基于随机森林算法与数据增强策略双驱动的多尺度疲劳裂纹扩展预测

刁圣轩 肖金雍 陈永保 单康中 杨 杰

(上海理工大学 能源与动力工程学院 上海市动力工程多相流动与传热重点实验室 上海 200093)

摘 要 为了实现多尺度疲劳裂纹扩展过程的准确预测，并增强预测模型的可解释性和泛化能力，本工作选用304奥氏体不锈钢为研究对象，首先对K近邻(KNN)、支持向量机回归(SVR)和随机森林(RF)三种算法的疲劳裂纹扩展预测能力进行了对比，筛选出预测性能最优的算法。随后，选择最优算法，分别基于长裂纹、短裂纹和多尺度裂纹三种疲劳裂纹扩展模型添加数据增强策略进行数据增强，以提高疲劳裂纹扩展预测的准确性。最后，基于添加数据增强策略的双驱动模型框架，在不同载荷下进行多尺度疲劳裂纹扩展预测。结果表明，与SVR和KNN算法相比，RF算法更适合于疲劳裂纹扩展预测，其基于多尺度疲劳裂纹扩展模型的数据增强效果明显优于其余两种模型。基于RF算法与数据增强策略的双驱动模型具有很好的泛化能力，且相比于基于RF算法的纯数据驱动模型，双驱动模型预测结果更加准确。

关键词 多尺度疲劳裂纹扩展，随机森林算法，数据增强策略，双驱动模型

中图分类号 TG111.8 **文章编号** 0412-1961(2026)09-1606-09

Multi-Scale Fatigue Crack Propagation Prediction Based on the Dual Drive of Random Forest Algorithm and Data Augmentation Strategy

DIAO Shengxuan, XIAO Jinyong, CHEN Yongbao, SHAN Kangzhong, YANG Jie

Shanghai Key Laboratory of Multiphase Flow and Heat Transfer in Power Engineering, School of Energy and Power Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China

Correspondent: YANG Jie, professor, Tel: (021)55272320, E-mail: yangjie@usst.edu.cn

Supported by National Natural Science Foundation of China (Nos.52375154 and 52311530067)

Manuscript received 2024−08−28, in revised form 2025−10−24

ABSTRACT Data-driven methods based on machine learning have been employed to predict fatigue crack propagation. However, existing studies have largely overlooked the multi-scale nature of this process. Relying solely on macroscopic data for long-crack prediction often fails to capture the complete crack growth process, potentially resulting in non-conservative predictions. Moreover, purely data-driven models often lack interpretability, exhibit limited generalization capabilities, and struggle to adhere to physical laws. The challenge of integrating modeling with reasonable interpretations, driven by both data and mechanisms, remains a significant issue for researchers. In this study, we selected 304 austenitic

资助项目 国家自然科学基金项目Nos.52375154和52311530067

收稿日期 2024-08-28 **定稿日期** 2025-10-24

作者简介 刁圣轩，男，2001年生，硕士生

通讯作者 杨 杰，yangjie@usst.edu.cn，主要从事重要装备寿命及可靠性等研究

引用格式 刁圣轩, 肖金雍, 陈永保等. 基于随机森林算法与数据增强策略双驱动的多尺度疲劳裂纹扩展预测 [J]. 金属学报, 2026, 62: 1606−1614.

Diao S X, Xiao J Y, Chen Y B, et al. Multi-scale fatigue crack propagation prediction based on the dual drive of random forest algorithm and data augmentation strategy [J]. Acta Metall. Sin., 2026, 62: 1606−1614. DOI: 10.11900/0412.1961.2024.00302

stainless steel as the research object. To identify the algorithm with the best predictive performance, firstly, the fatigue crack propagation prediction capabilities of three algorithms were compared: K-nearest neighbor (KNN), support vector machine regression (SVR), and random forest (RF). The most effective algorithm and implemented data augmentation strategies based on three fatigue crack propagation models were selected, namely, long cracks, short cracks, and multi-scale cracks, to enhance prediction accuracy. Finally, using a dual-drive model framework that incorporated the data augmentation strategy, predictions of multi-scale fatigue crack propagation under different loads were conducted. Compared to the SVR and KNN algorithms, the results indicated that the RF algorithm had a lower root mean square error (RMSE) value and higher coefficient of determination (R^2), making it more suitable for predicting fatigue crack propagation. Under a load of 370 MPa, the prediction accuracy for the training set ranked in the order of RF > KNN > SVR. By contrast, the accuracy for the test set was in the order of RF > SVR > KNN. The multi-scale fatigue crack propagation model effectively captured the entire process of crack growth, whereas the long- and short-crack models accurately represented only parts of it. Data enhancement based on the multi-scale crack model demonstrated significantly better results than the other two models, with increases in accuracy of 25.76% and 71.74% for the training and test sets, respectively. The dual-drive model based on the RF algorithm and data enhancement strategy exhibited strong generalization capabilities. Under a load of 350 MPa, the RMSE values for the training and test sets were 0.046 and 0.111, respectively, and R^2 reached 0.995 and 0.961. Under a load of 330 MPa, the RMSE values for the training and test sets were 0.171 and 0.081, respectively, and the R^2 reached 0.911 and 0.721. Finally, compared with the pure data-driven model based on the RF algorithm, the predictions from the dual-drive model were found to be significantly more accurate.

