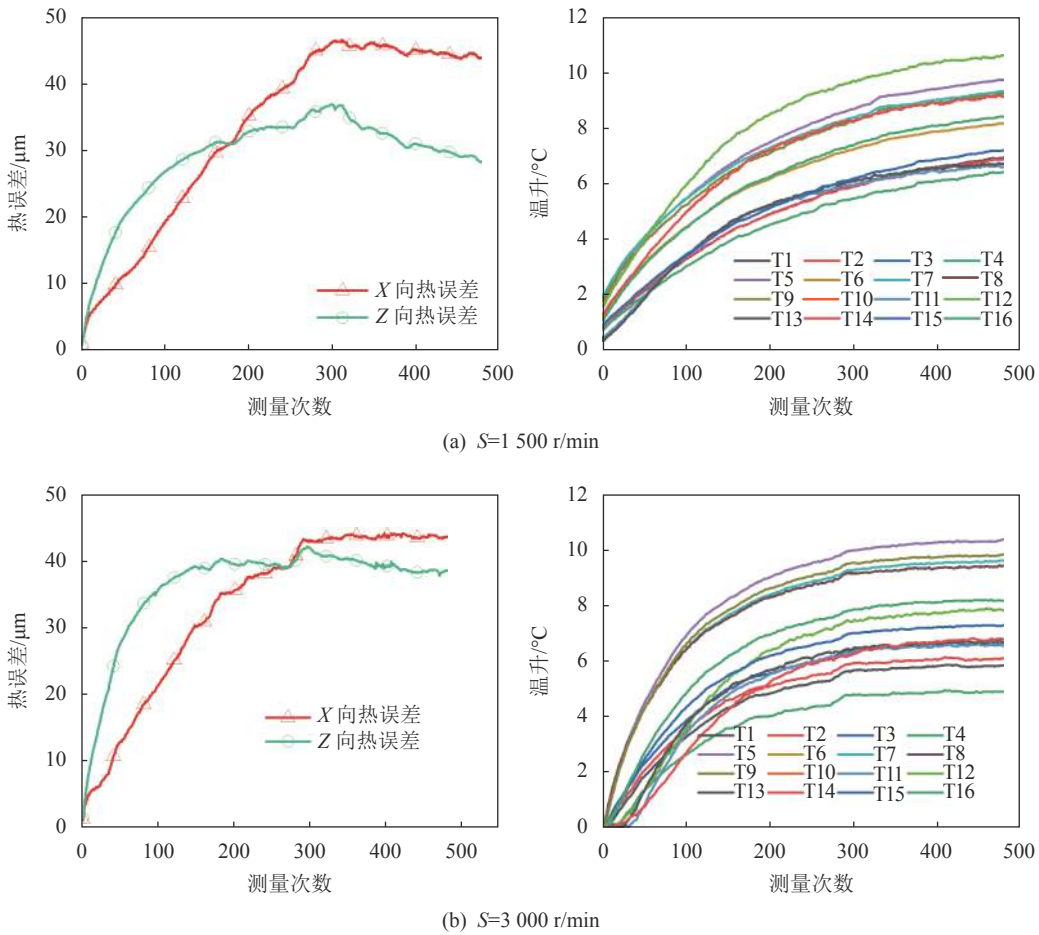


图 7 热误差比较、温升比较

Fig. 7 Comparison of temperature rise and thermal error

$$R_{\text{adj}}^2 = 1 - \frac{(1 - R^2)(N - 1)}{N - P - 1} \quad (12)$$

$$f(\text{value}) = \frac{\text{value}_1 - \text{value}_2}{\text{value}_1} \quad (13)$$

式中,  $N$  为观测值的个数,  $P$  为预测值数量,  $y_i$  为观测值,  $\hat{y}$  为预测值,  $\bar{y}$  为  $N$  个观测值的平均数。

泛化性指标反映了模型在受不确定因素影响下, 模型的预测精度对输入参数变化的敏感程度。机床转速变化时, 泛化性指标为残余标准差  $R_{\text{st}}$  的相对增量  $R_{\text{rpm}}$ , 其计算方法如下:

$$R_{\text{st}} = \sqrt{\sum_{i=1}^N \frac{|y_i - \hat{y}_i|}{N - P - 1}} \quad (14)$$

$$R_{\text{rpm}} = \frac{|R_{\text{st}} - R'_{\text{st}}|}{R_{\text{st}}} \quad (15)$$

