



Cite this: *RSC Adv.*, 2025, **15**, 15781

Fatigue life predictor: predicting fatigue life of metallic material using LSTM with a contextual attention model†

Hongchul Shin, ^{ab} Taeyoung Yoon ^{*b} and Sungmin Yoon ^{*b}

Low-cycle fatigue (LCF) data involve complex temporal interactions in a strain cycle series, which hinders accurate fatigue life prediction. Current studies lack reliable methods for fatigue life prediction using only initial-cycle data while simultaneously capturing both temporal dependencies and localized features. This study introduces a novel deep-learning-based prediction model designed for LCF data. The proposed approach combines long short-term memory (LSTM) and convolutional neural network (CNN) architectures with an attention mechanism to effectively capture the temporal and localized characteristics of stress-strain data from acquisition through a series of cycle strain-controlled tests. Among the models tested, the LSTM-contextual attention model demonstrated superior performance ($R^2 = 0.99$), outperforming the baseline LSTM and CNN models with higher R^2 values and improved statistical metrics. The analysis of attention weights further revealed the model's ability to focus on critical timesteps associated with fatigue damage, highlighting its effectiveness in learning key features from LCF data. This study underscores the potential of deep-learning-based methods for accurate fatigue life prediction in LCF applications. This study provides a foundation for future research to extend these approaches to diverse materials with varying fatigue conditions and advanced models capable of incorporating non-linear fatigue mechanisms.

Received 5th March 2025
 Accepted 5th May 2025

DOI: 10.1039/d5ra01578b
rsc.li/rsc-advances

1. Introduction

Accurate prediction of fatigue life is critical for ensuring the safe and efficient use of materials in various industrial applications.^{1–3} Metallic materials such as stainless steel, which is known for its high corrosion resistance and relative strength-strain ratio under extreme conditions, are widely used in industries such as power generation, petrochemicals, marine engineering, and aerospace.^{4–6} However, repeated cyclic loading can lead to fatigue failure, posing risks to structural integrity and operational reliability. Fatigue life prediction, especially for materials operating under extreme stress-strain conditions, remains a significant challenge owing to the complex mechanisms involved in failure processes.^{7,8}

Recent advancements in material modeling have aimed to improve the accuracy of fatigue life predictions.⁹ Many studies have used cumulative data from multiple loading cycles to predict fatigue life, often relying on empirical relationships or machine learning models.^{10–12} While these approaches provide

valuable insights, they face limitations in capturing local stress-strain behavior during high strain cycles, which are crucial for understanding material failure. In addition, these models often require extensive datasets, which may not always be available in industrial settings.¹³

Historically, empirical models have been used to relate fatigue life to mechanical properties such as damage parameters under monotonic loading.^{10–12} However, these models often oversimplify the intricate relationships between cyclic loading, material microstructure, and fatigue behavior. Machine learning methods have shown potential to overcome these limitations by learning complex, nonlinear relationships directly from data.^{14,15} However, the interpretability of these models and their reliance on large datasets are significant obstacles.

Convolutional Neural Networks (CNNs) have been associated with image analysis but have also demonstrated effectiveness in time-series data analysis.^{16,17} Unlike Recurrent Neural Networks (RNNs), which process data sequentially and rely on hidden states to capture dependencies, CNNs utilize convolutional layers to extract local patterns and features directly from the input data.¹⁸ This approach allows CNNs to identify localized temporal patterns, such as stress or strain peaks within a cycle, without being constrained by sequential processing.

A key advantage of CNNs over RNNs is their resilience to issues like the vanishing gradient problem.¹⁹ In RNNs,

^aDepartment of Mechanical Engineering, Korea University, Seoul, 02841, Republic of Korea

^bDepartment of Mechanical Engineering, Changwon National University, Changwon, 51140, Republic of Korea. E-mail: tyoonyo@changwon.ac.kr; yoonsm@changwon.ac.kr

† Electronic supplementary information (ESI) available. See DOI: <https://doi.org/10.1039/d5ra01578b>



gradients can diminish exponentially during backpropagation over time, leading to the loss of information from earlier timesteps and limiting the model's ability to capture long-term dependencies. In contrast, CNNs leverage hierarchical feature extraction through multiple convolutional layers, enabling the model to capture both short- and long-range temporal dependencies effectively. Additionally, the use of pooling layers in CNNs helps reduce the dimensionality of the data while preserving critical features, further mitigating the computational challenges associated with deep network architectures.²⁰ The long Short-Term Memory (LSTM) networks address this limitation by introducing cell states and gating mechanisms, specifically designed to retain important information over extended sequences.^{21,22} The forget gate determines which information to discard, the input gate decides which information to add to the cell state, and the output gate regulates the flow of information to the next time step. These mechanisms allow LSTM networks to prioritize and maintain relevant temporal features, making them more robust in modeling long-term dependencies compared to standard CNNs.²³ In the context of fatigue life prediction, where stress-strain behavior over time plays a critical role, LSTMs provide an effective way to capture the progression of material degradation within a cycle.

Furthermore, the addition of contextual attention mechanisms enhances LSTM networks by focusing on the most critical temporal features within a sequence while maintaining the data's temporal order.²⁴ First, unlike self-attention, which models all pairwise interactions within a sequence, contextual attention respects the sequential structure of time-series data.^{25,26} This is particularly important in fatigue life analysis, where the cyclic stress-strain relationship depends on the ordered progression of states within a single cycle. Second, contextual attention emphasizes localized features, such as stress or strain peaks, in their immediate context.²⁷ These points, which are highly indicative of fatigue life behavior, are often overlooked by self-attention's distributed focus across the entire sequence.²⁸ Third, self-attention mechanisms are better suited for tasks requiring global dependencies, such as natural language processing, where temporal order is less critical.^{29,30} However, in fatigue life analysis, the localized interactions within the stress-strain cycle, particularly at extreme conditions, are the key drivers of material failure and life prediction.³¹ Fourth, the use of contextual attention directly aligns with the nature of the dataset by highlighting these critical localized dynamics while preserving the temporal flow.^{32,33} This makes contextual attention a better fit for predicting fatigue life from a single cycle, where local stress-strain patterns are the most significant contributors to accurate life predictions.³⁴

Recent advancements in fatigue life prediction of metallic materials have leveraged machine learning techniques to address the challenges of accurately reflecting mechanical properties and complex loading conditions. Chen *et al.* proposed a multi-view neural network model that integrates frequency domain analysis with CNN and LSTM to predict the fatigue life of metallic materials under multiaxial loading, demonstrating robust performance across diverse experimental datasets.¹¹ Similarly, Min *et al.* explored the interplay between

processing parameters, post-treatment conditions, and fatigue properties of additively manufactured metals, utilizing machine learning models, including CNNs, to predict fatigue life and crack growth rates.¹² Moon *et al.* focused on the influence of surface roughness and pore characteristics on the fatigue life of laser powder bed fusion Ti-6Al-4V alloys, employing neural networks to establish quantitative relationships with high predictive accuracy.¹⁰

Unlike other methods that rely on large datasets with multiple cycles or various material parameters, our model takes a unique approach by using a contextual attention mechanism and focusing on data from just the initial single cycle.³² The contextual attention mechanism improves accuracy and interpretability by highlighting the most important features within the initial cycle, helping the model identify early signs of fatigue that might be missed when analyzing many cycles. Using only the initial cycle makes the model faster and more efficient, as it doesn't need to process large datasets, while still capturing critical material properties like microstructure or surface conditions. However, our study is the first to apply a contextual attention mechanism for fatigue life prediction, highlighting its novelty and contribution to the field.

In this study, we propose an LSTM-based fatigue life prediction framework that incorporates a contextual attention mechanism. Empirical fatigue dataset was used to assess our prediction model. Unlike traditional models that require multiple cycles of data, this approach predicts fatigue life using only the stress-strain data from a single initial cycle. By focusing on local stress-strain interactions during extreme strain events, the model provides improved accuracy while maintaining interpretability. To ensure accessibility and reproducibility, the model is deployed as a Python, allowing seamless integration into industrial workflows. This study sheds light on the AI-based fatigue field with predicting final fracture from experimental data.

2. Methods

2.1. LCF experiments

The dataset used in this study consists of stress-strain data for Type 316 stainless steel, obtained from low-cycle fatigue (LCF) experiments (Fig. 1).³⁵ The material used was Type 316 austenitic stainless steel, manufactured according to JIS G 4305 standards.³⁴ Table 1 presents the chemical composition range for the material.³⁵ All specimens underwent full annealing in a heat treatment furnace at 900 °C. The annealing process involved raising the temperature over 4 hours, maintaining it for 10 minutes, and allowing the material to cool in the furnace. Following the annealing process, surface treatment was performed to ensure the reliability of the stress-strain dataset obtained from strain-controlled fatigue tests.³⁶ Specifically, the specimens were mechanically polished with emery papers of grades ranging from 180 to 1000 and then finished with a diamond suspension (DiaPro 3 µm, Struers) to obtain a uniform and smooth surface.

