

# 基于TCN-BiGRU-SE两阶段特征提取 与多特征融合的注塑质量预测方法

邓晓强 战韬阳 项薇\* 林文文 余军合 郑志鹏

宁波大学机械工程与智能制造学院, 宁波, 315211

**摘要:** 注塑成形过程中, 塑件尺寸易受多种复杂因素的耦合影响。为提高预测精度, 提出一种基于时间卷积网络(TCN)-双向门控循环单元(BiGRU)-SE注意力机制(SE)的注塑质量预测方法(TCN-BiGRU-SE)。采用TCN-BiGRU-SE网络提取时序数据的深层特征, 表征注塑过程中的动态变化; 提取注射和保压阶段的定量特征值及量纲一值, 堆叠形成三维矩阵, 通过卷积神经网络(CNN)进行降维, 捕捉关键阶段的变化趋势。通过融合高频数据、统计特征与机器状态信息, 构建了一个端到端的深度预测模型, 以实现塑件尺寸的预测。在富士康注塑成形数据集上进行了模型对比、消融实验和稳定性检验, 并在三类注塑实验数据集上进行了泛化性检验, 结果表明, 所建模型在多项评价指标上优于其他方法, 具有良好的鲁棒性和泛化能力。

**关键词:** 注塑成形; 质量预测; 时序数据; 多特征融合; 深度学习

**中图分类号:** TP181

DOI:10.3969/j.issn.1004-132X.2026.02.017

开放科学(资源服务)标识码(OSID):



## Injection Molding Quality Prediction Method Based on Two-stage Feature Extraction and Multi-feature Fusion Using TCN-BiGRU-SE Model

DENG Xiaoqiang ZHAN Taoyang XIANG Wei\* LIN Wenwen YU Junhe ZHENG Zhipeng  
School of Mechanical Engineering and Intelligent Manufacturing, Ningbo University, Ningbo,  
Zhejiang, 315211

**Abstract:** During the injection molding processes, the dimensions of molded parts were easily affected by the coupling of various complex factors. To improve prediction accuracy, a quality prediction method was proposed based on temporal convolutional networks (TCN), Bidirectional gated recurrent units (BiGRU), and squeeze-and-excitation (SE) attention mechanism (TCN-BiGRU-SE). The TCN-BiGRU-SE network was utilized to extract deep features from time-series data, characterizing the dynamic changes during the injection molding processes. Quantitative feature values and dimensionless values from the injection and holding phases were extracted and stacked into a three-dimensional matrix, which was then dimensionally reduced using convolutional neural networks (CNN) to capture the changing trends at critical phases. By integrating high-frequency data, statistical features, and machine state information, an end-to-end deep prediction model was constructed for the prediction of molded part size. Comparative, ablation, and stability tests were conducted on the Foxconn injection molding dataset, along with generalization tests on three types of injection experimental datasets. The results show that the model outperforms other methods on multiple evaluation metrics, demonstrating strong robustness and generalization capability.

**Key words:** injection molding; quality prediction; time-series data; multi-feature fusion; deep learning

### 0 引言

注塑成形是塑料制品制造中最常见的工艺之一, 全球约40%的塑料制品通过注塑机加工完成<sup>[1]</sup>。通过更换模具, 注塑机能够生产出各种复杂结构的制品, 广泛应用于汽车、家电、航空等领

域。注塑件的质量直接关系到制品的整体性能和使用寿命, 注塑成形过程中, 各种工艺参数、环境和人工操作等因素往往相互耦合, 呈现出显著的非线性和复杂性。这些因素的波动可能导致制品尺寸偏差、形状不稳定等问题, 严重影响制品质量。目前, 制品质量检测主要依赖于生产完成后的人工抽样检测。这种事后检测方式存在滞后性, 难以在生产过程中发现潜在的质量问题, 导致企业生产成本增加<sup>[2]</sup>。因此, 对注塑件质量进行

收稿日期: 2024-11-29

基金项目: 国家重点研发计划(2019YFB1707101, 2019YFB1707103)

在线监测至关重要。

现代注塑机通常配备了大量传感器,能够实时监测生产过程中的关键工艺参数,记录注塑过程数据,为数据驱动的预测模型奠定了基础。目前,已有研究使用机器学习的方法构建注塑质量预测模型。刘永兴等<sup>[3]</sup>利用轻量级梯度提升机(light gradient boosting machine, LightGBM)建立注塑成形尺寸预测模型,并通过加权修正方法提高了对超规尺寸的预测精度。LOCKNER等<sup>[4]</sup>使用六种机器学习算法的集合进行特征选择,克服了单一模型的局限性。陈昱等<sup>[5]</sup>采用三段式特征选择方法,通过特征重要度排序筛选关键特征,再使用最大信息系数(maximum information coefficient, MIC)评估特征相关性,最后结合 APRIORI 算法挖掘与注塑制品质量最相关的特征输入分类器,有效提高了注塑制品异常检测的准确性。然而,上述研究主要依赖于离散的统计值作为特征,未能充分考虑时间依赖性和各阶段工艺参数的动态交互,在处理高维和动态数据时存在一定的局限性。

成形过程中,模内压力、注射压力和螺杆速度等随时间动态变化,其时序特征一定程度上反映了制品的质量变化情况。如果仅使用统计特征或静态参数,忽略时序特征,可能无法准确捕捉工艺过程中的动态变化,导致预测精度下降<sup>[6]</sup>。深度学习在处理多维时序数据方面展现出显著优势。不同于传统机器学习,深度学习模型能够自动从原始数据中提取多层次特征,在捕捉复杂的时序依赖和非线性关系时表现优异<sup>[7]</sup>。研究表明,利用深度学习模型对时序数据进行建模,可以更有效地捕捉工艺参数的动态变化,提高质量预测的准确性。例如,ZHOU等<sup>[8]</sup>提出了一种基于长短期记忆网络(long short-term memory network, LSTM)的刀具剩余寿命预测方法,通过 Hilbert-Huang 变换提取磨损特征,捕捉变工况下的复杂时空关系,实现了高精度的寿命预测。HUA<sup>[9]</sup>结合 CNN 和 LSTM 开发了一种电能质量预测模型,能够提取时序数据的局部和全局特征,实现了对主动配电网电能质量稳态指标的预测。针对注塑质量预测,MUAZ等<sup>[10]</sup>提出了一种基于多任务编码-解码(multi-task encoding-decoding, MED)的深度学习模型,通过融合时间序列和非时间序列数据,结合注意力机制显著提高了质量特性预测精度,其平均均方误差较基准模型降低了一个数量级。以上研究表明,深度学习模型在捕捉复

杂时序特征及其动态交互方面具有显著优势,为复杂工业场景中的质量预测提供了有效解决方案。

此外,借助特征融合方法,深度学习能将不同来源的特征进行有效整合,进一步提升模型的预测性能。XIE等<sup>[11]</sup>提出了一种基于多传感器融合和 CNN 的机械故障诊断方法,通过主成分分析(principal component analysis, PCA)将多传感器信号融合为 RGB 图像,用于诊断任务,显著提高了故障分类的准确性。DUAN等<sup>[12]</sup>提出了一种基于多信息融合轴(multi-information fusion axis attention mechanism, MFA)注意力机制和改进的多尺度卷积神经网络(improved multi-scale convolutional neural network, IMSCNN)的滚动轴承故障诊断方法,提高了小样本和噪声环境下的特征提取能力,在复杂工况和强噪声条件下具备更高的分类准确性和抗干扰能力。MA等<sup>[13]</sup>提出了一种基于多传感器和多层信息融合的压缩机叶片裂纹检测方法,利用循环平稳性相关度(cyclically dependent correlation stability, CDCS)进行数据融合,并结合多尺度注意力模块和双分支 1D-CNN 进行特征提取与融合,在复杂工况下显著提高了裂纹检测的准确性。钱庆杰等<sup>[14]</sup>提出了一种基于双层双向门控循环单元网络(DL-BiGRU)的多特征融合模型,用于注塑件尺寸预测,该模型融合了高频时序特征、瞬时特征和成形机状态特征,显著提高了预测精度和稳定性。以上研究证实了深度学习结合特征融合的有效性。