KEYWORDS multi-scale fatigue crack propagation, random forest algorithm, data augmentation strategy, dual drive model

疲劳是影响结构寿命与可靠性的重要因素。据统计，80%以上的工程结构破坏与疲劳有关[1~3]。为了实现长寿命、高可靠性的结构设计，需要对疲劳问题进行深入研究。

自20世纪50年代以来，受益于线弹性断裂力学的发展和实验技术的巨大进步，疲劳裂纹扩展的力学模型、疲劳裂纹扩展速率表达式、疲劳寿命的预测模型先后建立。Irwin[4]于1957年提出使用应力强度因子(K)来表征线弹性裂尖力学场，为后续疲劳裂纹扩展模型的发展奠定了理论基础。在此基础上，Paris和Erdogan[5]总结出了描述疲劳裂纹扩展速率的幂指数定律，即Paris公式，为后续预测疲劳裂纹扩展寿命提供了一种新的估算方法。为了更精确地描述疲劳裂纹扩展行为，Forman等[6]考虑了平均应力对裂纹扩展速率($\mathrm{d}a/\mathrm{d}N$，其中，a为裂纹扩展长度，N为循环周次)及应力强度因子范围(ΔK)的影响，通过修正Paris公式提出了Forman公式。Forman公式和Paris公式已经被广泛应用于宏观疲劳裂纹扩展之中。然而，实际的疲劳裂纹扩展过程是一个从微观到宏观的多尺度发展过程，Forman公式和Paris公式并不适用于描述短裂纹疲劳扩展规律[7,8]。为此，Deng等[9]提出了一个描述短裂纹疲劳扩展规律的经验公式。Ye等[10]探究了微观组织小裂纹(microstructurally small crack，MSC)、介观物理小裂纹(physically small crack，PSC)和宏观长裂纹(long crack，LC)阶段的疲劳裂纹扩展规律，并提出了一个多尺度疲劳裂纹扩展模型。该模型基于微观结构非相似性原理，认为微观结构非相似性主导了裂纹在MSC阶段的扩展行为，而在PSC和LC阶段，线弹性断裂力学主导了裂纹扩展行为。Ye等[10]还提出了通过局部屈服强度控制不同尺度下裂纹尖端塑性区内的局部应力分布以及材料的塑性变形能力和微观结构等，用转变参数控制不同工况下MSC阶段到PSC/LC阶段的转变速率。

上述研究使用的方法主要基于经验科学、理论科学和计算科学，如今，科学研究已迈入数据驱动科学范式[11,12]，以机器学习为主的数据驱动方法正在被广泛应用于各个领域[13,14]。在疲劳裂纹扩展预测领域，Wang等[15]利用径向基函数网络和遗传算法优化反向传播网络(genetic algorithm optimized backpropagation，GABP)对2024-T451铝合金的疲劳裂纹扩展进行了预测，其预测的准确性和有效性均优于经典的双参数模型(K方法)。Younis等[16]和Zafar等[17]基于多种神经网络方法预测了多种铝合金的疲劳裂纹扩展行为，其预测结果与实验数据吻合良好。Mortazavi和Ince[18]开发了一种径向基函

数人工神经网络(radial basis function artificial neural network，RBF-ANN)模型，用以预测Ti-6Al-4V钛合金等的疲劳裂纹扩展行为，模型展现出了良好的预测精度。Zhang等[19]提出了一种基于机器学习的新数据驱动方法，使用反向传播神经网络预测了疲劳裂纹扩展速率与有效应力强度因子的关系以及表面裂纹试样的疲劳寿命。结果表明，该方法能够准确且快速地获取材料数据和预测疲劳寿命。Barbosa等[20]基于人工神经网络提出了一种生成金属恒定寿命图(constant life diagram，CLD)的新方法，可以根据平均应力和应力幅度，在拉伸载荷占主导地位的区域中估计高周疲劳状态的安全区域。Gan等[21]提出了两种基于机器学习的模型用以作为半经验模型的替代方案，以提高平均应力效应的预测精度。Yang等[22]提出了一种基于深度学习的方法以预测多轴疲劳寿命，打破了传统模型局限于特定材料和载荷条件的限制。目前，已有较多关于机器学习在疲劳裂纹扩展预测中的研究，但其均未考虑疲劳裂纹的多尺度裂纹扩展。如前文所述，实际的疲劳裂纹扩展过程是一个多尺度发展过程，若仅使用宏观长裂纹的扩展数据进行预测，将无法全面反映整个裂纹扩展过程，甚至可能会得到非保守的预测结果，因此需要在预测中纳入多尺度裂纹的影响。

此外，虽然数据驱动方法正在被广泛应用，但囿于决策的透明度和结果的可解释性，该方法目前难以部署到一些性能要求高的关键领域。最近，部分学者尝试将数据驱动与机理驱动相结合，通过增加边界条件对模型参数进行约束，以增强数据驱动可解释性、提高泛化能力，进而得到符合物理规律的预测结果。Liu等[23]将局部屈曲理论和弹性屈曲理论引入到人工神经网络模型中，显著提升了在轴向压缩圆柱壳屈曲设计中屈曲载荷的预测效果。Zhang等[24]通过物理信息引导的特征扩展，提高了机器学习模型在蠕变疲劳寿命方面的预测精度。Karpatne等[25]论述了利用物理机理提高数据驱动模型在解决复杂科学问题时的准确性和可靠性。为了解决噪声处理、复杂网格生成和高维参数化等成本密集型问题，Kevrekidis等[26]提出了一种物理信息机器学习模型，该模型融合了噪声数据与数学模型，并通过神经网络或基于核的回归网络实现，达到了简化流程和提高准确性的目的。Yang等[27]针对尼龙-6 (polyamide-6，PA6)与速率有关的多轴疲劳寿命问题，在半经验寿命预测模型的基础上加入了临界面方法以及两个物理约束条件来引导神经网络的训练，改进后混合模型的泛化性能得到了显著提高。Jiang等[28]提出了一种基于物理信息的多层嵌套神经网络框架，用于预测316不锈钢在不同温度和应变率下的低周疲劳寿命。该框架将物理约束集成到多层嵌套神经网络中，以引入物理约束来提高机器学习方法的预测精度。结果表明，所有疲劳寿命预测结果的误差均在2倍以内，大部分在1.5倍以内。进一步，针对316不锈钢在高温高压水环境中的疲劳寿命预测难题，Jiang等[29]提出了一种基于条件生成对抗网络(conditional generative adversarial networks，cGAN)的新模型，该模型以应变幅值、应变速率、温度、溶解氧含量和层错能为条件，实现了环境疲劳寿命的预测与概率评估。然而，目前结合机理驱动的方法仍处于探索阶段，且对边界条件的要求较为苛刻，数据机理双驱动下的融合建模和合理解释仍然是学者们面临的一个难题。

