Engineering Fracture Mechanics 327 (2025) 111420

Contents lists available at ScienceDirect

Engineering Fracture Mechanics

journal homepage: www.elsevier.com/locate/engfracmech

Fretting fatigue crack propagation lifetime estimation using Long Short-Term Memory network

Can Wang [a], Qiqi Xiao [a], Yingjian Guo [a], Gang Chen [b], Lihua Wang [c,*], Magd Abdel Wahab [a,d,*]

[a] *Soete Laboratory, Department of Electrical Energy, Metals, Mechanical Constructions and Systems, Faculty of Engineering and Architecture, Ghent University, Belgium*
[b] *School of Chemical Engineering and Technology, Tianjin University, Tianjin 300072, China*
[c] *School of Aerospace Engineering and Applied Mechanics, Tongji University, Shanghai 200092, China*
[d] *Parallel and Distributed Systems Laboratory, Jožef Stefan Institute, Jamova cesta 39, SI-1000 Ljubljana, Slovenija*

A R T I C L E I N F O

Keywords:
Fretting fatigue
Crack propagation
Finite element method
Long short-term memory networks

A B S T R A C T

Fretting fatigue is a complex phenomenon occurring in mechanical structures under cyclic loading with small relative displacements between contacting surfaces. It presents significant challenges in crack propagation lifetime prediction due to its dependence on multiple influencing factors. Traditional methods, such as Finite Element Method (FEM) and experimental approaches, are widely recognized but suffer from high computational and economic costs. This study proposes an innovative Long Short-Term Memory (LSTM) network to address these limitations and provides a detailed comparative evaluation with classical Deep Neural Networks (DNNs). Leveraging a comprehensive dataset of fretting fatigue crack propagation behavior generated through FEM simulations across diverse materials, geometries, and loading conditions. To reduce the structural complexity of DNNs, we optimize input parameters by integrating geometric dimensions, and material constants, into a key mechanical parameter, namely the maximum compressive stress at the contact surface. For both the proposed LSTM framework and the traditional DNN, hyperparameter optimization is systematically conducted using Bayesian Optimization (BO) with the Tree-structured Parzen Estimator (TPE) algorithm implemented in the Hyperopt library, ensuring optimal model performance. The proposed LSTM method enables accurate estimation of the remaining 80% of crack propagation data using only 20% of the FEM-generated dataset. Results demonstrate that LSTM significantly outperforms DNN in prediction accuracy. Additionally, the approach reduces the computational costs by 91.04% compared to traditional FEM. This study also investigates the impact of the sliding window size in LSTM on the neural network framework, particularly its effects on convergence rate and prediction accuracy, providing valuable insights for improving fretting fatigue crack propagation lifetime prediction.

1. Introduction

Fretting fatigue occurs in mechanical structures subjected to cyclic loading, where two contacting surfaces experience small

* Corresponding authors at: Soete Laboratory, Department of Electrical Energy, Metals, Mechanical Constructions and Systems, Faculty of Engineering and Architecture, Ghent University, Belgium (M. Abdel Wahab).

E-mail addresses: lhwang@tongji.edu.cn (L. Wang), magd.abdelwahab@ugent.be (M. Abdel Wahab).

https://doi.org/10.1016/j.engfracmech.2025.111420
Received 17 March 2025; Received in revised form 27 May 2025; Accepted 11 July 2025
Available online 17 July 2025

relative displacements due to micro-vibrations. This phenomenon is commonly observed in aerospace, energy, and biomedical applications, where structural components undergo repeated contact and frictional interactions. Notable examples include aircraft fuselage lap joints, steel wires [1], bolted and riveted connections [2], turbine engine dovetail joints [3], and oil drill pipe connections [4]. In biomedical engineering, fretting fatigue is also a critical concern in artificial joints and bone implants [5], where micro-motions at contact interfaces can lead to premature failure.

Predicting fretting fatigue life is challenging due to the complex interplay of multiple factors, including contact stress distribution, slip amplitude, service environment, wear effects, and the geometric configuration of the components [6]. These variables significantly influence fatigue life estimation, making it difficult to establish a universal predictive model. Generally, fretting fatigue failure progresses through two distinct stages, i.e. crack initiation and crack propagation. While the initiation stage has been extensively studied, the propagation mechanism remains intricate, as crack growth is influenced by localized stress fields, material behavior, and environmental conditions. Experimental investigation of fretting fatigue crack growth is both time-consuming and cost-prohibitive, necessitating alternative predictive approaches. FEM is widely employed to model fretting fatigue behavior, allowing for the evaluation of stress intensity factors (SIFs) and crack growth rates under various loading conditions.

Linear Elastic Fracture Mechanics (LEFM) and eXtended Finite Element Method (XFEM) are two widely utilized techniques for analyzing crack propagation behavior in fretting fatigue. In a previous study, we [7] employed the Maximum Tangential Stress (MTS) and extended MTS criteria within the LEFM framework to predict crack propagation lifetime and paths under out-of-phase loading conditions. Similarly, XFEM has been applied to estimate crack propagation trajectories and lifetimes, as demonstrated in Pinto's work [8]. Nikfam et al. [9] further explored the application of XFEM in fatigue analysis of steel weldments with complex crack propagation patterns, evaluating its effectiveness and limitations. Han et al. conducted in-situ fretting fatigue experiments to measure crack growth rates under different loading conditions, which were then combined with FEM to compute the J-integral and establish a crack growth rate model. While this approach yielded promising results, it is highly costly and challenging to generalize for engineering applications across different materials and loading conditions [10].

To reduce time and cost, researchers have increasingly adopted machine learning methods for predicting crack propagation. Regarding the prediction of crack propagation paths, Yan et al. [11] employed a machine learning approach combined with phase-field simulation to predict crack paths. Martinez et al. [12] utilized machine learning-assisted XFEM, decoupling potential stochastic fracture mechanics problems into multiple classical fracture mechanics problems, thereby achieving effective prediction of crack propagation paths. However, research on crack propagation remains limited. Liao et al. [13] established a relationship between crack length and fatigue life using neural networks, incorporating Paris' law as a physically meaningful loss function into the overall loss function. Notably, the material constants C and m in Paris' law were adjusted in real-time based on the loss function values. While this approach yielded satisfactory prediction results, its physical significance is questionable since C and m should be constant for a given material. Nevertheless, this study provided a valuable approach for linking crack length and life in crack propagation life prediction.

This inspired us to consider Recurrent Neural Networks (RNNs) [14], which are particularly suitable for sequence data, as crack propagation is closely related to previous crack lengths and propagation rates. We propose using RNNs to predict crack propagation data. Given the limited experimental data on fretting fatigue crack propagation, we combine this approach with FEM to expand the dataset, enabling accurate prediction of fretting fatigue crack propagation life while reducing computational time. Unlike previous studies that primarily focus on crack path prediction or general fatigue damage [15,16], this work is one of the first to target the prediction of fretting fatigue crack lifetime using a data-driven approach.

In summary, we conduct finite element simulations to establish the relationship between fretting fatigue parameters and crack growth path and life. RNN is employed to predict crack propagation path and life, and its performance was compared with conventional DNNs. Notably, the hybrid FEM-LSTM approach significantly reduced the computational time while maintaining satisfactory prediction accuracy. To the best of our knowledge, this is the first application of a FEM-assisted LSTM model for predicting fretting fatigue crack propagation lifetime, filling a significant gap in current literature. To demonstrate the generalization capability of RNNs, we utilize different experimental materials, specimen structures, and loading conditions for fretting fatigue tests.

In this study, the crack propagation angle is not included as a network output because previous analyses showed minimal deviation in crack path under the tested conditions [7]. As fatigue life is more sensitive and critical, the model is designed to prioritize accurate life prediction over path estimation.

In our case, the primary varying input during crack propagation is the crack length, while other parameters (e.g., material constants, geometry) remain fixed for a given specimen. As a result, the dataset resembles a time series, where the output (remaining life) depends on the evolution of a single variable over time. DNNs, which assume independence between inputs, struggle to capture this sequential dependency. They tend to underperform in this context because they cannot utilize the history of crack growth effectively to infer future trends.

In contrast, RNNs (especially LSTM networks) are specifically designed to handle sequence-dependent data. They maintain internal memory and learn from sequential patterns, making them far better suited for modeling the dynamic nature of crack growth over time. In our LSTM model, the main is the crack growth life at each stage, enabling the model to learn the relationship between former and the later crack.

This paper is organized as follows. Section 2 details the fretting fatigue experiments and materials used. Section 3 describes the Finite Element (FE) model for predicting crack propagation behavior. Section 4 presents the specific setup of the FE model and provides comprehensive data for fretting fatigue crack propagation behavior. Section 5 introduces the neural network framework and parameter setup employed in this study. Section 6 discusses the prediction results of the two different models. Finally, Section 7 presents the conclusions of this research.

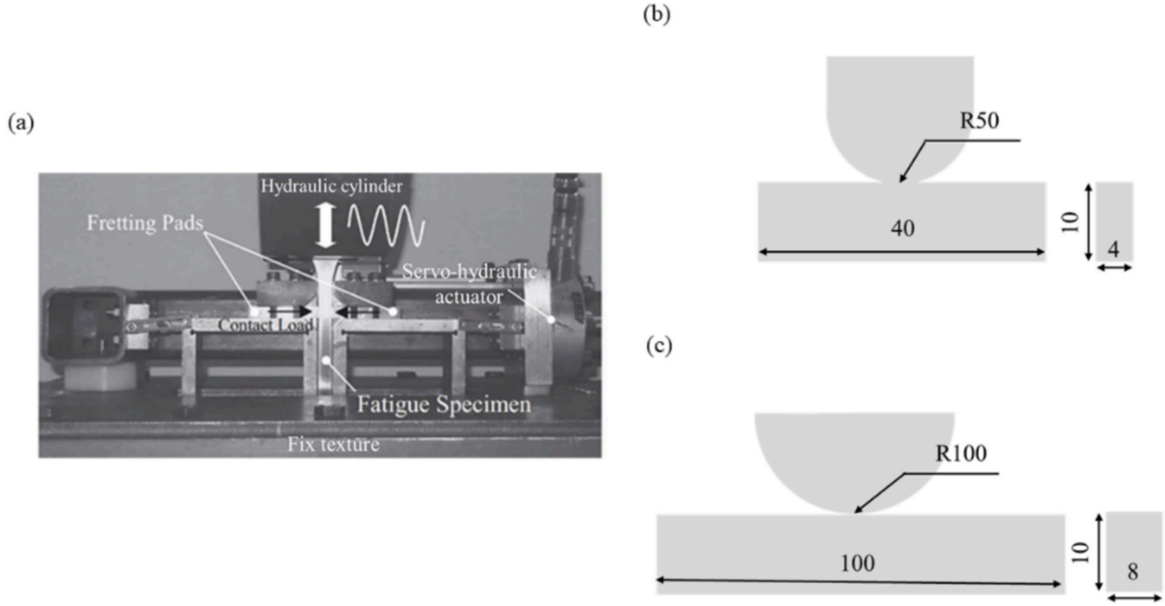

Fig. 1. Fretting fatigue experiment [17,18]: (a) Experimental setup, (b) AA2024-T3 specimen and (c) AA7075-T651 specimen.