已有研究表明,注射和保压阶段对注塑件质量有显著影响。CHEN等<sup>[15]</sup>提出了一种基于拉杆伸长信号的质量监控系统,通过调整注射速度、V/P 切换点和保压压力来提高注塑件的质量稳定性,发现注射和保压阶段对注塑件质量影响较大。LÓPEZ等<sup>[16]</sup>采用设计实验方法研究了注塑参数对复杂零件质量的影响,实验结果表明,保压时间、保压压力和注射温度显著影响零件质量。因此,提取注射和保压阶段的统计特征作为关键补充特征,有助于提高多特征融合模型的预测精度。

综上,深度学习模型对回归预测问题存在一定优势。考虑到深度学习模型对输入特征有着较高的质量要求,不同于以往相关特征工程研究中的高频时序特征、统计特征、状态特征的简单拼接融合思路,为了更高效捕捉关键注塑过程中复杂的动态变化的影响,本文提出了一种基于 TCN-

BiGRU-SE的注塑质量预测模型,引入两阶段特征提取的多源特征融合思路,第一阶段从高频传感器数据、注射和保压阶段统计特征及机器状态信息中构建高质量输入;第二阶段分别对输入数据进行深度特征提取,保留每类数据的独特信息;然后通过特征融合输入非线性映射层进行回归预测。为系统全面地验证方法的有效性,本文在富士康成形数据集上进行了模型性能对比和稳定性检验,此外,设计两类消融实验分别验证特征融合策略和各网络层对提升模型性能的作用,并扩展到多个注塑实验数据集上进行泛化性检验。这些系列实验均证实该方法在多源数据融合和时序特征建模中的优势,模型具有较强的稳定性和广泛的适应性。

### 1 注塑过程分析

#### 1.1 注塑过程基本原理

注塑成形是一种高效的塑料加工工艺,通常通过注塑机完成。图1所示为注塑机的关键结构。注塑机主要由料斗、螺杆、加热带、喷嘴、模具等部件组成,注塑机通过这些部件实现塑料的加热、熔融、注射、冷却及制品脱模等完整的成形过程。注塑成形的工作原理可以分为以下几个步骤:①具有热塑性的塑料颗粒从料斗进入螺杆系统,螺杆在加热带的作用下不断加热塑料,使其逐渐熔融;②螺杆通过旋转和轴向前移将熔融状态的塑料推送至喷嘴,在高压的作用下从喷嘴注入模具,逐渐填充模腔;③熔融塑料在模具内逐渐冷却固化,形成与模具型腔一致的塑料制品;④冷却完成后,模具打开,成形的制品通过机械装置或顶针从模具中取出,制品脱模,重复下一个注塑循环。整个注塑成形过程实现了高效的制造。

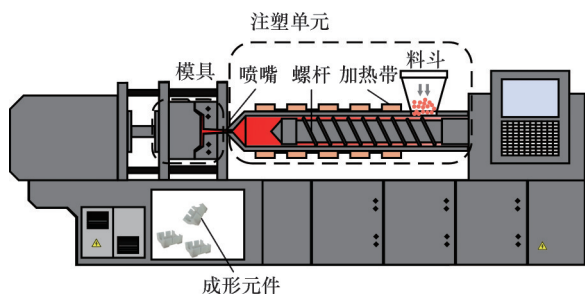


图1 注塑机的关键结构

Fig.1 Main structure of the injection molding machine

#### 1.2 数据来源及特征分析

##### 1.2.1 数据来源

数据来源于第四届工业大数据竞赛——注塑成形工艺虚拟测量,该数据集由富士康工业互联网

股份有限公司提供。本文将对注塑成形过程中产生的数据及特征进行具体分析。

现代注塑机通常会在模具内部安装多个传感器,以实时监测生产过程中的关键工艺参数。这些传感器通常具有0.5%~1%的测量精度,能够准确测量模内压力、温度等信息,并记录大量的生产过程数据。具体可分为以下三类:高频数据,机器状态,成形制品的质量(标签)。不同类型的部分特征说明见表1。高频数据反映成形过程的实时变化情况,一般由高频传感器通过一定频率进行采集,值是连续且高维的时间序列。机器状态数据通常反映设备的运行情况和生产周期中的关键节点,帮助分析生产过程是否顺畅并检测潜在的设备故障,值是离散的。质量标签(如尺寸、质量等)则直接体现了制品的精度和质量,是评估注塑工艺的核心标准,值是离散的。

表1 注塑成形过程中的部分特征说明

Tab.1 Description of partial features in the injection molding process

特征类型	变量名称	符号	数值范围	单位	说明
高频时序数据	模内压力	$p$	5~40	MPa	反映熔体填充模腔时模具内部的压力
	模内温度	$T$	90~240	°C	模腔内部的温度
	实际螺杆位置	$D$	32~105	mm	螺杆在注射筒中的实际物理位置
	冷却水流量	$Q$	12~20	L/min	冷却系统中循环水的流动速率
...	...	...	...	...	...
机器状态	保压时间	$T_b$	11~13.5	s	保压阶段的持续时间
	熔胶背压	$P_b$	5~11	MPa	液压系统施加在螺杆上的反向压力
	...	...	...	...	...
质量标签	尺寸(Size)	$S$	...	mm	成形制品某个部位的尺寸

虽然机器状态等非时间序列的数据在质量建模中也能提供较好的准确性,但它们无法捕捉生产周期内的动态变化。生产过程中的微小波动常常会对制品的最终质量产生深远影响,忽视这些动态可能导致模型错失关键信息,从而影响预测的精度。因此,深入挖掘和分析高频数据能构建更为精准的质量预测模型,优化生产过程,提高制品质量和生产稳定性。

##### 1.2.2 传感器高频特征分析

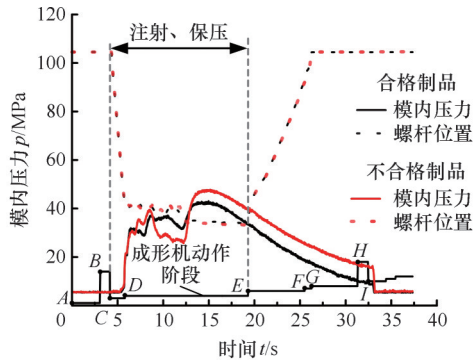
在注塑成形过程中,高频传感器记录了聚合物从固态到熔融状态的动态转变。压力和温度曲线反映了熔体从喷嘴注入模腔的全过程,对模型

预测微小偏差的质量标签至关重要。注塑数据集提供了制品三个关键部位的尺寸及上下偏差,见表2。随机抽取一个合格制品和不合格制品,绘制温度、压力随工序和时间变化的过程,对注塑过程进行具体分析,如图2所示。

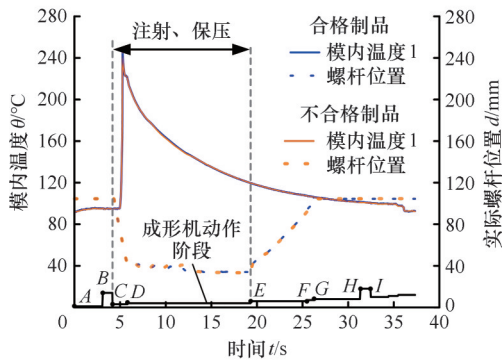
表2 注塑件尺寸及公差范围

Tab.2 Dimensions and tolerance range of injection parts

尺寸名称	尺寸及上下偏差/mm	公差范围 T/mm
Size1	300±0.150	0.30
Size2	200±0.075	0.15
Size3	200±0.075	0.15