鉴于此，本工作选用304奥氏体不锈钢为研究对象，首先对K近邻回归(KNN)、支持向量机回归(SVR)和随机森林回归(RF)三种算法的疲劳裂纹扩展预测能力进行对比，筛选出预测性能最优的算法。随后，选择该算法，分别基于长裂纹、短裂纹和多尺度裂纹三种疲劳裂纹扩展模型添加数据进行数据增强，以提高疲劳裂纹扩展预测的准确性。最后，基于添加数据增强策略的双驱动模型框架，在不同载荷下进行了多尺度疲劳裂纹扩展预测。与现有研究相比，本工作同样基于数据驱动与机理驱动相结合的思想，但采用了数据增强策略，并聚焦于多尺度疲劳裂纹扩展的预测。

1 数据来源、机器学习算法及数据增强

1.1 数据来源及其训练

文献[9]以304不锈钢为研究对象，采用单边缺口拉伸(single edge notch tension，SENT)试样，得到了当应力比$R=0.1$、载荷分别为330、350和370 MPa时，该材料的多尺度疲劳裂纹扩展实验数据，如图1[9]所示。根据文献[9,10]，MSC阶段向PSC阶段转变时，304不锈钢的裂纹长度为0.2 mm；PSC阶段向LC阶段转变时，裂纹长度为1 mm。本工作以上述数据为基础，进行训练与测试，流程如图2所示。

针对不同载荷条件，随机选取图1[9]中80%的实验数据作为训练集，并将另外20%的数据作为测试集，用以检验模型的预测精度。在模型训练阶段，分别采用KNN、SVR和RF三种机器学习算法；训练结束后，采用训练好的模型对测试集进行回归预测，并通过与实验数据的对比来评估每种机器学习算法预测疲劳裂纹扩展的准确性。

1.2 数据预处理

为了确保模型训练的数值稳定性并促进模型快速收敛，采用 Z 得分归一化(Z-score normalization)法对输入特征进行标准化处理，以消除特征尺度和分布的差异，如式(1)所示[30]。数据预处理可以限制特征的尺度，使其具有 0 均值和单位方差，有助于提升模型的预测性能和泛化能力。

$$z = \frac{x - \mu}{\sigma} \tag{1}$$

式中，z 为标准化后的特征值，μ 为所有数据的均值，σ 为所有数据的标准差，x 为原始数据点。变化后的数据均值为 0，标准差为 1，确保所有数据处于统一尺度，便于基于相应算法进行学习和预测。

1.3 机器学习算法

本工作采用 KNN、SVR 和 RF 三种机器学习算法。其中，KNN 算法是一种监督学习算法，主要用于回归预测[31]。该算法假设新输入数据对应的输出值在各个样本对应的输出值空间中相似，最相似的 K 个数据点及其各个特征具有相同权重。故输入新数据后，提取样本中与新数据最相似的 K 个数据，通过取 K 个数据对应输出值的均值作为新数据的输出值。KNN 算法通常采用 Euclidean 距离对数据点和新数据点进行距离度量，新数据点 $x_0 = [x_{0,1}, x_{0,2}, \cdots, x_{0,j}, \cdots, x_{0,n}]$ 和样本点 $x_i = [x_{i,1}, x_{i,2}, \cdots, x_{i,j}, \cdots, x_{i,n}]$ 之间的距离($d(x_0, x_i)$)为[31]：

$$d(x_0, x_i) = \left[\sum_{j=1}^{n}\left(x_{i,j} - x_{0,j}\right)^2\right]^{1/2} \tag{2}$$

SVR 算法也是一种监督学习算法，用于分析回归问题[32]。该算法的目标是在特征空间中找到一个超平面，使得不同类别的数据点被有效分开，同时最大化这个超平面与最近的数据点(即支持向量)之间的间隔，从而增强模型的鲁棒性和泛化能力。SVR 算法形式为[33]：

$$f(\boldsymbol{x}) = <\boldsymbol{\omega}, \boldsymbol{x}> + b \tag{3}$$

式中，$\boldsymbol{x}$ 为输入的特征向量，$\boldsymbol{\omega}$ 为权重向量，b 为偏置项。目标是在误差容限内找到最优的 $\boldsymbol{\omega}$ 和 b，以最小化模型的损失函数[33]：

$$J(\boldsymbol{\omega}, \xi_i, \xi_i^*) = \min_{\boldsymbol{\omega}, \xi_i, \xi_i^*}\left[\frac{1}{2}\|\boldsymbol{\omega}\|^2 + \lambda\sum_{i=1}^{N'}\left(\xi_i + \xi_i^*\right)\right] \tag{4}$$