Table 1
Mechanical properties of AA2024-T3 and AA7075-T651 [7,19–21].

Material mechanical properties	AA2024-T3	AA7075-T651
E (Young's modulus)	72.1×10^3 MPa	71×10^3 MPa
ν (Poisson's modulus)	0.33	0.33
σ_{yields} (Yield strength)	383 MPa	503 MPa
σ_{UTS} (Tensile strength)	506 MPa	572 MPa
C (Paris' crack growth coefficient)	2.73×10^{-11}	4.83×10^{-11}
m (Paris' crack growth exponent)	2.6526	3.517

2. Materials and experiments

In this study, two materials, AA2024-T3 and AA7075-T651 alloys, were used to investigate fretting fatigue crack propagation under various loading conditions and specimen sizes, aiming to obtain a broader dataset. The experimental setup [17,18] consists of two fretting pads, which apply a stable normal load (F) while experiencing a tangential force (Q). Meanwhile, cyclic fatigue loading was applied to both ends of the specimen. A 100 kN servo-hydraulic actuator generated the cyclic force, ensuring controlled loading conditions (See Fig. 1).

The device features movable supports, allowing for adjustments in stiffness, which enables the modification of the tangential load amplitude (Q) without altering the axial load amplitude (σ_{axial}). This flexibility makes it possible to achieve multiple fretting-load combinations (σ_{axial}, Q, F) within a single setup. To ensure precise monitoring of all fretting loads, the test device is equipped with load cells that measure these forces. The signals from the load cells are transferred to a signal conditioner and then processed via a data-acquisition card, enabling real-time display and recording of σ_{axial}, Q, and F on a PC [17,18].

In this study, two different sizes of fretting fatigue specimens were used, both adopting a cylinder-to-flat contact configuration. The first specimen has a contact arc radius of 50 mm, while the second specimen has a contact arc radius of 100 mm. The specific specimen shapes and contact configurations are shown in Fig. 1(b) and Fig. 1(c), respectively.

The material properties of AA2024-T3 and AA7075-T651, including elastic modulus and Poisson's ratio, are provided in Table 1.

The fretting fatigue experimental conditions and parameters are summarized in Table 2. It can be seen that the loading ranges for the two types of specimens are different. For the AA2024-T3 material, constant normal pressure is maintained while varying the cyclic load and tangential stress, with a total of 9 tests conducted. For the AA7075-T651 material, three distinct loads are used, and the tests are distributed discretely across these load levels, with a total of 23 tests conducted. The load ratios (R) for the axial load are 0.1, −1 in AA2024-T3 and AA7075-T651 tests, respectively. It is worth mentioning that the lifetime shown in the list are the average values. In Tests No. 1 to 9, the normal load was kept constant while the axial stress and tangential load amplitude varied. The axial stress had a stress ratio of 0.1. In Tests No. 10 to 32, four different axial stress levels (70 MPa, 110 MPa, 150 MPa, and 175 MPa), four different tangential load amplitude (971 N, 1257 N, 1543 N, and 2113 N), and four different normal loads (3006 N, 4217 N, 5429 N, and 6629 N) were applied. The stress ratio for the axial stress in these tests was −1.

Table 2
Experimental data [5,18,19].

Material	Test Number	σ_{axial} (MPa)	Q (N)	F (N)	R	$N_{failure}$
AA2024-T3	1	100	155.165	543	0.1	1,407,257
	2	115	186.25	543	0.1	1,105,245
	3	135	223.7	543	0.1	358,082
	4	135	195.55	543	0.1	419,919
	5	160	193.7	543	0.1	245,690
	6	190	330.15	543	0.1	141,890
	7	205	322.1	543	0.1	114,645
	8	220	267.15	543	0.1	99,607
	9	220	317.845	543	0.1	86,647
AA7075-T651	10	70	971	6629	−1	241,150
	11	110	971	5429	−1	119,331
	12	110	1257	5429	−1	117,231
	13	110	1543	4217	−1	88,796
	14	110	1543	5429	−1	85,020
	15	150	971	3006	−1	59,637
	16	150	971	4217	−1	64,032
	17	150	971	5429	−1	49,655
	18	150	1543	3006	−1	29,315
	19	150	1543	4217	−1	44,685
	20	150	1543	5429	−1	44,735
	21	150	2113	3006	−1	37,953
	22	150	2113	4217	−1	37,360
	23	150	2113	5429	−1	34,385
	24	175	971	3006	−1	29,201
	25	175	971	4217	−1	30,284
	26	175	971	5429	−1	32,136
	27	175	1543	3006	−1	30,689
	28	175	1543	4217	−1	34,839
	29	175	1543	5429	−1	30,677
	30	175	2113	3006	−1	21,438
	31	175	2113	4217	−1	27,792
	32	175	2113	5429	−1	28,145

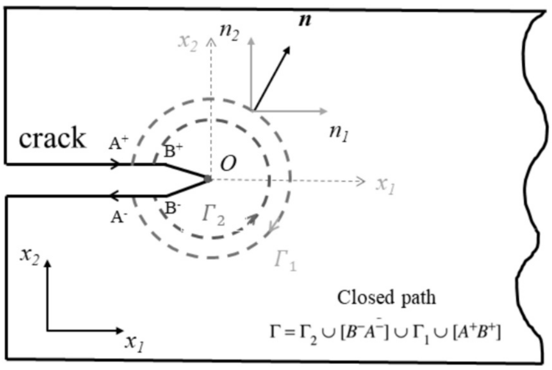

Fig. 2. The interaction integral method.

3. Crack propagation method adopted in FEM

In the crack propagation analysis, we adopt the approach of the path-independent J-integral and interaction integral to obtain the stress intensity factors around the crack tips. Erdogan and Sih [22] initially devised the Maximum Tangential Stress (MTS) criterion to ascertain the crack propagation path during proportional loadings, later expanded upon by Ribeaucourt et al. [23].

The J-integral expression is given as:

$$\int_{\Gamma}\left(Wn_1 - \sigma_{ij}n_k\frac{\partial u_j}{\partial x_1}\right)ds \tag{1}$$

The unit outward vector components of Γ are denoted as n_k, where σ_{ij} and u_j represent the stress tensor and displacement, respectively. W signifies the density of strain energy [24]:

$$W = \int_0^{\varepsilon} \sigma_{ij} d\varepsilon_{ij} \tag{2}$$

where ε is equal to $[\varepsilon_{ij}]$, i.e. the infinitesimal strain tensor.

Yau et al. [25] proposed the interaction integral $M^{(1,2)}$ to derive the hybrid SIF K_I and K_{II}, where the contact stress on the crack surface is taken into account. The method exploits two independent states of elastic isotropy and homogeneous crack body, denoted as (1) and (2), each characterized by its own exponent (1) and (2), respectively.

State (1) represents the stress, strain, and displacement fields of the actual problem, while state (2) serves as an arbitrary auxiliary state. To implement it in the FE calculations, the contour integral is converted to a domain integral [26]. By multiplying the integrand of J or $M^{(1,2)}$ by a sufficiently smooth weighting function $q(x)$, it is ensured that it remains unified on the open set around the crack tip while maintaining uniformity on the prescribed outer contour Γ_2. Then, for each contour Γ_1 as illustrated in Fig. 2, assuming that the crack plane is straight in the region A bounded by the contours Γ_1 and Γ_2, the interaction integral can be expressed as:

$$M^{(1,2)} = -\int_{\Gamma} \left[\sigma_{ij}^{(1)} \int_{ij}^{(2)} \delta_{1j} - \sigma_{ij}^{(1)} \frac{\partial u_i^{(2)}}{\partial x_1} - \sigma_{ij}^{(2)} \frac{\partial u_i^{(1)}}{\partial x_1} \right] q n_j d\Gamma - \int_{[A^+B^+]\cup[B^-A^-]} \left[\sigma_{i2}^{(1)} \frac{\partial u_i^{(2)}}{\partial x_1} + \sigma_{i2}^{(2)} \frac{\partial u_i^{(1)}}{\partial x_1} \right] q n_2 d\Gamma \tag{3}$$

Utilizing the divergence theorem on the closed-form integral over the profile Γ, and accounting for the scenario where the profile Γ_1 nears the crack tip, we derive the subsequent equation for the interaction integral in domain form:

$$\begin{aligned} M^{(1,2)} = &-\int_{A} \left[\sigma_{ij}^{(1)} \int_{ij}^{(2)} \delta_{1j} - \sigma_{ij}^{(1)} \frac{\partial u_i^{(2)}}{\partial x_1} - \sigma_{ij}^{(2)} \frac{\partial u_i^{(1)}}{\partial x_1} \right] \frac{\partial q}{\partial x_j} dS \\ &- \int_{[A^+B^+]\cup[B^-A^-]} \left[\sigma_{i2}^{(1)} \frac{\partial u_i^{(2)}}{\partial x_1} + \sigma_{i2}^{(2)} \frac{\partial u_i^{(1)}}{\partial x_1} \right] q n_2 d\Gamma \end{aligned} \tag{4}$$

In fretting fatigue problems, the crack behavior is intricate due to non-proportional loading conditions [27,28]. To accurately determine the crack path, SIFs k^*_I and k^*_{II} are essential and is expressed as a linear function of K_I and K_{II} [23]. FEM permits the computation of K_I and K_{II} at the crack tip as a function of time t. Here, k^*_I and k^*_{II} are correlated to the angle θ and time t:

$$k_I^*(\theta, t) = K_{11}(\theta) \cdot K_I(t) + K_{12}(\theta) \cdot K_{II}(t) \tag{5}$$

where $K_{11}(\theta)$, $K_{12}(\theta)$ are given by:

$$K_{11}(\theta) = R\left(\cos\theta - \frac{1}{2\pi}\sin L\theta\right) \tag{6}$$

$$K_{12}(\theta) = R\left(-\frac{3}{2}\sin\theta\right) \tag{7}$$

where $R = \left(\frac{1-m}{1+m}\right)^{\frac{m}{2}}$, $m = \frac{\theta}{180}$ and $L = ln\left(\frac{1-m}{1+m}\right) - 2\left(\frac{m}{1-m^2}\right)$.