(a) 压力-动作曲线



(b) 温度-动作曲线

图2 合格与不合格制品的温度、压力动作曲线对比

Fig.2 Comparison of temperature and pressure action curves between qualified and unqualified products

点A到点B为合模阶段,此时螺杆位置和模腔内压力、温度保持不变,动模和定模逐步合拢直至完全闭合;点B到点C为型芯入位阶段,型芯伸入模腔并定位,以形成复杂的制品几何结构,螺杆开始移动,模腔内压力、温度逐渐上升;点C到点D为注射填充阶段,螺杆按照给定速度移动,熔体由喷嘴迅速进入浇口并填充模腔,模腔内压力传感器读数开始变化,由于熔体接触模具壁,温度传感器读数迅速升高,达到最大值;点D到点E为保压阶段,螺杆以给定速度缓慢移动,模内压力在前半段呈现先升后降再升的趋势,这是由于熔融塑料快速填充模具型腔,当型腔接近充满时,流动速

度变慢,压力有所减小,随着塑料开始冷却和收缩,模具内需要增大保压压力来补偿收缩和变形,确保制品的形状和尺寸稳定,之后制品逐渐成形,当保压压力达到最大设定值后,螺杆停止移动直到保压结束,模内压力开始减小,由于长时间的保压过程,熔体逐渐冷却导致模内温度下降;点E到点F为熔胶阶段,螺杆逐渐回退并开始新一轮熔融塑料的准备,此时模腔内压力、温度迅速减小;点F到点G为后松退阶段,螺杆回退到初始位置,释放压力,为下一次注塑循环做准备;点G到点H为冷却阶段,模腔内的塑料冷却固化,形成最终制品,模腔内压力逐渐减小;点H到点I为中子返回阶段,冷却完成后中子回到初始位置;点I之后为开模-顶出阶段,模具打开,顶出系统将塑料制品顶出模腔,完成一个注塑循环。

通过对比合格与不合格制品的压力与温度曲线,可以明显看出合格制品和不合格制品的压力、温度差异。在注射和保压阶段,合格制品的模内压力变化曲线更加稳定,模内压力的增减较为平稳,温度曲线呈均匀的降温过程,有助于材料的均匀冷却,确保成形的精度。不合格制品的压力则波动较大,而模内压力的异常变化可能导致材料流动不均,进而影响制品尺寸的稳定性。除注射、保压阶段外的其余动作阶段的压力则无明显差异。抽取合格制品的3个部位尺寸分别为300.026、199.962、199.979 mm,而不合格制品3个部位的尺寸分别为300.3、199.915、199.082 mm,两者差值分别为-0.274、0.047、0.897 mm,表明注射、保压阶段对制品质量有直接影响。

全动作阶段的时序特征提取有助于捕捉生产过程中的动态变化,并准确反映整个生产周期的质量变化。而以上分析表明注射和保压阶段对制品质量的影响显著,且两阶段加工时间约占整个注塑循环的40%。因此,提取注射、保压阶段的统计特征作为补充特征,可提升模型对细微质量差异的敏感性,使模型对微小差异标签的预测更为准确。

## 2 基于 TCN-BiGRU-SE 的注塑质量预测模型

为了实现注塑质量的预测,提出了一种基于时间卷积网络(temporal convolutional network, TCN)-双向门控循环单元(bidirectional gated recurrent unit, BiGRU)-SE注意力机制(squeeze-and-excitation, SE)的注塑质量预测模型(TCN-

BiGRU-SE)。如图3所示,该模型主要由三部分  
组成:第一阶段特征提取、第二阶段深度特征提取  
以及输出层。第一阶段的特征提取主要由高频数  
据抽样、统计特征提取、机器状态数据清洗构成。

第二阶段通过深度学习模型对各类特征进行深入  
提取,之后,将不同来源的特征进行融合。最终,  
深度融合特征经过多层感知机(multilayer percep-  
tron, MLP)进行非线性映射,得到预测输出。

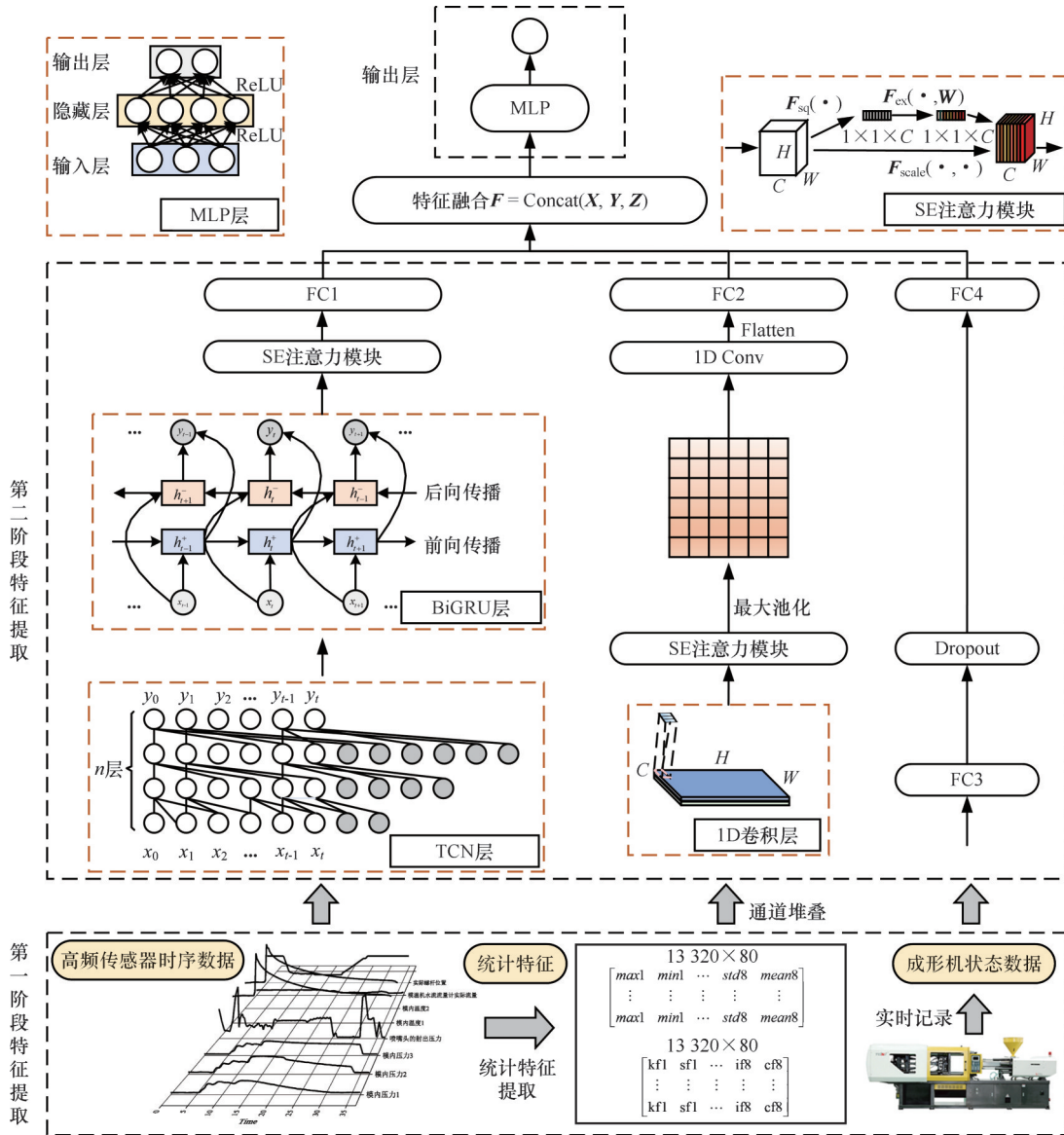


图3 基于TCN-BiGRU-SE多特征融合的注塑质量预测模型

Fig.3 Injection quality prediction model based on TCN-BiGRU-SE multi-feature fusion