式中，ξ_i 和 ξ_i^* 为误差项，λ 为正则化参数，N' 为样本数量。

RF 算法是一种集成机器学习算法，通过构建多个决策树并综合它们的预测结果以提高整体的预测性能[34]。其中，每棵决策树的分割节点数目均根据样本特征数随机确定，这种随机性也使得集合决策树具备更优的预测性能。RF 算法模型表示为[34]：

$$R_{\mathrm{F}} = \max_{i,K}\prod_{k=1}^{K}\eta\left(C_k, S_i\right) \tag{5}$$

式中，R_{F} 为预测输出，S_i 为第 i 个样本集，C_k 为第 k 棵决策树，$\eta(C_k, S_i)$ 为 C_k 对 S_i 的决策值，$\max_{i,K}$ 为投票机制的函数式，K 为决策树数量。

1.4 模型评价标准

在评价标准方面，采用均方根误差(root mean

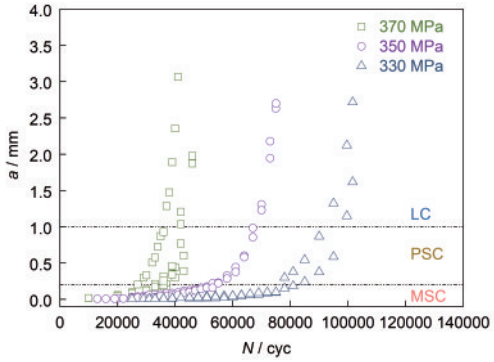


图 1 不同应力下 304 不锈钢的多尺度疲劳裂纹扩展实验数据[9]

Fig.1 Multi-scale fatigue crack propagation test data of 304 stainless steel under different stress levels[9] (a—crack length, N—number of cycles; LC—long crack, PSC—physically small crack, MSC—microstructurally small crack)

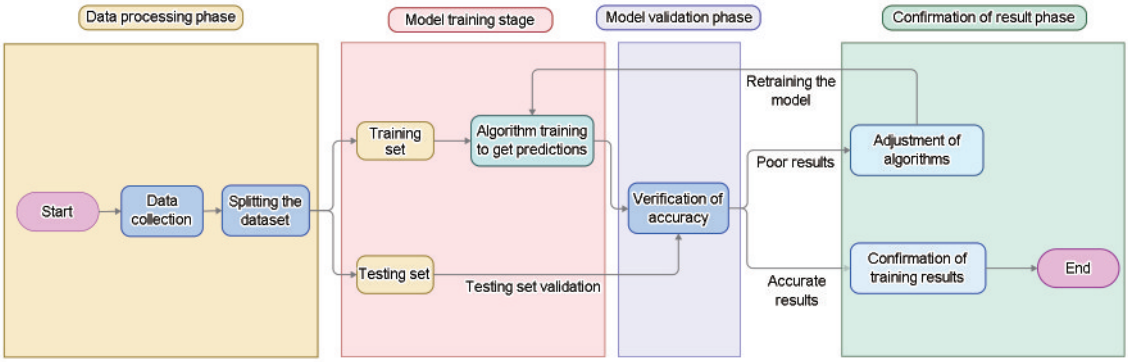


图 2 数据的训练与测试流程

Fig.2 Processes of data training and testing

square error，RMSE)及决定系数(coefficient of determination，R^2)量化模型的预测精度。采用RMSE度量预测值与实际观测值之间的差异，如式(6)所示[35]，RMSE越小则预测结果越准确。R^2用于衡量模型对数据的拟合程度，如式(7)所示[35]，R^2越接近1表明预测结果与实验结果吻合度越高。二者相互补充，可以更加全面地反映模型的预测性能。

$$\mathrm{RMSE}=\sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i-\hat{y}_i)^2} \tag{6}$$

$$R^2=1-\frac{\sum_{i=1}^{n}(y_i-\hat{y}_i)^2}{\sum_{i=1}^{n}(y_i-\bar{y})^2} \tag{7}$$

式中，y_i为第i个样本的实测值，$\hat{y}_i$为模型预测值，$\bar{y}$为实际值的平均值，n为训练集的数据大小。

1.5 基于不同裂纹扩展模型的数据增强

为了增强数据的可解释性、提高泛化能力以及使预测结果符合物理规律，分别选择长裂纹、短裂纹和多尺度裂纹三种疲劳裂纹扩展模型，将其作为边界条件对模型参数进行约束，并通过在训练集中增加满足疲劳裂纹扩展模型的数据进行数据增强，如图3所示。其中长裂纹扩展模型为Paris模型[5]，其表达式如下：

$$\frac{\mathrm{d}a}{\mathrm{d}N}=C(\Delta K)^m \tag{8}$$

式中，C、m为材料参数。

短裂纹扩展模型由Deng等[9]提出，其定义如式(9)所示：

$$\frac{\mathrm{d}a}{\mathrm{d}N}=A\times a^B \tag{9}$$

式中，A、B为拟合参数。

多尺度裂纹扩展模型由Ye等[10]提出，其表达式如下：

$$\frac{\mathrm{d}a}{\mathrm{d}N}=C(\Delta K\cdot q-\Delta K_{\mathrm{ths}})^m \tag{10}$$

式中，q为修正参数，ΔK_{ths}为应力强度因子范围的短裂纹阈值。三种模型中的各参数如表1[36]所示。数据增强时，分别基于长裂纹、短裂纹和多尺度裂纹三种疲劳裂纹扩展模型的拟合曲线增加数据。从$a=0.05$ mm开始，在MSC阶段，每隔0.025 mm增加1个数据点；在PSC阶段，每隔0.15 mm增加1个数据点；在LC阶段，每隔0.5 mm增加1个数据点，从而生成符合三种疲劳裂纹扩展模型的数据增强集。数据增强集数量约为实验数据集数量的30%，如表2所示。

