Computers, Materials & Continua

Tech Science Press

https://doi.org/10.32604/cmc.2026.079732

ARTICLE

Ratcheting Behavior and Intelligent Prediction Algorithms for Inner Liner Welds of Multi-Layered Pressure Vessels

Linbin Li[1], **Ruiyuan Xue**[1,*], **Juyin Zhang**[2,*], **Xueping Wang**[2] **and Tiantian Chu**[1]

[1]College of Petrochemical Engineering, Lanzhou University of Technology, Lanzhou, China
[2]Gansu Province Special Equipment Inspection and Testing Institute, Lanzhou, China

*Corresponding Authors: Ruiyuan Xue. Email: xrywyy@163.com; Juyin Zhang. Email: zjytejian@163.com

Received: 27 January 2026; Accepted: 11 March 2026; Published: 08 May 2026

ABSTRACT: The plastic strain accumulation results of the multi-layered wrapped pressure vessel liner during long-term service are an important basis for its safety performance evaluation. However, the complex welds distributed on the liner bring challenges to the calculation of plastic cumulative strain. To this end, a novel hybrid deep learning framework is proposed for the efficient and precise prediction of ratcheting behavior in the liner welds of multilayered pressure vessels. By employing a BiLSTM network to extract bidirectional temporal dependencies from the strain history and incorporating a Multi-Head Attention (MHA) mechanism for adaptive feature weighting, the proposed method effectively addresses the difficulty of modeling cumulative effects in long-sequence ratcheting data. Firstly, asymmetric cyclic loading experiments were conducted on various welded joints and base metals to reveal the evolutionary laws of ratcheting behavior and construct a training dataset. The results show that the ratcheting strain evolution of different structural specimens shows the typical 'two-stage' characteristics, and the ratcheting strain accumulation on the base metal is significantly higher than that of the weld structure specimen. The proposed deep learning model can not only accurately capture the 'two-stage' evolution of ratcheting strain, but also directly use the base metal data to accurately predict the ratcheting strain accumulation in different weld parts, avoiding the complex parameter calibration process of the traditional constitutive model. It provides technical support for the integrity assessment and online monitoring of the complex weld structure of multi-layered pressure vessels.

KEYWORDS: Multi-layered pressure vessels; ratcheting behavior; inner tank weld; BiLSTM network; multi-head attention mechanism

1 Introduction

With the outstanding advantages of reliable structure and safe storage, the integral multi-layered pressure vessels have gradually become the core pressure-bearing equipment in the field of high-pressure gaseous hydrogen storage [1,2]. However, due to the superposition of the pre-tightening force of the laminate and the working internal pressure during the manufacturing process, the stress distribution of the multi-layered pressure vessels after loading shows a significant non-uniformity. During the loading-unloading cycle of the equipment operation cycle, some locations not only maintain a high average stress for a long time, but also bear a higher amplitude of alternating stress [3]. This asymmetric cyclic load coupled with high average stress and high stress amplitude, can easily induce the accumulation of plastic deformation inside the material, that is, the ratchet effect. The continuous accumulation of plastic strain will not only lead to structural size instability, but also significantly accelerate the low-cycle fatigue damage process of materials,

Copyright © 2026 The Authors. Published by Tech Science Press. This work is licensed under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

which seriously threatens the life cycle safety of equipment [4]. However, due to the particularity of the multi-layer wrapping structure, it is difficult for conventional non-destructive testing methods to effectively capture the degradation of liner performance [5]. Online monitoring based on digital intelligence has become an effective means to ensure the safe service of multi-layered pressure vessels. Accurate analysis and prediction of the ratcheting behavior of the liner structure in complex service environments is a key prerequisite for achieving this goal [6].

The inner cylinder of the multi-layered pressure vessels is assembled by multiple cylinder sections, and there are a large number of staggered longitudinal welds and circumferential welds inside the structure. Experimental studies have shown that due to the differences in grain morphology and dislocation evolution mechanism between weld metal, heat-affected zone, and base metal, each region exhibits distinct cyclic hardening or softening characteristics, resulting in significantly different ratcheting behavior of weld structure from that of base metal [7–10]. The spatial heterogeneity of this material property makes the mechanical response of the liner weld extremely complex, which brings great challenges to the prediction of the ratcheting behavior of the multi-layered pressure vessels.

At present, the prediction of ratcheting behavior of metal structures mainly depends on the traditional constitutive model combined with the finite element analysis method. Zeng et al. [11] and Janulionis and Dundulis [12] simulated the ratcheting behavior of welded joints of different materials using the Chaboche model. Liu et al. [13] carried out a series of uniaxial ratcheting tests on super duplex SAF2507 stainless steel and developed the AF-OW II constitutive model based on Armstrong-Frederick (A-F) hardening criterion and Ohno-Wang (O-WII) model to accurately predict its ratcheting behavior. Dong et al. [14] proposed an improved constitutive model framework that combines micro-mechanism and macro-mechanics, which significantly improves the adaptability of the model to cyclic response under different loading paths, and realizes the accurate prediction of fatigue failure location and life of welded joints.

However, the reliability of numerical simulation is highly dependent on the selection of the material constitutive model and the accuracy of its parameters. Although the traditional phenomenological constitutive model is relatively mature in theory, it still faces difficulties in practical engineering applications. For example, as pointed out by Kalnins et al. [15], the parameter calibration for the Chaboche nonlinear kinematic (NLK) hardening model is a tedious, multi-step iterative process. It requires not only dividing the stress-plastic strain curve into multiple segments for initial calculation but also frequently necessitates a secondary manual calibration to iteratively minimize the deviation between the predicted and experimental curves. Secondly, for welded structures, the existing methods usually simplify the base metal, heat-affected zone, and weld as independent uniform materials for modeling and parameter calibration, respectively. However, this method ignores the performance change at the boundary of the heterostructure. This traditional modeling strategy not only increases the experimental workload and computational cost, but also makes it more difficult to achieve unified characterization and rapid prediction of the ratcheting behavior of different types of welds.

In recent years, with the rapid development of data-driven technology, deep learning has provided a new paradigm for the modeling of complex materials. Among them, the long short-term memory network (LSTM), with its unique gating mechanism, can effectively overcome the gradient disappearance problem of traditional recurrent neural networks (RNN), and has been widely used in residual life prediction and structural response analysis. Ding et al. [16] constructed a CNN-BiLSTM-Attention hybrid neural network, which realized the accurate correction of the nonlinear model parameters of the bridge tower structure under seismic load, and provided a feasible algorithm basis for replacing the traditional complex constitutive integral operation. The above research shows the potential of deep learning in mechanical prediction. However, the traditional one-way LSTM neural network is difficult to fully exploit the potential temporal characteristics in the ratchet evolution process, and when dealing with long sequence data, it lacks the weight distribution mechanism for key historical moments, resulting in a significant reduction in prediction

accuracy. In addition, the current research mostly focuses on a single material or a single loading path, and there are still few reports on the unified prediction of ratcheting behavior of multi-type welds under asymmetric load in welded structures.

In order to solve the problem that it is difficult to predict the ratcheting behavior of the inner liner weld of the multi-layered pressure vessels, this paper first carried out a systematic asymmetric cyclic loading test on different weld structures and base metals of the inner liner material, and deeply studied the evolution law and difference of the weld structure and the ratcheting behavior of the base metal, and proposed a prediction method of the ratcheting behavior of the weld structure based on the Bi-directional Long Short-Term Memory (BiLSTM) and Multi-Head Attention (MHA). The BiLSTM network can capture the forward and backward long-term dependencies when extracting the time features of the data [17]. At the same time, the MHA mechanism is introduced, which can give different weights to the features of different load stages, thus enhancing the model's ability to extract and invert the evolution of ratchet behavior. The significance of this study is to effectively solve the problem that the weld structure is difficult to accurately model and the constitutive parameters are difficult to calibrate, which provides a new theoretical support for the integrity evaluation and digital intelligent online monitoring of the complex weld structure of the multi-layered pressure vessels.

The following chapters of this paper are arranged as follows: Section 2 introduces the asymmetric cyclic loading test and data extraction process of different types of weld specimens. Section 3 introduces the basic algorithm theory and the construction of the proposed model. Section 4 shows the performance comparison and prediction results of the proposed models. Section 5 is the conclusion of this paper.