### 2.1 第一阶段特征提取

在第一阶段,针对三种不同来源的数据进行相应的预处理,为第二阶段的深度特征提取提供高质量的特征输入。具体步骤如下:

1)高频传感器时序特征。由于传感器采集的数据是长时间序列,将其完整输入模型处理会导致计算复杂度过高,故采用等间隔采样方法,选取128个关键数据点作为样本,以保留生产过程中的动态变化,同时减少计算量。

2)统计特征。针对注射和保压这两个关键阶段,提取统计特征。除了定量特征值(最大值、

最小值、均方根、平均值和标准差)外,还提取了量纲一指标(峰度因子、偏度因子、波形因子、冲量因子和间隙因子)。与传统特征拼接方式不同,本研究将定量统计特征与量纲一指标进行通道拼接,构建三维特征矩阵,以便通过CNN进行高效的降维处理。

3)成形机状态特征。状态特征为离散值,只需进行常规的数据清洗处理,如异常值剔除和缺失值填补,以确保数据质量。

### 2.2 第二阶段特征提取

在第二阶段,通过构建多种深度特征提取模

型,对第一阶段预处理后的数据进行进一步的特征提取。具体步骤如下:

1)时序特征由TCN-BiGRU-SE模块提取。首先,时序数据输入TCN。TCN由两种卷积形式组合:因果卷积和膨胀卷积。通过因果卷积保证模型在预测时仅依赖过去信息,避免未来信息干扰;同时,膨胀卷积扩大感受野,捕捉长时间的依赖特征,使每个时序特征能够完整表达独特的时序关系。然后,TCN提取的特征进一步输入至BiGRU,通过双向处理路径同时捕捉前向和后向的依赖信息,增强对复杂时序特征的理解与表达。随后,BiGRU提取的特征传递至SE注意力机制,SE注意力机制通过全局加权对各通道特征进行自适应调整,突出关键通道特征,优化特征表示的有效性,使模型能学习到最重要的特征。最后,将提取的时序特征通过一层全连接层(fully connected,FC)进行非线性降维,确保高频传感器时序特征中的关键动态变化被精准捕捉,同时便于后续的特征融合操作。

2)统计特征的提取主要通过CNN和SE注意力机制进行处理。由于不涉及时序信息,且两类统计特征在通道上堆叠形成三维矩阵,故使用简单的CNN即可完成特征提取操作。首先,特征经过一维卷积层(1D Conv)高效提取局部模式并减少维度,确保计算效率。然后,特征进入SE注意力机制,通过自适应调整通道权重,增强关键特征的表达。随后,再使用一层1D Conv细化局部模式的提取,增强对特征的建模表现。最后,特征被展平并传递至FC2层进行处理。此流程避免了特征直接展平并通过全连接层降维带来的高计算开销,同时保留统计特征的核心信息。

3)成形机状态特征由于其离散性和低维性,采用两层FC(FC3, 4)进行降维,随后传递至特征融合阶段。

各深度学习模型的机制及介绍将在后文阐述。此外,模型采用Huber损失函数,以兼顾小误差下的精确拟合和大误差下的鲁棒性。

### 2.2.1 TCN网络

TCN<sup>[17]</sup>是一种针对序列数据设计的神经网络结构,主要特点是结合了因果卷积和膨胀卷积。TCN层的计算公式如下:

$$y_t = \text{ReLU}\left(\sum_{i=0}^{K-1} W_i \cdot x_{t-d+i} + b\right) \quad (1)$$

式中: $x_t$ 为输入序列; $y_t$ 为输出序列; $W_i$ 为卷积核的权重; $b$ 为偏置项; $K$ 为卷积核的大小; $d$ 为膨胀因子;ReLU为非线性激活函数。

TCN通过引入随机失活(Dropout)和权重归一化(WeightNorm)技术,提高了模型的稳定性与泛化能力。Dropout随机丢弃神经元以减少对特定数据点的依赖,防止过拟合;WeightNorm则解决梯度爆炸问题,加速训练过程。如图4所示,TCN由多层残差连接的膨胀卷积层组成,每层的膨胀因子逐步增大,扩展感受野,从而捕捉更长的时间依赖信息。残差连接用于缓解深层网络的梯度消失问题,确保信息跨层传递,并加速模型的收敛。

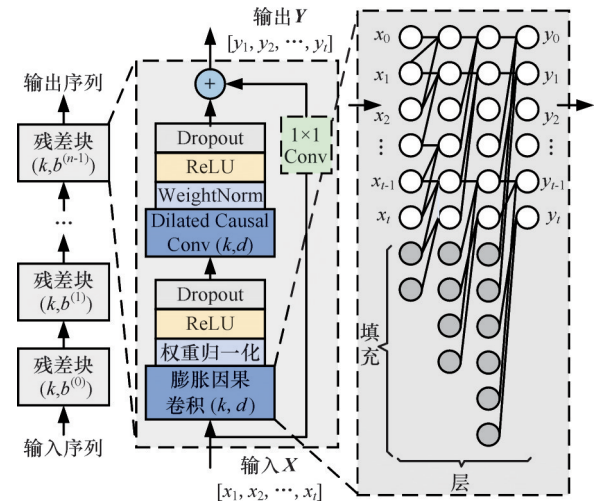


图4 TCN网络

Fig.4 TCN network

### 2.2.2 BiGRU网络

GRU<sup>[18]</sup>是一种高效的循环神经网络,通过更新门和重置门来保留长依赖信息,同时降低计算复杂度。BiGRU进一步增强了模型对正反向依赖关系的捕捉能力,通过同时处理正向和反向的输入序列,能够有效地获取序列的前向和后向信息,如图5所示。通过将正向和反向GRU单元的输出拼接,BiGRU能够在每个时间步上综合前后文信息,从而提高模型对复杂时序特征的建模能力。其工作机制可表示为

$$\vec{h}_t = \text{GRU}(\vec{h}_{t-1}, x_t) \quad (2)$$

$$\overleftarrow{h}_t = \text{GRU}(\overleftarrow{h}_{t+1}, x_t) \quad (3)$$

$$h_t = [\vec{h}_t; \overleftarrow{h}_t] \quad (4)$$

式中: $\vec{h}_t$ 为正向GRU在时间步 $t$ 的隐藏状态; $\overleftarrow{h}_t$ 为反向GRU在时间步 $t$ 的隐藏状态; $x_t$ 为输入序列在时间步 $t$ 的输入; $h_t$ 为正向和反向隐藏状态的拼接结果。

### 2.2.3 SE注意力机制

SE注意力机制<sup>[19]</sup>通过对卷积网络的特征图进行通道的重标定,有效提高了模型对关键特征的捕捉能力。如图6所示,该机制通过全局平均池化操作将每个通道的空间信息压缩为通道描述

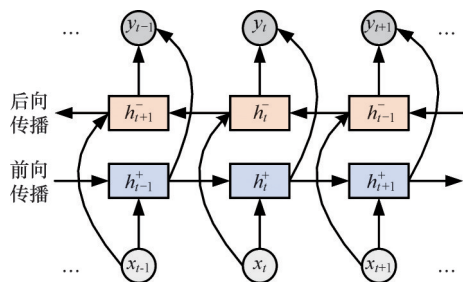


图5 BiGRU结构

Fig.5 BiGRU structure

符 \$z\_c\$, 随后利用两层全连接网络对 \$z\_c\$ 进行非线性变换: 第一层通过 ReLU 函数降维, 第二层通过 Sigmoid 函数恢复维度并生成每个通道的权重。通过这一机制, SE 自适应地对各通道的重要性进行加权, 从而增强模型对显著特征的响应能力。其数学计算公式如下:

$$z_c = \frac{1}{HW} \sum_{i=1}^H \sum_{j=1}^W x_{i,j,c} \quad (5)$$

$$s = \sigma(W_2 \cdot \text{ReLU}(W_1 \cdot z + b_1) + b_2) \quad (6)$$

式中: \$x\_{i,j,c}\$ 为特征图在通道位置 \$(i, j)\$ 处的值; \$H, W\$ 分别为特征图的高和宽; \$W\_1, W\_2\$ 为全连接层的权重矩阵, \$b\_1, b\_2\$ 为偏置项; \$\sigma\$ 为 Sigmoid 函数; ReLU 为非线性激活函数。

