

基于 TCN-Transformer 的 Q195 钢应力-应变曲线预测

柯贤伦, 孟政臣, 罗毅恒

武汉理工大学 汽车工程学院, 湖北 武汉 430070

DOI:10.61369/ME.2026040037

摘 要 : 应力-应变曲线是金属板材成形仿真、工艺参数设计和成形质量评价的重要基础, Q195 钢在不同加载速率下呈现非线性强化特征。传统 Johnson-Cook (J-C) 模型形式简洁, 但在室温小样本条件下难以充分描述弹塑性过渡、加工硬化及速率效应耦合关系。针对上述问题, 基于 9 组 Q195 钢室温单向拉伸实验数据 (样本级平均应变速率 $1.00 \times 10^{-3} \sim 1.70 \times 10^{-1} \text{ s}^{-1}$), 提出融合时序卷积网络 (Temporal Convolutional Network, TCN) 与 Transformer 编码器的序列回归模型 (TCN-Transformer), 用于预测不同应变速率下的完整工程应力-应变曲线。模型以工程应变、瞬时应变速率和因果平均应变速率为输入, 通过 TCN 提取局部曲线形态, 并利用 Transformer Encoder 建立全局加载历史依赖; 采用嵌套留一交叉验证 (nested leave-one-out cross-validation, nested LOOCV) 进行超参数选择与泛化评估。结果表明, TCN-Transformer 的 MAE、RMSE 和 R^2 分别为 3.40 MPa、5.22 MPa 和 0.9849, 优于 J-C 模型和 SVR; 其极限抗拉强度 (UTS) 应力 MAE 为 1.91 MPa。特征消融与 Gradient SHAP 分析表明, 速率相关特征对峰值强度和曲线重建均具有重要贡献。研究结果可为小样本条件下低碳钢应力-应变曲线的数据驱动建模提供参考。

关 键 词 : Q195 钢; 应力-应变曲线; TCN-Transformer; 嵌套留一交叉验证; Johnson-Cook 本构模型

Prediction of Stress-strain Curves for Q195 Steel Based on the TCN-Transformer Model

Ke Xianlun, Meng Zhengchen, Luo Yiheng

School of Automotive Engineering, Wuhan University of Technology, Wuhan, Hubei 430070

Abstract : Stress-strain curves serve as a fundamental basis for metal sheet forming simulation, process parameter design, and forming quality evaluation. Q195 steel exhibits nonlinear strengthening characteristics under varying loading rates. While the traditional Johnson-Cook (J-C) model is concise in structure, it struggles to adequately describe elastic-plastic transitions, work hardening, and rate-effect coupling under small-sample conditions at room temperature. Addressing these limitations, this study proposes a sequential regression model (TCN-Transformer) integrating a Temporal Convolutional Network (TCN) with a Transformer encoder, utilizing nine sets of uniaxial tensile test data from Q195 steel at room temperature (with sample-level average strain rates ranging from 1.00×10^{-3} to $1.70 \times 10^{-1} \text{ s}^{-1}$). The model inputs include engineering strain, instantaneous strain rate, and causal average strain rate; the TCN extracts local curve morphology, while the Transformer encoder captures global loading history dependencies. Hyperparameter selection and generalization performance were evaluated using nested leave-one-out cross-validation (LOOCV). The results demonstrate that the TCN-Transformer achieves MAE, RMSE, and R^2 values of 3.40 MPa, 5.22 MPa, and 0.9849, respectively, outperforming both the J-C model and SVR; its mean absolute error (MAE) for ultimate tensile strength (UTS) stress is 1.91 MPa. Feature ablation and gradient SHAP analysis reveal that rate-dependent features significantly contribute to both peak strength prediction and curve reconstruction. These findings provide valuable insights for data-driven modeling of low-carbon steel stress-strain curves under small-sample conditions.