2 Experimental Campaign

2.1 Sample Preparation

In order to study the difference in ratcheting behavior between the weld structure and base metal under asymmetric cyclic loading, a systematic uniaxial ratcheting fatigue experiment was carried out in this chapter. The experimental material is S31603 austenitic stainless steel. Firstly, according to the same welding process of the multi-layer wrapped pressure vessel liner, several S31603 stainless steel plates with a thickness of 8 mm are spliced. Then, according to the cutting method shown in Fig. 1, four types of samples, including circumferential weld (CW), longitudinal weld (LW), T-type weld (TW), and base metal (BM), are manufactured. After the processing of all kinds of samples, 100% ray detection (RT) was carried out to eliminate the influence of initial defects such as welding porosity and slag inclusion on the subsequent ratcheting experiment.

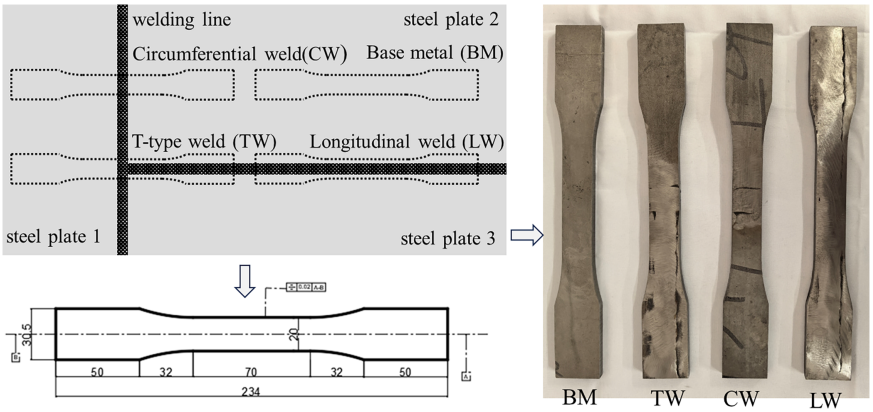

Figure 1: Sample size and preparation method.

2.2 Experimental Process

The ratcheting tests of the four types of specimens were carried out on the SDS-100 electro-hydraulic servo fatigue testing machine. In order to consider the influence of stress mean and stress amplitude on the ratcheting behavior, a total of three cyclic loads were designed. The specific test parameters are shown in Table 1. During the experiment, the axial strain of the sample under cyclic load was measured by KFEM-5-120-C1L1M2R strain gauge (maximum range 30%). The acquisition frequency was 100 Hz, and the measuring points of each sample were the same, all located in the center of the gauge section. Before the formal loading, all the strain gauge channels are cleared when the sample is installed, and the load of the test machine is zero, to eliminate the influence of the initial clamping state of the sample, the welding residual stress, and the electrical deviation of the measurement system on the initial reading. The experimental process is shown in Fig. 2.

Table 1: Test parameter settings.

Type	Mean of Stressσ_m/MPa	Stress Amplitudesσ_a/MPa	FrequencyHz	Cycle IndexN_f
Condition A	200	200	1	2000
Condition B	250	200	1	2000
Condition C	200	250	1	2000

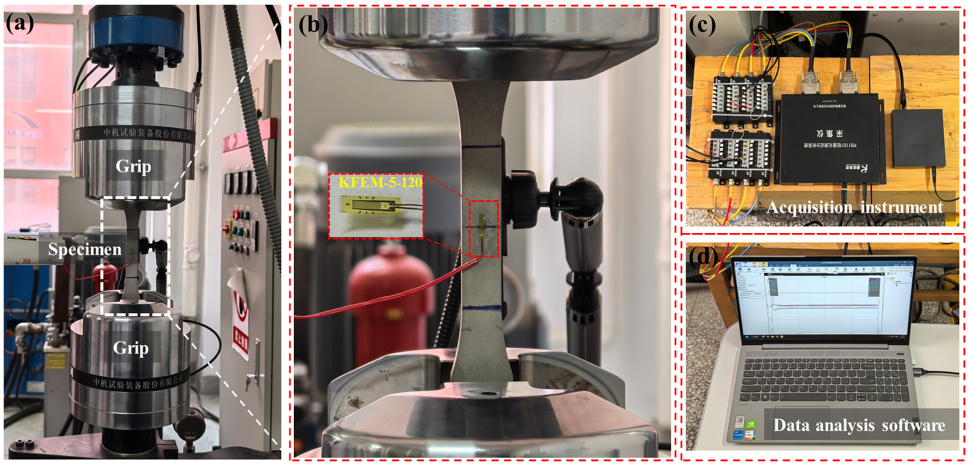

Figure 2: Experimental process. (**a**) Overall view of the testing machine; (**b**) Partial enlarged detail; (**c**) Strain acquisition instrument; (**d**) Data analysis software interface.

2.3 Discussion of Experimental Results

Fig. 3 shows the strain-time curves of the four samples under condition B. Although the stress levels of condition A and condition C are different, the four specimens show consistent ratcheting behavior characteristics under all test conditions. Therefore, only condition B is taken as an example for detailed analysis.

In Fig. 3, with the increase of the number of cycles, the four samples have obvious strain accumulation, that is, the ratcheting effect. The strain growth of all kinds of samples showed a trend of 'fast first and then slow', and the ratcheting evolution process showed typical 'two-stage' characteristics: at the initial stage of loading, the strain accumulation rate was fast, and the slope of the curve was significant, indicating that the dislocation structure inside the material was undergoing severe plastic adjustment. Subsequently, the

material enters the cyclic hardening stage, the strain accumulation rate gradually decreases and tends to be stable, and the structural response gradually reaches a stable state.

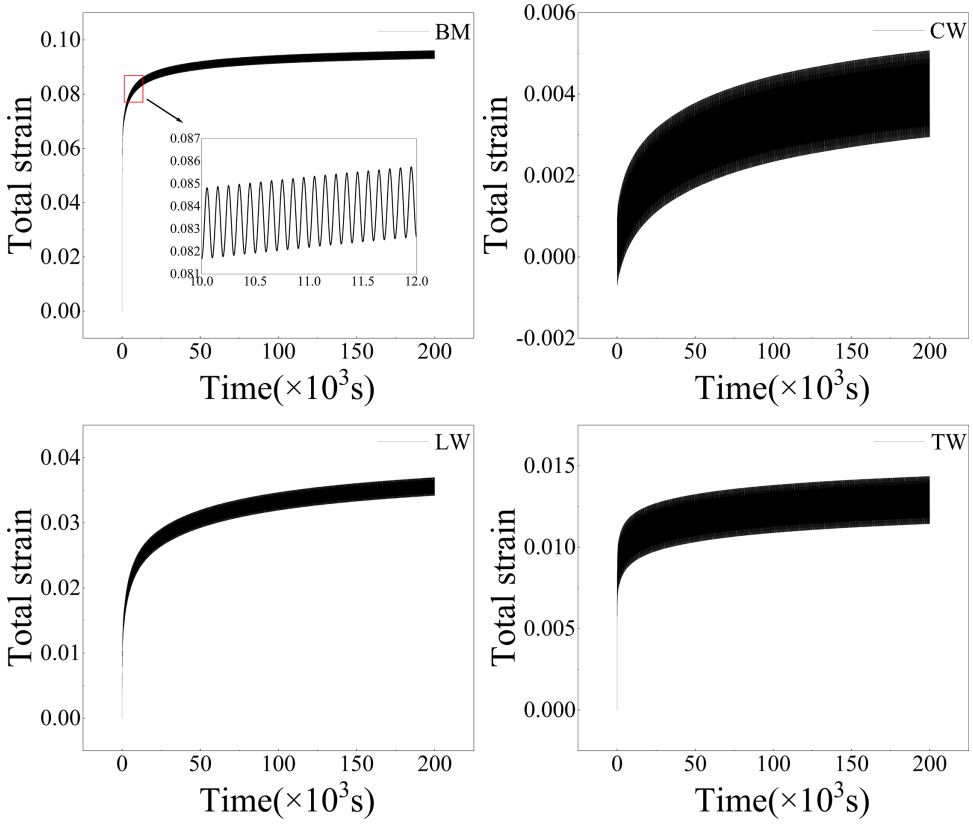

Figure 3: The measured strain-time curve of four samples under condition B.