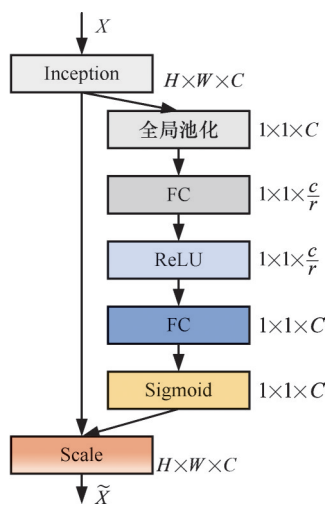


图6 SE注意力机制结构

Fig.6 SE attention mechanism structure

### 2.2.4 Huber 损失函数

Huber 损失函数特别适用于于标签值差距较小的回归任务。在注塑工艺中, 制品质量的变化范围通常较小, 要求模型具有较高的预测精度。Huber 损失函数在误差较小时采用平方损失, 以确保模型对小误差的精细拟合; 当误差较大时, 损失函数切换为线性损失形式, 从而减少异常值的影响。其计算公式为

$$L_\delta(y, \hat{y}) = \begin{cases} (y - \hat{y})^2 / 2 & |y - \hat{y}| \leq \delta \\ \delta(|y - \hat{y}| - \delta/2) & |y - \hat{y}| > \delta \end{cases} \quad (7)$$

式中: \$y\$ 为真实值; \$\hat{y}\$ 为预测值; \$\delta\$ 为阈值。

### 2.3 多特征融合与输出层

完成时序特征、统计特征以及成形机状态特征的深度提取后, 通过拼接操作融合不同类型特征的深度表征:

$$F = \text{Concat}(X, Y, Z) \quad (8)$$

式中: \$X, Y, Z\$ 分别为提取的时序、统计和成形机状态特征。

在输出层中, 融合后的深度特征矩阵传递到 MLP, 通过若干全连接层结合 ReLU 激活函数进行非线性映射, 最终生成预测结果。

## 3 实验结果与分析

为验证所提方法的有效性, 通过实验进行系统评估。

### 3.1 数据集介绍

本研究使用两类注塑数据集, 分别用于模型验证和泛化性检验。模型验证使用第四届工业大数据竞赛的注塑成形数据集, 共 16 600 模次的数据记录。该数据集包含 24 个高频传感器采集的 1562~1672 维时序信息、71 维机器状态数据和质量标签。经过数据清洗, 最终只保留了 8 个变化明显的高频时序数据和 41 维机器状态数据。数据集按照 8:2 的比例划分为训练集和测试集, 用于模型训练和评估。

泛化性检验使用来自 BOGEDALE 等<sup>[20]</sup>提供的注塑实验数据。实验在德国 Allrounder 520E 1500-800 注塑机上进行, 采用人为干扰变量(如料筒温度、模具温度、注射流量等)诱导不同的过程状态, 以模拟实际的工业生产环境。实验共收集了三个不同制品的注塑数据, 用于评估模型在多变生产环境下的泛化能力。其中一个为外壳制品(1167 模次), 其余为堆叠箱制品(829、1332 模次), 两个堆叠箱数据集是在不同工艺条件下生产的。数据集包含 2 个(注射流量、注射压力)高频传感器采集的 2048 维时序信息、机器状态和质量标签。同样以 8:2 的比例划分为训练集和测试集, 用于模型训练和评估。

### 3.2 模型超参数设置

所有程序均基于 TensorFlow 2.10 深度学习框架在 Python 平台上实现, 运行环境为 Windows 11 操作系统, 硬件配置包括 3.40 GHz 的 Intel Core i7-14700KF CPU、8 GB 显存的 NVIDIA GeForce RTX 4060 GPU 以及 32 GB 的运行内存。采用 Adam 优化器训练模型, 模型的 batch\_size 设置为 64, epoch 为 400, 学习率为 0.001, 损失函数选用 Huber loss。超参数优化通过贝叶斯算法完成, 具体模型参数设置见表 3。

表 3 各网络层参数设置

Tab.3 Parameter settings of each network layer

输入特征	网络层名称	设置参数	输出维度
传感器时序数据	Input(sensor_input)	(128,8)	(None,128,8)
	TCN	(filters=64,kernel=3,dilation_rate=[1,2,4,8,16],dropout_rate=0.2)	(None,128,64)
	BiGRU	(units=64,return_sequences=True)	(None,128,128)
	SE Block	(reduction_ratio=16)	(None,1,128)
	FC1	(units=512,activation='relu')	(None,512)
统计特征	Input(tj_input)	(2,80)	(None,2,80)
	Conv1D	(filters=64,kernel=3,padding='same',activation='relu')	(None,2,64)
	SE Block	(reduction_ratio=16)MaxPooling1D(pool_size=2)	(None,1,64)
	Conv1D	(filters=64,kernel=3,activation='relu',padding='same')Flatten	(None,64)
	FC2	(units=160,activation='relu')	(None,160)
成形机状态数据	Input(state_input)	(,41)	(None,41)
	FC3,4	(units=41,activation='relu')*2,dropout_rate=0.2	(None,41)
输出层 MLP	Concatenate	-	(None,713)
	FC5,6	(units=512,activation='relu')*2,dropout_rate=0.2	(None,512)
	Out	units=1	(None,1)

3.3 模型对比与稳定性检验

3.3.1 评价指标

本研究以注塑件尺寸为预测目标,属于典型的回归任务。为全面评估模型性能,选择以下 3 个评价指标:决定系数( $R^2$ )、均方误差(MSE)以及平均绝对误差(MAE),计算公式略。

3.3.2 对比实验

为了验证所提 TCN-BiGRU-SE 多特征融合模型在注塑件尺寸预测中的性能,选取常用的传统机器学习模型,即支持向量回归(support vector regression, SVR)、极限梯度提升(XGBoost)、LightGBM,以及常见的时序提取模型(MLP、LSTM、GRU、Transformer、CNN-GRU 和 CNN-LSTM)模型作为对比基准。

由于传统机器学习模型不能直接提取时序特征,故提取时序特征的统计值(均值、标准差等),

使特征粒度统一,并通过相关性系数和多模型平均重要度进行两步特征筛选,选取最重要的 12 维特征(模内温度均值、压力标准差等)作为传统机器学习模型的输入。模型进行多次独立训练和测试,结果取均值。各模型在 Size1~Size3 的具体预测结果见表 4。

由表 4 可看出,TCN-BiGRU-SE 模型表现优异。所提模型在 3 个尺寸的预测上均取得了最高的  $R^2$  值和最低的 MSE、MAE 值。与表现最好的传统机器学习模型 XGBoost 相比,所提模型在 Size1~Size3 上的 MSE 值分别降低了 20.0%、11.2% 和 19.3%;与常用时序提取模型 GRU 对比,分别降低了 31.3%、22.4% 和 13.9%;与一些较为先进的模型(CNN-GRU、Transformer 等)相比也有所提升,这表明所提出的模型具有良好的预测性能。TCN-BiLSTM-SE 与 TCN-BiGRU-