数据增强后的温度图像大小对模型精度会产生的影响如图8所示。从图8中可以看出, 选择将图像增强为  $12 \times 12$  大小时, 模型训练后的输入模型训练后得到的模型各项指标均为最优。故采用  $12 \times 12$  输入图像大小作为模型输入尺寸时能获得最高的预测精度。

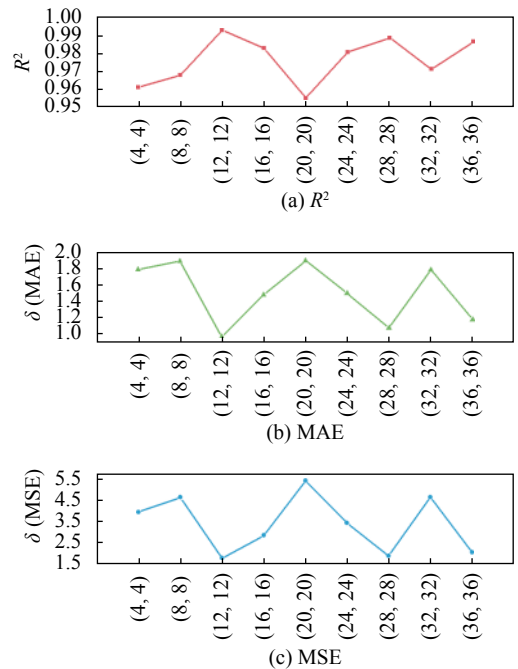


图 8 数据增强图像尺寸对模型精度的影响

Fig. 8 Influence of image size on model accuracy in data enhancement

带注意力机制CNN模型的建立基于TensorFlow框架,集成开发环境为Pycharm以及GTX 1050 Ti显卡。表4为模型训练参数的设置,在模型的全连接层到输出层之间再设置2个回归层,用于将传统的卷积神经网络输出的分类模式变成回归预测模式。2个回归层都使用Dropout正则化优化方法,分别设置Dropout值为0.4和0.3。模型的输出层激活函数选择回归预测模型中常用的Sigmoid激活函数,以得到具有连续性的热误差预测值。

表 4 CNN热误差模型参数设置

| 参数        | 数值或详情   |
|-----------|---------|
| 每批大小      | 86      |
| 迭代次数      | 280     |
| 初始学习率     | 0.01    |
| Dropout值1 | 0.40    |
| Dropout值2 | 0.30    |
| 激活函数      | Sigmoid |

选择最大值池化函数减少估计值方差。卷积层与池化层之间的激活函数为ReLU激活函数,卷积核的初始化采用正态分布的随机数,结合Adam优化算法对卷积核进行优化,模型训练结果如图9所示。基于深度卷积神经网络的热误差模型设置有4个卷积层,深度不断加深。卷积层的卷积核大小通常优先选择3×3或更小的卷积核,在每个卷积核后进行非线性化,使得在具有相同感受野大小的同时,可以节省参数,减少过拟合。而且由于网络更深,非线性更强,效果往往更好[25–28]。

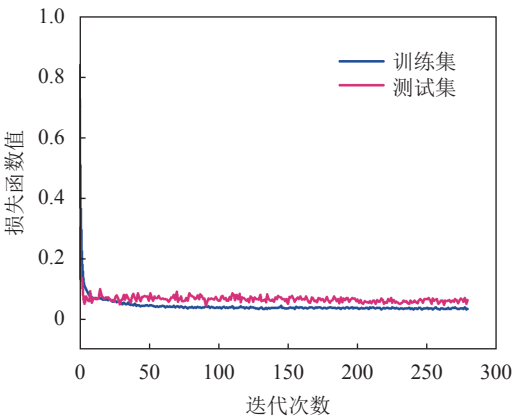


图 9 CNN模型训练集和测试集的损失函数曲线

Fig. 9 Loss function value curves of CNN model training set and test set

选用应用改进优化的BP神经网络热误差模型[29]、MLR热误差模型和非改进CNN热误差模型用于验证本文AM–CNN热误差模型的预测效果,并进行对比。将不同工况、不同季节的数据合并作为训练数据,对