Keywords : Q195 steel; stress-strain curve; TCN-Transformer; nested leave-one cross-validation; Johnson-Cook constitutive model

基金项目: 省级大学生创新创业训练计划项目“基于深度学习的智能驱动形函数构建关键技术研究”(项目编号: S202510497115)。

作者简介: 柯贤伦 (2005—), 男, 汉族, 湖北十堰人, 本科生, 研究方向为材料成形与深度学习。

引言

应力-应变关系是金属材料成形工艺设计、有限元仿真和质量控制的重要基础。Q195钢成本低、塑性较好，常用于结构件、冲压件及一般工程零部件。实际板材成形过程中，材料不同区域会经历拉伸、弯曲和局部塑性变形，应变速率也会随变形位置和成形阶段发生变化。若材料模型不能准确描述不同加载速率下的完整应力-应变响应，将影响载荷预测、回弹分析、破裂风险判断和工艺参数优化。受实验成本与周期限制，工程中通常仅能获取有限应变速率下的少量拉伸曲线，因此建立适用于小样本实验条件的 Q195 钢应力-应变曲线预测方法具有实际工程意义。

Johnson-Cook (J-C) 等传统本构模型表达简洁、参数具有物理意义^[1]，在金属材料流动应力描述中应用广泛；但该类模型依赖预设函数形式，应变硬化与速率强化项相互解耦，在弹塑性过渡段和峰值邻域易产生较大局部偏差，对室温小样本条件下多速率曲线的整体拟合精度有限。近年来 SVR、人工神经网络等数据驱动方法在材料流动应力和应力-应变曲线预测中得到了应用^{[2-3][4]}，但逐点回归范式难以利用曲线上前后位置的状态关联。TCN 和 Transformer 等序列建模方法在处理时序数据方面展现出优势^{[5][6]}。TCN 擅长局部序列特征提取，Transformer 可通过自注意力机制捕捉长程依赖，二者结合有利于同时捕获曲线的局部形态特征和全局依赖关系。

基于上述问题，本文围绕 9 组 Q195 钢室温拉伸数据，提出 TCN-Transformer 序列回归模型。模型以工程应变、瞬时应变速率和因果平均应变速率为输入，输出完整工程应力序列；通过 TCN 局部编码与 Transformer 全局注意力建模相结合，以提升对弹塑性过渡段、加工硬化段和峰值邻域的重建能力。为降低小样本条件下模型选择带来的性能估计偏差，本文采用嵌套留一交叉验证 (nested LOOCV) 进行超参数选择与泛化评估^[7]。在此基础上，将 TCN-Transformer 与 J-C 本构模型、支持向量回归 (Support Vector Regression, SVR)、Transformer 和 TCN 进行系统对比，并结合特征消融与 Gradient SHAP 分析各输入特征对预测的贡献。

一、实验与方法

(一) 实验材料与方法

实验材料为 I 级 Q195 钢热轧板材，化学成分符合 GB/T 700—2019^[8]。试样按 GB/T 228.1—2021 加工为哑铃状薄片试样^[9]，在 Instron 5982 电子万能材料试验机上进行室温单向拉伸，采样频率为 10 Hz。实验共获得 9 组曲线，各试样的平均应变速率范围为 $1.00 \times 10^{-3} \sim 1.70 \times 10^{-1} \text{ s}^{-1}$ 。典型曲线如图 1 所示，Q195 钢随应变速率升高表现出强度提高、峰值应变略减小的趋势。

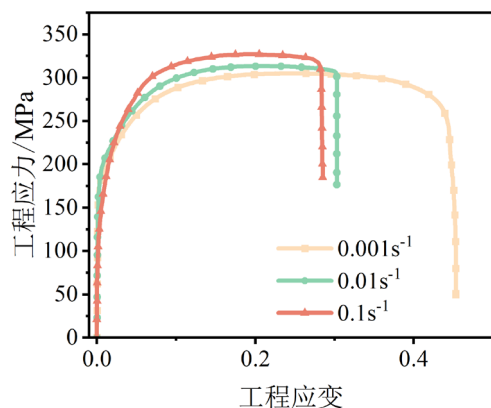


图 1 Q195 钢在典型应变速率下的工程应力-工程应变曲线

(二) J-C 模型与数据预处理

室温下忽略热软化项，J-C 模型可写为

$$\sigma = (A + B\varepsilon_p^N) \left[1 + C \ln \left(\frac{\dot{\varepsilon}_p}{\dot{\varepsilon}_0} \right) \right]$$