However, the difference in the structure of the weld specimen leads to a significant difference in its strain response. Fig. 4 shows the ratcheting strain evolution curves of the four specimens under three working conditions, with the number of cycles. The ratchet strain is defined as:

$$\varepsilon_r = \frac{1}{2}(\varepsilon_{max} + \varepsilon_{min}) \tag{1}$$

Among them, ε_{max} and ε_{min} are the maximum strain and minimum strain of the response in each stress cycle, respectively.

Observed in Fig. 4, the ratchet strain of the polished rod under the three working conditions is significantly higher than that of the weld sample, and the circumferential weld is the lowest, followed by the T-type weld and the longitudinal weld. Taking the working condition A as an example, the ratcheting strain accumulation of the polished rod sample is 0.0539, the cumulative strain of the girth weld is the smallest, about 0.0013, and the cumulative strain of the longitudinal weld and the T-joint is between the two, reaching 0.0226 and 0.0072, respectively. The relatively small ratchet strain indicates that the mechanical strength of the welded joint is much higher than that of the polished rod sample. On the one hand, the grain refinement at the weld seam limits the dislocation slip, thereby improving the strength. On the other hand, the welding process is due to thermal stress. Residual stress and residual deformation play a role in work hardening [18].

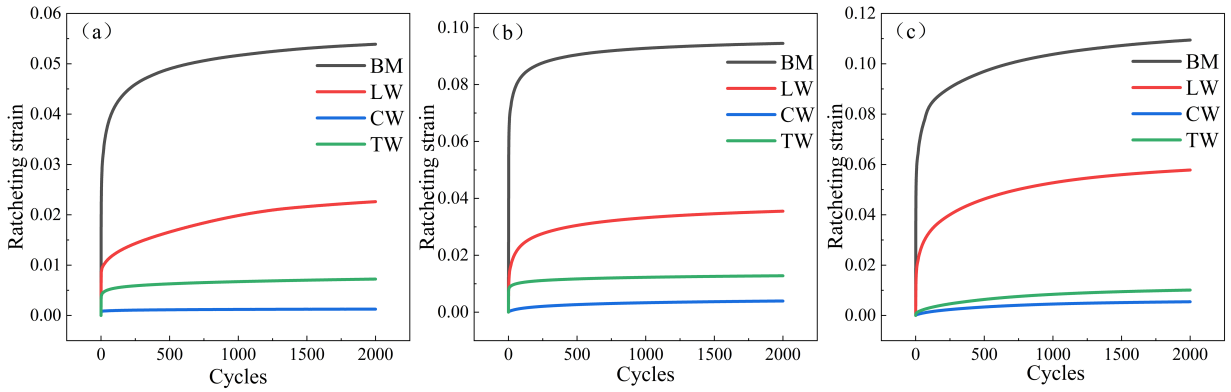

Figure 4: The evolution curve of ratcheting strain with the number of cycles of the four samples under different working conditions. (**a**) Condition A; (**b**) Condition B; (**c**) Condition C.

Table 2 shows the ratcheting strain accumulation results of each sample under different working conditions in detail, and the relative growth rates under working conditions B and C are calculated based on working condition A. Whether increasing the mean stress (condition B) or increasing the stress amplitude (condition C), the ratcheting deformation will be significantly aggravated. However, the strain growth rate induced by condition C was substantially higher than that of condition B in all samples. Taking the light bar as an example, the strain increase under condition C is as high as 102.97%, while that under condition B is only 75.32%. This trend is more evident in the longitudinal weld (LW), which is 155.75% and 57.08%, respectively. The above data show that compared with the mean stress, the stress amplitude is the dominant factor driving the evolution of the ratcheting effect, and its impact on the cumulative plastic deformation is more intense.

Table 2: Ratcheting strain of four specimens under the 2000th cycle under three working conditions.

Sample Shape	Condition A	Condition B	Condition C
BM	0.0539	0.0945 (↑75.32%)	0.1094 (↑102.97%)
CW	0.0013	0.0039 (↑200.00%)	0.0054 (↑315.38%)
LW	0.0226	0.0355 (↑57.08%)	0.0578 (↑155.75%)

Through the research in this section, the difference in ratcheting behavior between the weld structure and the light bar under asymmetric cyclic loading is revealed, and a time-series strain data set including various weld structures and working conditions is constructed. This not only clarifies the complex mechanical response characteristics of the weld area but also provides the necessary experimental basis and data source for the construction and training of the ratcheting behavior prediction model based on deep learning.

3 The Establishment of a Ratcheting Behavior Prediction Model Based on LSTM-MHA

Aiming at the problems of cumbersome parameter calibration and difficult quantification of different structural differences in the prediction of ratcheting behavior of weld structure, a data-driven nonlinear mapping strategy is proposed. The core of this strategy is to construct a high-dimensional mapping model from 'base metal reference response + weld type attribute label' to 'actual response of weld structure'. In order to process the above input data with long-time series dependence and accurately capture the evolution law of the ratchet effect, BiLSTM and a multi-head attention mechanism are used to construct the core network.

3.1 Discussion of Experimental Results

3.1.1 BiLSTM Model

The traditional Recurrent Neural Networks (RNN) model has efficient performance in processing sequential data. However, due to the disappearance of gradients in multiple backpropagation processes, the ability of RNNs to learn long-term dependencies is weak. Long Short-Term Memory (LSTM) deals with long-term sequential data by introducing ignore gates, input gates, and output gates that control access to unit states [19], which effectively solves the limitations of traditional RNN and enables it to maintain the persistence and accuracy of information when processing long-term sequence data. The neuron structure of the LSTM network is shown in Fig. 5.

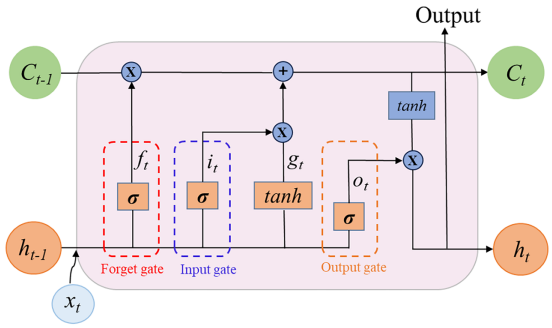

Figure 5: LSTM model construction.

The specific operation process is as follows:

Firstly, the external neuron information h_{t-1} and the current input x are read by the forgetting gate f_t, and the forgetting gate decides to discard the information. The output is controlled between [0, 1], 1 means 'complete retention', and 0 means 'complete abandonment'.

$$f_t = \sigma\left(W_f x_t + W_f h_{t-1} + b_f\right) \tag{2}$$

The input gate controls how much information the candidate state needs to be saved. It consists of two layers. The Sigmoid layer acts as the 'input gate layer' to determine the Value i to be updated; the tanh layer creates a new candidate value $\tilde{c}_t$ to replace the old information that needs to be forgotten.

$$i_t = \sigma\left(W_i x_t + W_i h_{t-1} + b_i\right) \tag{3}$$

$$\tilde{c}_t = \tanh\left(W_c x_t + W_c h_{t-1} + b_c\right) \tag{4}$$

After completing the extraction set of the above input information, the cell state c_{t-1} can be updated to add new candidate values $\tilde{c}_t$. The calculation is as follows:

$$c_t = f_t \odot c_{t-1} + i_t \odot \tilde{c}_t \tag{5}$$

The output gate is calculated as:

$$o_t = \sigma\left(W_o x_t + W_o h_{t-1} + b_o\right) \tag{6}$$

$$h_t = o_t \odot \tanh\left(c_t\right) \tag{7}$$

The Sigmoid activation function and the activation function together determine which part of the cell state c_t is output to complete the production of h_t.

Among them: f_t, i_t and o_t for the three doors to control the path of information transmission, f_t for the forgotten door, i_t for the input door, o_t for the output door; σ and tanh is the activation function; W_f, W_i, W_c and W_o are the recursive connection weights of their corresponding thresholds; c_{t-1} denotes the memory unit of the previous moment; $\tilde{c}_t$ is a candidate state obtained by a nonlinear function; h_t for the current neuron information; h_{t-1} preneuronal information; b_f, b_i, b_o and b_c are the bias matrix between layers; x_t is the current input value; $\odot$ represents the product of vector elements.