不同转速下测试集进行预测[30]。

根据前述方法进行数据预处理及降噪后,一共得到23组数据,其中17组作为训练集合,选择不同典型环境温度下的6组数据进行热误差预测验证。根据前述建模方法与优化方法,建立热误差预测模型,6组数据均有较高的预测准确性,其中,转速条件分别为1 500、3 000、6 000 r/min的K<sub>8</sub>、K<sub>13</sub>、K<sub>23</sub>组预测效果如图10所示(热误差观测值为X向数据)。

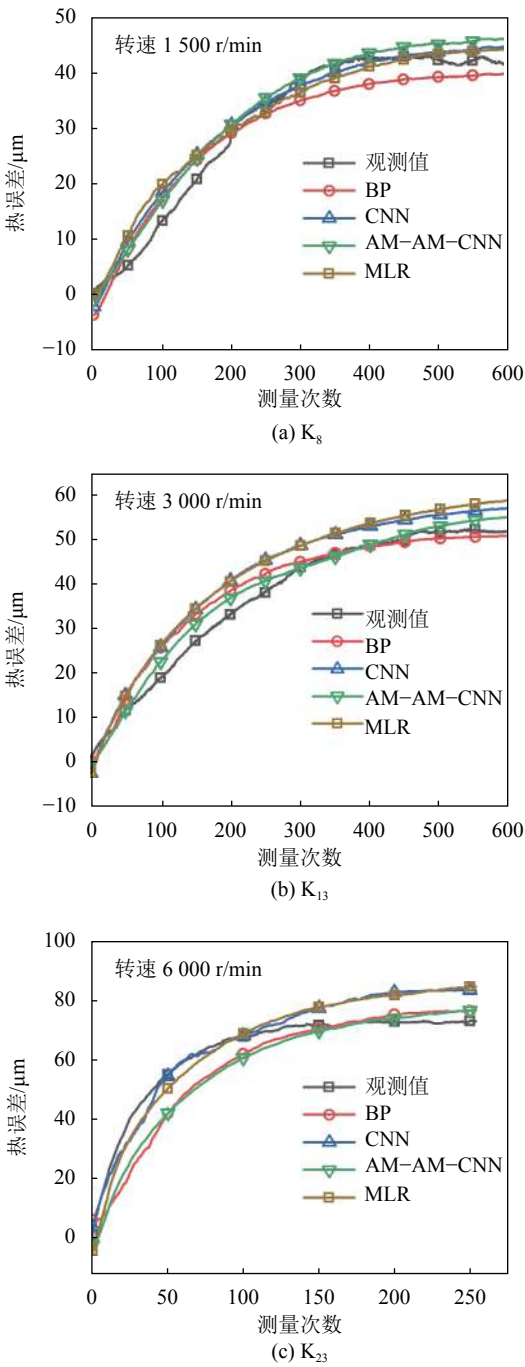


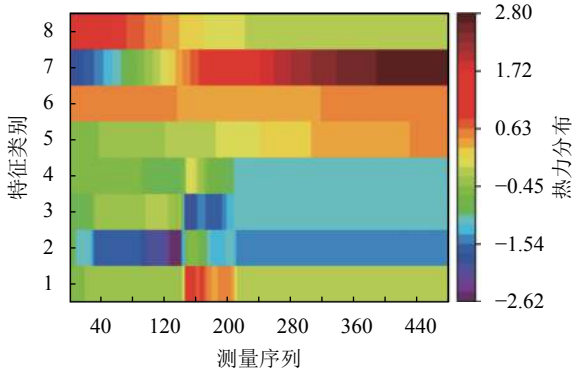
图 10 不同混合转速下预测模型预测结果对比

Fig. 10 Comparison of prediction results of different prediction models under mixed speed