2 结果与讨论

2.1 不同算法对比

首先选用370 MPa下的数据作为训练集和测试集，基于三种算法的训练集的训练结果和测试集预测结果如图4所示(a_{pre}为裂纹长度预测值，a_{exp}为裂纹长度实测值)。由图可见，与SVR和KNN算法相比，RF算法训练结果和测试结果的预测精度均为最佳，这表明与SVR和KNN算法相比，RF算法更适合疲劳裂纹扩展预测。进一步，通过对比三种算法的RMSE和R^2 (图5)可知，相较于其他两种算法，RF算法的RMSE更低，R^2更高，即RF算法的预测更准确。从训练集训练结果来看，三种算法的精度排序为

表1 三种疲劳裂纹扩展模型中的参数[36]

Table 1 Parameters in the three fatigue crack propagation models[36]

Model	Parameter	Value
Long crack	C	1.615×10^{-10}
	m	2.58
Short crack	A	$1.000\times10^{-10.49}$
	B	1.19
Multi-scale crack	C	3.630×10^{-10}
	q	1.37

Note: C, m—material constants; A, B—fitting constants; q—modified parameter

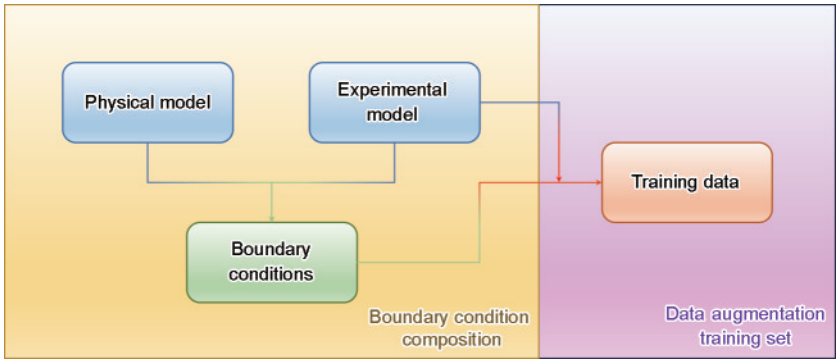


图3 数据增强流程

Fig.3 Process of data augmentation

RF > KNN > SVR;从测试集预测结果来看,三种算法的精度为RF > SVR > KNN。鉴于此,本工作将选用RF算法进行后续研究。

2.2 不同疲劳裂纹扩展模型数据增强效果对比

针对370 MPa下的数据,在RF算法的基础上,分别基于长裂纹、短裂纹和多尺度裂纹三种疲劳裂纹扩展模型进行数据增强,数据增强后训练集的训练结果和测试集预测结果如图6所示。同时,数据增强前(即采用纯RF算法)的训练结果和预测结果也在图中显示。由图6可见,综合考虑MSC、PSC和LC阶段的多尺度裂纹扩展模型的数据增强效果优于其余两种模型,数据增强后,其训练集训练结果精度提升了25.76%,测试集预测结果精度提升了71.74%。

基于三种模型数据增强后的RMSE及R^2如图7所示。可以看出,经多尺度裂纹扩展模型数据增强后,RMSE明显降低、R^2显著提高。这说明多尺度疲劳裂纹扩展模型完整地反映了裂纹扩展的全过程,

表2 数据集数量

Table 2 Numbers of datasets

Load / MPa	Number of experimental dataset	Number of data augmentation dataset
330	61	18
350	50	15
370	45	14

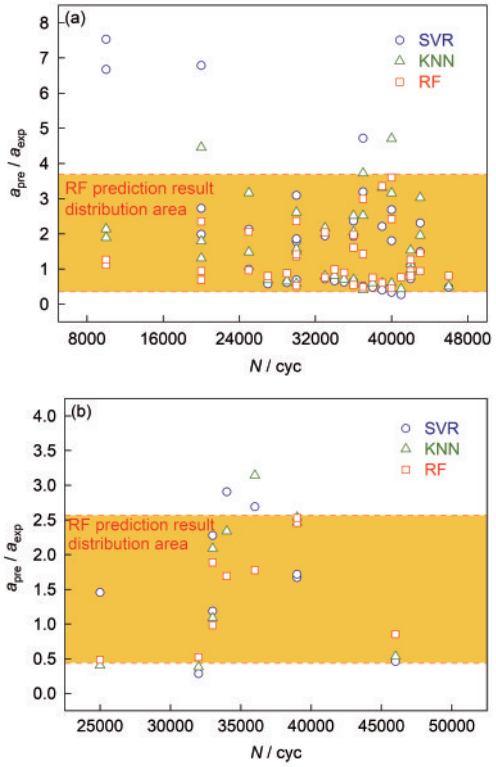


图4 三种算法训练集训练结果和测试集预测结果对比

Fig.4 Comparisons of training results of training sets (a) and prediction results of testing sets (b) for three algorithms (SVR—support vector machine regression, KNN—nearest neighbor regression, RF—random forest regression, a_{pre}—predicted crack length, a_{exp}—experimental crack length)

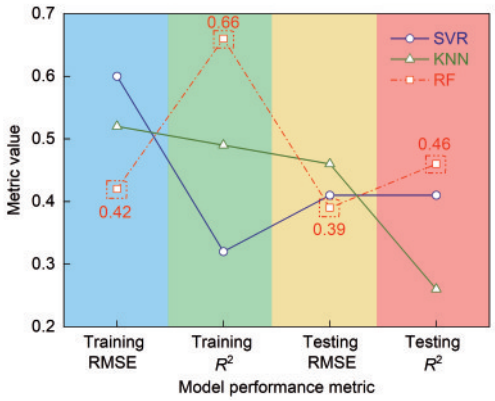


图5 三种算法的均方根误差(RMSE)及决定系数(R^2)对比

Fig.5 Comparisons of root mean square errors (RMSE) and coefficient of determinations (R^2) among the three algorithms