Based on this approach, Hourlier et al. [29] introduced an extended MTS criterion, where the crack propagation angle is determined by the difference between $k^*_I(\theta)$ at the instant of maximum load and at minimum load, $\Delta k^*_{I,\max}$ at the peak value:

$$\Delta k_I^*(\theta) = \max k_I^*(\theta, t) - \min k_I^*(\theta, t) \tag{8}$$

The propagation lifetime is estimated through integration of a Paris' Law equation, depicted in Eq. (9):

$$N_p = \int_{a_i}^{a_f} \frac{da}{C \Delta K^m} \tag{9}$$

This formula describes the number of cycles N_p, required for the crack to propagate from an initial length a_i to a final a_f.

Herein, da represents a small increment in crack length, the propagation lifetime can be estimated accurately by assuming the initial crack length da as a fixed value of 0.05 mm for all increments based on previous studies [7,21]. C and m are material constants, which are listed in Table 1. ΔK is the increment in the stress intensity factor at the crack tip. Under mixed-mode situations, researchers [30] assumed that the value of ΔK was equal to ΔK_{eff}, which could be obtained by:

$$\Delta K_{eff} = \sqrt{\Delta K_I^2 + \Delta K_{II}^2} \tag{10}$$

where $\Delta K_I = K_{I,maxload} - K_{I,minload}$ and $\Delta K_{II} = K_{II,maxload} - K_{II,minload}$.

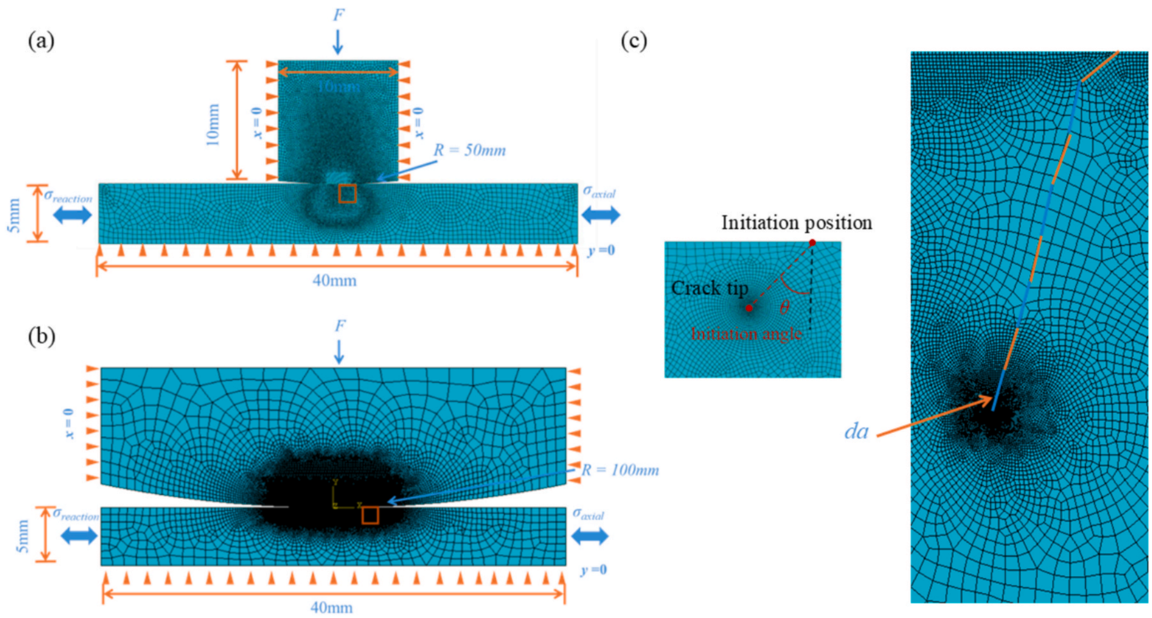

Fig. 3. Schematic representation of fretting fatigue models and crack propagation details: (a) Model I, (b) Model II and (c) sketch of crack propagation part.

4. Dataset from FEM

4.1. Illustration of FEM models

Based on the symmetric experimental configuration, two-dimensional FE models are established (the load parameters are listed in Table 2 and specimens are shown in Fig. 1). By leveraging geometric symmetry, the model was simplified to include only half of the specimen-pad system, the geometric configuration is shown in Fig. 3 (a) and (b).

Regarding boundary conditions, the bottom of the specimen is constrained in the *y*-direction, while the axial cyclic stress and the contact reaction stress, determined by Eq. (11), are applied to its left and right sides, respectively. The top of the pad is subjected to a normal load *F*, while *x*-direction displacement constraints are imposed on both sides to ensure contact stability.

To accurately characterize the contact interface behaviour, considering the unique contact mechanics challenges in fretting fatigue, the Augmented Lagrange Method is employed. The coefficient of friction is set to 0.65 and 0.75, respectively, based on experimental measurements for the two types of material contact pairs [17,21]. To ensure the computational accuracy of the contact stress field, a gradient mesh refinement strategy is applied in the contact region, with a minimum element size of 2 μm. All components were discretized using four-node plane strain reduced integration elements (CPE4R). Although higher-order elements such as 8-node quadrilaterals offer improved stress field accuracy near crack tips, the stress intensity factors in this study were extracted using the MTS criterion based on angular stress fitting. With sufficient mesh refinement near the crack front as detailed in Resende Pereira's work [31], CPE4R elements provide reliable results while maintaining computational efficiency.

It is particularly noteworthy that the contact pairs are established strictly following the master–slave surface algorithm. The lower surface of the pad, which has higher stiffness, is defined as the master contact surface, while the upper surface of the specimen is set as the slave contact surface. This configuration is consistent with fundamental contact mechanics principles and improves the convergence efficiency of the iterative calculations. The reaction stress is given by:

$$\sigma_{reaction} = \sigma_{axial} - \frac{Q}{A_S} \tag{11}$$

where A_s is the half cross-section area of the specimen.

Fig. 3(c) presents in detail the numerical modeling method of cracks. In this study, the pre-set notch (seams) technique is adopted to construct the geometric shape of the cracks. This pre-set notch allows the crack surfaces to open freely during the analysis. For the contact interaction between the crack surfaces, the tangential contact characteristics are defined based on the penalty formulation, and the friction coefficient is set to 0.8 [21]. To accurately represent the stress singularity at the crack tip, a ring-shaped singular element is constructed in the crack tip region. Collapsed linear quadrilateral elements are used to form a 1/4 node singularity structure, and the minimum element size at the crack tip is controlled at about the order of 1 μm through a gradient encryption strategy to ensure the calculation accuracy of the stress intensity factors.

Additionally, the second subfigure in Fig. 3(c) illustrates the step-by-step construction of the crack propagation model. The crack

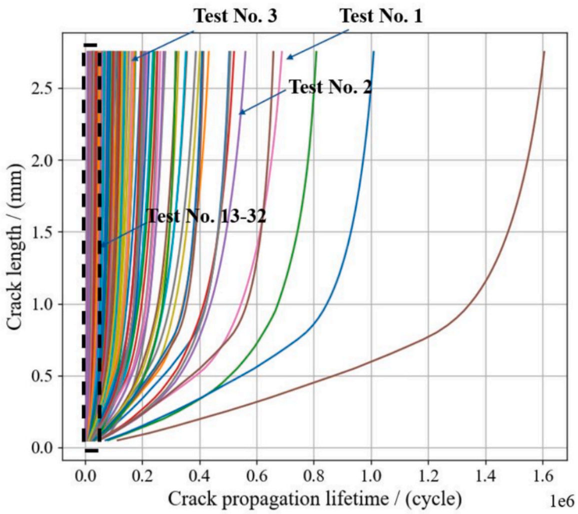

Fig. 4. Dataset obtained from FEM.

Table 3
Selected input and output variables of the dataset.

Input					Output
P_{max} (MPa)	σ_{axial} (MPa)	Modified R (Stress ratio)	Q (N)	a (Crack length/mm)	Crack propagation lifetime(cycle)
96.6	100	0.1	155.165	0.05	12,202
96.6	100	0.1	155.165	0.10	12,199
96.6	100	0.1	155.165	0.15	10,993
96.6	100	0.1	155.165	0.20	10,077
96.6	100	0.1	155.165	0.25	8669
96.6	100	0.1	155.165	0.30	7132
…	…	…	…	…	…
96.6	100	0.1	155.165	3.05	325
142.17	205	0.1	322.1	0.05	2153
142.17	205	0.1	322.1	0.10	1928
142.17	205	0.1	322.1	0.15	1315
142.17	205	0.1	322.1	0.20	1371
142.17	205	0.1	322.1	0.25	1052
142.17	205	0.1	322.1	0.30	1009
…	…	…	…	…	…
142.17	205	0.1	322.1	3.05	186
121.54	190	1	330.15	0.05	2999
121.54	190	1	330.15	0.10	2306
121.54	190	1	330.15	0.15	1839
121.54	190	1	330.15	0.20	1651
121.54	190	1	330.15	0.25	1413
121.54	190	1	330.15	0.30	1356
…	…	…	…	…	…
121.54	190	1	330.15	3.05	306
…					

initiation position and direction are first determined through a crack initiation analysis by using the critical plane method [32]. Following this, incremental crack growth is modeled in discrete steps, with each crack segment length denoted as da. This stepwise modeling approach aligns with the methodology introduced in the previous section, ensuring consistency in analyzing the crack evolution process.