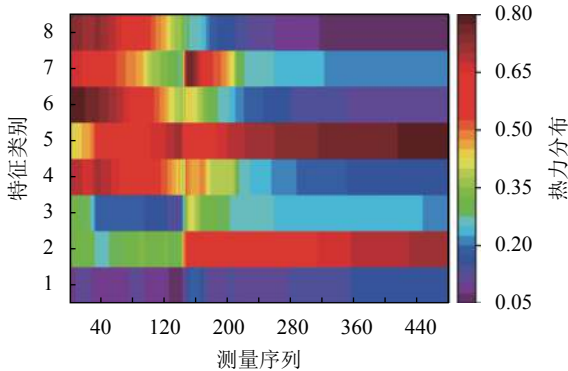


MLR模型、BP模型均采用灰色关联度与聚类分析方法筛选出4个温度敏感点作为模型的输入,CNN与本文AM-CNN模型则采用温度场转化方法,将温度图像作为模型的输入。BP模型的网络超参数采用遗传算法对其进行优化,优化后其网络结构为4-9-3-1。从图10可以看出,基于注意力机制的CNN热误差模型具有较好的预测精度与准确性,对于混合转速下的数控机床热误差做出预测的同时也表现出了良好的泛化性,能够在不同的条件下对机床热误差做出准确的预测。

AM-CNN模型采用的注意力网络能自动对机床温度场敏感特征进行识别,进行一次建模即能取得和其他模型针对性建模的预测效果。注意力机制对温度敏感点特征重构如图11所示,在不同的时间,机床温度特征重要性发生了变化,注意力机制很好地抓住了这个变化并分配不同的重要性权重,与传统CNN方法比较,本文模型的预测精度更好,具有更高的鲁棒性。



(a) 注意力机制分配前特征热力图

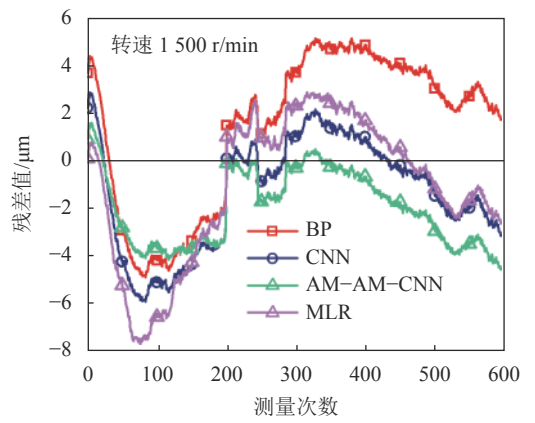


(b) 注意力机制分配后特征热力图

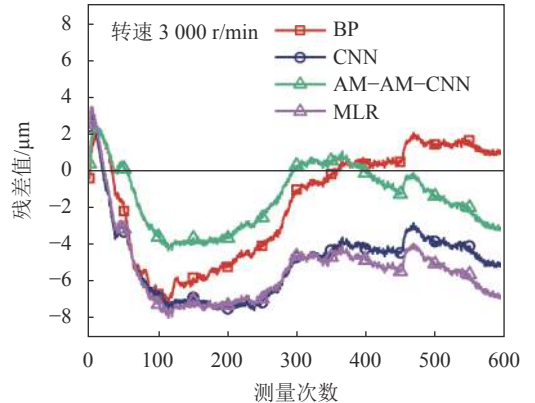
图11 注意力机制对温度敏感点特征重构

Fig. 11 Reconstruction of temperature sensitive point features by attention mechanism

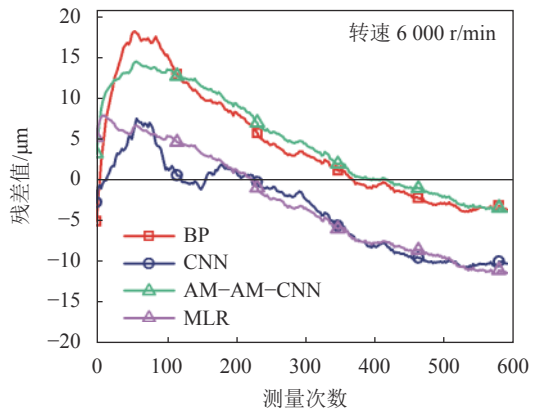
图12为BP方法、CNN方法、MLR方法、本文基于注意力机制的CNN方法3种数据残差对比。由图12可以看出,与CNN模型、BP和MLR模型相对比,本文AM-CNN模型的残差值波动最小,预测表现更加稳定。