The experiments followed the ASTM E606-21 Standard Test Method for Strain-Controlled Fatigue Testing and compared



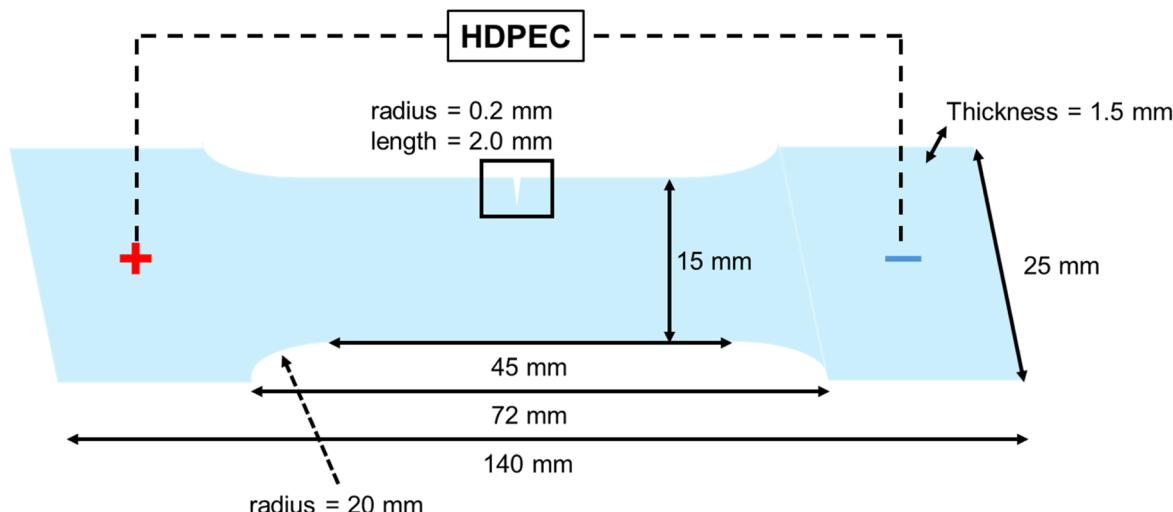


Fig. 1 Three-dimensional schematic showing the geometry of the specimens.

Table 1 Chemical composition of the sample material in wt%

| Fe | C | Si | Mn | P | S | Ni | Cr | Mo |
|---------|-------|-------|-------|--------|-------|-------------|-------------|-----------|
| Balance | ≤0.08 | ≤1.00 | ≤2.00 | ≤0.045 | ≤0.03 | 10.00–14.00 | 16.00–18.00 | 2.00–3.00 |

two treatment types: (1) Type 316 stainless steel subjected to heat treatment only and (2) Type 316 stainless steel subjected to additional high-density pulsed electric current (HDPEC) processing after heat treatment.^{7,37} Dogbone-shaped fatigue specimens, with a thickness of 1.5 mm and a shape radius of 20 mm, were prepared, as shown in Fig. 1. A notch with a radius of 0.2 mm and length of 2 mm was machined at the center of one side to induce mode I fatigue cracking.³⁶ The study focused on evaluating one-way cracking to assess the effects of HDPEC on damage evolution during fatigue. An electric current was applied to the HDPEC-treated specimens, and their fatigue behavior during crack initiation and propagation was compared to that of heat-only specimens.¹³

Each sample in the dataset represents a single cycle with 50 timesteps, where stress and strain values were recorded at regular intervals.^{13,38,39} The dataset spans various strain ranges (ϵ_{\min} to ϵ_{\max}), including ranges such as [0.18–0.36], [0.27–0.54], [0.198–0.396], [0.216–0.432], [0.234–0.468], [0.252–0.504], [0.108–0.216], [0.126–0.252], [0.144–0.288], and [0.162–0.324]. Each sample also includes its fatigue life (N_f), defined as the number of cycles until material failure, which ranges from 2000 to 90 000 cycles. Data with incomplete cycles (cycle 1) or irregular stress-strain behavior were excluded to ensure consistency and reliability in the analysis.

2.2. Dataset and data preprocessing

To prepare the dataset for modeling, stress and strain values were normalized using z-score normalization to standardize their distributions.⁴⁰ This transformation ensures that the stress and strain values have a mean of zero and a standard deviation of one, as calculated by:

$$z = \frac{x_{\text{data}} - u}{\text{SD}} \quad (1)$$

where x_{data} is the raw value, u is the mean, and SD is the standard deviation.

The fatigue life value N_f was normalized separately using MinMax scaling, which scaled N_f values to the range [0, 1] according to:

$$N_f^{\text{scaled}} = \frac{N_f - N_f^{\min}}{N_f^{\max} - N_f^{\min}} \quad (2)$$

where N_f^{\min} and N_f^{\max} correspond to the minimum and maximum fatigue life values in the dataset (2000 and 90 000 cycles, respectively).

K-fold cross-validation and stratified K-fold cross-validation differ in how they handle randomization and data distribution. K-fold relies on random shuffling to ensure diversity in each fold, whereas stratified K-fold explicitly maintains the proportional representation of features or labels across folds, making it useful for datasets with imbalanced distributions because it prevents underrepresented groups from being excluded. However, K-fold may introduce an imbalance unless the dataset is uniform. In this study, a stratified K-fold was initially applied to preserve the distribution of material treatments and strain ranges. Despite this, K-fold with shuffling ultimately provided superior predictive accuracy and stability because the added diversity from randomization enhanced the model's ability to generalize to unseen data.⁴¹ Therefore, K-fold cross-validation with shuffling was chosen for constructing the final model.

For K-fold cross-validation, the dataset D was divided into k folds, denoted as $D = D_1, D_2, \dots, D_k$. For each fold i , training set

T_i was defined as $T_i = D - D_i$ and the validation set V_i was defined as $V_i = D_i$. The model was trained on T_i and evaluated on V_i , and the performance scores across all folds were aggregated to compute the overall validation score:

$$\text{Score}_{\text{K-fold}} = \frac{1}{k} \sum_{i=1}^k \text{Score}(V_i) \quad (3)$$

where $\text{Score}(V_i)$ represents the performance metric evaluated on the validation set V_i .

2.3. Mechanical analysis of feature relevance

To evaluate the plastic deformation energy (E_{plastic}) per cycle, the area of the stress-strain hysteresis loop was calculated by integrating the loop curve.⁴³

$$E_{\text{plastic},n} = \int_{\varepsilon_{\min}}^{\varepsilon_{\max}} \sigma d\varepsilon \quad (4)$$

where $E_{\text{plastic},n}$ is the plastic deformation energy for cycle n (in J units), σ is the stress value, ε is the strain value, ε_{\max} and ε_{\min} are maximum and minimum strain values, respectively.

The cumulative energy $E_{\text{cumulative}}$ absorbed by the material across all cycles was calculated as the sum of the plastic deformation energy for all cycles

$$E_{\text{cumulative}} = \sum_{n=1}^{N_{\text{cycles}}} E_{\text{plastic},n} \quad (5)$$

where N_{cycles} is the total number of cycles for the material in the dataset.

To analyze the behavior of stress and strain over time, the differences and variances of stress and strain values at each timestep were computed. These metrics provide insights into the dynamic changes and variability of stress and strain throughout the fatigue cycles.

The differences in stress and strain between consecutive timesteps were calculated as:

$$\Delta\sigma_t = \sigma_{t+1} - \sigma_t \quad (6)$$

$$\Delta\varepsilon_t = \varepsilon_{t+1} - \varepsilon_t \quad (7)$$

where σ_t and ε_t represent stress and strain values at timestep t , respectively. The mean differences across all samples were visualized to identify trends over the timesteps.

To evaluate the variability in stress and strain at each time step, we computed the variance across all samples as follows:

$$\text{Var}(\sigma_t) = \frac{1}{N_{\text{samples}}} \sum_{i=1}^n (\sigma_{i,t} - \bar{\sigma}_t)^2 \quad (8)$$

$$\text{Var}(\varepsilon_t) = \frac{1}{N_{\text{samples}}} \sum_{i=1}^n (\varepsilon_{i,t} - \bar{\varepsilon}_t)^2 \quad (9)$$

where N_{samples} is the number of samples, $\bar{\sigma}_t$ and $\bar{\varepsilon}_t$ are the mean stress and strain values at timestep t respectively.

Correlation analysis and flattened principal component analysis (PCA) were conducted to assess whether the time-series stress and strain data could effectively predict the fatigue life N_f .

The correlation analysis was aimed at identifying the timesteps in the stress-strain cycle that showed a significant relationship with N_f , while flattened PCA was used to evaluate whether the dataset could be meaningfully separated according to N_f values.

The purpose of the temporal correlation analysis was to examine the relationship between N_f and each feature at individual timesteps, identifying specific points in the cycle that are most predictive of fatigue life.⁴³ To accomplish this, we computed a temporal correlation coefficient at each timestep, denoted as r_t to capture the linear association between N_f and the stress or strain values at timestep t . This approach allows us to quantify the relationship between N_f and each specific point within the cycle. The temporal correlation coefficient r_t for a given timestep t is defined as:

$$r_t = \frac{\sum_{i=1}^{N_{\text{samples}}} (x_{\text{data},i,t} - \bar{x}_t)(N_{f,i} - \bar{N}_f)}{\sqrt{\sum_{i=1}^{N_{\text{samples}}} (x_{\text{data},i,t} - \bar{x}_t)^2} \cdot \sqrt{\sum_{i=1}^n (N_{f,i} - \bar{N}_f)^2}} \quad (10)$$

where $x_{\text{data},i,t}$ is the stress or strain value of sample i at time step t , \bar{x}_t is the mean of the stress or strain values across all samples at time step t , $N_{f,i}$ is the fatigue life of sample i , and \bar{N}_f is the mean fatigue life across all samples.

The purpose of the flattened PCA was to determine if N_f values could be meaningfully separated in the reduced dimensional space, thus suggesting that the stress-strain data contain patterns indicative of different fatigue life outcomes.⁴⁴ In this approach, each stress-strain cycle was flattened into a single vector, capturing the entire cycle's stress and strain values. PCA was then applied to these flattened vectors to reduce dimensionality and extract the principal components that account for the most variance in the dataset.