4.2. Dataset generation

One major challenge in this study is the limited amount of experimental data, as only 32 tests were originally available. However, for data-driven models such as RNN and DNN, this dataset is far from sufficient to ensure reliable training and accurate predictions. To address this limitation, we expanded the dataset by increasing the number of loading conditions for the two types of materials. This augmentation process significantly increased the total number of experimental groups to 125.

Each experimental group contains about 70 data points, where each data point represents the number of fatigue cycles required for a given crack length. This refined dataset provides a more comprehensive and diverse representation of crack propagation behavior under different loading conditions, thereby improving the robustness and generalizability of the machine learning models. The overall distribution of the expanded dataset is presented in Fig. 4, highlighting the increased data availability and its potential to enhance predictive performance.

The overall dataset is visualized in Fig. 4, where the crack propagation lifetime ranges up to a maximum of approximately 1.6×10^6 cycles. A significant variation in crack growth rates can be observed, with most curves exhibiting a characteristic pattern. An initial slow growth phase followed by an accelerating propagation stage, where the cycles required for crack growth over the same length gradually decrease.

A large portion of the data is concentrated within 0.4×10^6 cycles, with only a few cases extending beyond this range, indicating longer fatigue lifetimes. This extensive dataset provides a solid foundation for machine learning applications, ensuring sufficient variability and representation of different crack growth behaviours.

The selection of key parameters for DNN and RNN, as well as the specific neural network architectures, is discussed in detail in the next section.

5. Architecture of DNN and LSTM

5.1. DNN

DNN is a computational model inspired by biological neural networks and is widely used in pattern recognition, regression analysis and prediction tasks. DNN builds nonlinear mappings through a large number of neurons and weights, learns complex relationships from input data, and improves prediction performance by continuously optimizing weights and biases. The basic structure of DNN includes input layer, hidden layer and output layer. The input layer receives raw data, the hidden layer performs feature extraction and nonlinear transformation, and the output layer provides the final prediction results. The core of DNN lies in the activation function and loss function. The former introduces nonlinear characteristics, and the latter is used to measure prediction errors and guide the optimization process. We have successfully applied DNNs with various algorithms to predict fretting fatigue crack initiation behavior, achieving promising results [33].

In this study, since the input data contains five factors and the value range is large (from 10^{-2} to 10^3 as shown in Table 3), ReLU (Rectified Linear Unit) is selected as the activation function instead of Sigmoid or Tanh. The mathematical expression of ReLU is:

$$f(z) = \max(0, z) \tag{12}$$

Its advantage is that it does not cause the gradient to vanish when the value is large, like Sigmoid and Tanh, so that deep networks can be trained effectively. In addition, ReLU has strong adaptability to data of different scales. It can retain large numerical information while suppressing inputs less than zero, thereby enhancing the sparsity of the network and improving training efficiency.

In terms of loss function, this study uses Smooth L1 Loss:

$$L(y, \hat{y}) = \begin{cases} \frac{1}{2}(y - \hat{y})^2, if|y - \hat{y}| < 1 \\ |y - \hat{y}| - \frac{1}{2}, otherwise \end{cases} \tag{13}$$

Compared to Mean Squared Error (MSE), Smooth L1 Loss behaves similarly to MSE when the error is small, i.e. the loss value increases quadratically with the error, ensuring smooth gradient changes and avoiding gradient oscillations. When the error is large, it behaves like the Absolute Error (L1 Loss), where the loss value grows linearly, reducing sensitivity to outliers. Since crack propagation data often exhibit significant uncertainty and are influenced by local stress fluctuations, parameters such as crack growth rate may experience severe oscillations, leading to instability in the loss function during training. By using a piecewise function, Smooth L1 Loss provides smooth gradients for small errors, ensuring a stable optimization process, while mitigating the impact of outliers for larger errors. This makes it more robust to noisy data and better suited for crack propagation prediction tasks.

This study uses AdamW (Adam algorithm with Weight Decay) [34] instead of the standard Adam algorithm. Adam optimizes parameters through momentum estimation and adaptive learning rate:

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1)\nabla L \tag{14}$$

$$v_t = \beta_2 v_{t-1} + (1 - \beta_2)(\nabla L)^2 \tag{15}$$

$$w \leftarrow w - \frac{\eta}{\sqrt{v_t} + \epsilon} m_t \tag{16}$$

However, due to its adaptive learning rate, Adam algorithm has a weak regularization ability and is prone to overfitting. AdamW improves this by adding a weight decay term when updating weights.

$$w \leftarrow w - \eta\left(\lambda w + \frac{m_t}{\sqrt{v_t} + \epsilon}\right) \tag{17}$$

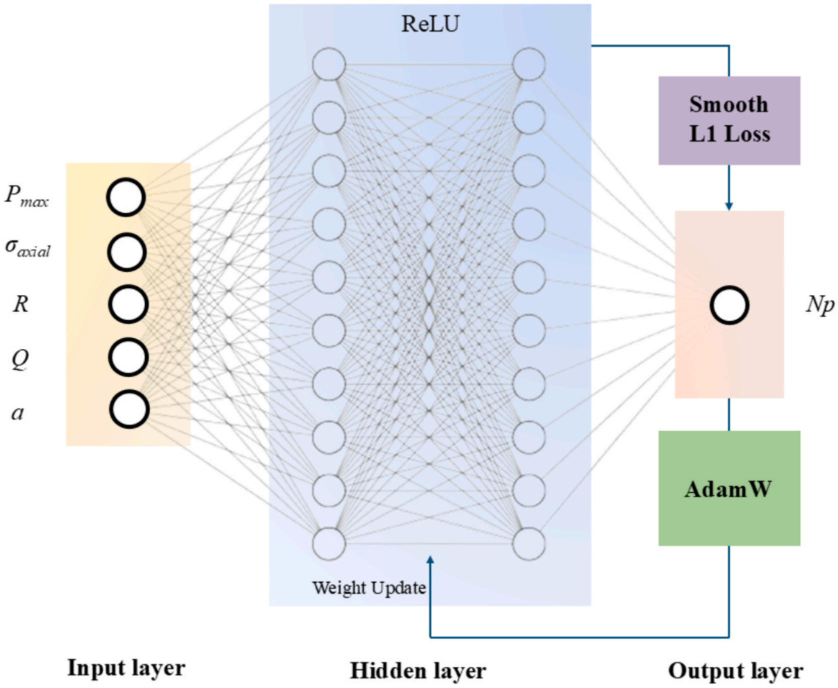

Fig. 5. DNN architecture.

Compared to standard Adam algorithm, which incorporates the L_2 regularization term directly into the gradient computation, AdamW decouples weight decay from the gradient update process. This separation prevents interference with the first-moment estimation, allowing the model to better regulate weight magnitudes, reduce overfitting, and improve generalization performance. The DNN architecture is shown in Fig. 5.

Table 3 presents a selected portion of the dataset to illustrate the specific input and output variables. Due to the large volume of data, only a subset is shown to avoid excessive length. Additionally, detailed crack lifetime and crack length data are already provided in Fig. 4.

Table 3 presents a partial dataset from three crack propagation groups, out of a total of 125 groups of crack propagation data. Among the input parameters, $P_{\max}$ represents a composite physical quantity that integrates the effects of contact sketch, pad radii (R_{pad}), material properties and constant normal load (see Eq. (18)). This approach helps reduce the number of input parameters while retaining essential information for the analysis.

$$p_{max} = \sqrt{\frac{FE^*}{t\pi R^*}} \tag{18}$$

where F means normal load, t is the thickness and E^*, R^* are equivalent material properties and radius [35].

Besides, when $R = 0.1$, we retain its value as 0.1; however, when $R = -1$, and we change it to 1 before inputting it into the DNN. Since ReLU activation is used in the network, negative input values such as $R = -1$ would be mapped to zero, causing loss of input variability and hindering learning. Therefore, R is transformed to a positive range to maintain meaningful activations and improve training stability. Regarding the selection of crack length a, it represents the length of the crack. Physically, different crack lengths indicate varying states of stress, which in turn affect the crack's propagation behavior and its associated lifespan. To accurately capture the complete crack propagation lifespan across multiple crack lengths, we base this approach on LEFM. Consequently, we divide each crack into several segments, with each segment having a length of 0.05 mm. This segmentation allows for a more detailed analysis of the crack's behavior under different conditions, ensuring that the model can effectively learn the relationship between crack length and its propagation lifespan.

Unlike traditional DNN applications that predict the total fatigue life based solely on the initial configuration, the approach adopted in this study focuses on modelling the progressive crack growth process. This paradigm shift enables a more flexible and detailed prediction framework, where the model inputs include not only the applied loading and geometry but also the current crack length, thus allowing stage-wise prediction of fatigue damage.

5.2. LSTM

RNNs are a type of neural network designed to process sequential data by leveraging temporal information. However, standard

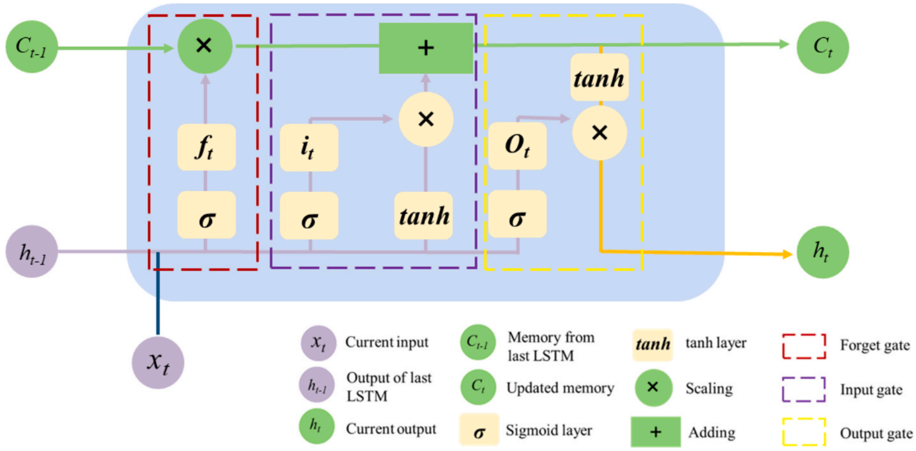

Fig. 6. Illustration diagram of LSTM.