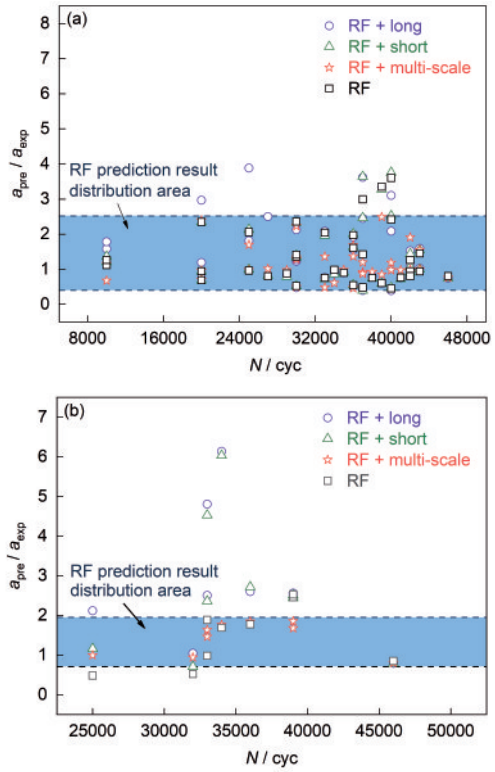


图6 三种疲劳裂纹扩展模型对训练集和测试集的数据增强效果

Fig.6 Data augmentation effects of three fatigue crack propagation models on training results of training sets (a) and prediction results of test sets (b)

而长裂纹扩展模型和短裂纹扩展模型仅能准确反映裂纹扩展的部分过程，其预测精度反而较数据增强前有所降低。

2.3 不同载荷下多尺度疲劳裂纹扩展预测

沿用370 MPa载荷条件下的数据增强策略，对350和330 MPa条件下的多尺度疲劳裂纹扩展进行预测，结果如图8和9所示。由图可见，在350及330 MPa载荷条件下，预测结果与实验数据之间的误差分别处于±0.25与±0.50误差带内，这说明基于RF算法与数据增强策略的双驱动模型具有很好的泛化能力。且与基于RF算法的纯数据驱动模型相比，双驱动模型的预测结果明显更加准确。

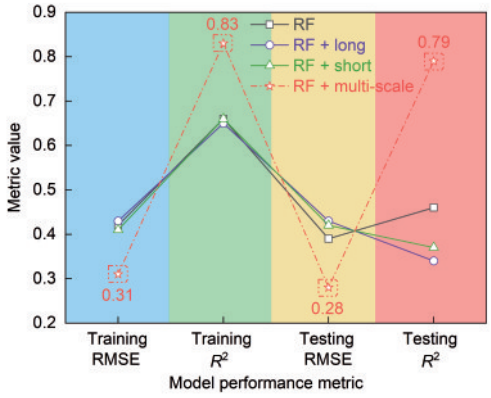


图7 基于三种模型数据增强后RMSE及R^2对比

Fig.7 Comparison of RMSE and R^2 after data augmentation based on the three models

350和330 MPa载荷条件下，RMSE与R^2如图10所示。可以看出，数据增强后，350 MPa下基于多尺度裂纹扩展模型的训练集与测试集的RMSE分别为0.046与0.111，R^2达到了0.995与0.961；330 MPa下训练集与测试集的RMSE分别为0.171与0.081，R^2达到了0.911与0.721，均优于纯数据驱动模型。在MSC阶段的预测结果优于PSC阶段与LC阶段。

在实际工程中，由于各种条件限制，无法得到大量的实验数据，也很难获得不同尺度下的疲劳裂纹扩展结果。本工作通过数据增强方法增加了实验数据数量，并基于多尺度疲劳裂纹扩展模型进行预测，获得了较为准确的不同尺度下的疲劳裂纹扩展数据，有效地解决了这一难题，为实际工程应用提供了便利。需要说明的是，虽然本工作以304奥氏体不锈钢为研究对象，但所提出的方法具备一定的普适性，可适用于其他金属材料。

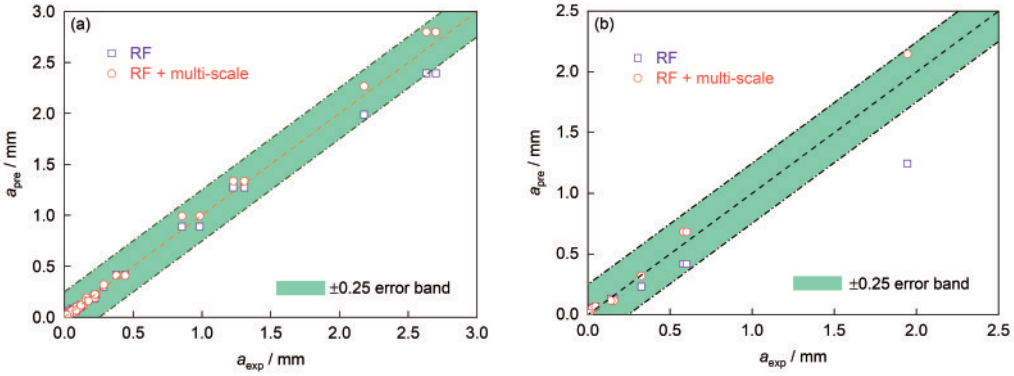


图8 350 MPa下基于双驱动模型和随机森林(RF)算法模型的多尺度疲劳裂纹扩展预测结果

Fig.8 Prediction results of multi-scale fatigue crack propagation at 350 MPa based on dual-drive model and RF model (a) training results of training sets (b) prediction results of testing sets