The first two principal components (PC_1 and PC_2) were plotted against each other, with each point colored according to its N_f value.⁴⁵ This visualization revealed potential clusters in the data, showing whether samples with similar N_f values tended to group together in PCA space. Such clustering would indicate that the primary components capture significant features of the stress-strain data that are relevant to fatigue life prediction. The transformation for each principal component PC_k is given by:

$$\text{PC}_k = \sum_{j=1}^p w_{kj} \cdot x_j \quad (11)$$

where w_{kj} are the weights of each original variable, x_j for the j -th principal component. This projection allowed us to visualize the separation of N_f values in the PCA space.

2.4. Model architecture

To build the prediction framework, several neural network architectures were explored, including CNN, LSTM, CNN with attention, and LSTM with attention.⁴⁶ The CNN and LSTM models were implemented to evaluate their ability to capture sequential dependencies in the cyclic stress-strain data recorded by a series of strain-controlled fatigue tests. We then



focused on the stress-strain datasets to reach our goal. However, these models treat all timesteps equally, limiting their focus to critical regions such as peak stress or strain points. To address this limitation, attention mechanisms were integrated into the CNN and LSTM architectures (Fig. 2 and 3). These mechanisms assign weights to timesteps based on their contribution to fatigue life prediction, effectively emphasizing key features while reducing the influence of less critical points. The contextual attention mechanism used in this study was designed to highlight local patterns within the cyclic data, making it particularly effective in capturing fatigue-related behaviors.

To systematically identify the optimal architecture for each model, we performed hyperparameter optimization tailored to both the LSTM and CNN based frameworks.⁴⁷ For the LSTM with contextual attention model, we utilized Optuna, a Bayesian optimization framework, to minimize validation MAE. The search space included the number of LSTM layers (ranging from 2 to 6), the number of units per layer (32 to 256), dropout rates (0.1 to 0.5), dense layer configurations (32 to 128 units), and learning rates (from 1×10^{-5} to 1×10^{-2} , log-uniform). The best-performing model consisted of five stacked LSTM layers with 128 units per layer, intermediate dropout of 0.3, a 64-unit dense layer, and an Adam optimizer with a learning rate of 0.001.

For the CNN-based models, we applied Keras Tuner with the Hyperband algorithm, which efficiently searched the space of candidate architectures. The hyperparameters included the number of convolutional layers (2 to 4), number of filters (16 to 128), kernel sizes (3 to 7), dense layer units (32 to 128), dropout

rates (0.1 to 0.4), and learning rate (1×10^{-4} to 1×10^{-2}). The final CNN model selected through this process employed three convolutional layers with kernel size 5, max pooling, 0.2 dropout, and a 64-unit dense layer.

Final model selection across all architectures was based on quantitative performance metrics, including R^2 , MAE, and RMSE averaged over five independent training iterations. As reported in ESI Table S2,† the LSTM-contextual attention model consistently demonstrated statistically significant improvements over all other models ($p < 0.05$), supporting its suitability for fatigue life prediction using single-cycle time series data.

The final model, which combined LSTM with contextual attention, consisted of five stacked LSTM layers followed by a custom attention layer. Each LSTM layer was accompanied by a dropout layer to mitigate overfitting. The attention mechanism computed weights for each timestep (α_t) as:

$$\alpha_t = \frac{\exp(e_t)}{\sum_{i=1}^T \exp(e_i)}, e_t = \tanh(Wx_t + b) \quad (12)$$

where x_t is the input vector at timestep t , and W and b are learnable weight matrix and learnable bias vector, respectively. The output of the attention layer ($h_{\text{attention}}$) was a weighted sum of inputs:

$$h_{\text{attention}} = \sum_{t=1}^T \alpha_t x_t \quad (13)$$

This mechanism enabled the model to focus on critical stress-strain regions within a single cycle.

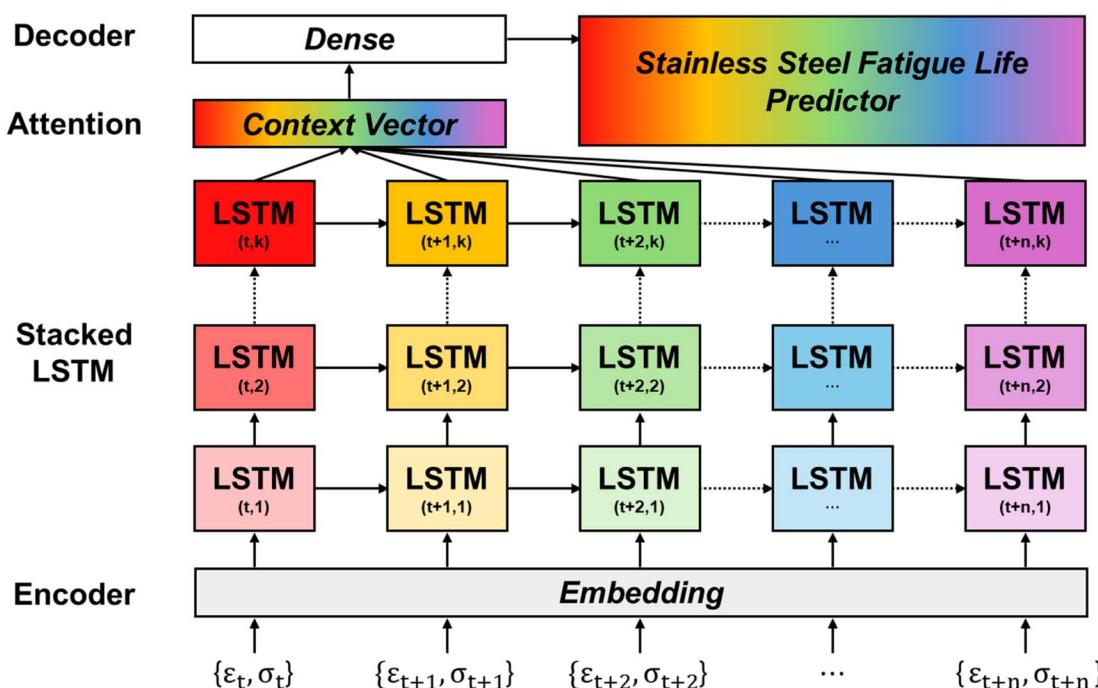


Fig. 2 Architecture of the LSTM-based fatigue life prediction model with an attention mechanism. Each LSTM block processes sequential stress-strain features, and the outputs are aggregated through an attention mechanism. The rainbow-colored context vector represents the attention-weighted sum of LSTM outputs, enabling the model to focus on the most informative timesteps.



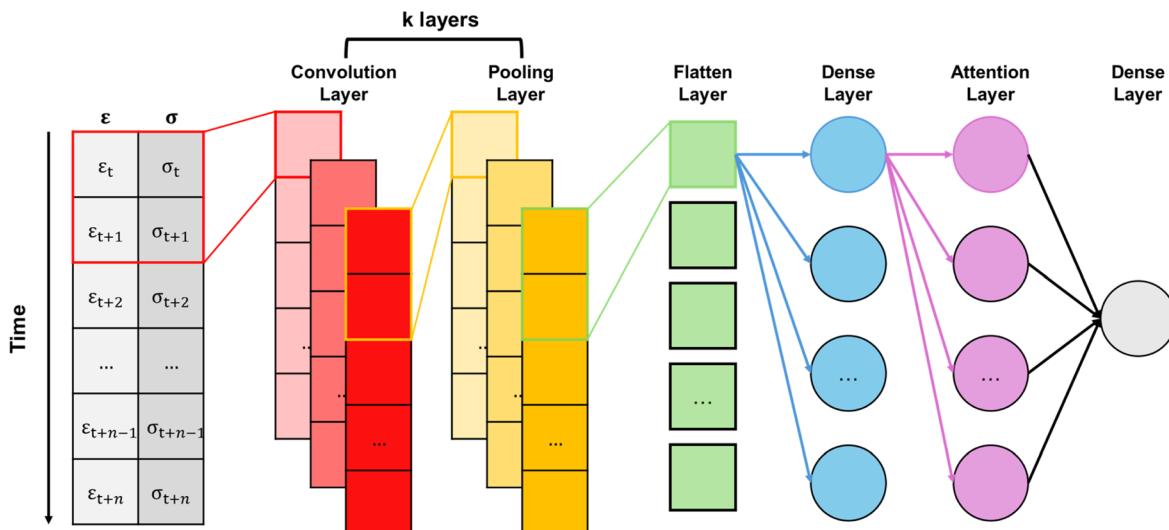


Fig. 3 Architecture of the CNN-based fatigue life prediction model with an attention mechanism. Sequential stress–strain inputs are passed through multiple convolutional layers to extract hierarchical features. The outputs are then aggregated via an attention mechanism, where the colored circular nodes represent the attention-weighted feature representation.

The coefficient of determination (R^2) measures how well the predicted values align with the true values and is defined as:¹⁴

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y}_i)^2} \quad (14)$$

where y_i is the true value, \hat{y}_i is the predicted value, and \bar{y}_i is the mean of the true values.