The LSTM network can only consider the time series characteristics of structural response from front to back. The utilisation rate of historical information in the network training process is low, and the internal factors of the data cannot be fully explored. BiLSTM is an extension of LSTM, which combines forward and backward LSTM layers to realise the flow mechanism of information forward and backward propagation, which is conducive to improving the utilisation of feature data and the accuracy of model prediction. The BiLSTM network structure is shown in Fig. 6, where x_i represents the input data at different times; A_i is the hidden state of the forward LSTM; B_i is the secret state of the backward LSTM; h_i represents the corresponding output data; w_i denotes the weight of different levels.

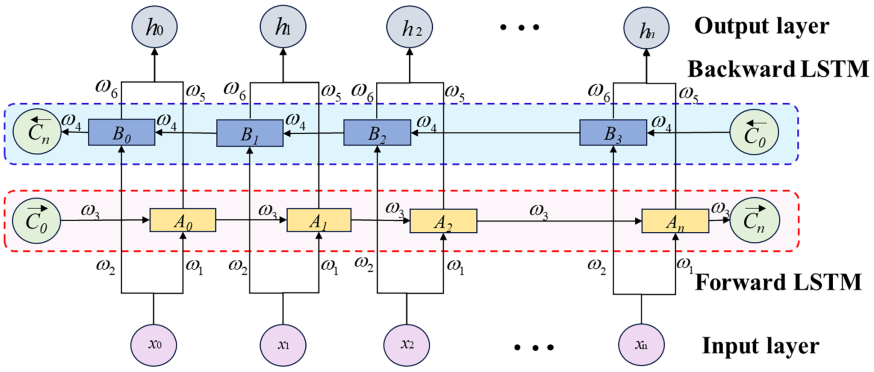

Figure 6: BiLSTM model construction.

The mathematical equations for the bi-directional information flow fusion task in the BiLSTM network are explained as follows:

$$A_i = f_1\left(\omega_1 x_i + \omega_2 A_{i-1}\right) \tag{8}$$

$$B_i = f_2\left(\omega_3 x_t + \omega_4 A_{i+1}\right) \tag{9}$$

$$h_i = f_3\left(\omega_5 A_i + \omega_6 B_i\right) \tag{10}$$

In the Eqs. (8)–(10), f_1, f_2 and f_3 are the activation functions of each layer.

3.1.2 Multi-Headed Attention Mechanism

The basic principle of the attention mechanism is to assign weights according to the different effects of input features on output results, and to improve the feature extraction ability of the network model by differentiating weights. The multi-head attention mechanism solves the problem of excessive concentration of attention and decreased information grasping ability by generating multiple sets of learnable parameters for the same output. The general attention mechanism calculation process is shown in Fig. 7. First, a query is given, and the correlation between the Query and Key is calculated. Then the most appropriate input data

value (Value) is found according to the correlation size, so as to realise the weight distribution of Value and generate the output result. In the multi-head attention mechanism, Query, Key, and Value are transformed separately through the linear projection of h-group independent learning, and the transformed Query, Key and Value are input into the attention mechanism in parallel, and the parallel output of h-group is obtained after the respective aggregation processing. Finally, these outputs are integrated through a learnable linear projection transformation, and the output results of the model are finally generated. The calculation process is as follows:

(1) For the same input x, generate h groups of learnable parameters, and, calculate and generate h groups of different queries, keys and values Q_i, K_i and V_i;

(2) For each group of Q_i and K_i, the dot product method is used to calculate the correlation between the two input vectors, and the attention score S_i of each group of vectors is calculated, which can be expressed as:

$$S_i = similarity\left(Q_i, K_i\right) = Q_i \cdot K_i \tag{11}$$

(3) Scaling the S_i obtained by dividing the square root of its dimension, and then through the softmax activation function, an attention weight matrix ai is finally received, which can be expressed as:

$$a_i = soft\max\left(S_i\right) \tag{12}$$

(4) Multiply ai and V_i, calculate the weighted sum, get a vector representation Z_i, repeat h times to get the vector representation Z_1, Z_2, . . ., Z_h of vector x, which can be expressed as:

$$Z_i = \sum a_i \cdot V_i \tag{13}$$

(5) The h groups of Z_i are spliced and multiplied by a learnable linear projection W_0 to obtain the final output Z, which can be expressed as:

$$Z = concat\left(Z_1, Z_2, \ldots Z_h\right) \cdot W_0 \tag{14}$$

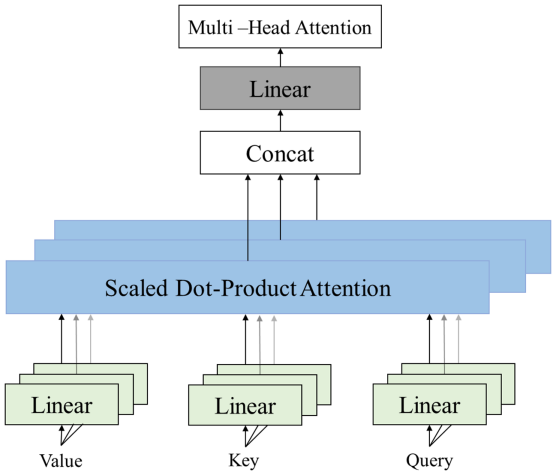

Figure 7: The calculation process of the multi-head attention mechanism.

3.1.3 BiLSTM-MHA Hybrid Neural Network Model

The BiLSTM-Attention neural network structure is shown in Fig. 8, which is mainly composed of an input layer, BiLSTM layer, Attention layer and output layer.

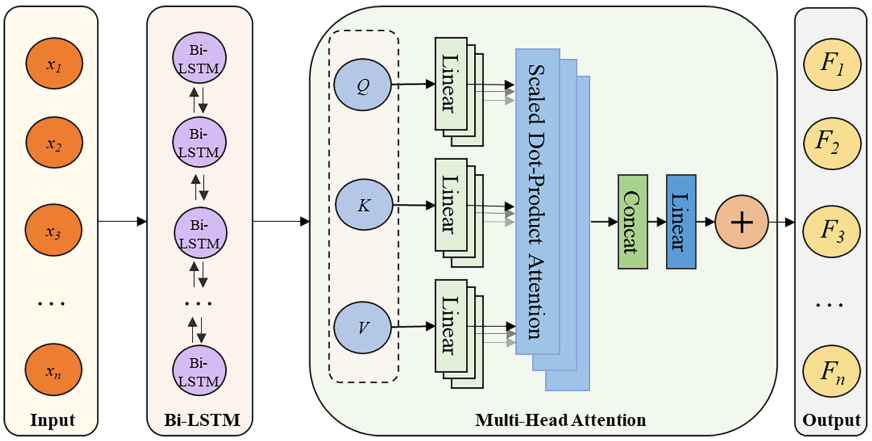

Figure 8: BiLSTM-MHA model construction.

The input layer is used to receive the measured strain data of the base metal and the characteristics of the weld type as the network input. The BiLSTM layer extracts the time series features of the LSTM unit with forward and backward chain connection for the input information, and introduces the MHA mechanism. The time information extracted by the BiLSTM hidden layer is weighted to distinguish the degree of importance, and the deep time series features of the response data are mined. Then the weighted feature vector is input into the fully connected layer for linear transformation, and finally the predicted ratchet strain scalar Value of the weld is output.

3.2 BiLSTM-MHA Model Training Process

The weld ratcheting strain prediction process based on BiLSTM-MHA constructed in this paper is shown in Fig. 9. It can be divided into three stages:

(1) Data preprocessing. Firstly, the collected original strain data is cleaned to eliminate the noise and outliers caused by equipment vibration. Subsequently, the Z-score normalisation method is used to scale the original time series data uniformly, and the physical quantities of different dimensions are converted into dimensionless values that obey the standard normal distribution, so as to eliminate the scale difference between features and accelerate the convergence speed of gradient descent. The equation is:

$$x_i' = \frac{x_i - \mu}{\sigma} \tag{15}$$

Among them, x_i is the strain value at the original moment, and μ and σ are the mean and standard deviation of the sequence, respectively. After standardisation, the mean Value of the data is 0 and the standard deviation is 1, which effectively prevents the gradient disappearance or gradient explosion problem in the neural network training process, and improves the feature extraction efficiency of the BiLSTM network. After the model outputs the prediction results, the normalised values are restored to the real physical strain values by inverse transformation for error analysis.