式中： σ 为流动应力，A 为屈服强度；B、n 描述应变硬化；C 为应变速率敏感系数； ε_p 为真塑性应变 $\dot{\varepsilon}_p$ 、 $\dot{\varepsilon}_0$ 分别为塑性应变速率和参考应变速率。

将颈缩前工程应力-工程应变曲线转换为真应力-真塑性应变曲线，取 $\dot{\varepsilon}_0 = 0.001 \text{ s}^{-1}$ ，得到 $A = 175.93 \text{ MPa}$ 、 $B = 502.91 \text{ MPa}$ 、 $n = 0.5537$ 、 $C = 0.010591$ 。J-C 模型可反映总体强化趋势，但应变硬化项与速率强化项相互解耦，易在过渡段和高应变速率样本中产生偏差。数据驱动模型输入为工程应变、瞬时应变速率和因果平均应变速率；各外层折的截断阈值和标准化参数均由训练样本确定。

(三) TCN-Transformer 模型

TCN-Transformer 由 TCN 局部时序编码器、线性投影层、位置编码模块、Transformer Encoder 和逐点回归输出头组成，结构如图 2 所示。输入序列由工程应变、瞬时应变速率和因果平均应变速率构成，输出为对应的工程应力序列。其中，因果平均应变速率为当前的应变点之前所有瞬时应变速率的逐点累积平均，保证不引入未来信息。TCN 前端通过因果卷积和扩张卷积扩大感受野，逐点提取邻近应变区间的局部序列模式；Transformer Encoder 通过多头自注意力进一步建立长程依赖关系，使模型能够同时利用局部形态变化和全局加载历程。训练以标准化工程应力序列均方误差为损失函数，采用 Adam 优化，并通过外层 LOOCV 评估泛化性能、内层 LOOCV 选择超参数。为解释速率特征影响，采用 Gradient SHAP 进行归因分析^[10]。

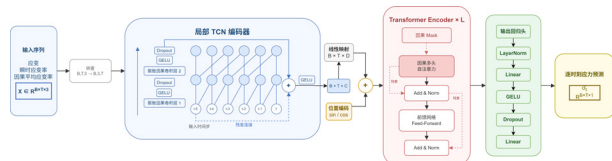


图2 TCN-Transformer模型结构图

二、结果与分析

（一）模型性能对比

为系统考察不同建模范式的性能差异，本文选取 J-C 本构模型（物理基线）、SVR（经典机器学习基线）以及 Transformer、TCN 和 TCN-Transformer（深度学习序列模型）进行逐级对比。表1为各模型的预测性能。J-C本构模型整体预测误差较大；SVR较 J-C 明显改善，表明直接学习非线性映射更适合当前小样本任务。深度学习序列模型进一步降低了误差：其中 TCN 的 MAE 最低（3.25 MPa），TCN-Transformer 的 RMSE 最低、 R^2 最高（5.22 MPa，0.9849）。TCN 的低 MAE 反映了其局部卷积编码在此任务中的强基线地位，而 TCN-Transformer 在 RMSE 和 R^2 上的优势表明，全局自注意力能够降低大偏差样本上的预测波动，提升了模型对困难样本的鲁棒性。