RNNs suffer from the vanishing gradient and exploding gradient problems, which significantly hinder their performance on long sequences.

LSTM networks are an improved version of RNNs that address the vanishing gradient issue through a gating mechanism. This allows the model to effectively capture long-term dependencies. This capability is particularly crucial for predicting crack propagation life, as crack growth is inherently a time-dependent process where the current state is influenced by past conditions. Fig. 6 shows the LSTM architecture in detail.

The structure of a LSTM unit consists of three essential gates: the forget gate, input gate, and output gate. These gates regulate the flow of information through the network, ensuring that important past information is retained while irrelevant information is discarded. The forget gate plays a crucial role in determining which historical information should be retained or discarded. It receives both the previous hidden state (h_{t-1}) and the current input (x_t) as inputs, processes them through a sigmoid activation function, and subsequently generates a forget coefficient (f_t) ranging between 0 and 1. This coefficient serves as a weighting factor that regulates the extent to which past information is preserved or eliminated from the memory cell:

$$f_t = \sigma\left(W_f[h_{t-1}, x_t] + b_f\right) \tag{19}$$

where W_f and b_f are the weight matrix and bias term, respectively. If f_t is close to 0, past information is largely forgotten, whereas values closer to 1 indicate retention of past information.

Input gate aims to decide how much new information will be added to the unit. There are two components included:

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i) \tag{20}$$

$$\widetilde{C}_t = \tanh(W_C[h_{t-1}, x_t] + b_C) \tag{21}$$

The new cell state is updated as a combination of the previous state and the newly computed candidate state, by using Eq. (22):

$$C_t = f_t C_{t-1} + i_t \widetilde{C}_t \tag{22}$$

For the output gate, h_t can be determined by the following equations:

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o) \tag{23}$$

$$h_t = o_t \tanh(C_t) \tag{24}$$

In this study, a sliding window mechanism is implemented to enhance the predictive capability of the LSTM model for crack propagation prediction. This approach allows the model to process a fixed-length segment of historical crack length measurements, effectively capturing the temporal patterns of past crack growth states to forecast future crack propagation life. The sliding window technique ensures a robust and data-driven methodology for fatigue analysis, enabling the model to adaptively learn from sequential data patterns. The selection of optimal sliding window parameters plays a crucial role in model performance. A comprehensive investigation into the impact of different window sizes has been conducted, with detailed analysis and findings presented in subsequent sections of this study.

Based on the dataset summarized in Table 3, which spans multiple load levels and crack length increments, 80 % of the data points were randomly selected across all combinations for training and validation, while the remaining 20 % were used for testing. This random sampling approach ensures that the model learns general features across different loading and crack growth scenarios. This conventional approach ensures that the model learns the general mapping between input parameters and crack propagation life from a sufficiently large dataset before being tested on unseen data.

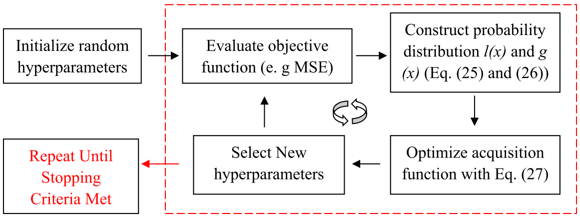

Fig. 7. Flowchart for Hyperparameter Optimization with TPE.

Table 4
Summary of optimized neural network architectures.

Model	Architecture	Hidden layers	Activation functions	Training details
DNN	Fully connected feedforward network	4 layers:72–88-43-output	ReLU	1800 epochs with learning rate 0.00128
LSTM	Stacked LSTM network	2 LSTM: 140–70	Implicit (Tanh and Sigmoid in LSTM)	Trained for 500 epochs with batch size 20

For LSTM, 80 % of the entire crack growth dataset is used for training and validation, allowing the model to learn the crack propagation trend from 0 mm to 3 mm. Then, for the remaining 20 % of the dataset, only the first 20 % (0 mm to 0.6 mm) of the sequence is used as input during training. The trained RNN model is then tasked with predicting the subsequent 80 % of the sequence, effectively simulating real-world scenarios where future crack growth needs to be estimated based on limited prior observations.

This approach enables LSTM to learn from existing datasets and gain experience with crack growth patterns. When new crack growth data is available, the LSTM model only needs a small portion of the initial crack growth sequence, such as from 0 mm to 0.06 mm (or even less), to accurately predict the remaining crack growth life. This significantly reduces the computational cost of crack growth modelling, making it a more efficient prediction tool. This advantage is analysed in detail in Section 6.3.

6. Results and discussion

6.1. Hyperparameter optimization

To ensure that the Neural Network (NN) models achieve optimal predictive performance for crack propagation life estimation, hyperparameter optimization is employed. The selection of NN architecture parameters, such as the number of hidden layers, the number of neurons per layer, the batch size, and the number of epochs, significantly impacts model accuracy and generalization. Instead of using manually predefined values, we utilize Bayesian Optimization (BO) with the Tree-structured Parzen Estimator (TPE) algorithm implemented in the Hyperopt library to systematically explore the hyperparameter space.

The objective function is defined based on the Mean Squared Error (MSE) of the validation dataset. During optimization, different sets of hyperparameters are evaluated, and the model is trained using the best-performing configuration. The final optimized model is then retrained using the selected hyperparameters to ensure optimal performance.

TPE models the probability of hyperparameters leading to good results by constructing two probability densities:

$$l(x) = P(x|f(x) < \gamma) \tag{25}$$

$$g(x) = P(x|f(x) \geq \gamma) \tag{26}$$

where $l(x)$ models hyperparameters leading to better-than-threshold performance and $g(x)$ models worse-than-threshold performance. To steer clear of aimless searching, the optimization process hinges on maximizing Eq. (27). This strategic approach effectively narrows the search scope to hyperparameter regions where optimal results are more likely to be achieved. γ is a quantile threshold of the objective function value (typically set to the top 10–20 % of the best values). The optimal hyperparameter combination, x^*, is given by:

$$x^* = \underset{\{x \in X\}}{\operatorname{argmax}} \left(\frac{l(x)}{g(x)} \right) \tag{27}$$

Overall, this approach (shown in Fig. 7) not only automates hyperparameter selection but also prevents the inefficiencies of manual tuning, ultimately leading to a more robust and accurate crack propagation life prediction model. By ensuring that both the DNN and LSTM architectures are optimized under the same conditions, this method allows for a fair and representative comparison of their predictive accuracy in crack propagation. It eliminates the randomness introduced by manual hyperparameter tuning, ensuring that performance differences are due to the models' inherent capabilities rather than arbitrary parameter choices. The best architecture for both models are list in Table 4.

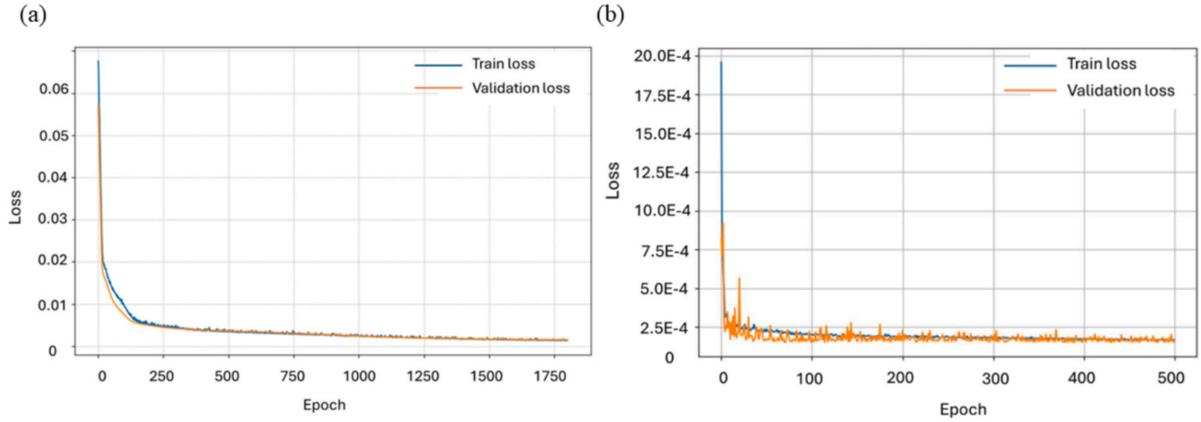

Fig. 8. Loss curve for: (a) DNN and (b) LSTM.

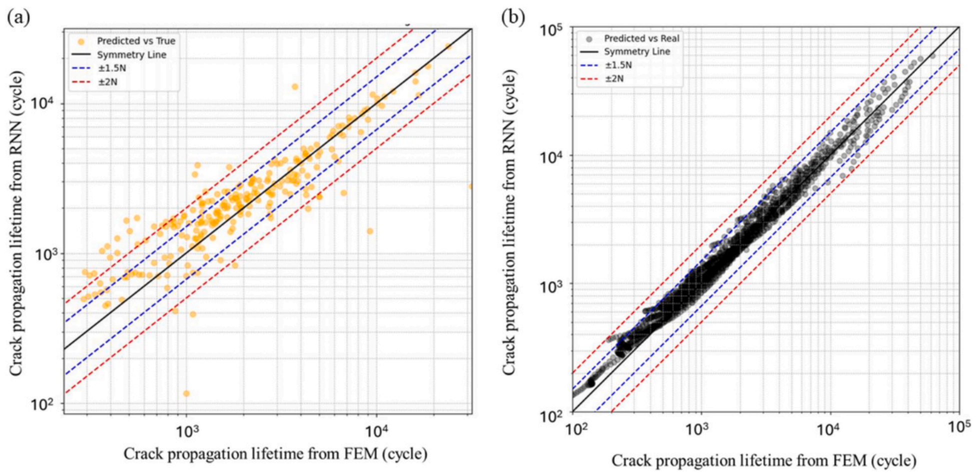

Fig. 9. Lifetime prediction comparison between: (a) DNN and (b) LSTM.

6.2. Performance comparison

In this section, we present a detailed comparison of the performance of the two models. By examining the loss curves in Fig. 8, we can observe distinct patterns in the training processes of the DNN and LSTM architectures.