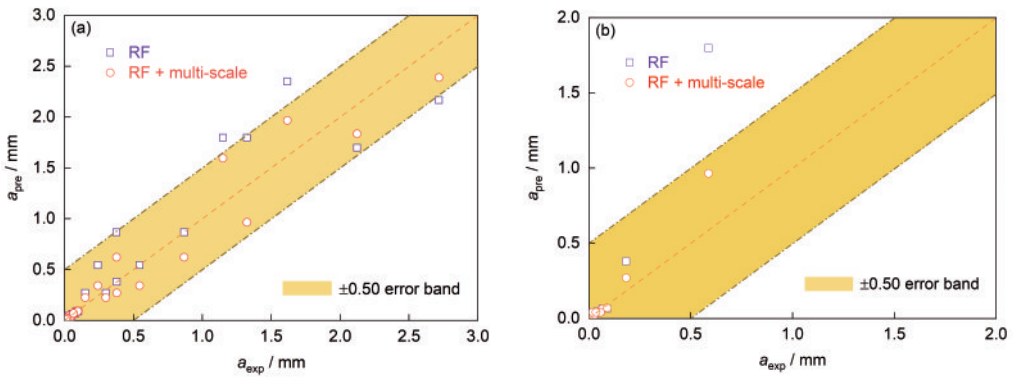


图9 330 MPa下基于双驱动模型和RF算法模型的多尺度疲劳裂纹扩展预测结果

Fig.9 Prediction results of multi-scale fatigue crack propagation at 330 MPa based on dual-drive model and RF model (a) training results of training sets (b) prediction results of testing sets

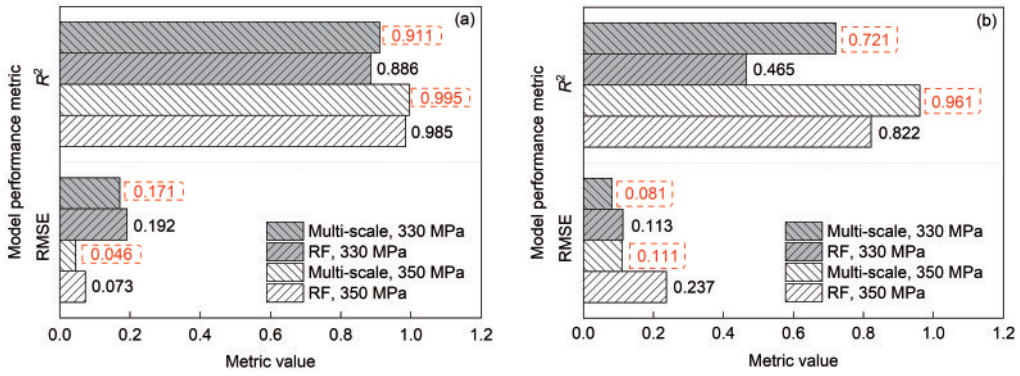


图10 350和330 MPa下基于双驱动模型和RF算法模型的RMSE及R^2

Fig.10 RMSE and R^2 at 350 and 330 MPa based on dual-drive model and RF model

(a) training results of training sets　　(b) prediction results of testing sets

3 结论

(1) 与SVR和KNN算法相比，RF算法的RMSE更低，R^2更高，更适合疲劳裂纹扩展预测。370 MPa载荷条件下，从训练集训练结果来看，三种算法的精度为RF > KNN > SVR；从测试集预测结果来看，三种算法的精度为RF > SVR > KNN。

(2) 多尺度疲劳裂纹扩展模型完整地反映了裂纹扩展全过程，长裂纹和短裂纹扩展模型仅能准确反映裂纹扩展的部分过程，基于多尺度疲劳裂纹扩展模型的数据增强效果明显优于其余两种模型。数据增强后，其训练集训练结果精度提升了25.76%，测试集预测结果精度提升了71.74%。

(3) 基于RF算法与数据增强策略的双驱动模型具有很好的泛化能力，350 MPa下基于多尺度裂纹扩展模型的训练集与测试集的RMSE分别为0.046与0.111，R^2分别达到了0.995与0.961；330 MPa下训练集与测试集的RMSE分别为0.171与0.081，R^2分别达到了0.911与0.721。此外，与基于RF算法的纯数据驱动模型相比，双驱动模型预测结果明显更加准确。

参 考 文 献

[1] Xu W, Huang M H, Wang J L, et al. Review: Relations between metastable austenite and fatigue behavior of steels [J]. Acta Metall. Sin., 2020, 56: 459
(徐 伟, 黄明浩, 王金亮等. 综述: 钢中亚稳奥氏体组织与疲劳性能关系 [J]. 金属学报, 2020, 56: 459)

[2] Wei L S, Liu H L, Tang C G, et al. Investigation of novel type of cylindrical lithium-ion battery heat exchangers based on topology optimization [J]. Energy, 2024, 304: 131886

[3] Wang H J, Li B, Gong J G, et al. Machine learning-based fatigue life prediction of metal materials: Perspectives of physics-informed and data-driven hybrid methods [J]. Eng. Fract. Mech., 2023, 284: 109242

[4] Irwin G R. Analysis of stresses and strains near the end of a crack traversing a plate [J]. J. Appl. Mech., 1957, 24: 361

[5] Paris P, Erdogan F. A critical analysis of crack propagation laws [J]. J. Basic Eng., 1963, 85: 528

[6] Forman R G, Kearney V E, Engle R M. Numerical analysis of crack propagation in cyclic-loaded structures [J]. J. Basic Eng., 1967, 89: 459

[7] Pearson S. Initiation of fatigue cracks in commercial aluminium alloys and the subsequent propagation of very short cracks [J]. Eng. Frac. Mech., 1975, 7: 235

[8] Okazaki M, Yada T, Endoh T. Surface small crack growth behavior in low-cycle fatigue at elevated temperature and application limit of macroscopic crack growth law [J]. Nucl. Eng. Des., 1989, 111: 123