表 4 不同模型的预测性能对比

Tab.4 Comparison of prediction performance of different models

模型	Size1			Size2			Size3		
	$R^2$	MSE 值	MAE 值	$R^2$	MSE 值	MAE 值	$R^2$	MSE 值	MAE 值
SVR	0.53	0.000 825	0.023 123	0.81	0.000 509	0.017 658	0.69	0.000 406	0.015 551
XGBoost	0.70	0.000 519	0.017 469	0.81	0.000 507	0.017 422	0.68	0.000 415	0.015 847
LightGBM	0.60	0.000 719	0.020 475	0.70	0.000 800	0.022 579	0.56	0.000 566	0.018 641
MLP	0.58	0.000 728	0.020 926	0.73	0.000 729	0.021 019	0.61	0.000 507	0.017 850
LSTM	0.69	0.000 552	0.018527	0.73	0.000 752	0.021 378	0.67	0.000 430	0.016 478
GRU	0.66	0.000 604	0.018 805	0.78	0.000 580	0.018 711	0.70	0.000 389	0.015 510
Transformer	0.74	0.000 454	0.016 421	0.83	0.000 464	0.016 628	0.72	0.000 363	0.014 904
CNN-GRU	0.75	0.000 440	0.015 839	0.79	0.000 580	0.018 447	0.72	0.000 366	0.014 663
CNN-LSTM	0.75	0.000 434	0.015 640	0.79	0.000 562	0.018 172	0.72	0.000 364	0.014 595
TCN-BiLSTM-SE	0.76	0.000 420	0.015 388	0.83	0.000 457	0.016 202	0.72	0.000 359	0.014 527
TCN-BiGRU-SE	0.77	0.000 415	0.014 983	0.83	0.000 450	0.015 973	0.74	0.000 335	0.014 078

SE的效果相近,LSTM可以处理可变量长输入,在处理复杂时序数据时有优势,但计算成本较高。在资源有限的情况下,TCN-BiGRU-SE是更高效的选择。

所提模型的预测效果如图7所示。可以看出,预测值曲线与真实值曲线较为吻合,模型具有较高的预测准确性。此外,预测曲线的波动与真实曲线的波动趋势基本一致,表明工艺变动导致的尺寸变化能被模型准确捕捉,验证了模型在处理数据动态变化方面的有效性。

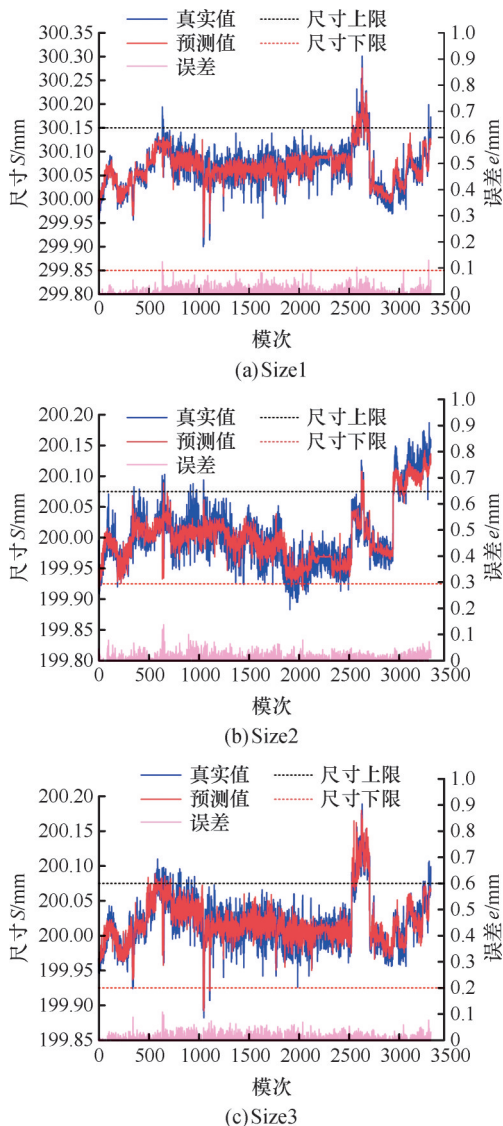


图7 TCN-BiGRU-SE模型预测效果  
Fig.7 Prediction performance of the TCN-BiGRU-SE model

### 3.3.3 稳定性检验

深度学习模型由于其复杂的结构和高维参数空间,容易受到随机初始化和数据分布变化等因素的影响,从而导致模型性能在不同训练过程中出现波动。因此,进行稳定性检验是必要的。为

了验证本文提出的TCN-BiGRU-SE模型的稳定性,将预测效果较好的一些模型进行了10次独立的训练与测试,绘制箱线图,如图8所示。可看出,所提模型的波动范围最小,稳定性最好。

## 3.4 消融实验

### 3.4.1 特征融合消融

为了评估不同特征组合对注塑件尺寸预测模型的影响,本研究进行了消融实验,测试了统计特征、高频传感器特征和机器状态特征单独使用及其组合的效果。不同特征组合对模型预测性能的影响见表5。相比单独使用统计特征时,采用高频+统计+机器状态特征的组合后,Size1的MSE和MAE值分别降低了91.9%和59.2%,Size2的MSE和MAE值分别减少了81.1%和48.7%,Size3的MSE和MAE值分别降低了76.7%和41.8%。与任意两种特征组合相比,三特征融合在MSE和MAE值上的性能表现均显著优越。结果表明,多特征融合显著提升了模型的预测性能,有效降低了预测误差。

### 3.4.2 模块消融

为了验证所提TCN-BiGRU-SE多特征融合模型中各模块的有效性,需要进行模块消融实验。在消融实验中,分别测试了移除或替换TCN、BiGRU和SE模块后的模型性能,以评估每个模块的独立贡献。对不同消融模型进行了多次独立的训练和测试,结果取均值,见表6。在Size1上,TCN-BiGRU相比仅使用BiGRU进行时序特征提取时,MSE值降低了4.6%,表明TCN模块在捕捉长时间依赖关系上发挥了重要作用。TCN通过卷积操作,能够更高效地提取全局时序特征。当增加SE注意力后,MSE值有一定幅度的下降。这表明SE注意力通过自适应调整各特征通道的权重,能够增强重要特征的表达,削弱无关特征,从而进一步优化了整体模型的表现。在Size2和Size3上,完整模型的MSE值比仅使用BiGRU分别减少了6.6%和8.5%,进一步证明了TCN、BiGRU和SE模块的协同作用。整体来看,多模块的合理结合提高了模型的预测精度,在模型计算效率与复杂度之间找到了较好的平衡。

此外,通过对比TCN-BiGRU-SE, FC, FC和TCN-BiGRU-SE, CNN, FC模型,发现使用CNN降维的误差比FC降维略低,且FC参数多、计算复杂度高,因此,采取将不同类型的统计特征进行通道堆叠后再使用CNN进行降维的策略,可