For DNN architecture, the loss value experiences a sharp decline after approximately 250 epochs. Subsequently, it stabilizes after 1000 epochs, reaching a relatively low error value. This behavior can be attributed to the complex structure of DNN. With multiple input parameters, the model tends to disperse its learning efforts across various dimensions. As a result, it may struggle to focus on the most crucial features and extract the key information effectively, leading to a slower convergence rate during the training process.

In contrast, LSTM architecture demonstrates a significantly faster decrease in the loss value. It attains stability within the range of 100 to 200 epochs, achieving a lower error compared to DNN. This remarkable performance can be explained by the inherent characteristics of LSTM. One of the major advantages of LSTM is its high sensitivity to data having non-linear relationships. Unlike DNN, which deals with multiple input parameters simultaneously, RNN typically has a single input parameter. This simplicity allows LSTM to better capture the sequential and temporal dependencies in crack propagation pre-lifetime prediction, enabling it to learn more efficiently and reach optimal performance at a much faster pace.

For fretting fatigue crack growth life prediction, researchers usually plot the data points in a logarithmic scale graph. The prediction results of DNN and LSTM are shown in Fig. 9 (a) and (b), respectively. Fig. 9 consists of 2 subplots (a and b), each depicting a comparison between the crack propagation lifetime predicted by the NN models and the corresponding values obtained from FEM simulations. The x-axis represents the crack propagation lifetime calculated using FEM (in cycles), while the y-axis shows the corresponding values predicted by the NN models (in cycles). Both axes are logarithmic, which helps visualize the data across a wide range of values. The solid black diagonal line in each subplot represents the ideal 1:1 correlation, the blue dashed lines indicate a certain tolerance range around this ideal correlation (±1.5 N), while the red dashed lines represent a wider tolerance band (±2.0 N).

In Fig. 9(a), the distribution of data points is notably more scattered. While the majority of the data points fall within the ± 1.5 N

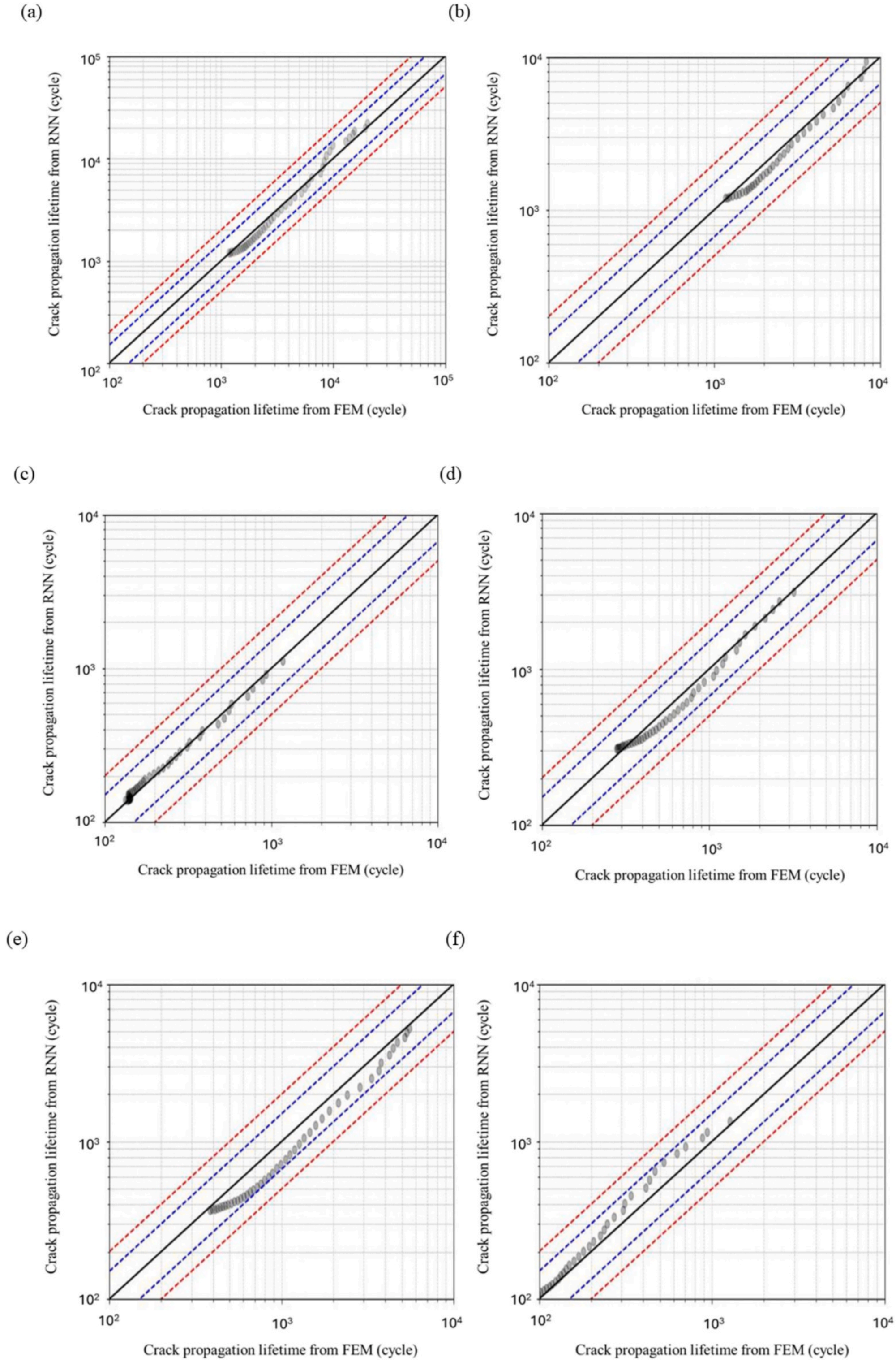

Fig. 10. LSTM prediction for a single crack propagation under six different conditions: (a) Test No. 2, (b) Test No. 3, (c) Test No. 25, (d) Test No. 17, (e) Test No. 12 and (f) Test No. 24.

Table 5
Computational Cost Comparison.

Method	Crack length range (mm)	Time per simulation (min)	Number of simulation	Total time (min)
FEM	0–1.5	20	30	600
FEM	1.5–2.5	40	20	800
FEM	2.5–3.5	60	20	1200
Total (FEM)				**2600**
FEM	0–0.7	20	14	280
LSTM	Training			11
LSTM	Prediction			2.33
Total (hybrid)				**293.33**

range, approximately 10 % of them lie outside the ± 2 N boundary. A closer look reveals that the prediction accuracy of the model deteriorates when dealing with low lifetime data, specifically for data points corresponding to lifetimes of less than 10^3 cycles, the dispersion is even more pronounced. This phenomenon can be attributed to the characteristics of the model and the nature of the data. The DNN model, which has a relatively large number of input parameters, struggles to handle highly non-linear data, especially when there are abrupt changes. The large number of input variables can lead to overfitting or an inability to capture the underlying patterns accurately. Additionally, referring to the original data, a high crack rate might also contribute to this inaccurate prediction, as DNN fails to adapt well to the complex and rapidly changing relationships within the data. Mortazavi and Ince [36] attributed the low accuracy to the limited data. However, in this study, the primary reason is the model itself.

LSTM demonstrates remarkable prediction performance. Almost 99 % of the predicted data from the LSTM falls within the ± 1.5 N range, regardless of whether the lifetimes are less than 10^2 cycles or more than 10^4 cycles. The LSTM architecture is well suited for analysing sequential data with non-linear relationships. It can effectively capture long-term dependencies and remember past information, enabling it to understand the overall behaviour of the data. This allows LSTM to better adapt to sudden changes and complex patterns in the data, resulting in more accurate predictions across a wide range of lifetimes.

In conclusion, the comparison between the DNN and RNN predictions in Fig. 9 highlights the limitations of the DNN in handling non-linear and dynamic data, while emphasizing the superiority of LSTM in terms of prediction accuracy and robustness.

Due to the excessive number of data points, it is challenging to effectively illustrate the LSTM's prediction of the propagation life for a single crack. Therefore, we present the overall prediction results for a selected few cracks, as shown in Fig. 10.

In each crack propagation case, the overall predicted data points exhibit a clear trend rather than a random distribution, demonstrating that the LSTM model successfully learned the sequential patterns from the original data. Moreover, a noticeable trend variation can be observed in the predictions, further indicating that the model captures the progressive changes in crack propagation behaviour.

Furthermore, the accuracy of LSTM predictions is not only related to the neural network structure but also significantly influenced by the definition of the step size due to the use of a sliding window. The sliding window mechanism enables the model to capture temporal dependencies by sequentially processing overlapping segments of data, ensuring that information from previous states is retained and utilized in future predictions. This becomes particularly crucial when learning data with strong correlations between past and future states. We also discovered some intriguing phenomena related to the step size selection, highlighting its impact on the model's learning process. These findings will be further explored in the following section.

6.3. Comparison of computational cost between FEM and FEM-LSTM

To comprehensively evaluate the computational efficiency of the conventional FEM versus the proposed hybrid prediction framework based on LSTM network, a detailed comparison of the time required to simulate a single crack growth trajectory was conducted (see Table 5).

In the traditional FEM framework, the crack propagation process must be discretized into multiple small growth increments (0.05 mm per step in this study), with each increment requiring an independent simulation to extract key fracture mechanics parameters. As a result, simulating the full crack growth path from initiation to failure typically requires approximately 70 separate simulations.

Due to the increasing complexity of the model and mesh refinement needs as the crack grows, the FEM simulations were divided into three stages: (a) For crack lengths in the range of 0–1.5 mm, the model remains relatively simple and each simulation takes approximately 20 min, totaling 30 simulations. (b) For the range of 1.5–2.5 mm, a finer mesh is required, increasing the simulation time to 40 min per step with 20 simulations needed. (c) When the crack length extends to 2.5–3.5 mm, the number of mesh elements increases significantly, and each simulation takes about 60 min, with another 20 simulations required.

Thus, the total FEM computation time for a single complete crack growth path amounts to 2600 min. Considering that this study includes 125 full crack growth cases, the cumulative simulation time exceeds 5000 h, which highlights the considerable computational burden of the conventional approach.