The mean absolute error (MAE), which quantifies the average absolute difference between the predicted and true values, is given as:⁴⁸

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (15)$$

The root mean square error (RMSE), a metric that penalizes larger errors more heavily, is defined as:⁴⁸

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (16)$$

3. Results and discussions

3.1. Plastic deformation energy and fatigue behavior

The changes in the plastic deformation energy with cycle progression were analyzed through the stress-strain hysteresis loops and cumulative energy comparisons. As shown in Fig. 4a, the hysteresis loop area is large during the initial cycles, particularly in cycle 1, where significant energy consumption and active plastic deformation are observed. However, the irregular behavior in cycle 1 reduced its reliability in the energy analysis. To address this, cycle 20 was selected as the starting point for dataset construction to calculate fatigue life (Fig. S1†).

With increasing cycle numbers, the hysteresis loop area progressively decreased, and the stress values declined as the material underwent fatigue damage. Near the fatigue failure, as observed in cycles 19 260 and 15 640, the loop area became minimal, indicating a sharp reduction in the energy required for plastic deformation. This trend suggests that material strength weakens over time owing to cumulative fatigue damage, and in the final cycles, the energy drops sharply, indicating the inability of the material to sustain further deformation.

The stress-strain envelope curves in Fig. 4b provide further insight into the fatigue progression. In cycle 20, the material exhibited the highest stress and strain values and consumed most of the deformation energy. However, in intermediate cycles, such as cycles 7820 and 9640, the stress values gradually decrease, reflecting the progressive strength reduction of the material.¹³ Near the end of fatigue life, the envelope curves narrowed significantly and the stress values became minimal, indicating that the material approached its failure point. This behavior highlights the limited capacity of the material to accommodate additional deformation as fatigue progresses.¹

To quantify the energy changes further, the cumulative plastic deformation energy was calculated by integrating the hysteresis loop areas. Fig. 5 compares the cumulative energies of the heat- and heat-HDPEC-treated specimens across various strain ranges. The results show that cumulative energy decreases with an increasing number of cycles, which is consistent with the trends observed in Fig. 4.

In low strain ranges, HDPEC-treated specimens exhibit higher cumulative energy compared to heat-treated specimens, suggesting that HDPEC treatment improves the material's ability to sustain deformation energy over a longer period. In contrast, at high strain ranges, cumulative energy decreases rapidly for both specimen types, indicating accelerated fatigue



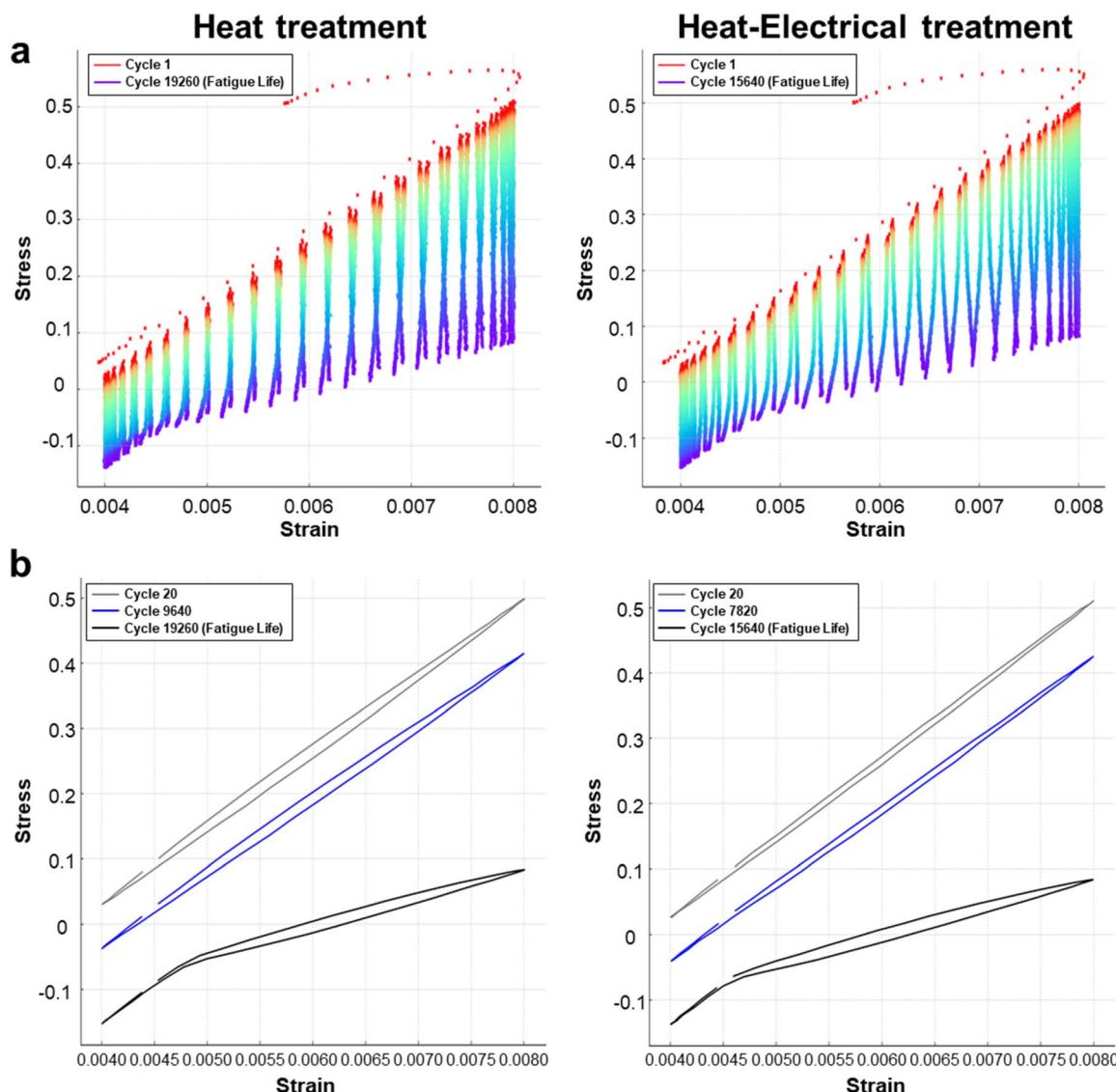


Fig. 4 Stress-strain behavior over cycles for type 316 stainless steel. (a) Stress-strain hysteresis loops for cycles at the beginning (cycle 1) and near fatigue life (cycle 19 260 and cycle 15 640). (b) Stress-strain envelope curves at intermediate cycles (cycle 7820, cycle 9640) and at fatigue life (cycle 19 260, cycle 15 640).

damage and shorter fatigue lives. Nevertheless, the HDPEC-treated specimens retain relatively higher cumulative energy levels compared to the heat-treated specimens, even in high strain ranges.

These results suggest that heat-HDPEC treatment enhances fatigue resistance by reducing the plastic deformation energy during cyclic loading.^{3,13} The improved energy retention in the heat-HDPEC-treated specimens indicates that the treatment delayed the onset of significant fatigue damage under low strain conditions. Conversely, under high-strain conditions, energy depletion occurs more rapidly, accelerating fatigue failure.

The combination of hysteresis loop analysis (Fig. 4) and cumulative energy comparison (Fig. 5) demonstrated that heat-HDPEC treatment improved the fatigue resistance of the material. The heat-HDPEC-treated specimens showed a slower decline in plastic deformation energy and sustained more

energy through later cycles than the heat-treated specimens. These findings confirm that heat-HDPEC treatment extends fatigue life by delaying fatigue damage accumulation and maintaining material strength over a greater number of cycles.

3.2. Statistical approach for fatigue life predicting AI model

The plots of the mean stress/strain differences and their variances across timesteps highlight the dynamic nature of the stress and strain during the fatigue cycles. However, these analyses alone do not provide clear conclusions or direct insights into their predictive relationship with the fatigue life. As shown in Fig. S2,† the raw data obtained without pre-processing resulted in low and irregular correlations, making it difficult to capture meaningful patterns. Therefore, data scaling is used to improve the analysis. Flattened temporal correlation analysis and PCA were performed to extract more



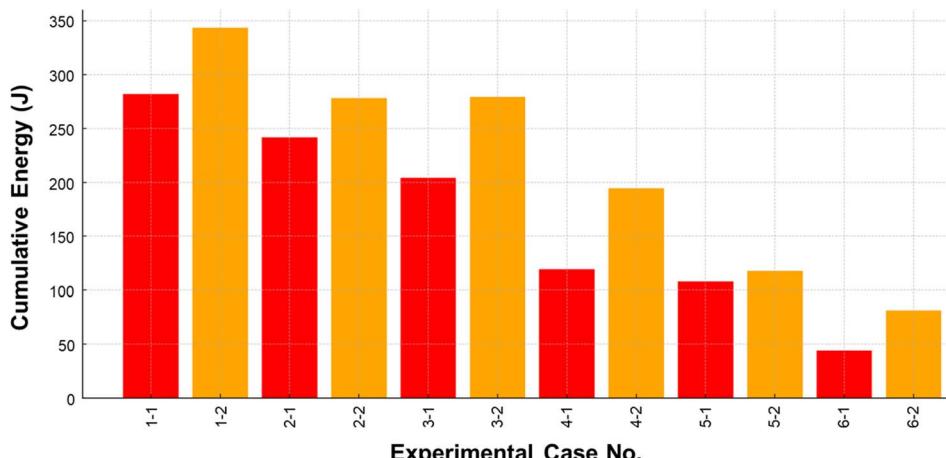


Fig. 5 Comparison of cumulative energy between heat-treated and heat-HDPEC-treated Type 316 stainless steel specimens across various strain ranges (Table S1, ESI dataset†). The red bars represent heat-treated specimens, while the orange bars represent heat-HDPEC-treated specimens.

meaningful patterns. Fig. 6 illustrates the relationship between the stress/strain differences and fatigue life across 50 timesteps, whereas Fig. 7 presents the PCA on the stress and strain time series data.