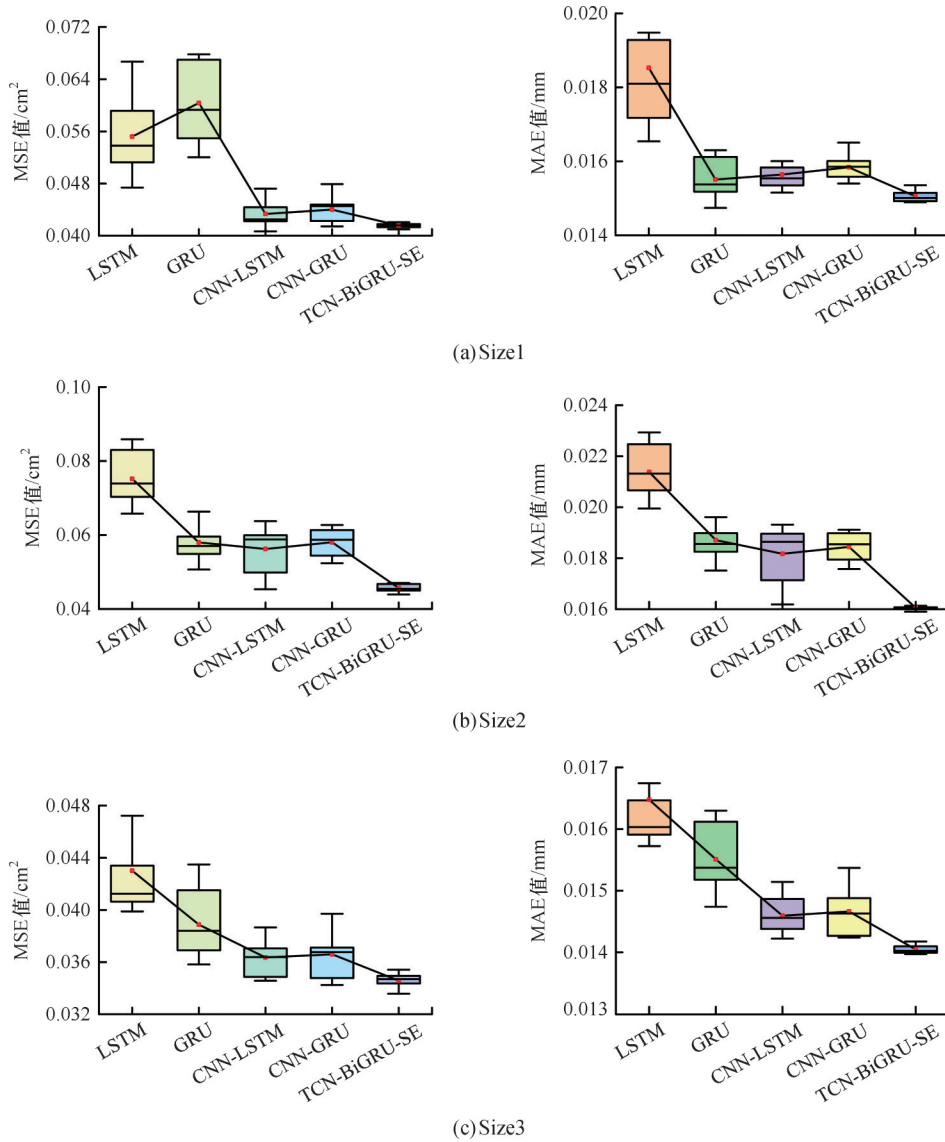


图8 模型稳定性箱线图

Fig.8 Box plot of model stability

表5 不同特征组合对模型预测性能的影响

Tab.5 Impact of different feature combinations on model prediction performance

使用的特征组合	Size1		Size2		Size3	
	MSE	MAE	MSE	MAE	MSE	MAE
统计	0.005 135	0.036 829	0.002 384	0.031 139	0.001 437	0.024 214
高频	0.002 273	0.030 285	0.002 503	0.031 555	0.000 656	0.019 644
机器状态	0.001 961	0.029 203	0.001 759	0.029 362	0.002 351	0.033 088
统计+机器状态	0.004 637	0.037 401	0.003 282	0.034 318	0.001 315	0.023 194
高频+统计	0.003 726	0.034 728	0.001 016	0.023 633	0.001 450	0.023 932
高频+机器状态	0.001 025	0.023 320	0.001 360	0.026 253	0.002 147	0.029 931
高频+统计+机器状态	0.000 415	0.015 034	0.000 450	0.015 973	0.000 335	0.014 078

以减少模型复杂度,优化模型性能。

### 3.5 泛化性检验

为验证所提模型的泛化性,在三个注塑实验数据集上进行了实验。以线性回归(LR)、SVR、K近邻算法(KNN)和XGBoost作为对比基准,以确保实验结果的代表性。同样对尺寸进行预测,

模型的表现见表7。Dataset1为注塑外壳制品数据集, Dataset2和Dataset3则为注塑堆叠箱制品数据集,两者的区别在于不同的注塑工艺条件。实验结果表明,TCN-BiGRU-SE模型在三类不同制品的尺寸预测上同样表现出色。在Dataset1和Dataset2上,其 $R^2$ 、MSE和MAE值均优于对比

表6 消融不同模块对预测性能的影响

Tab.6 Impact of ablation different modules on model prediction performance

不同特征使用的模块 (时序,统计,机器状态)	Size1		Size2		Size3	
	MSE	MAE	MSE	MAE	MSE	MAE
(TCN, CNN, FC)	0.000 427	0.015 352	0.000 452	0.016 157	0.000 367	0.014 370
(BiGRU, CNN, FC)	0.000 438	0.015 703	0.000 482	0.016 759	0.000 366	0.014 383
(TCN-SE, CNN, FC)	0.000 430	0.015 411	0.000 508	0.016 943	0.000 365	0.014 653
(TCN-BiGRU, CNN, FC)	0.000 418	0.015 350	0.000 470	0.016 404	0.000 377	0.014 749
(BiGRU-SE, CNN, FC)	0.000 421	0.015 234	0.000 455	0.016 283	0.000 351	0.014 250
(TCN-BiGRU-SE, FC, FC)	0.000 426	0.015 642	0.000 452	0.016 444	0.000 370	0.015 117
(TCN-BiGRU-SE, CNN, FC)	0.000 415	0.015 034	0.000 450	0.015 973	0.000 335	0.014 078

表7 所提模型在三个注塑实验数据集上的表现情况

Tab.7 Performance of the proposed model on three types of injection molding experimental datasets

模型	Dataset1			Dataset2			Dataset3		
	$R^2$	MSE	MAE	$R^2$	MSE	MAE	$R^2$	MSE	MAE
LR	0.35	0.000 608	0.018 520	0.72	0.001 798	0.033 474	0.81	0.002 897	0.031 546
SVR	0.52	0.000 455	0.016 998	0.69	0.001 975	0.033 688	0.74	0.004 123	0.029 324
KNN	0.48	0.000 490	0.017 588	0.78	0.001 380	0.028 473	0.85	0.002 392	0.027 357
XGBoost	0.51	0.000 460	0.017 041	0.77	0.001 512	0.029 644	0.87	0.001 957	0.025 745
TCN-BiGRU-SE	0.54	0.000 435	0.016 753	0.80	0.001 249	0.028 315	0.83	0.002 756	0.027 607

模型;在Dataset3中,尽管所提模型未能取得最优 $R^2$ 值,但其误差控制效果依然良好,接近最佳结果,展现了较强的泛化性。

#### 4 结语

1)描述了注塑过程的基本原理及特征数据来源,选取典型注塑数据集,分析注射和保压阶段对制品质量的影响。通过提取注射、保压阶段时序特征的定量特征值和量纲一值,将其在通道上进行堆叠,形成三维的补充特征集,为模型提供数据支持。

2)提出了基于TCN-BiGRU-SE的多特征融合预测模型。通过两阶段的特征提取策略,充分提取高频传感器数据、统计特征和机器状态等不同来源的深度特征并进行有效融合,提高了模型的动态捕捉能力和预测精度。

3)将模型与传统机器学习和常见深度学习模型进行比较,结果表明,TCN-BiGRU-SE在多项评价指标( $R^2$ 、MSE、MAE值)上均表现更佳,精度提升约10%;通过10次独立的稳定性检验,箱线图显示模型在不同训练过程中保持了较高的稳定性;通过模块消融和特征融合消融实验,验证了TCN、BiGRU、SE模块各自的独立贡献及其协同作用,展现了多特征融合在提升模型性能上的效果。