In contrast, the LSTM-based model learns from a pre-existing FEM database and requires only the initial 20 % of the crack growth data (approximately 0–0.7 mm) as input to predict the remaining 80 % of the new propagation path. Although this initial segment still requires FEM simulation, only about 14 steps are needed, with each step taking 20 min, resulting in a total of 280 min. On top of that, LSTM training takes approximately 11 min, and the prediction phase takes just 2.33 min. Therefore, the total time required to predict a single crack growth path using the LSTM framework is only 293.33 min, i.e. just 11.3 % of the time required by the FEM-only approach.

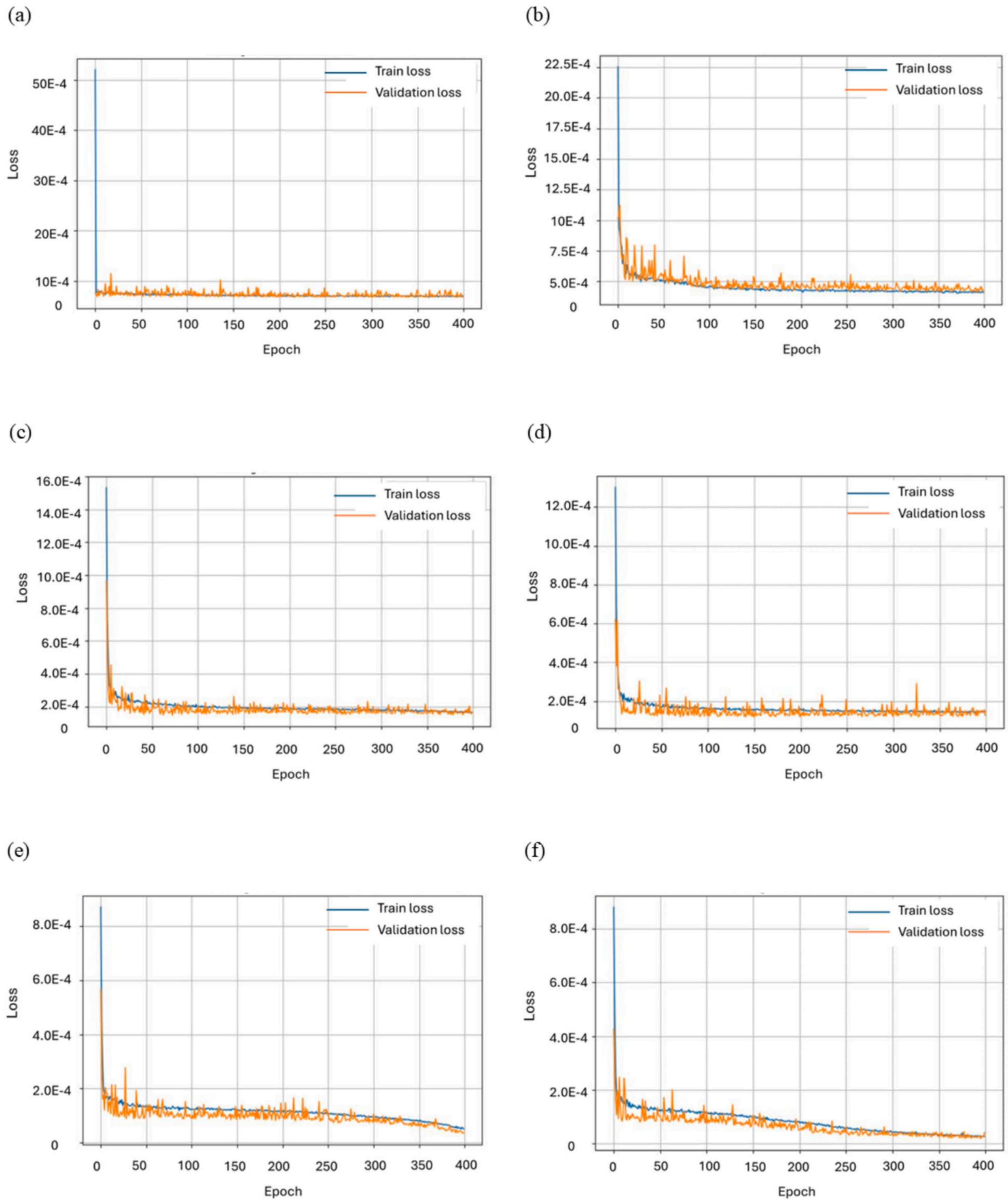

Fig. 11. Performance comparison of LSTM models with different sliding window values: (a) 1, (b) 2, (c) 3, (d) 4, (e) 5, and (f) 6.

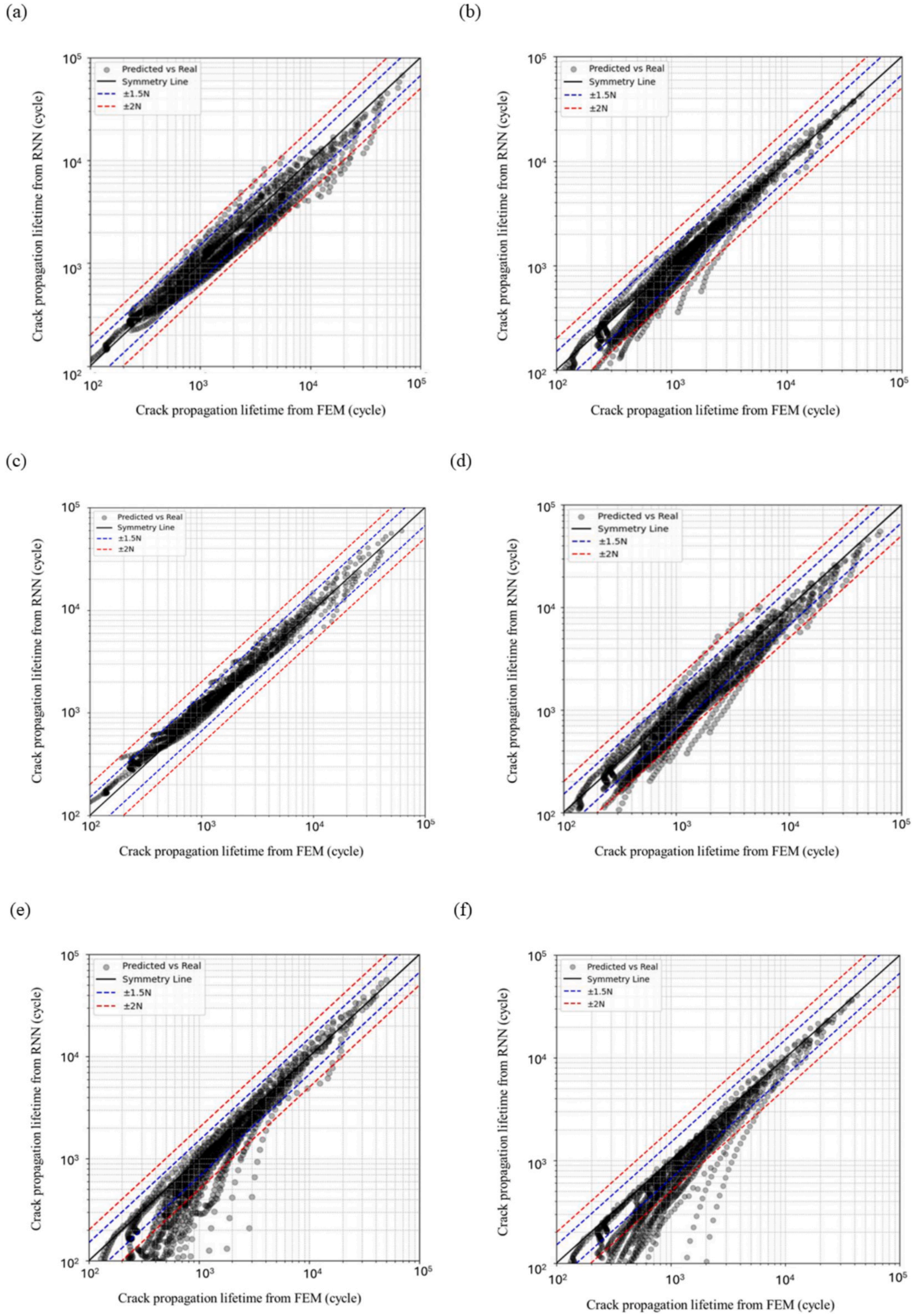

Fig. 12. Lifetime prediction comparison in different sliding window values: (a) 1, (b) 2, (c) 3, (d) 4, (e) 5, and (f) 6.

These results demonstrate that the proposed hybrid prediction framework significantly reduces computational cost while maintaining high prediction accuracy, making it highly suitable for large-scale parametric studies, fatigue life assessments under variable loading, and near real-time applications in structural health monitoring.

6.4. Impact of sliding window size on LSTM performance

To further evaluate the influence of the sliding window size on the performance of the LSTM model, we tested five different values (window sizes = 1, 2, 3, 4, 5). The sliding window determines the number of historical time steps used as input to predict the next point in the sequence, thereby affecting the model's ability to capture temporal dependencies in crack growth behavior. In general, larger window sizes enable the model to learn longer-term patterns, while smaller windows offer faster training and lower risk of overfitting. Fig. 11 illustrates the convergence behavior of the model under different window sizes, while Fig. 12 reflects the impact of window size on prediction accuracy.

Fig. 11 illustrates how varying the sliding window size impacts the model's performance, providing insights into the optimal window selection for improved predictive accuracy.

In the first few graphs ((a), (b), (c) and (d)), the loss drops rapidly within the initial 50–100 epochs, indicating fast convergence. However, as the sliding window increases (especially in (e) and (f)), convergence slows down, meaning the model requires more epochs to stabilize. This suggests that larger windows make the training process more complex, as the model has to process more dependencies.

Although larger sliding windows ((e) and (f)) achieve lower loss values overall, more fluctuations are observed towards the end. This suggests that while more data is captured, the model's weights are still being adjusted, and full stabilization has not yet been reached. In contrast, smaller windows ((a), (b), (c), and (d)) exhibit much faster convergence, with (a) stabilizing the quickest. Similar convergence speeds are observed in (b), (c), and (d), all reaching stability around 100 epochs. However, in (e) and (f), the convergence process is significantly slower, and stabilization is not fully achieved even after 400 epochs, indicating that additional training time may be required.