Fig. 6 shows that both the stress–fatigue life and strain–fatigue life correlations exhibit periodicity across timesteps, with significant peaks and troughs corresponding to the critical phases within the stress–strain cycles. Positive correlation peaks, nearing +1, highlight time intervals where changes in stress and strain are highly predictive of fatigue life, whereas negative correlation peaks, nearing -1, indicate intervals where these changes are inversely correlated with fatigue life. The similarity between the stress and fatigue life and strain–fatigue life correlation curves suggests that both features provide complementary and essential information for predicting fatigue life. These findings emphasize the necessity of capturing localized temporal dynamics within stress–strain cycles for accurate predictive modeling.

Fig. 7 shows the distribution of data under the two conditions. The left panel includes fatigue life values ranging from 0 to 20 000, whereas the right panel incorporates the entire dataset, spanning fatigue life values from 0 to 90 000. In the left panel, the data points align along a linear trend, as denoted by the red dashed lines, suggesting that the primary variance in the data is explained by the linear relationships between the first two principal components. This linearity indicates that the stress–strain patterns in the lower fatigue life regimes are uniform and predictable. In contrast, the right panel shows a more complex distribution when the full dataset is considered. While most data points conformed to the linear trend observed in the left panel, a distinct cluster emerged for fatigue life values exceeding 60 000, as highlighted by the red circle. This deviation suggests the presence of non-linear dynamics or unique stress–strain behaviors associated with high fatigue life.

Combining temporal correlation analysis (Fig. 6) and flattened PCA (Fig. 7) provided significant insights into the behavior of stress and strain data across different fatigue life

regimes. Fig. 6 emphasizes the critical importance of specific time intervals within the stress–strain cycles, supporting the necessity of temporal attention mechanisms in prediction models. Meanwhile, although majority of the data in Fig. 7 can be described by linear trends, high fatigue life cases exhibit distinct non-linear complexities, necessitating tailored modeling approaches for these regimes.

To apply these findings, both LSTM and CNN models were employed to predict the fatigue life, capturing the linear and non-linear dynamics revealed by Fig. 6 and 7, respectively. This combined approach aims to improve predictive accuracy by integrating the temporal and spatial features inherent in the stress–strain data.

3.3. Architecture and performance analysis of LSTM and CNN models for fatigue life prediction

Two types of deep learning architectures, LSTM and CNN, with and without attention mechanisms, were employed to capture the temporal and spatial characteristics of the stress–strain time series data and to predict the fatigue life under LCF conditions. The LSTM model (Fig. S3†) processes sequential stress and strain values through stacked LSTM layers to learn the long- and short-term dependencies. However, in the LSTM with attention model (Fig. 2), the attention mechanism enhances this capability by identifying and emphasizing critical timesteps, allowing the model to focus on the regions most relevant to fatigue initiation and progression.

Similarly, the CNN model (Fig. S4†) extracts localized spatial patterns from the input stress–strain data through convolutional layers. In a CNN with attention model (Fig. 3), the attention mechanism assigns importance to specific local features, further refining the predictive process. While the baseline CNN effectively captures local dependencies, it lacks the ability to model the long-term temporal relationships inherent in the LCF data.

The results demonstrated the superior performance of the LSTM with attention model for fatigue life prediction (Fig. 8 and



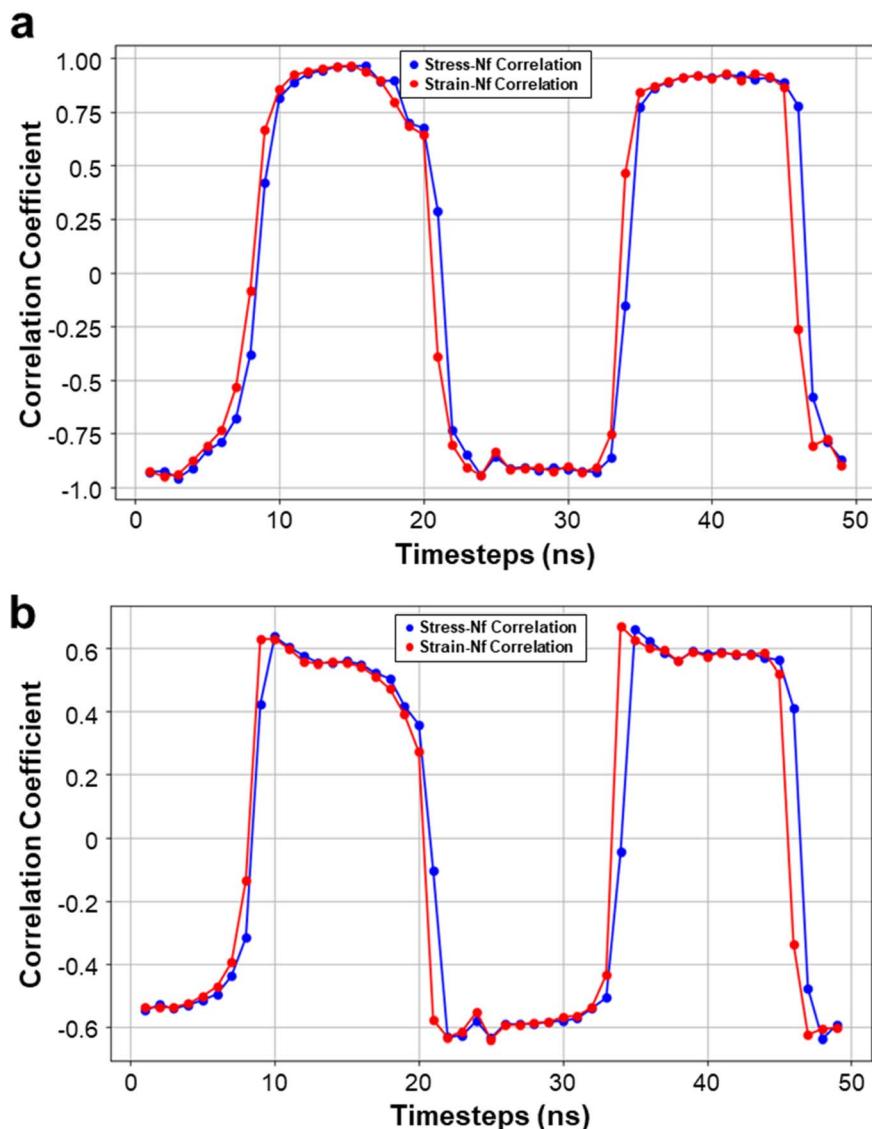


Fig. 6 Temporal correlation analysis between stress (blue), strain (red), and fatigue life. Correlation analysis for specimens with fatigue life in the range of (a) 0–20000 and (b) 0–80000 cycles.

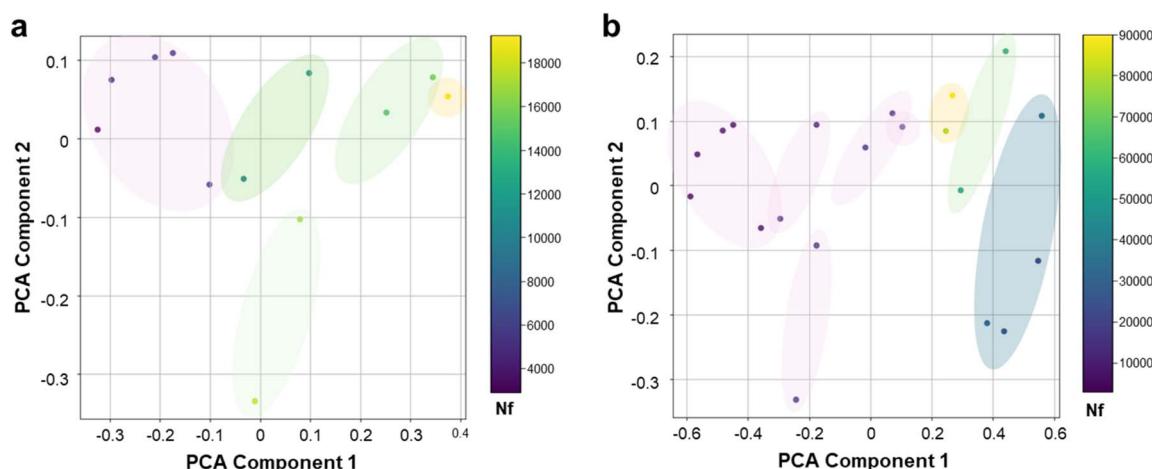


Fig. 7 Flattened PCA analysis of stress–strain data for type 316 stainless steel specimens. Analysis for specimens with fatigue life in the range of (a) 0–20000 cycles and (b) 0–80000 cycles.

S5†). Statistically significant improvements across all performance metrics (R^2 , MAE, and RMSE) were observed when comparing the LSTM with attention model to the baseline LSTM model (Table S2,† p -values: $R^2 = 0.0001$, MAE = 0.0003, and RMSE = 0.0000). The attention mechanism allows the model to focus on critical timesteps within the cyclic data, leading to improved predictive accuracy and reliability.