In order to enable the supervised learning model to capture the time series evolution law of the ratchet effect, the sliding window technique is also used to convert continuous long-time series data into 'input-output' sample pairs that the model can train. The time sliding window is shown in Fig. 10. Where: x_i is the data at time t_i; L is the window size, and l is the moving increment.

(2) Model training optimisation. The sample data of working conditions B and C are divided into a training set and a verification set according to the time series in a ratio of 4:1 for neural network model training. In the process of model optimisation, the mean square error (MSE) is used as the loss function, and the AdamW (Adam with Weight Decay) optimiser is used to update the model parameters. AdamW introduces the weight attenuation term, which can more effectively suppress the model over-fitting and improve its performance on the unknown weld structure. In order to dynamically adjust the convergence step size, the Reduce LR On Plateau learning rate attenuation strategy is introduced. When the validation set Loss does not decrease within 5 consecutive cycles, the learning rate is automatically halved. At the same time, the Early Stopping mechanism is used to monitor the error of the validation set, and the training is terminated in advance when the performance of the validation set is not improved for several rounds, and the optimal model weight with the minimum verification loss is automatically saved for subsequent testing.

(3) Model evaluation. After the model outputs the standardised Value of the prediction, the mean and standard deviation of the training set are used for Inverse standardisation. Finally, the restored prediction curve is compared with the experimental measured data, and the Coefficient of Determination (R^2), Root Mean Square Error ($RMSE$) and Mean Absolute Error (MAE) are used as quantitative indicators. Comprehensively evaluate the prediction accuracy of the model on the ratcheting behavior of the weld structure. The calculation equation of each index is shown as Eqs. (16)–(18):

$$R^2 = 1 - \frac{\sum_{j=1}^{N} \left[\hat{y}(i) - y(i)\right]^2}{\sum_{j=1}^{N} \left[y(i) - \overline{y}\right]^2} \tag{16}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n} \left(y(i) - \hat{y}(i)\right)^2} \tag{17}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n} \left|y(i) - \hat{y}(i)\right| \tag{18}$$

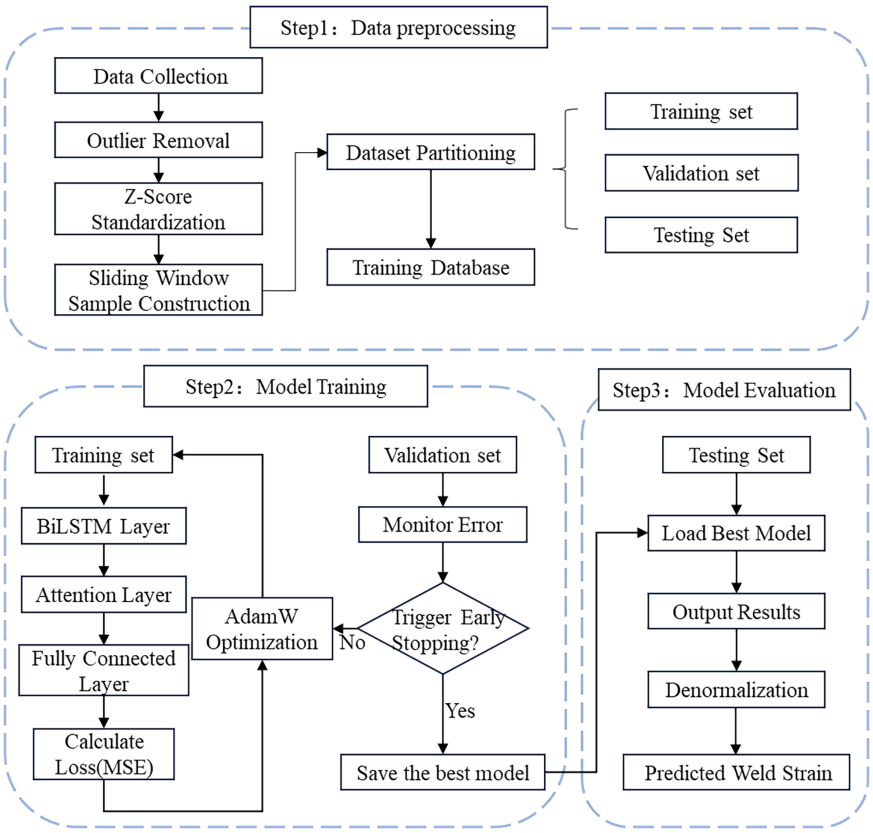

Figure 9: Model training process.

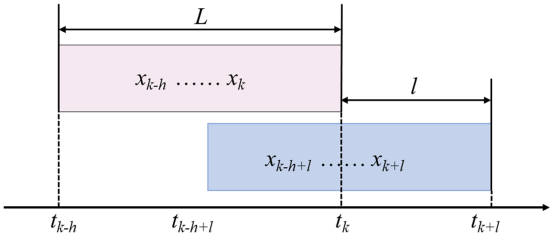

Figure 10: Sliding window model.

4 The Prediction Results and Discussion of BiLSTM-Attention Model

4.1 Data Processing and Model Training

Considering the significant differences in mechanical properties of BM, CW, LW and TW, it is not enough to establish an accurate mapping relationship by simply using strain history as input. Therefore, the principle of binary and One-Hot Encoding is used to label different types of samples. Although One-Hot Encoding is a discrete label, combined with the memory ability of BiLSTM, the model actually implicitly learns the differences in physical properties behind different weld areas (CW, LW, TW) through labels. This approach avoids the complexity of direct measurement of microscopic physical parameters, while ensuring the model's ability to distinguish heterogeneous structures. One-Hot coding results are shown in Table 3.

Table 3: One-hot labels of different types of samples.

Sample Shape	One-Hot Tag
BM	1000
CW	0100
LW	0010
TW	0001

Since the ratcheting behavior of the sample is evident at the initial stage of the cycle and is stable for a long time, only the strain data of the 1000 cycles before the ratcheting experiment are extracted as the sample data. Therefore, each sample contains 100,000 original data points under each working condition. At the same time, the time window length of the time sliding window model is set to 100, and the step size is set to 1. This scale can cover a complete cycle to ensure that the model can learn the complete 'loading-unloading' cycle characteristics so that each sample can generate 99,901 sets of 'input-output' sample pairs. This study contains three working conditions. In order to verify the generalisation ability of deep learning under unknown working conditions, all the data of working conditions B and C are used as the training set and test set. After mixing the data, the time series is divided into a training set and a verification set according to the ratio of 4:1. At this time, the training set contains about 639,360 samples, and the verification set contains 159,840 samples. All weld data under condition A are selected as independent test sets, including about 399,660 samples. Subsequently, the Z-score method is used to standardise the input data, and the calculation equation is shown in Eq. (15).

The AdamW optimiser is used to update the model parameters. The initial learning rate (Learning Rate) is set to 0.001, the batch size (BatchSize) is set to 128, the maximum iteration round (Epochs) is set to 100, and the number of heads of the multi-head attention mechanism is 4. The mean square error (MSE) of the 100-round iterative neural network model on the training set and the validation set, with the change curve of the iteration rounds (Epoch), is shown in Fig. 11. It can be seen from the diagram that the model loss value shows a sharp downward trend at the initial stage of training. This indicates that the model can quickly capture the main linear trends and periodic characteristics in the strain data. As the number of iterations increases, the slope of the loss curve gradually slows down and enters the stage of steady fine-tuning. After 50 training rounds, the two curves gradually stabilise and approach 0, indicating that the established BiLSTM-AM neural network model has high parameter fitting and generalisation ability.