4)通过对三个注塑实验数据集进行泛化性检验,结果显示TCN-BiGRU-SE模型具有良好的

泛化能力,在不同工艺条件下仍能稳定捕捉工艺变化,表现出较强的适应性。

#### 参考文献:

- [1] 王新铭,党开放,马艺涛,等. 注塑机故障诊断技术进展[J]. 塑料工业, 2023, 51(11): 15-20.  
WANG Xinming, DANG Kaifang, MA Yitao, et al. Development of Fault Diagnosis Technology for Injection Molding Machines[J]. China Plastics Industry, 2023, 51(11): 15-20.
- [2] KHOSRAVANI M R, NASIRI S. Injection Molding Manufacturing Process: Review of Case-based Reasoning Applications [J]. Journal of Intelligent Manufacturing, 2020, 31(4): 847-864.
- [3] 刘永兴,唐小琦,钟靖龙,等. 基于轻量级梯度提升机的非对称风险注塑成形产品尺寸预测模型[J]. 中国机械工程, 2022, 33(8): 965-969.  
LIU Yongxing, TANG Xiaoqi, ZHONG Jinglong, et al. Asymmetric Risk Injection Molding Product Size Prediction Model Based on LightGBM[J]. China Mechanical Engineering, 2022, 33(8): 965-969.
- [4] LOCKNER Y, HOPMANN C, ZHAO Weibo. Transfer Learning with Artificial Neural Networks between Injection Molding Processes and Different Polymer Materials [J]. Journal of Manufacturing Processes, 2022, 73: 395-408.
- [5] 陈昱,项薇,龚川. 基于数据挖掘的注塑产品质量在线故障检测及预测[J]. 中国机械工程, 2023, 34(14): 1749-1755.  
CHEN Yu, XIANG Wei, GONG Chuan. Online Diagnostic Inspection and Prediction of Product Quality in Injection Molding Intelligent Factories Based on Data Mining [J]. China Mechanical Engineering, 2023, 34(14): 1749-1755.
- [6] LU Fanlei, GUI Weihua, QIN Liyang, et al. A Novel Semi-supervised Prediction Modeling Method Based on Deep Learning for Flotation Process with Large Drift of Working Conditions [J]. Advanced Engineering Informatics, 2024, 62: 102934.

- [7] SOBRIE L, VERSCHELDE M, HENNEBEL V, et al. Capturing Complexity over Space and Time *via* Deep Learning: an Application to Real-time Delay Prediction in Railways[J]. *European Journal of Operational Research*, 2023, 310(3): 1201-1217.
- [8] ZHOU Jingtao, ZHAO Xu, GAO Jing. Tool Remaining Useful Life Prediction Method Based on LSTM under Variable Working Conditions[J]. *The International Journal of Advanced Manufacturing Technology*, 2019, 104(9): 4715-4726.
- [9] HUA Liang. Comprehensive Early Warning of Power Quality in Distribution Network Based on Deep Learning [J]. *Wireless Networks*, 2024, 30(7): 6371-6384.
- [10] MUAZ M, YU Hanxin, SUNG W L, et al. A Multitask Encoder - Decoder Model for Quality Prediction in Injection Moulding [J]. *Journal of Manufacturing Processes*, 2023, 103: 238-247.
- [11] XIE Tingli, HUANG Xufeng, CHOI S K. Intelligent Mechanical Fault Diagnosis Using Multisensor Fusion and Convolution Neural Network[J]. *IEEE Transactions on Industrial Informatics*, 2022, 18(5): 3213-3223.
- [12] DUAN Xiaoyan, XUE Linlin, LEI Chunli, et al. Rolling Bearing Fault Diagnosis Method Based on Multi-information Fusion Characteristics under Complex Working Conditions[J]. *Applied Acoustics*, 2023, 214: 109685.
- [13] MA Tianchi, SHEN Junxian, SONG Di, et al. Multi-sensor and Multi-level Information Fusion Model for Compressor Blade Crack Detection [J]. *Measurement*, 2023, 222: 113622.
- [14] 钱庆杰, 余军合, 战洪飞, 等. 基于 DL-BiGRU 多特征融合的注塑件尺寸预测方法[J]. *浙江大学学报(工学版)*, 2024, 58(3): 646-654.  
QIAN Qingjie, YU Junhe, ZHAN Hongfei, et al. Dimension Prediction Method of Injection Molded Parts Based on Multi-feature Fusion of DL-BiGRU [J]. *Journal of Zhejiang University (Engineering Science)*, 2024, 58(3): 646-654.
- [15] CHEN Jianyu, ZHUANG Jiaxiang, HUANG M S. Enhancing the Quality Stability of Injection Molded Parts by Adjusting V/P Switchover Point and Holding Pressure[J]. *Polymer*, 2021, 213: 123332.
- [16] LÓPEZ A, AISA J, MARTINEZ A, et al. Injection Moulding Parameters Influence on Weight Quality of Complex Parts by Means of DOE Application: Case Study[J]. *Measurement*, 2016, 90: 349-356.
- [17] BAI S, KOLTER J Z, KOLTUN V. An Empirical Evaluation of Generic Convolutional and Recurrent Networks for Sequence Modeling [J]. *Arxiv Preprint Arxiv:1803.01271*, 2018.
- [18] CHUNG J, GULCEHRE C, CHO K H, et al. Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling [J]. *Arxiv Preprint Arxiv:1412.3555*, 2014.
- [19] HU Jie, SHEN Li, SUN Gang. Squeeze-and-excitation Networks[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, 2018: 7132-7141.
- [20] BOGEDALE L, DOERFEL S, SCHRODT A, et al. Online Prediction of Molded Part Quality in the Injection Molding Process Using High-resolution Time Series[J]. *Polymers*, 2023, 15(4): 978.

(编辑 陈勇)

**作者简介:**邓晓强,男,1999年生,硕士研究生。研究方向为制造系统工程。项薇(通信作者),女,1971年生,教授。研究方向为制造系统工程/工业工程、人工智能在制造及服务系统管理中的应用等。发表论文50余篇。E-mail: xiangwei@nbu.edu.cn。

**本文引用格式:**

邓晓强,战韬阳,项薇,等. 基于 TCN-BiGRU-SE 两阶段特征提取与多特征融合的注塑质量预测方法[J]. *中国机械工程*, 2026, 37(2):416-427.

DENG Xiaoqiang, ZHAN Taoyang, XIANG Wei, et al. Injection Molding Quality Prediction Method Based on Two-stage Feature Extraction and Multi-feature Fusion Using TCN-BiGRU-SE Model[J]. *China Mechanical Engineering*, 2026, 37(2):416-427.

(上接第 415 页)

- [22] 武春龙, 朱天明, 张鹏, 等. 基于功能模型和层次分析法的智能产品服务系统概念方案构建[J]. *中国机械工程*, 2020, 31(7): 853-864.  
WU Chunlong, ZHU Tianming, ZHANG Peng, et al. Conceptual Scheme Construction of Smart PSS Based on Functional Model and AHP [J]. *China Mechanical Engineering*, 2020, 31(7): 853-864.
- [23] 张水. 基于 TRIZ 面向功能的机械产品概念设计模型研究 [D]. 哈尔滨: 东北林业大学, 2019.  
ZHANG Shui. Research for TRIZ-based Function-oriented Mechanical Product Concept Design Model [D]. Harbin: Northeast Forestry University, 2019.
- [24] HAINES-GADD L. TRIZ for Dummies[M]. New York: John Wiley & Sons, Ltd., 2016.
- [25] 李弘, 肖玉, 颜惠庚. 技术系统进化法则在九屏幕法中的应用[J]. *统计与管理*, 2013, 28(2): 145-146.  
LI Hong, XIAO Yu, YAN Huigeng. Application of Evolu-

tion Law of Technical System in Nine-screen Method[J]. *Statistics and Management*, 2013, 28(2): 145-146.

(编辑 陈勇)

**作者简介:**武春龙(通信作者),男,1986年生,副教授、博士。研究方向为功能设计、创新设计、数字孪生、智能产品等。E-mail: wuchunlong@hebut.edu.cn。

**本文引用格式:**

武春龙,姚迈,王兴旺,等. 基于专利规避设计的疲劳驾驶预警系统扩展概念设计[J]. *中国机械工程*, 2026, 37(2):406-415.  
WU Chunlong, YAO Mai, WANG Xingwang, et al. Extended Conceptual Design of Driver Fatigue Monitor Systems Based on Patent Circumvention Design[J]. *China Mechanical Engineering*, 2026, 37(2):406-415.