When compared with the CNN with attention model, the LSTM with attention also achieved statistically significant improvements (Table S2,† p -values: $R^2 = 0.0320$, MAE = 0.0474, RMSE = 0.0006). These results highlight the ability of LSTM to effectively capture the long- and short-term temporal dependencies inherent in the LCF data. While the attention mechanism benefits both models, the LSTM with attention better combines global temporal patterns and localized critical features, which are essential for accurate fatigue life prediction. By contrast, the addition of attention to CNN did not result in statistically significant improvements (Table S2,† p -values: $R^2 = 0.9379$, MAE = 0.2364, RMSE = 0.9999). CNN, which excels in local feature extraction, struggles to model the sequential nature of LCF data. Even with the inclusion of attention, CNN processes sequences as discrete segments, limiting its ability to capture broader temporal dependencies.⁴⁹

The LCF data exhibited intricate stress-strain interactions over multiple cycles, often involving nonlinear evolution. The LSTM with attention model effectively addresses this complexity by capturing long-term dependencies across cycles while focusing on key phases within each cycle. In contrast, the local feature extraction of CNN cannot fully leverage the sequential nature of LCF data, leading to a lower performance. These findings underscore the importance of selecting models that align with the inherent characteristics of data, particularly for time series problems such as low-cycle fatigue.⁵⁰ The

combination of LSTM and attention provides a robust solution for capturing both global and localized temporal patterns, making it well-suited for fatigue life prediction in LCF datasets.

3.4. Analysis of attention weights for fatigue life predictor

In this study, we employed only the stress-strain curve measured at the 20th cycle to determine whether a single stabilized loop could reliably predict fatigue life. Early cycles (1st–10th cycle) exhibited rapid strain hardening or softening, and the hysteresis loop typically stabilized within the 10th–20th cycle. Therefore, the 20th-cycle stress-strain curve represents the characteristic fatigue behavior of each specimen at a relatively steady state. Nevertheless, because we did not track the entire process up to the final fatigue stage, certain specimens displayed a wide loop with prominent plastic deformation (short-life specimens), whereas others remained narrow and more elastic (long-life specimens) (Table 2 and Fig. S1†). This contrast is especially evident when comparing short and long fatigue life specimens, and the LSTM with the contextual attention model detects these differences by focusing on distinct segments in the time or stress-strain domains. As recorded in the ESI dataset,† the test specimens exhibited fatigue lives ranging from a few thousand to more than 90 000 cycles. In particular, many heat-treated specimens reached short or moderate fatigue life values (under 20 000 cycle), whereas heat-HDPEC specimens could endure far longer fatigue life values, sometimes exceeding 90 000 cycles. Hence, even for the same 20th cycle dataset, the heat-HDPEC group may show either advanced plastic deformation or only minimal strain, depending on individual microstructural or surface enhancement effects.^{1,7,13,35}

Fig. 9 presents the attention weight distributions computed by our model when predicting the fatigue life solely from these

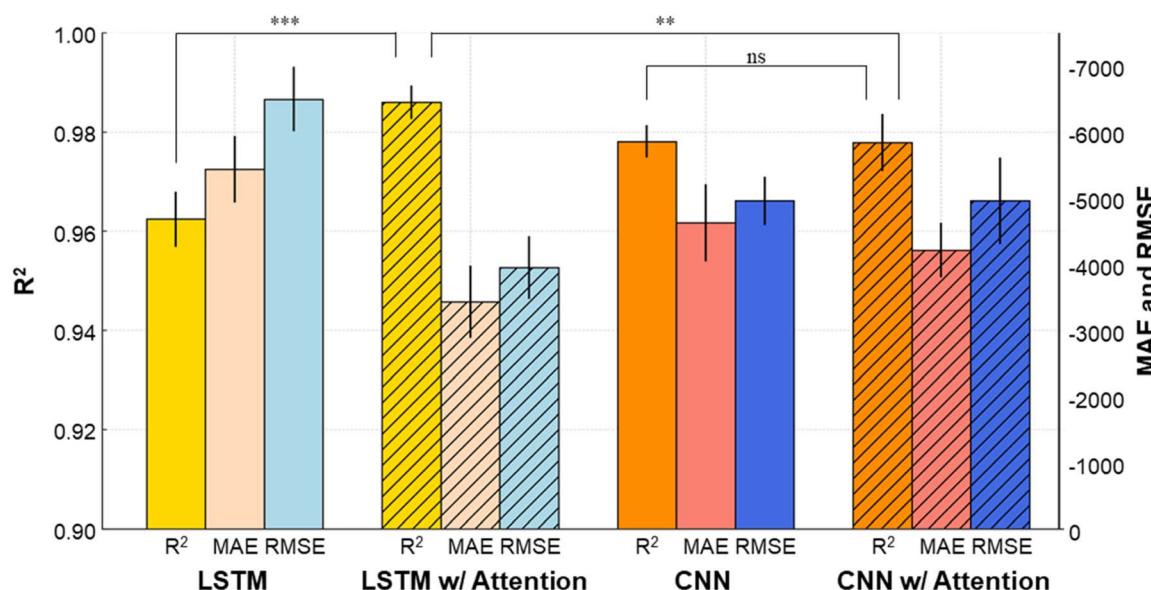


Fig. 8 Performance comparison of fatigue life prediction models (LSTM, LSTM with attention, CNN, and CNN with attention) based on R^2 , MAE, and RMSE metrics. Error bars represent the standard deviation calculated from the results of the 5 model training iterations (statistically significant differences are indicated with star annotations * $p < 0.1$, ** $p < 0.05$, and *** $p < 0.01$).



Table 2 Comparison of hysteresis loop characteristics and attention mechanism focus between short fatigue life and long fatigue life for heat-treated and heat-HDPEC-treated conditions

| | Short fatigue life | Long fatigue life |
|-------|--|---|
| Heat | Hysteresis loop tends to be quite wide, indicating substantial plastic deformation The attention mechanism primarily focuses on the tensile peak (red). | Even though some plasticity emerges, the attention remains mostly near the tensile peak The attention mechanism primarily focuses on the tensile peak (red). |
| HDPEC | Hysteresis loop is wide, but more local variations in deformation may appear The attention mechanism focuses not only tensile peak (red) but also transitions between tension and compression (purple). | Cyclic behavior in-between elastic and plastic regimes may prominent, in which plastic proportion lost its effect in this mechanical response The attention mechanism focuses transitions between tension and compression, not peak stress |

20th-cycle hysteresis loops, where the red line denotes the global maximum attention, the blue line denotes the global minimum, and the purple line denotes the local maximum. The red line appears near the tensile peak or regions where the plastic strain is prominent, particularly in the short-life specimens. For heat-HDPEC specimens with a long fatigue life, attention often shifts toward the midrange stress or strain, reflecting the cyclic behavior dominated by elastic responses and the reduced influence of plastic deformation. Such specimens tended to exhibit pronounced plastic deformation relatively early, causing the model to devote maximum attention. The blue line represents regions of low informational value,

occurring in near-elastic unloading segments or areas with minimal stress-strain variations. While it can appear in both short- and long-life materials under either heat or heat-HDPEC treatment, it is often more noticeable in long-life specimens that have not accumulated substantial plasticity and therefore have extended elastic-like intervals. The purple line represents additional local maxima that are not as dominant as the red line and can appear near transition points, such as tension-compression transitions, or when minor asymmetry or ratcheting emerges. Notably, heat-HDPEC specimens often exhibit cyclic behavior transitioning between the elastic and plastic regimes, where the plastic proportion loses its significance in

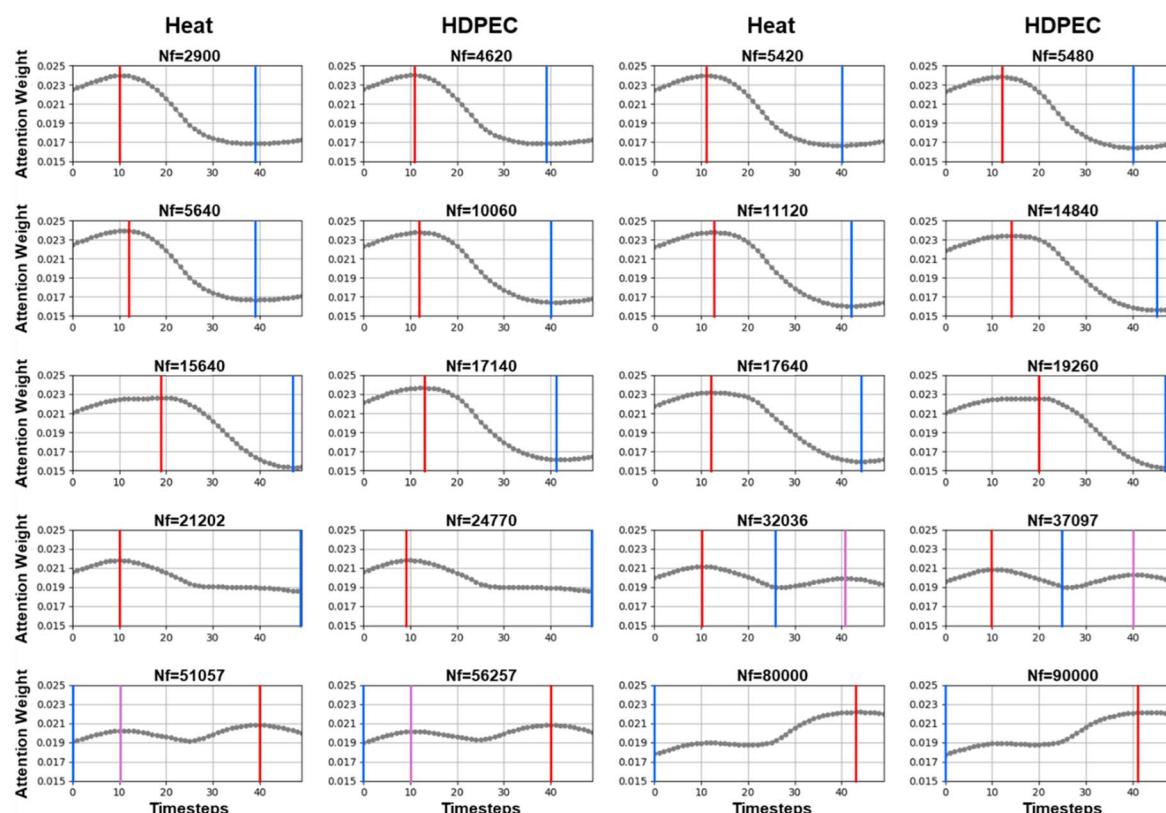


Fig. 9 Attention weights from the best-performing LSTM-contextual attention model across different fatigue life (red line: global maximum, blue line: global minimum, purple line: local maximum).