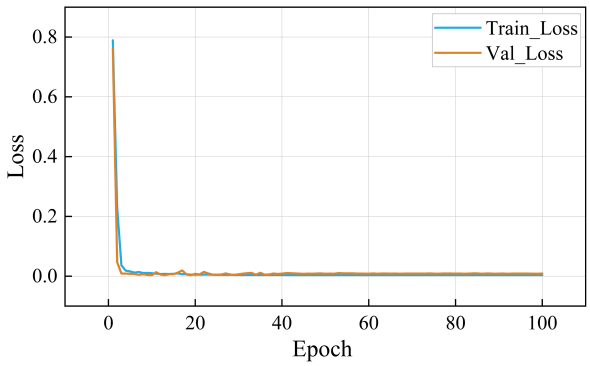

Figure 11: BiLSTM-MHA model training curve.

In order to evaluate the advantages of the BiLSTM-MHA model in predicting the ratcheting strain of the weld, the ablation experiment was carried out. Four models of LSTM, LSTM-MHA, BiLSTM and BiLSTM-MHA were selected. Taking the predicted ratcheting strain of the girth weld under condition A as an example, the error scatter diagram of the four models is shown in Fig. 12, and the specific error index is shown in Table 4. In the table, '↑' indicates that the evaluation index shows an increasing trend, and '↓' indicates that the evaluation index shows a growing trend.

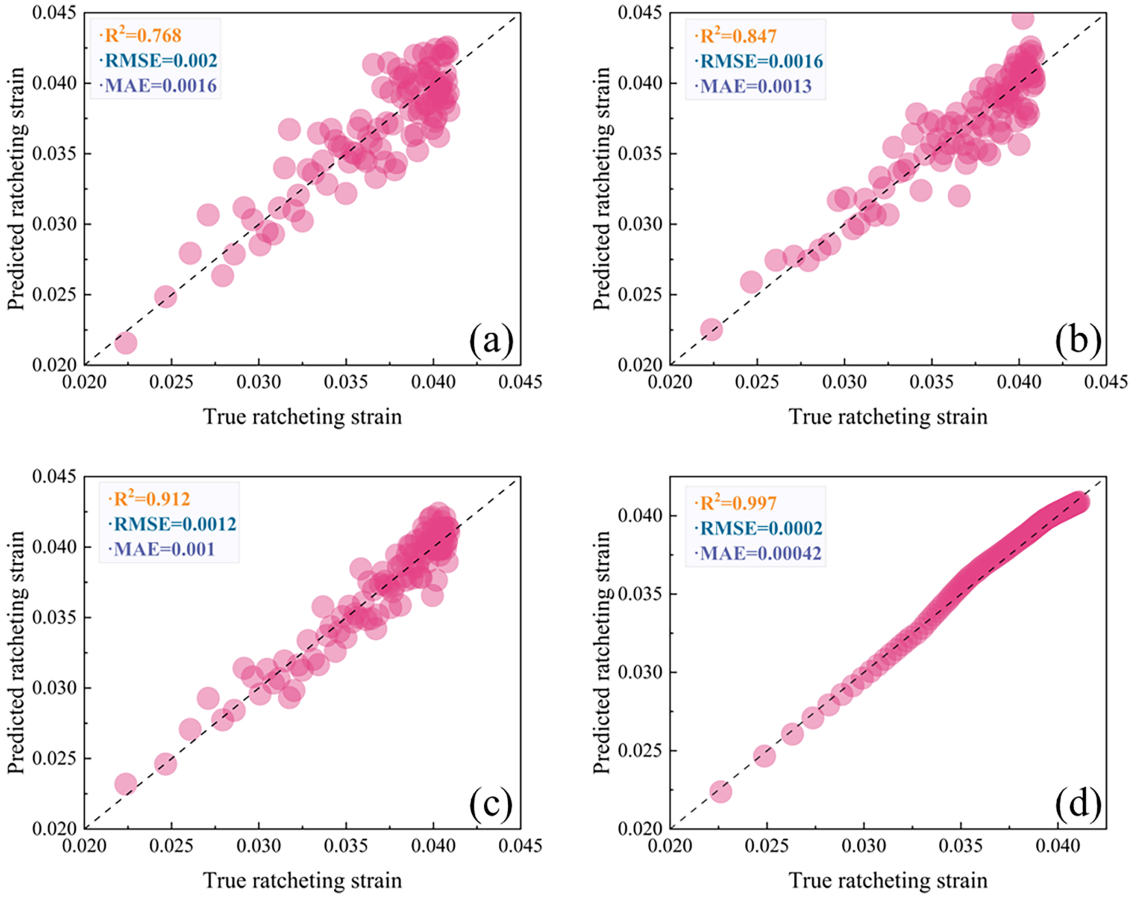

Figure 12: Scatter diagram of prediction error of girth weld strain under different models under working condition A. (**a**) LSTM; (**b**) BiLSTM; (**c**) LSTM-MHA; (**d**) BiLSTM-MHA.

Table 4: Comparison of error indexes under different models.

Model	R^2↑	RMSE↓	MAE↓
LSTM	0.768	0.002	0.0016
BiLSTM	0.847	0.0016	0.0013
LSTM-MHA	0.912	0.0012	0.001
BiLSTM-MHA	0.997	0.0002	0.00042

According to the R^2 index, BiLSTM-MHA is as high as 0.992, which is 29.8%, 17.7% and 9.3% higher than LSTM (0.768), LSTM-MHA (0.847) and BiLSTM (0.912), respectively. From the RMSE and MAE indicators,

BiLSTM-MHA has the smallest Value in the four models, indicating that the BiLSTM-MHA model has the most minor prediction error compared with the other three models. Through comprehensive comparison, it can be seen that the BiLSTM-MHA model has a better effect on the prediction of ratcheting strain of welds than the single model LSTM and BiLSTM. It not only has the highest prediction accuracy, but also shows strong robustness, which can predict the ratcheting strain behavior of welds most accurately.

4.2 Comparison of Prediction Results

In order to visually demonstrate the detail capture ability of the BiLSTM-MHA model, Fig. 13 takes the circumferential weld (CW) specimen under condition A as an example to draw a complete 'predicted strain-real strain' time domain comparison curve. From a macro point of view, the predicted curve and the actual curve have a high degree of coincidence throughout the loading history, and the model perfectly reproduces the transition process from the rapid growth of ratcheting strain at the initial stage of the cycle to the stable growth at the later stage. It can be clearly seen from the local magnification map that the model can not only predict the overall ratchet evolution trend, but also highly restore the detailed features of each cycle. The peaks and troughs of each cyclic curve are accurately predicted, and there is no obvious lag at the turning point of loading and unloading.

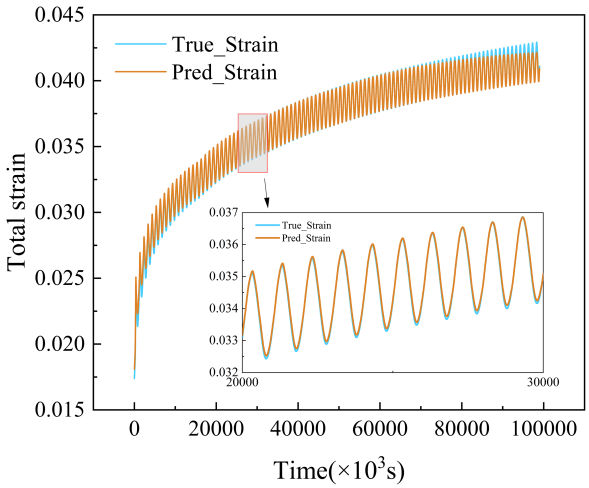

Figure 13: The strain curve predicted by the BiLSTM-MHA model.

In order to comprehensively evaluate the model's ability to predict the ratcheting behavior of the weld, Fig. 14 shows the ratcheting strain prediction curves of the BiLSTM-MHA model under three working conditions. It can be seen from Fig. 14 that the predicted ratcheting strain shows the ratcheting behavior of each sample in two stages: the ratcheting strain accumulates rapidly in the first stage. At the same time, as the number of cycles increases, each sample undergoes cyclic hardening, resulting in a decrease in the ratcheting strain accumulation rate and entering a stable stage. This shows that the model can not only predict the ratcheting behavior of the weld specimen, but also capture the cyclic hardening mechanism hidden in the time series data by learning.