From Fig. 12, it is evident that in predicting crack growth life using LSTM, the choice of sliding window value significantly affects the prediction accuracy. Through extensive experiments and data analysis, it is found that different sliding window values lead to varying levels of prediction accuracy.

When the sliding window value is set to 3 (see Fig. 12 (c)), the prediction shows the highest accuracy. However, when the sliding window value is greater than 3, the performance of the prediction becomes more complex. For long-cycle crack growth scenarios, the prediction can still maintain a certain level of accuracy. But in the case of short-cycle crack growth, the predicted values gradually deviate from the actual values and become lower.

Specifically, as the crack growth progresses, especially in the later stage of crack growth, the prediction trend captured by LSTM with a larger sliding window is not as accurate as that with a smaller sliding window. This may be because a larger sliding window smooths out some of the short-term and local features of crack growth data, causing the model to lose sensitivity to the rapid changes in the short-cycle crack growth process.

7. Conclusions

This study utilized LSTM for the prediction of fretting fatigue crack propagation lifetime. In addition to the available experimental data, FEM is used to acquire additional fretting fatigue crack propagation data. The following conclusions are drawn.

1. DNN demonstrated limited predictive capability for crack propagation behavior, showing substantial deviations. This limitation stems from its architectural constraints, i.e. despite processing numerous input parameters, the single-output structure impedes effective learning of complex parameter relationships.
2. LSTM exhibited superior performance in capturing the correlation between crack length and lifetime in fretting fatigue crack propagation scenarios.
3. LSTM achieved exceptional accuracy in predicting fretting fatigue crack propagation behavior. When coupled with FEM, this hybrid approach (FEM-LSTM) substantially reduced computational time for crack propagation lifetime estimation, achieving a remarkable 91.04 % reduction in processing time.
4. The selection of LSTM's sliding window parameter significantly influences model performance. Concerning convergence efficiency, smaller window sizes facilitate faster stabilization of the loss function, while larger windows prolong convergence time. Regarding predictive accuracy, smaller window sizes, particularly the three values employed in this study, effectively capture short-term and localized crack growth characteristics, yielding enhanced accuracy, especially during later crack propagation stages.

This paper presents an effective method for predicting fretting fatigue crack propagation life under various loading conditions while minimizing computational costs. With a sufficiently large experimental dataset, this approach can be further enhanced by incorporating more physically representative data, thereby reducing the time required for experimental testing.

CRediT authorship contribution statement

Can Wang: Conceptualization, Methodology, Investigation, Writing – original draft. **Qiqi Xiao:** Data curation, Investigation.

Yingjian Guo: Methodology, Validation. **Gang Chen:** Supervision, Writing – review & editing. **Lihua Wang:** Supervision, Writing – review & editing. **Magd Abdel Wahab:** Conceptualization, Investigation, Writing – review & editing, Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors would like to acknowledge the financial support of the grants from the China Scholarship Council (202008130124). This work is supported by the National Natural Science Foundation of China (Project no. 12272270) and the Shanghai Pilot Program for Basic Research.

Data availability

Data will be made available on request.

References

[1] Llavori I, Zabala A, Mendiguren J, Gómez X. A coupled 3D wear and fatigue numerical procedure: application to fretting problems in ultra-high strength steel wires. Int J Fatigue 2021;143:106012.
[2] Venugopal Poovakaud V, Jiménez-Peña C, Talemi R, Coppieters S, Debruyne D. Assessment of fretting fatigue in high strength steel bolted connections with simplified Fe modelling techniques. Tribol Int 2020;143:106083.
[3] Croccolo D, De Agostinis M, Fini S, Olmi G, Robusto F, Scapecchi C. Fretting fatigue in mechanical joints: a literature review. Lubricants 2022.
[4] Ozguc O. Analysis of fatigue behaviour of drill pipe on pin-box connection. Proc Instit Mech Eng, Part M: J Eng Maritime Environ 2020;235(1):68–80.
[5] Vadiraj A, Kamaraj M. Effect of surface treatments on fretting fatigue damage of biomedical titanium alloys. Tribol Int 2007;40(1):82–8.
[6] Bhatti NA, Abdel Wahab M. Fretting fatigue crack nucleation: a review. Tribol Int 2018;121:121–38.
[7] Wang C, Pereira K, Wang D, Zinovev A, Terentyev D, Abdel Wahab M. Fretting fatigue crack propagation under out-of-phase loading conditions using extended maximum tangential stress criterion. Tribol Int 2023;187:108738.
[8] Pinto AL, Cardoso R, Talemi R, Araújo JA. Fretting fatigue under variable amplitude loading considering partial and gross slip regimes: numerical analysis. Tribol Int 2020;146:106199.
[9] Nikfam MR, Zeinoddini M, Aghebati F, Arghaei AA. Experimental and XFEM modelling of high cycle fatigue crack growth in steel welded T-joints. Int J Mech Sci 2019;153–154:178–93.
[10] Han Q, Lei X, Su Y, Rui S, Ma X, Cui H, et al. In-situ observation and finite element analysis of fretting fatigue crack propagation behavior in 1045 steel. Chin J Aeronaut 2021;34(11):131–9.
[11] Yan H, Yu H, Zhu S, Yin Y, Guo L. Machine learning based framework for rapid forecasting of the crack propagation. Engng Fract Mech 2024;307:110278.
[12] Martínez ER, Chakraborty S, Tesfamariam S. Machine learning assisted stochastic-XFEM for stochastic crack propagation and reliability analysis. Theor Appl Fract Mech 2021;112:102882.
[13] Liao W, Long X, Jiang C. A physics-informed neural network method for identifying parameters and predicting remaining life of fatigue crack growth. Int J Fatigue 2025;191:108678.
[14] Mienye ID, Swart TG, Obaido G. Recurrent neural networks: a comprehensive review of architectures, variants, and applications. Information 2024.
[15] Nguyen-Le DH, Tao QB, Nguyen V-H, Abdel-Wahab M, Nguyen-Xuan H. A data-driven approach based on long short-term memory and hidden Markov model for crack propagation prediction. Engng Fract Mech 2020;235:107085.
[16] Tran TV, Nguyen-Xuan H, Zhuang X. Investigation of crack segmentation and fast evaluation of crack propagation, based on deep learning. Front Struct Civ Engng 2024;18(4):516–35.
[17] Martín V, Vázquez J, Navarro C, Domínguez J. Fretting-fatigue analysis of shot-peened Al 7075-T651 test specimens. Metals 2019;9(5).
[18] Hojjati-Talemi R, Abdel Wahab M, De Pauw J, De Baets P. Prediction of fretting fatigue crack initiation and propagation lifetime for cylindrical contact configuration. Tribol Int 2014;76:73–91.
[19] Vázquez J, Navarro C, Domínguez J. Experimental results in fretting fatigue with shot and laser peened Al 7075-T651 specimens. Int J Fatigue 2012;40:143–53.
[20] Giner E, Sabsabi M, Ródenas JJ, Javier Fuenmayor F. Direction of crack propagation in a complete contact fretting-fatigue problem. Int J Fatigue 2014;58:172–80.
[21] Pereira K, Abdel Wahab M. Fretting fatigue crack propagation lifetime prediction in cylindrical contact using an extended MTS criterion for non-proportional loading. Tribol Int 2017;115:525–34.
[22] Erdogan F, Sih G. On the crack extension in plates under plane loading and transverse shear. J Basic Engng 1963;85(4):519–27.
[23] Ribeaucourt R, Baietto-Dubourg M, Gravouil A. A new fatigue frictional contact crack propagation model with the coupled X-FEM/Latin method. Comput Methods Appl Mech Engng 2007;196(33–34):3230–47.
[24] Rice JR. A path independent integral and the approximate analysis of strain concentration by notches and cracks. J Appl Mech 1968;35(2):379–86.
[25] Yau JF, Wang SS, Corten HT. A mixed-mode crack analysis of rectilinear anisotropic solids using conservation laws of elasticity. Int J Fract 1980;16(3):247–59.
[26] Moran B, Shih C. Crack tip and associated domain integrals from momentum and energy balance. Engng Fract Mech 1987;27(6):615–42.
[27] Zerres P, Vormwald M. Review of fatigue crack growth under non-proportional mixed-mode loading. Int J Fatigue 2014;58:75–83.
[28] Bold PE, Brown MW, Allen RJ. A review of fatigue crack growth in steels under mixed-mode I and II loading. Fatigue Fract Engrg 1992;15:965–7.
[29] Hourlier F. Fatigue crack path behavior under polymodal fatigue. Multiaxial Fatigue 1982.
[30] Liu L. Modeling of mixed-mode fatigue crack propagation. Vanderbilt University; 2008.
[31] Pereira KFR. Numerical modelling of fretting fatigue crack initiation and propagation using extended finite element method and cyclic cohesive zone model. Ghent, Belgium: Ghent University, Faculty of Engineering and Architecture; 2019. Ph.D. dissertation.
[32] Wang C, Fan K, Li C, Abdel Wahab M. Prediction of the effect of shot peening residual stress on fretting fatigue behaviour. Int J Fatigue 2023;176:107909.
[33] Wang C, Li Y, Tran NH, Wang D, Khatir S, Wahab MA. Artificial neural network combined with damage parameters to predict fretting fatigue crack initiation lifetime. Tribol Int 2022;175:107854.

[34] Llugsi R, Yacoubi SE, Fontaine A, Lupera P. Comparison between Adam, AdaMax and Adam W optimizers to implement a Weather Forecast based on Neural Networks for the Andean city of Quito. In: IEEE Fifth Ecuador Technical Chapters Meeting (ETCM) 2021; 2021. p. 1–6.
[35] Wang C, Li C, Ling Y, Abdel Wahab M. Investigation on fretting fatigue crack initiation in heterogenous materials using a hybrid of multiscale homogenization and direct numerical simulation. Tribol Int 2022;169:107470.
[36] Mortazavi SNS, Ince A. An artificial neural network modeling approach for short and long fatigue crack propagation. Comput Mater Sci 2020;185:109962.