表1 各模型整体曲线预测误差

模型	MAE/MPa	RMSE/MPa	R^2	评价方式
J-C 本构	10.06 ± 4.16	13.28 ± 5.20	0.8851 ± 0.0558	全数据标定（有效段）
SVR (RBF)	5.15 ± 2.74	11.86 ± 1.81	0.9222 ± 0.0259	nested LOOCV
Transformer	4.35 ± 2.50	6.47 ± 2.91	0.9763 ± 0.0213	nested LOOCV
TCN	3.25 ± 1.26	5.67 ± 2.43	0.9820 ± 0.0130	nested LOOCV
TCN-Transformer	3.40 ± 2.10	5.22 ± 1.99	0.9849 ± 0.0175	nested LOOCV

为展示模型在不同预测难度样本上的曲线重建能力，依据 TCN-Transformer 在 9 个外层 LOOCV 测试样本上的 RMSE 排序，选取最低、中位和最高的 3 个代表性样本绘制预测曲线，如图3所示。

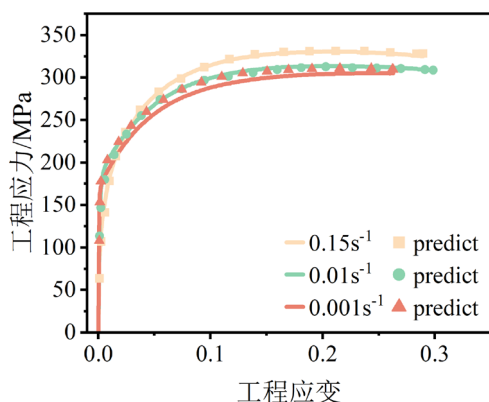


图3 代表性外层测试样本的预测曲线对比

（二）特征消融实验

为验证输入特征贡献，本文对 TCN-Transformer 开展特征消融实验。在保持模型结构、nested LOOCV 流程和搜索预算不变的前提下，分别移除因果平均应变速率、瞬时应变速率以及仅保留应变通道，结果如表2所示。

表2 特征消融实验结果（TCN-Transformer）

特征组	MAE/MPa	RMSE/MPa	R^2	Δ RMSE/MPa
应变 + 瞬时应变速率 + 因果平均应变速率	3.40	5.22	0.9849	0
应变 + 瞬时应变速率	3.79	5.57	0.9827	+0.36
应变 + 因果平均应变速率	3.86	7.01	0.9725	+1.79
仅应变	8.55	10.51	0.9390	+5.30

消融结果表明，移除瞬时应变速率后 RMSE 增加 1.79 MPa，高于移除因果平均应变速率后的 0.36 MPa；仅保留应变通道时 RMSE 增加 5.30 MPa，说明速率相关特征不可或缺。瞬时应变速率反映当前局部加载状态，因果平均应变速率描述加载历程信息，二者具有互补作用。

（三）UTS 特征点预测结果

除整条曲线预测精度外，本文进一步考察各模型对 UTS 应力和 UTS 应变的预测能力。分别从真实曲线和预测曲线中提取最大应力及其对应应变，在 9 个外层测试样本上计算误差，结果如表3所示。

表3 各模型 UTS 特征点预测误差

模型	UTS 应力 MAE/MPa	UTS 应力 RMSE/MPa	UTS 应变 MAE	UTS 应变 RMSE
J-C 本构	7.96	8.80	0.0781	0.0786
SVR (RBF)	5.18	6.38	0.0246	0.0335
Transformer	3.48	4.04	0.0120	0.0158
TCN	2.22	2.61	0.0111	0.0139
TCN-Transformer	1.91	2.54	0.0186	0.0214

综合表1和表3可见，TCN-Transformer 在完整曲线重建和峰值强度预测方面表现较好，其 UTS 应力 MAE 和 RMSE 分别为 1.91 MPa 和 2.54 MPa，均低于其他模型；但其 UTS 应变误差高于 TCN 和 Transformer，说明峰值位置对局部曲线形态更敏感。原因在于峰值附近应力曲线趋于平坦，微小的应力预测波动即可能导致峰值位置偏移，因此完整曲线误差较低并不必然对应 UTS 应变预测最优。