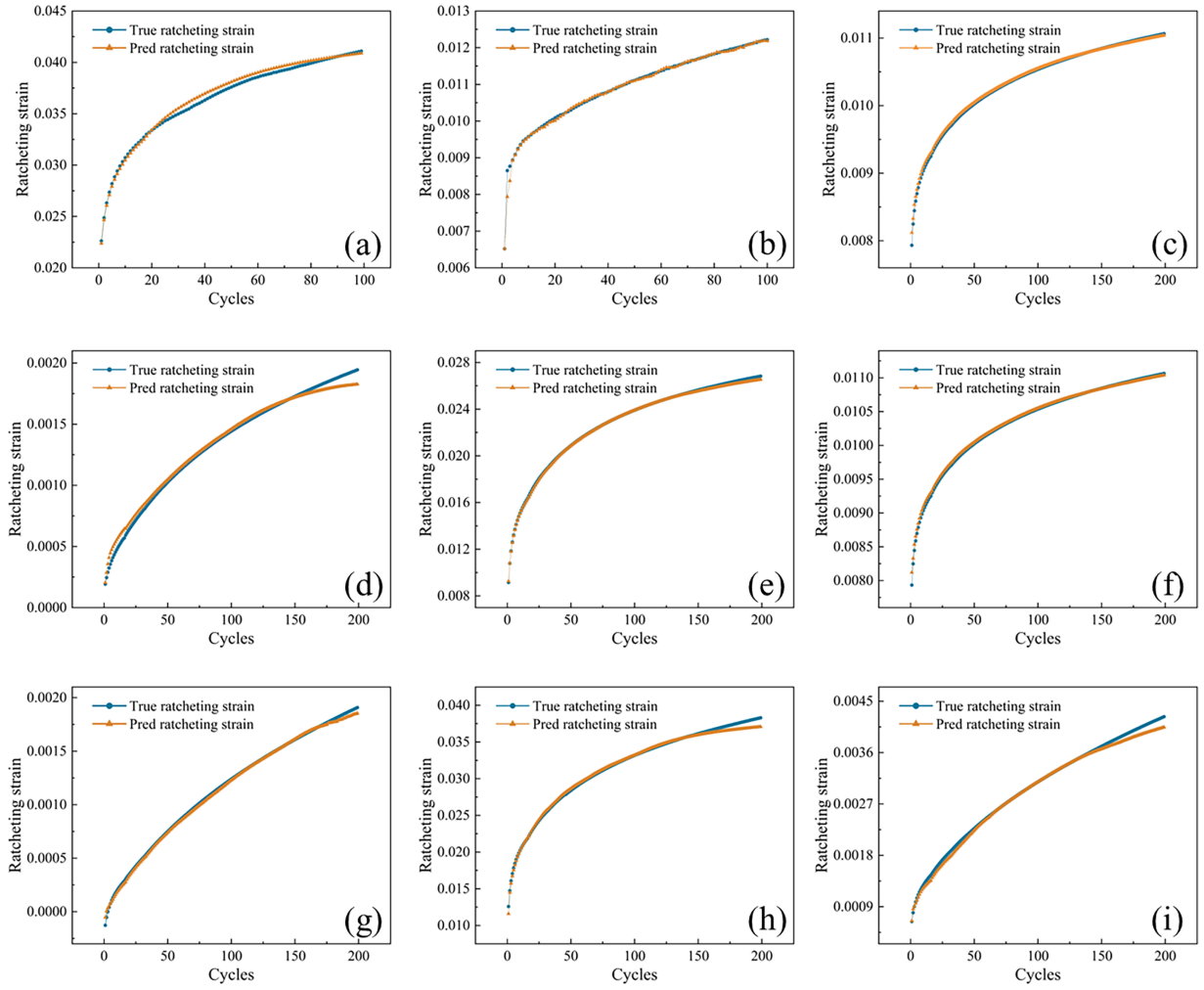

Figure 14: Predicted ratcheting strains of different specimens under various working conditions: (**a**) CondA-CW; (**b**) CondA-LW; (**c**) CondA-TW; (**d**) CondB-CW; (**e**) CondB-LW; (**f**) CondB-TW; (**g**) CondC-CW; (**h**) CondC-LW; (**i**) CondC-TW.

In order to evaluate the accuracy of the model prediction, Fig. 15 shows the error scatter diagram of each sample, each scatter represents the 'true strain-predicted strain' of each cycle, and the accuracy of the model is evaluated by R^2, RMSE and MAE. It can be seen from Fig. 15 that all the scatter points are concentrated near the diagonal, and the error does not increase with the increase of the number of cycles. At the same time, all the determination coefficients R^2 of the model are above 0.98, and the maximum values of root mean square error RMSE and mean absolute error MAE are 0.0004 and 0.243, respectively, indicating that the model has a high prediction accuracy for the ratcheting strain of the whole cycle of the weld specimen, which proves the feasibility of this method in the prediction of the ratcheting behavior of the weld.

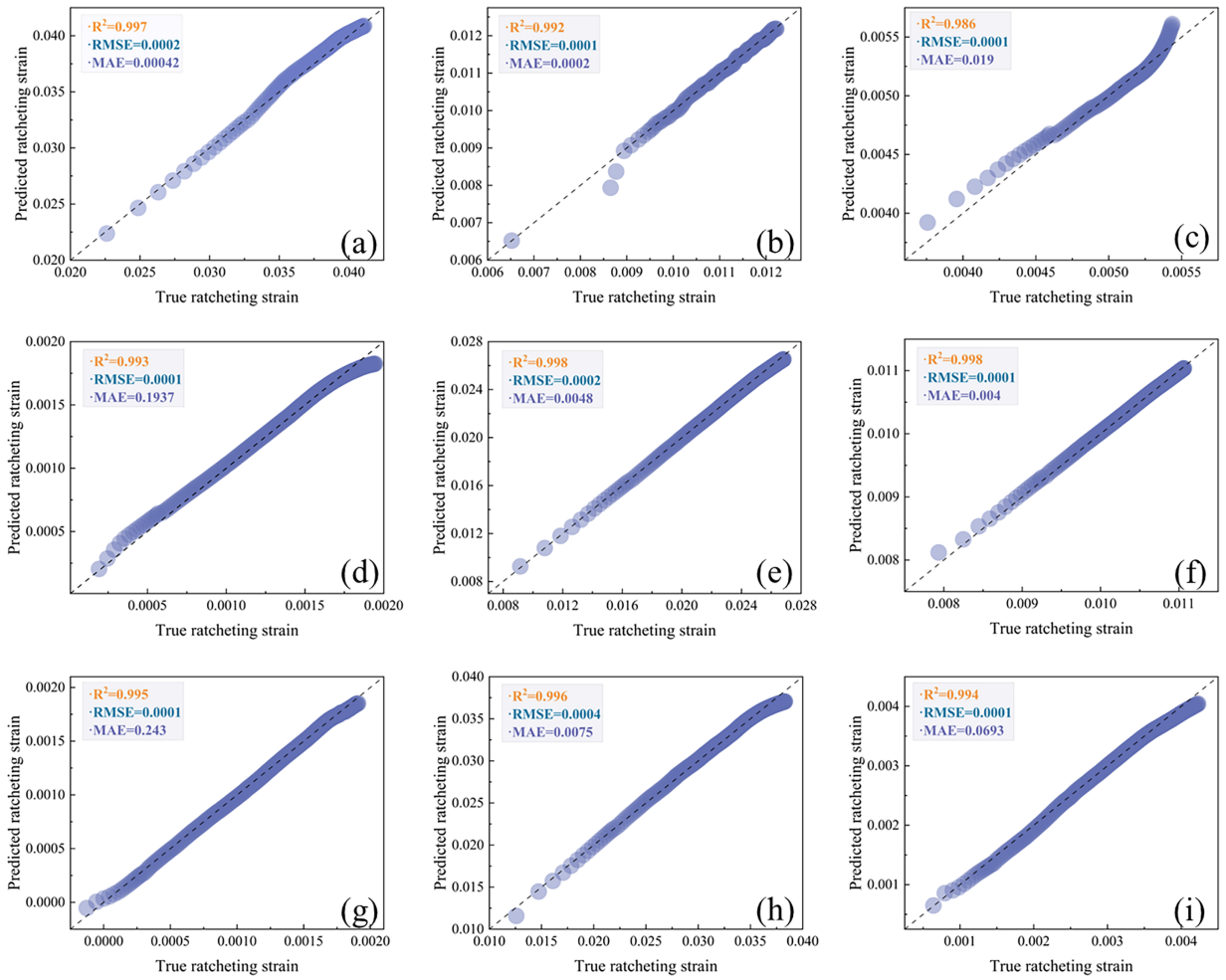

Figure 15: Prediction of ratcheting strain scatter error of different specimens under different working conditions: (**a**) CondA-CW; (**b**) CondA-LW; (**c**) CondA-TW; (**d**) CondB-CW; (**e**) CondB-LW; (**f**) CondB-TW; (**g**) CondC-CW; (**h**) CondC-LW; (**i**) CondC-TW.