(a)  $K_8$



(b)  $K_{13}$



(c)  $K_{23}$

图12 模型残差对比

Fig. 12 Comparison of model residuals

表5~7为各数据集的预测指标。从表5~7可以看出:当数控机床的工况发生变化时,各个模型的预测指标均有一定程度的变化;对于 $K_{13}$ 数据集,CNN模型的 $R_{adj}^2$ 下降幅度最大,AM-CNN模型的 $R_{adj}^2$ 却有小幅度的上升。由公式(14)、(15)可得:以 $K_8$ 数据集转速为基准,AM-CNN模型的泛化性指标 $R_{rpm}$ 为0.1347与0.0760,BP模型的泛化性指标为0.1852与0.1671,CNN模型的泛化性指标为0.1374与0.0733,MLR模型的泛化性指标 $R_{rpm}$ 为0.0989与0.2824。



表 5 K<sub>8</sub>数据集下各模型预测指标

| 模型     | MAE/ $\mu\text{m}$ | MSE/ $\mu\text{m}^2$ | $R_{\text{st}}$ | $R_{\text{adj}}^2$ |
|--------|--------------------|----------------------|-----------------|--------------------|
| BP     | 3.633 6            | 15.165 0             | 0.482 0         | 0.914 9            |
| CNN    | 2.480 5            | 10.074 5             | 0.398 2         | 0.943 5            |
| AM-CNN | 2.175 0            | 7.257 0              | 0.372 9         | 0.953 0            |
| MLR    | 2.072 5            | 7.125 6              | 0.359 7         | 0.949 8            |

表 6 K<sub>13</sub>数据集下各模型预测指标

| 模型     | MAE/ $\mu\text{m}$ | MSE/ $\mu\text{m}^2$ | $R_{\text{st}}$ | $R_{\text{adj}}^2$ |
|--------|--------------------|----------------------|-----------------|--------------------|
| BP     | 2.654 2            | 11.946 8             | 0.571 4         | 0.914 9            |
| CNN    | 5.107 5            | 28.829 9             | 0.343 5         | 0.861 7            |
| AM-CNN | 1.846 0            | 5.200 4              | 0.322 6         | 0.974 9            |
| MLR    | 1.542 1            | 2.845 8              | 0.324 1         | 0.962 1            |

表 7 K<sub>23</sub>数据集下各模型预测指标

| 模型     | MAE/ $\mu\text{m}$ | MSE/ $\mu\text{m}^2$ | $R_{\text{st}}$ | $R_{\text{adj}}^2$ |
|--------|--------------------|----------------------|-----------------|--------------------|
| BP     | 6.081 8            | 64.356 4             | 0.401 4         | 0.717 8            |
| CNN    | 5.823 8            | 41.249 4             | 0.343 5         | 0.819 1            |
| AM-CNN | 5.989 6            | 49.122 9             | 0.401 2         | 0.895 0            |
| MLR    | 6.100 1            | 46.912 5             | 0.410 5         | 0.884 6            |

从各个模型的残差对比(图12)中也可以看出,AM-CNN模型的残差值波动范围均为最小。MLR作为热误差经验模型中应用最广泛的模型,在针对不同工况下进行的建模,其模型的预测指标与AM-CNN表现较为接近。但是,AM-CNN模型建模与训练过程较为简单,且不需要人为地对温度敏感点进行数学分析与选择,采用全部原始机床温度场数据,利用注意力机制对其进行识别,很好地利用数据驱动模型的特点,增加了模型的泛化性。

综合上述分析,在变转速状况下,AM-CNN模型更能学习到机床热特性动态变化的特征,综合预测指标相较于BP、MLR和CNN模型更优,很好地体现出了基于深度卷积神经网络热误差模型对机床温度场信息的特征提取能力。

4 结 论

温度敏感点随着季节、工况、气候变化而变化,从而影响机床热误差补偿的鲁棒性和补偿精度。基于注意力机制和卷积神经网络探索建立了无温度敏感点预选的热误差预测模型,并在G460L型高速线轨刀塔式数控车床上进行了验证应用,提高效率的同时取得了良好的预测精度与鲁棒性,可在机床热误差建模与补偿中应用推广。

1)本文热误差模型预测准确,且泛化性好,得益

于深度卷积神经网络高层次特征提取能力。该模型可以学习温度与热误差之间的高层次映射关系,提高了模型的适用性。与BP模型、MLR和CNN模型相比较,AM-CNN模型具有更好的泛化性与预测准确性。

2)基于注意力机制的敏感点自动识别网络解决了温度测点人为选取对热误差模型的精度与鲁棒性的影响。模型的建立所采用的温度传感器数量为16个,无需对敏感点进行预选择,在保留尽可能多的温度特征信息的同时也能抑制各温度点的共线性弊端,解决不同工况、季节与气候等因素导致的温度敏感点变化对模型精度的影响,与传统线性模型和机器学习模型相比,优势明显。

参考文献:

[1] Bryan J. International status of thermal error research[J]. *CIRP Annals*,1990,39(2):645–656.