（四）SHAP 可解释性分析

各特征的平均绝对 SHAP 贡献值如表4所示。Gradient SHAP 结果显示，UTS 应力预测中因果平均应变速率贡献最大，瞬时应变速率次之，说明模型判断峰值强度时不仅关注当前加载速率，也利用了加载历史中的累计速率信息。UTS 应变预测中瞬时应变速率贡献略高，表明峰值位置对局部加载状态更敏感。该结果与 Q195 钢随应变速率升高而强度增加、峰值应变略有变化的实验趋势一致，也与特征消融结果相互印证。考虑到样本数仅为 9 组，SHAP 结果宜作为趋势性解释，不能替代独立泛化验证。

表 4 各输入特征的平均绝对 SHAP 贡献值

输入特征	UTS 应力 SHAP	UTS 应变 SHAP
因果平均应变率	6.35	0.0148
瞬时应变率	4.14	0.0153
应变	0.68	0.0125

三、结论

本文基于9组 Q195钢室温单向拉伸实验数据，提出 TCN-Transformer序列回归模型，并结合 nested LOOCV、特征消融和 Gradient SHAP进行了研究。主要结论如下：

（1）基于训练集 UTS 应变阈值的截断策略与三通道特征构造，在小样本条件下为序列模型提供了稳定的数据预处理流程，各模型均能在此基础上完成整条应力-应变曲线预测。

（2）深度学习序列模型在整体曲线重建和 UTS 应力预测上均优于物理基线（J-C）和经典机器学习基线（SVR）；其中 TCN 凭借局部卷积编码在此任务中已构成强基线，TCN-Transformer 通过全局自注意力进一步降低了大偏差样本的预测波动。

（3）特征消融和 Gradient SHAP 分别从曲线整体误差和 UTS 点贡献两个角度，共同表明瞬时应变速率和因果平均应变速率是当前建模中的关键输入；二者互补的解释路径提升了特征重要性结论的可靠性。

本文结论基于 Q195钢、室温、 $1.00 \times 10^{-3} \sim 1.70 \times 10^{-1} \text{ s}^{-1}$ 应变率范围内的单变量加载条件得出，所提流程对同类小样本应力-应变曲线建模具有参考价值。后续可围绕温度-应变率耦合工况、物理先验嵌入及模型不确定性量化开展进一步研究。

参考文献

[1]Johnson G R, Cook W H. A constitutive model and data for metals subjected to large strains, high strain rates and high temperatures[C]//Proc. 7th Int. Symp. Ballistics. 1983: 541-547.

[2]岳峰丽, 邵扬, 陈大勇, 等. 基于不同机器学习模型的 TP2铜材流动应力-应变曲线预测 [J]. 锻压技术, 2026, 51(2): 1-11.

[3]Desu R K, Guntuku S C, Aditya B, et al. Support vector regression based flow stress prediction in 304 steel[J]. Procedia Materials Science, 2014, 6: 368-375.

[4]Song S H. Constitutive equation, neural networks and SVR for hot deformation of 316L stainless steel[J]. Materials, 2020, 13(17): 3763.

[5]Bai S, Kolter J Z, Koltun V. An empirical evaluation of generic convolutional and recurrent networks for sequence modeling[EB/OL]. arXiv:1803.01271, 2018.

[6]Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need[C]//NeurIPS. 2017, 30: 5998-6008.

[7]Cawley G C, Talbot N L C. On over-fitting in model selection and selection bias in performance evaluation[J]. JMLR, 2010, 11: 2079-2107.

[8]GB/T 700—2019. 碳素结构钢 [S]. 北京：中国标准出版社，2019.

[9]GB/T 228.1—2021. 金属材料 拉伸试验 第1部分：室温试验方法 [S]. 北京：中国标准出版社，2021.

[10]Lundberg S M, Lee S I. A unified approach to interpreting model predictions[C]//NeurIPS. 2017, 30: 4765-4774.