5 Conclusion

Aiming at the problem that the ratcheting behavior of the weld in the multi-layered pressure vessels is difficult to predict, this paper carried out a systematic asymmetric cyclic loading test, revealed the ratcheting evolution law of the weld structure, and proposed a method for predicting the ratcheting behavior of welded joints based on BiLSTM-MHA. The main conclusions are as follows:

(1) The ratcheting behavior of S31603 base metal and welded structure under asymmetric cyclic loading is characterised by two stages. In the first stage, the ratcheting strain accumulates rapidly, but the ratcheting strain accumulation rate decreases gradually. In the second stage, the ratcheting strain gradually accumulates at a stable rate.

(2) The ratcheting behavior of the weld structure and base metal has an obvious difference. Under the same load, the ratcheting strain of the circumferential weld is the lowest, followed by the T-type weld and the longitudinal weld. The ratcheting strain of the base metal is the largest, and the accumulation rate is the fastest.

(3) The ratcheting behavior prediction method of welded joints based on BiLSTM-MHA is proposed, and the nonlinear mapping relationship from the base metal reference to the weld structure response is successfully established. Compared with the traditional constitutive model, this model can not only accurately capture the evolution characteristics of ratcheting behavior, but also does not need to calibrate the parameters of a complex constitutive model. This method can predict the ratcheting behavior of complex weld structures only by relying on the standard data of base metal, which provides a new idea of low cost and high precision for online safety monitoring and life evaluation of multi-layer wrapped pressure vessels.

Acknowledgement: The authors would like to thank all the members involved in this project for their help in developing this article.

Funding Statement: This work was supported by The Key Research and Development Program of Gansu Province—Industrial Category, China (Grant No. 24YFGA018); The Science and Technology Program of the State Administration for Market Regulation of China (Grant No. 2024MK129); and The Gansu Provincial Department of Education: Gansu Provincial Graduate Student "Innovation Star" Project, China (Grant No. 2025CXZX612).

Author Contributions: Linbin Li: Algorithm design, experiments, data processing, draft writing. Ruiyuan Xue: Research direction, supervision, funding, manuscript revision. Juyin Zhang: Experimental assistance, review. Xueping Wang: Model verification. Tiantian Chu: Experimental assistance, visualization, construction of data resources. All authors reviewed and approved the final version of the manuscript.

Availability of Data and Materials: The authors confirm that the data supporting the findings of this study are available within the article.

Ethics Approval: Not applicable.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Zheng JY, Ma K, Ye S, Gu CH, Hua ZL, Peng WZ. Development status and challenges of equipment for storage and transportation of high-pressure gaseous hydrogen in China. Press Vessel Technol. 2022;39(3):1–8. (In Chinese).
2. Hua Z, Gao W, Chi S, Wang X, Zheng J. Development status and challenges of high-pressure gaseous hydrogen storage vessels and cylinders in China. Renew Sustain Energy Rev. 2025;214:115567. doi:10.1016/j.rser.2025.115567.
3. Xue R, Zhang X, Zhang J, Wang X, Zhang Y, Li L, et al. Stress distribution characteristics and intelligent online monitoring methods for multilayer wound pressure vessel. Int J Press Vessels Pip. 2026;221(20):105742. doi:10.1016/j.ijpvp.2025.105742.
4. Yang D, Yang Z, Zhai Z, Chen X. Homogenization and localization of ratcheting behavior of composite materials and structures with the thermal residual stress effect. Materials. 2019;12(18):3048. doi:10.3390/ma12183048.
5. Li ZF, Wang XB, Yang LJ. The structure, damage and inspection of multi-layer wrapped pressure vessels. Process Equip Pip. 2022;59(3):21–6. doi:10.3969/j.issn.1009-3281.2022.03.003.
6. Khaled I, Vasiukov D, Shakoor M, Bennebach M, Chaki S. Digital twin for predicting progressive damage in operating pressure vessels. Procedia Struct Integr. 2024;57(5):280–9. doi:10.1016/j.prostr.2024.03.030.
7. Luo H, Kang G, Kan Q, Huang Y. Experimental investigation on the heterogeneous ratchetting of SUS301L stainless steel butt weld joint during uniaxial cyclic loading. Int J Fatigue. 2017;105:169–79. doi:10.1016/j.ijfatigue.2017.08.027.
8. Luo H, Qu K, Yu C, Kan Q, Kang G. Experimental study on multiaxial ratchetting-fatigue interaction of SUS301L stainless steel tubular welded joint. Int J Fatigue. 2024;186:108411. doi:10.1016/j.ijfatigue.2024.108411.
9. Wang H, Jing H, Zhao L, Han Y, Lv X, Xu L. Uniaxial ratcheting behaviour of 304L stainless steel and ER308L weld joints. Mater Sci Eng A. 2017;708:21–42. doi:10.1016/j.msea.2017.09.109.

10. Li N, Yang Y, Tian Y, Yue Y, Gao S, Yang L. Failure analysis of laser-welded Q690qENH steel joints under tensile and fatigue loading. Eng Fail Anal. 2025;181(10):109911. doi:10.1016/j.engfailanal.2025.109911.
11. Zeng Z, Li G, Liu SW, Huang DH, Liu YP, Chan SL. Low-cycle fatigue of laser-arc hybrid welded S960 high-strength steel joints. J Constr Steel Res. 2026;236(8):110040. doi:10.1016/j.jcsr.2025.110040.
12. Janulionis R, Dundulis G. Numerical simulation of low-cycle fatigue test of welded 9Cr-1Mo steel at 550°C. Eng Struct. 2025;326(4):119540. doi:10.1016/j.engstruct.2024.119540.
13. Liu H, Chen X, Zhang X, Cui X. Experimental and numerical analysis of ratcheting behavior of super duplex SAF2507 stainless steel under uniaxial loading. Appl Sci. 2025;15(3):1424. doi:10.3390/app15031424.
14. Dong Z, Xie XF, Jiang W, Niu R, Wan Y, Zhai X, et al. Fatigue failure mechanism of duplex stainless steel welded joints including role of heterogeneous cyclic hardening/softening: experimental and modeling. Int J Fatigue. 2024;178(4):108010. doi:10.1016/j.ijfatigue.2023.108010.
15. Kalnins A, Rudolph J, Willuweit A. Using the nonlinear kinematic hardening material model of chaboche for elastic-plastic ratcheting analysis. J Press Vessel Technol. 2015;137(3):31006. doi:10.1115/1.4028659.
16. Ding YJ, Wang ZC, Xin Y, Liu WL, Cheng M. Structural nonlinear model updating based on CNN-BiLSTM-Attention hybrid neural network. Eng Mech. 2026;1(13):1–13.
17. Ding M, Li J, Gao D, Zhou G, Wang B, Wu Z. Fatigue life prediction of composite materials based on BO-CNN-BiLSTM model and ultrasonic guided waves. Comput Mater Continua. 2025;85(1):597–612. doi:10.32604/cmc.2025.067907.
18. Kang G, Dong Y, Wang H, Liu Y, Cheng X. Dislocation evolution in 316L stainless steel subjected to uniaxial ratchetting deformation. Mater Sci Eng A. 2010;527(21–22):5952–61. doi:10.1016/j.msea.2010.06.020.
19. Hochreiter S, Schmidhuber J. Long short-term memory. Neural Comput. 1997;9(8):1735–80. doi:10.1162/neco.1997.9.8.1735.