[9] Deng G J, Tu S T, Wang Q Q, et al. Small fatigue crack growth mechanisms of 304 stainless steel under different stress levels [J]. Int. J. Fatigue, 2014, 64: 14

[10] Ye S, Zhang C C, Zhang P Y, et al. Fatigue life prediction of nickel-based GH4169 alloy on the basis of a multi-scale crack propagation approach [J]. Eng. Fract. Mech., 2018, 199: 29

[11] Agrawal A, Choudhary A. Perspective: Materials informatics and big data: Realization of the “fourth paradigm” of science in materials science [J]. APL Mater., 2016, 4: 053208

[12] Schleder G R, Padilha A C M, Acosta C M, et al. From DFT to machine learning: Recent approaches to materials science—A review [J]. J. Phys.: Mater., 2019, 2: 032001

[13] Zhao S K, Jiang C, Long X Y. Remaining useful life estimation of mechanical systems based on the data-driven method and bayesian theory [J]. J. Mech. Eng., 2018, 54(12): 115
(赵申坤, 姜 潮, 龙湘云. 一种基于数据驱动和贝叶斯理论的机械系统剩余寿命预测方法 [J]. 机械工程学报, 2018, 54(12): 115)

[14] Liu Y L, Niu C, Wang Z, et al. Machine learning in materials genome initiative: A review [J]. J. Mater. Sci. Technol., 2020, 57: 113

[15] Wang H X, Zhang W F, Sun F Q, et al. A comparison study of machine learning based algorithms for fatigue crack growth calculation [J]. Materials, 2017, 10: 543

[16] Younis H B, Kamal K, Sheikh M F, et al. Prediction of fatigue crack growth rate in aircraft aluminum alloys using optimized neural networks [J]. Theor. Appl. Fract. Mech., 2022, 117: 103196

[17] Zafar M H, Younis H B, Mansoor M, et al. Training deep neural networks with novel metaheuristic algorithms for fatigue crack

growth prediction in aluminum aircraft alloys [J]. Materials, 2022, 15: 6198

[18] Mortazavi S N S, Ince A. An artificial neural network modeling approach for short and long fatigue crack propagation [J]. Comput. Mater. Sci., 2020, 185: 109962

[19] Zhang J Q, Zhu J C, Guo W, et al. A machine learning-based approach to predict the fatigue life of three-dimensional cracked specimens [J]. Int. J. Fatigue, 2022, 159: 106808

[20] Barbosa J F, Correia J A F O, Júnior R C S F, et al. Fatigue life prediction of metallic materials considering mean stress effects by means of an artificial neural network [J]. Int. J. Fatigue, 2020, 135: 105527

[21] Gan L, Wu H, Zhong Z. Fatigue life prediction considering mean stress effect based on random forests and kernel extreme learning machine [J]. Int. J. Fatigue, 2022, 158: 106761

[22] Yang J Y, Kang G Z, Liu Y J, et al. A novel method of multiaxial fatigue life prediction based on deep learning [J]. Int. J. Fatigue, 2021, 151: 106356

[23] Liu F, Chen H F, Yang J, et al. Application of physics-informed machine learning methods in buckling design of axially compressed cylindrical shells [J]. Thin-Walled Struct., 2024, 200: 111963

[24] Zhang X C, Gong J G, Xuan F Z. A physics-informed neural network for creep-fatigue life prediction of components at elevated temperatures [J]. Eng. Fract. Mech., 2021, 258: 108130

[25] Karpatne A, Atluri G, Faghmous J H, et al. Theory-guided data science: A new paradigm for scientific discovery from data [J]. IEEE Trans. Knowl. Data Eng., 2017, 29: 2318

[26] Karniadakis G E, Kevrekidis I G, Lu L, et al. Physics-informed machine learning [J]. Nat. Rev. Phys., 2021, 3: 422

[27] Yang J Y, Kang G Z, Kan Q H. Rate-dependent multiaxial life prediction for polyamide-6 considering ratchetting: Semi-empirical and physics-informed machine learning models [J]. Int. J. Fatigue, 2022, 163: 107086

[28] Jiang L F, Hu Y N, Liu Y X, et al. Physics-informed machine learning for low-cycle fatigue life prediction of 316 stainless steels [J]. Int. J. Fatigue, 2024, 182: 108187

[29] Jiang L F, Hu Y N, Li H, et al. A cGAN-based fatigue life prediction of 316 austenitic stainless steel in high-temperature and high-pressure water environments [J]. Int. J. Fatigue, 2025, 190: 108633

[30] Jain A K, Duin R P W, Mao J C. Statistical pattern recognition: A review [J]. IEEE Trans. Pattern Anal. Mach. Intell., 2000, 22: 4

[31] Cover T, Hart P. Nearest neighbor pattern classification [J]. IEEE Trans. Inform. Theory, 1967, 13: 21

[32] Cortes C, Vapnik V. Support-vector networks [J]. Mach. Learn., 1995, 20: 273

[33] Smola A J, Schölkopf B. A tutorial on support vector regression [J]. Stat. Comput., 2004, 14: 199

[34] Breiman L. Random forests [J]. Mach. Learn., 2001, 45: 5

[35] Xie J X, Su Y J, Xue D Z, et al. Machine Learning for Materials Research and Development [J]. Acta Metall. Sin., 2021, 57: 1343
(谢建新, 宿彦京, 薛德祯等. 机器学习在材料研发中的应用 [J]. 金属学报, 2021, 57: 1343)

[36] Lu R S, Yang J, Wang J, et al. Probabilistic damage tolerance assessment method based on the multi-scale crack growth model [J]. Eng. Fract. Mech., 2023, 285: 109297

(责任编辑:梁 烨)