Table 3 Comparison of fatigue prediction models focus on datasets, methods, performance and interpretability

| Dataset | AI model | Accuracy | Interpretability | Code availability accessibility | Ref. |
|---|--------------------------------|--------------------------|---------------------------------|---------------------------------|--------------|
| Micro-CT images including pore size, pore density, surface roughness | Bayesian neural network | Not explicitly mentioned | Not explicitly mentioned | Not available | 11 |
| Entire cycles including additive manufacturing process parameters | FNN, CNN, random forest, ANFIS | $R^2 > 0.97$ | None explicitly mentioned | Not available | 12 |
| Entire multiaxial loading cycles including strain paths and material properties | Multi-view deep learning | $R^2 = 0.94$ | Self-attention mechanisms | Not available | 10 |
| Initial single cycle of hysteresis loop | LSTM | $R^2 = 0.99$ | Contextual-attention mechanisms | Python package provided | Our research |

the mechanical response.^{7,8,37} Even with high fatigue life, subtle changes in midrange stress or strain can serve as significant predictors (Fig. 9 purple line).

When examining how the attention of the model adapts to different fatigue lives, short-life specimens frequently show wide loops with significant plastic behavior, causing the red line to cluster around the maximum stress region. Midrange lives exhibit partial plasticity in cycle 20, and the model may highlight both the tensile peak (Fig. 9, red line) and tension-compression transitions (Fig. 9, purple line). Long-life specimens often exhibit cyclic behavior, predominantly in the transitional regime between the elastic and plastic responses. In these cases, the effect of the plastic proportion diminishes, leading to a more uniform attention distribution and smaller local peaks without a single dominant maximum. For the heat-treated group, the loop shape was often more predictably guided by sequential hardening/softening, making the maximum stress region an important area of attention. However, the shorter-life heat specimens can exhibit substantial plastic deformation, thus reinforcing the red line peak in the early cycles.⁵¹ In contrast, the heat-HDPEC specimens, which include many long-life samples, exhibit localized deformation onsets during mid-level stress or strain, yielding a stronger purple-line local maxima. This behavior implies that even an elastic loop can contain hidden features that the model considers critical for an accurate life prediction.⁵²

In summary, short-life specimens show prominent plastic deformation in the maximum stress regions, whereas long-life specimens exhibit transitional elastic-plastic behavior in the midrange stress or tension-compression transitions. The heat-treated specimens showed predictable hardening/softening patterns, and the heat-HDPEC specimens demonstrated an extended fatigue life with respect to localized midrange stress or strain changes.

3.5. Fatigue_life_predictor: python package for fatigue life predicting

Our model achieved a high prediction accuracy of $R^2 = 0.99$ by using only the initial cycle of the hysteresis loop. The contextual

attention mechanism allows the model to focus on the most important features within the first cycle, making it both efficient and highly interpretable compared to other models that rely on entire cycles or additional parameters. Among these studies, Random Forest was also employed in prior work to predict fatigue life and crack growth rate based on additive manufacturing AM processing parameters, stress states, and defect-related features.¹² While ensemble models such as RF achieved reasonably high accuracy ($R^2 > 0.85$), they often depend on large-scale feature sets and provide limited interpretability regarding temporal cyclic behavior. In contrast, our model offers early-stage fatigue life estimation based solely on the initial cyclic response, enabling transparent decision-making through attention-based interpretability.

Furthermore, unlike other studies, in which code availability and accessibility are unavailable, our research provides an open Python package (Table 3). The proposed fatigue life prediction model was developed as an open-source Python package to ensure practical usability and accessibility. The model predicts fatigue life using stress-strain data from the initial cycles (20 cycles), enabling efficient and early stage estimation. To enhance usability, the package includes automated data pre-processing functions that streamline the handling of raw stress-strain data, thereby eliminating the need for manual intervention. The optimized inference process ensured that predictions were generated in less than 5.0 seconds, as verified during multiple trials under standard testing conditions, including tests conducted on Google Colab without GPU acceleration. The tests were conducted on an Intel Xeon CPU with two vCPUs and 13 GB of RAM. The implementation and tutorial are available on GitHub (https://github.com/mschongchulshin/fatigue_life_predictor/tree/main) to ensure transparency, reproducibility, and ease of use. This package provides a robust and efficient solution for fast and reliable fatigue life predictions based on initial cycle data.

4. Conclusion

This study proposed an LSTM-contextual attention model to predict the fatigue life of type 316 stainless steel under LCF



conditions. The proposed model with attention mechanism effectively captured the temporal dependencies and localized critical features in the stress-strain time series data, achieving statistically significant improvements over baseline LSTM and CNN models. Compared with the CNN-contextual attention model, the LSTM-contextual attention model exhibited a superior capacity for handling temporal dependencies, effectively capturing the complex interplay between stress and strain in the LCF datasets. The developed prediction model was released as an open-source Python package to facilitate practical adoption.

Overall, this study demonstrates the potential of deep-learning-based approaches to capture the temporal properties of LCF data and accurately predict the fatigue life. Future studies should expand the applicability of the proposed methods to diverse materials and fatigue environments. Moreover, investigating advanced deep learning architectures, such as more refined attention mechanisms or transformer-based models, may further improve the capture of non-linear fatigue behaviors and cumulative damage processes, thereby enabling more robust and reliable fatigue life predictions in complex scenarios.

Abbreviations

| | |
|------------------------|--|
| AI | Artificial intelligence |
| ANFIS | Adaptive neuro-fuzzy inference system |
| ASTM | American society for testing and materials |
| CNN | Convolutional neural network |
| FNN | Feedforward neural network |
| HDPEC | High-density pulsed electric current |
| JIS | Japanese industrial standards |
| LCF | Low-cycle fatigue |
| LSTM | Long short-term memory |
| MAE | Mean absolute error |
| PCA | Principal component analysis |
| R^2 | Coefficient of determination |
| RMSE | Root mean square error |
| RNN | Recurrent neural network |
| x_{data} | Raw value |
| $x_{\text{data},i,t}$ | Stress or strain value of sample i at timestep |
| x_t | Raw value at timestep |
| \bar{x}_t | Input vector at timestep |
| \bar{x}_t | Mean values across all samples at timestep |
| x_j | j -th Raw value |
| y_i | True value in data |
| \hat{y}_i | Predicted value |
| \bar{y}_i | Mean of the true values |
| SD | Standard deviation |
| N_f | Fatigue life value |
| N_f^{\min} | Minimum fatigue life values |
| N_f^{\max} | Maximum fatigue life values |
| D | Dataset |
| T_i | Training set |
| V_i | Validation set |
| Score() | Performance metric |
| E_{plastic} | Plastic deformation energy |
| $E_{\text{plastic},n}$ | Plastic deformation energy for cycle |
| σ | Stress |

| | |
|-------------------------|---|
| σ_t | Stress values at timestep |
| $\bar{\sigma}_t$ | Mean stress values at timestep |
| ε | Strain |
| ε_t | Strain values at timestep |
| $\bar{\varepsilon}_t$ | Mean strain values at timestep |
| t | Timestep |
| T | Total number of timesteps |
| ε_{\max} | Maximum stress values |
| ε_{\min} | Minimum strain values |
| $E_{\text{cumulative}}$ | Cumulative energy |
| n | Total number of cycles |
| N_{samples} | The number of samples |
| r_t | Temporal correlation coefficient |
| $N_{f,i}$ | Fatigue life of sample i |
| \bar{N}_f | Mean fatigue life across all samples |
| PC_k | k -th Principal component |
| w_{kj} | PCA weights of each original variable x_j |
| α_t | Attention weight for timestep |
| e_t | Unnormalized attention score for timestep |
| W | Learnable weight matrix |
| b | Learnable bias vector |
| $h_{\text{attention}}$ | Output of the attention layer |
| R^2 | Coefficient of determination |

Data availability

The code and data set used in this study is available in the ESI† and GitHub (https://github.com/mschongchulshin/fatigue_life_predictor/).

Conflicts of interest

There are no conflicts of interest to declare.

Acknowledgements

This study was supported by the National Research Foundation of Korea (NRF), under grant numbers RS-2024-00349057, RS-2025-00555333. This research was funded by the 'New Faculty Research Support Grant' at Changwon National University in 2024. This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (No. 2018R1A6A1A03024509). This research was supported by Global – Learning & Academic research institution for Master's PhD students, and Postdocs (LAMP) Program of the National Research Foundation of Korea (NRF) grant funded by the Ministry of Education (No. RS-2024-00444460).