[2] Deng Xiaolei,Lin Huan,Wang Jianchen,et al. Review on thermal design of machine tool spindles[J]. *Optics and Precision Engineering*,2018,26(6):1415–1429.[邓小雷,林欢,王建臣,等. 机床主轴热设计研究综述[J]. *光学精密工程*,2018,26(6):1415–1429.]

[3] Tan Feng,Yin Guofu,Zheng Kai,et al. Thermal error prediction of machine tool spindle using segment fusion LSSVM[J]. *The International Journal of Advanced Manufacturing Technology*,2021,116(1/2):99–114.

[4] Guo Qianjian,He Lei,Yang Jianguo. Application of a new modeling method for CNC machine thermal error based on projection pursuit regression[J]. *Journal of Sichuan University(Engineering Science Edition)*,2012,44(2):227–230.[郭前建,贺磊,杨建国. 基于投影追踪回归的机床热误差建模技术[J]. *四川大学学报(工程科学版)*,2012,44(2):227–230.]

[5] Yao Xiaopeng,Yin Guofu,Fang Hui. A new synthesis error modeling and compensating algorithm of the five-axis machine tools[J]. *Journal of Sichuan University(Engineering Science Edition)*,2013,45(5):154–159.[要小鹏,殷国富,方辉. 五轴机床误差建模与补偿解析新算法[J]. *四川大学学报(工程科学版)*,2013,45(5):154–159.]

[6] Mayr J,Jedrzejewski J,Uhlmann E,et al. Thermal issues in machine tools[J]. *CIRP Annals(Manufacturing Technology)*,2012,61(2):771–791.

[7] Kim S K,Cho D W. Real-time estimation of temperature distribution in a ball-screw system[J]. *International Journal of Machine Tools and Manufacture*,1997,37(4):451–464.

[8] Yun W S,Kim S K,Cho D W. Thermal error analysis for a CNC lathe feed drive system[J]. *International Journal of Machine Tools and Manufacture*,1999,39(7):1087–1101.

[9] Xie Fei,Wang Ling,Tan Feng,et al. Pseudo-hysteresis modeling for machine tool thermal error based on metabolic theory[J]. *Journal of Harbin Institute of Technology*,2019,51(7):154–159.[谢飞,王玲,谭峰,等. 基于新陈代谢原理的机床热误差伪滞后建模[J]. *哈尔滨工业大学学报*,2019,51(7):154–159.]

[10] Tan Feng,Xiao Hong,Zhang Yi,et al. Thermal error modeling method of CNC machine tool based on unified frame-

- work[J].*Chinese Journal of Scientific Instrument*,2019,40(10):95–103.[谭峰,萧红,张毅,等.基于统一框架的数控机床热误差建模方法[J].*仪器仪表学报*,2019,40(10):95–103.]
- [11] Wei Xian,Gao Feng,Li Yan,et al.Optimization of thermal error model critical point for gantry machine tool feeding system[J].*Chinese Journal of Scientific Instrument*,2016,37(6):1340–1346.[魏弦,高峰,李艳,等.龙门机床进给系统热误差模型关键点优化[J].*仪器仪表学报*,2016,37(6):1340–1346.]
- [12] Zhang Wei,Ye Wenhua.Optimization of temperature measuring points for machine tools based on grey correlation and fuzzy clustering analysis[J].*China Mechanical Engineering*,2014,25(4):456–460.[张伟,叶文华.基于灰色关联和模糊聚类的机床温度测点优化[J].*中国机械工程*,2014,25(4):456–460.]
- [13] Fan Zhilei,Li Zhonghua,Yang Jianguo.NC machine tool temperature measuring point optimization and thermal error modeling based on partial correlation analysis[J].*China Mechanical Engineering*,2010,21(17):2025–2028.[凡志磊,李中华,杨建国.基于偏相关分析的数控机床温度布点优化及其热误差建模[J].*中国机械工程*,2010,21(17):2025–2028.]
- [14] Chen J S,Yuan J,Ni J.Thermal error modelling for real-time error compensation[J].*The International Journal of Advanced Manufacturing Technology*,1996,12(4):266–275.
- [15] 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.[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.]
- [16] Yan Jiayu,Zhang Hongtao,Liu Guoliang,et al.Application of a new optimizing method for the measuring points of CNC machine thermal error based on grey synthetic degree of association[J].*Journal of Sichuan University(Engineering Science Edition)*,2008,40(2):160–164.[闫嘉钰,张宏韬,刘国良,等.基于灰色综合关联度的数控机床热误差测点优化新方法及应用[J].*四川大学学报(工程科学版)*,2008,40(2):160–164.]
- [17] Lecun Y,Boser B,Denker J S,et al.Backpropagation applied to handwritten zip code recognition[J].*Neural Computation*,1989,1(4):541–551.
- [18] Hubel D H,Wiesel T N.Receptive fields,binocular interaction and functional architecture in the cat's visual cortex[J].*The Journal of Physiology*,1962,160(1):106–154.
- [19] Zhou Feiyan,Jin Linpeng,Dong Jun.Review of convolutional neural network[J].*Chinese Journal of Computers*,2017,40(6):1229–1251.[周飞燕,金林鹏,董军.卷积神经网络研究综述[J].*计算机学报*,2017,40(6):1229–1251.]
- [20] Vaswani A,Shazeer N,Parmer N,et al.Attention is all you need[C]//Advances in Neural Information Processing System.New York:CL,2017:5988–6008.
- [21] Poernomo A,Kang D K.Biased dropout and Crossmap Dropout:Learning towards effective Dropout regularization in convolutional neural network[J].*Neural Networks*,2018,104:60–67.
- [22] Kingma D,Ba J.Adam:A method for stochastic optimization[C]//3rd International Conference for Learning Representations.San Diego:CL,2014.
- [23] 彭东林.时栅位移传感器与新型机床动态检测系统[M].北京:科学出版社,2010.
- [24] Liu Ming,Zhang Qing.Thermal error model and compensation of machining tool using multibody system theory and neural network[J].*Journal of Vibration,Measurement & Diagnosis*,2010,30(6):657–661.[刘明,章青.运用多体理论和神经网络的机床热误差补偿[J].*振动、测试与诊断*,2010,30(6):657–661.]
- [25] Zhang Xiaomei,Yang Liu,Lou Ping,et al.Thermal error modeling for heavy duty CNC machine tool based on convolution neural network[C]//Proceedings of the 2019 IEEE 3rd Information Technology,Networking,Electronic and Automation Control Conference(ITNEC).Chengdu:IEEE,2019:665–669.
- [26] Su Chun,Li Le,Wen Zejun.Remaining useful life prediction via a variational autoencoder and a time-window-based sequence neural network[J].*Quality and Reliability Engineering International*,2020,36(5):1639–1656.
- [27] 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):051013.
- [28] 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.
- [29] Yang Yi,Yao Xiaodong,Yang Jianguo,et al.Thermo-drifting error modeling of spindle based on combination of principal component analysis and BP neural network[J].*Journal of Shanghai Jiaotong University*,2013,47(5):750–753.[杨漪,姚晓栋,杨建国,等.基于主成分分析与BP神经网络相结合的机床主轴热漂移误差建模[J].*上海交通大学学报*,2013,47(5):750–753.]
- [30] Li Fengchun,Wang Haitong,Li Tiemin.Research on thermal error modeling and prediction of heavy CNC machine tools[J].*Journal of Mechanical Engineering*,2016,52(11):154–160.[李逢春,王海同,李铁民.重型数控机床热误差建模及预测方法的研究[J].*机械工程学报*,2016,52(11):154–160.]

(编辑 黄小川)

引用格式: Du Liuqing,Li Renjie,Li Baochuan.Deep learning prediction for thermal error of CNC machine tools based on attention mechanism[J].*Advanced Engineering Sciences*,2021,53(6):194–203.[杜柳青,李仁杰,李宝钏.基于注意力机制的数控机床热误差深度学习预测方法[J].*工程科学与技术*,2021,53(6):194–203.]