References

- 1 S. Yoon, Y. Cui, Y. Kimura, S. Gu, Y. Toku and Y. Ju, *Int. J. Fatigue*, 2022, **156**, 106639.
- 2 R. Liu, Z. Zhang, P. Zhang and Z. Zhang, *Acta Mater.*, 2015, **83**, 341–356.
- 3 X. Yan, S. Yoon, S. Gu, Y. Kimura, D. Kobayashi, Y. Ju and Y. Toku, *Mech. Eng. J.*, 2024, **11**, 24–124.





- 4 E. Lee, M. Yoo, T. Byun, J. Hunn, K. Farrell and L. Mansur, *Acta Mater.*, 2001, **49**, 3277–3287.
- 5 A. Pardo, M. Merino, A. Coy, F. Viejo, M. Carboneras and R. Arrabal, *Acta Mater.*, 2007, **55**, 2239–2251.
- 6 G. T. Gray III, V. Livescu, P. Rigg, C. P. Trujillo, C. M. Cady, S.-R. Chen, J. S. Carpenter, T. J. Lienert and S. J. Fensin, *Acta Mater.*, 2017, **138**, 140–149.
- 7 S. Yoon, Y. Kimura, Y. Ju and Y. Toku, *Int. J. Pressure Vessels Piping*, 2024, **209**, 105178.
- 8 S. Yoon, M. Morii, Y. Kimura and Y. Toku, *Mech. Eng. J.*, 2024, 24–00015.
- 9 S. R. Yeratapally, M. G. Glavicic, M. Hardy and M. D. Sangid, *Acta Mater.*, 2016, **107**, 152–167.
- 10 S. Moon, R. Ma, R. Attardo, C. Tomonto, M. Nordin, P. Wheelock, M. Glavicic, M. Layman, R. Billo and T. Luo, *Sci. Rep.*, 2021, **11**, 20424.
- 11 S. Chen, X. Zhou and Y. Bai, *Int. J. Fatigue*, 2025, **190**, 108620.
- 12 Y. Min, X. Ming, C. Peihong, S. Yang, H. Zhang, W. Lingfeng, Z. Liucheng, L. Yinghong and G. Wanlin, *Chin. J. Aeronaut.*, 2024, **37**, 1–22.
- 13 J. Jung, S. Yoon, S. Gu, Y. Kimura, Y. Toku and Y. Ju, *Eng. Failure Anal.*, 2023, **150**, 107230.
- 14 H. Shin, T. Yoon, J. You and S. Na, *J. Mech. Behav. Biomed. Mater.*, 2024, 106643.
- 15 Y. Zhang, L. Fan, C. Su, Z. Shu and H. Zhang, *RSC Adv.*, 2024, **14**, 37737–37751.
- 16 A. Derry, M. Krzywinski and N. Altman, *Nat. Methods*, 2023, **20**, 1269–1270.
- 17 J. M. Jeong, M. Ra, J. Jeong and W. Lee, *RSC Adv.*, 2024, **14**, 18489–18500.
- 18 L. Salmela, N. Tsipinakis, A. Foi, C. Billet, J. M. Dudley and G. Genty, *Nat. Mach. Intell.*, 2021, **3**, 344–354.
- 19 Y. Cao, T. A. Geddes, J. Y. H. Yang and P. Yang, *Nat. Mach. Intell.*, 2020, **2**, 500–508.
- 20 N. Akhtar and U. Ragavendran, *Neural Comput. Appl.*, 2020, **32**, 879–898.
- 21 F. Landi, L. Baraldi, M. Cornia and R. Cucchiara, *Neural Network*, 2021, **144**, 334–341.
- 22 S. Iravani, A. Khosravi, E. N. Zare, R. S. Varma, A. Zarrabi and P. Makvandi, *RSC Adv.*, 2024, **14**, 36835–36851.
- 23 X. Jin, X. Yu, X. Wang, Y. Bai, T. Su and J. Kong, Prediction for Time Series with CNN and LSTM, in *Proceedings of the 11th International Conference on Modelling, Identification and Control (ICMIC2019)*, Lecture Notes in Electrical Engineering, ed. R. Wang, Z. Chen, W. Zhang and Q. Zhu, Springer, Singapore, 2020, vol. 582, DOI: [10.1007/978-981-15-0474-7_59](https://doi.org/10.1007/978-981-15-0474-7_59).
- 24 Z. Song, D. Huang, B. Song, K. Chen, Y. Song, G. Liu, J. Su, J. P. d. Magalhães, D. J. Rigden and J. Meng, *Nat. Commun.*, 2021, **12**, 4011.
- 25 B. Yang, J. Li, D. F. Wong, L. S. Chao, X. Wang and Z. Tu, Context-aware self-attention networks, in *Proceedings of the AAAI Conference on Artificial Intelligence*, 2019, vol. 33(1), pp. 387–394, DOI: [10.1609/aaai.v33i01.3301387](https://doi.org/10.1609/aaai.v33i01.3301387).
- 26 J. Yang, G. Kang and Q. Kan, *Int. J. Fatigue*, 2022, **162**, 106851.
- 27 S. Nerella, S. Bandyopadhyay, J. Zhang, M. Contreras, S. Siegel, A. Bumin, B. Silva, J. Sena, B. Shickel and A. Bihorac, *arXiv*, 2023, preprint, arXiv:2307.00067, DOI: [10.48550/arXiv.2307.00067](https://doi.org/10.48550/arXiv.2307.00067).
- 28 Y. Tao, S. Ren, M. Q. Ding, R. Schwartz and X. Lu, Predicting drug sensitivity of cancer cell lines via collaborative filtering with contextual attention, in *Machine Learning for Healthcare Conference*, PMLR, 2020, pp. 660–684.
- 29 C. Guan, X. Wang and W. Zhu, AutoAttend: Automated attention representation search, in *International Conference on Machine Learning*, PMLR, 2021, pp. 3864–3874.
- 30 Z. Chen, R. Zhou and P. Ren, *RSC Adv.*, 2024, **14**, 8053–8066.
- 31 S. A. A. Shams, J. W. Bae, J. N. Kim, H. S. Kim, T. Lee and C. S. Lee, *J. Mater. Sci. Technol.*, 2022, **115**, 115–128.
- 32 T. Anjum and N. Khan, *Neural Process. Lett.*, 2023, **55**, 7227–7257.
- 33 M. Wang, S. Li, J. Wang, O. Zhang, H. Du, D. Jiang, Z. Wu, Y. Deng, Y. Kang and P. Pan, *Nat. Commun.*, 2024, **15**, 10127.
- 34 S. K. Paul, *J. Mater. Res. Technol.*, 2019, **8**, 4894–4914.
- 35 S. Yoon, A Study on Improvement of Tensile and Fatigue Properties of Type 316 Austenitic Stainless Steel by High-Density Pulsed Electric Current, *Doctoral dissertation*, Nagoya University, 2022.
- 36 S. Yoon, Y. Kimura, Y. Cui, Y. Toku and Y. Ju, *Mater. Trans.*, 2021, **62**, 748–755.
- 37 S. Yoon, Y. Kimura, Y. Toku and Y. Ju, *Materialia*, 2023, **32**, 101922.
- 38 H. Dang, A. Liang, R. Feng, J. Zhang, X. Yu and Y. Shao, *Int. J. Fatigue*, 2022, **165**, 107187.
- 39 J. Gao, H. Zhou, J. Shen, H. Wei, X. Fang and Y. He, *J. Mater. Eng. Perform.*, 2024, **33**, 496–509.
- 40 S. Sun, H. Fang, Y. Li, X. Zhang, B. Zhu, X. Ding and R. Chen, *Mater. Sci. Eng. A*, 2024, **893**, 146113.
- 41 X. Zhang, A. Wang, C. Shao and H. Bao, *Acta Mater.*, 2024, **276**, 120160.
- 42 Q. Li, Y. Chen, Y. Liu, D. Jiang, Y. Yang, H. Yang, K. Yu, Y. Ren and L. Cui, *Acta Mater.*, 2024, **265**, 119625.
- 43 Y. Xiang, Q. Tang, W. Xu, S. Hu, P. Zhao, J. Guo and J. Liu, *Renewable Energy*, 2024, **223**, 119962.
- 44 A. Mehrabinezhad, M. Teshnehab and A. Sharifi, *Comput. Methods Biomed. Eng.*, 2024, **12**, 2379526.
- 45 S. K. Dewangan, R. Jain, S. Bhattacharjee, S. Jain, M. Paswan, S. Samal and B. Ahn, *J. Mater. Res. Technol.*, 2024, **30**, 2377–2387.
- 46 X. Wen and W. Li, *IEEE Access*, 2023, **11**, 48322–48331.
- 47 T. Akiba, S. Sano, T. Yanase, T. Ohta and M. Koyama, Optuna: A next-generation hyperparameter optimization framework, in *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, ACM, 2019, pp. 2623–2631.
- 48 T. Chai and R. R. Draxler, *Geosci. Model Dev. Discuss.*, 2014, **7**, 1525–1534.
- 49 T. H. Trojahn and R. Goularte, *Multimed. Tool. Appl.*, 2021, **80**, 17487–17513.
- 50 S. Song, C. Lan, J. Xing, W. Zeng and J. Liu, *IEEE Trans. Image Process.*, 2018, **27**, 3459–3471.

51 M. A. N. Ali, *Thermo-elastic-plastic Analysis for Elastic Component under High Temperature Fatigue Crack Growth Rate*, Sheffield Hallam University, United Kingdom, 2013.

52 S. Qian, J. Wang, J. Hu, Q. Fang and C. Xu, Hierarchical multi-modal contextual attention network for fake news

detection, in *Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval*, ACM, 2021, pp. 153–162.

