Gao *et al. J Mater Inf* 2024;4:25
DOI: 10.20517/jmi.2024.28

Journal of Materials Informatics

Research Article

Open Access

Check for updates

Knowledge assisted machine learning to clarify pore influence on fatigue life of forging/additive hybrid manufactured Ti-17 alloy

Shuailong Gao[1,2,3,#], Wenyuan Li[1,#], Yuting Ma[2], Baitao Wang[2], Xiaolin Dong[1], Shujun Li[1], Jianrong Liu[1,*], Yi Yang[2,*], Shen Qu[4], Zhenlin Chen[4], Hao Wang[1,*], Rui Yang[1,5]

[1]Institute of Metal Research, Chinese Academy of Sciences, Shenyang 110016, Liaoning, China.
[2]School of Materials and Chemistry, University of Shanghai for Science and Technology, Shanghai 200093, China.
[3]Beijing Key Laboratory of Lightweight Multi-functional Composite Materials and Structures, Institute of Advanced Structure Technology, Beijing Institute of Technology, Beijing 100081, China.
[4]AECC Shenyang Liming Areo-engine Co., Ltd., Shenyang 110043, Liaoning, China.
[5]Centre for Adaptative System Engineering, School of Creativity and Arts, ShanghaiTech University, Shanghai 201210, China.
[#]Authors contributed equally.

[*]**Correspondence to:** Prof. Hao Wang, Prof. Jianrong Liu, Institute of Metal Research, Chinese Academy of Sciences, 72 Wenhua Road, Shenyang 110016, Liaoning, China. E-mail: haowang@imr.ac.cn; jrliu@imr.ac.cn; Prof. Yi Yang, School of Materials and Chemistry, University of Shanghai for Science and Technology, 580 Jungong Road, Shanghai 200093, China. E-mail: yangyi@usst.edu.cn

How to cite this article: Gao S, Li W, Ma Y, Wang B, Dong X, Li S, Liu J, Yang Y, Qu S, Chen Z, Wang H, Yang R. Knowledge assisted machine learning to clarify pore influence on fatigue life of forging/additive hybrid manufactured Ti-17 alloy. *J Mater Inf* 2024;4:25. https://dx.doi.org/10.20517/jmi.2024.28

Received: 31 Jul 2024 **First Decision:** 11 Oct 2024 **Revised:** 16 Nov 2024 **Accepted:** 28 Nov 2024 **Published:** 30 Nov 2024

Academic Editors: Chaolin Tan, Qian Ma **Copy Editor:** Pei-Yun Wang **Production Editor:** Pei-Yun Wang

Abstract

Forging/additive hybrid manufactured Ti alloy parts suffer from relatively low fatigue life due to the existence of metallurgical defects in the transition zone, which also brings difficulty to fatigue life modeling. In this work, the synergistic effect of pore size and location on the rotating-bending fatigue life of hybrid manufactured Ti-5Al-2Sn-2Zr-4Mo-4Cr (Ti-17) samples was systematically investigated with the combination of machine learning approaches and physical knowledge. A machine learning framework with a back propagation neural network and generative adversarial network (GAN) was constructed and employed on sparse and limited datasets. A general and interpretable model was obtained with a high level of 90% confidence. In general, the fatigue life of hybrid manufactured Ti-17 alloys decreases with pore size and increases with its distance to surface. Specifically, critical sizes were obtained for near-surface and in-depth pores that have negligible influence on fatigue life of hybrid

manufactured samples with respect to pore-free samples. The present work thus provides a systematic platform for the evaluation of the fatigue performance of hybrid manufactured titanium alloys.

Keywords: Hybrid manufacturing, fatigue, titanium, defect, machine learning

INTRODUCTION

Owing to their high strength, low density, and superior corrosion resistance[1], titanium alloys, such as Ti-5Al-2Sn-2Zr-4Mo-4Cr (Ti-17), are widely used in aeronautical, nautical and biomedical applications, whereas such advantage is largely restricted due to the overall high manufacturing expense and long period of use. Forging/additive hybrid manufacturing is thus widely used on structural materials to extend service life with less expense, especially for the repair of aeroengine parts[2,3]. With the concern for both performance and efficiency, hybrid manufacturing combined with forging and wire arc additive manufacturing (WAAM) is currently popular for repairing Ti alloy parts, where the arc-fused wire is additively grown on the forged substrate. However, metallurgical defects, in particular pores, are still inevitable in additive manufactured titanium alloys, which results in significantly deteriorated fatigue performance compared with the wrought counterparts[4-7].

For the repaired parts, an appropriate performance decrease compared to brand-new parts is acceptable as long as their use remains economically viable; e.g., the overall service life of aeroengine blades that have experienced sudden failure may reach 80% of the original design after repair. However, such a decrease must be strictly evaluated and precisely predicted. Unfortunately, the randomness of the additive manufacturing defect features, e.g., size, shape, location, *etc.*, leads to large sparsity in the fatigue test results, and the prediction of fatigue life remains a challenge for additive manufactured metallic materials[8-11]. With the development of artificial intelligence[12], machine learning[13,14], including deep learning[15], and relevant techniques have emerged as important tools for materials science[16-18]. Currently, machine learning is one of the most popular research directions, which uses computational methods to obtain knowledge from data through association, classification, clustering, and regression. Based on a given training sample, machine learning aims to estimate dependencies between inputs and outputs of the system so that future outputs can be predicted as accurately as possible. As a result, machine learning methods can be readily adapted to a wide range of processes with little or no assumptions[19-21]. Random forest, support vector regression, and artificial neural network models are three commonly used regression algorithms for predicting fatigue life. Artificial neural networks, particularly, which have long been used to predict S-N (S: fatigue stress, N: fatigue life) relations[22], have demonstrated excellent performance in mapping complex nonlinear relations[23,24]. When compared to traditional statistical methods, machine learning has higher computational accuracy and efficiency for small-sample prediction and nonlinear regression analysis[25,26]. Machine learning techniques are beneficial in solving engineering problems because of their ability to recognize patterns in complex data[27-31].

Concerning the fatigue property, the influence of defects on the fatigue behavior has been investigated in additive fabricated Ti-6Al-4V[32] and machine learning with was used to study fatigue life and performance of additive manufactured parts[33]. Despite the success of machine learning in predicting fatigue performance, the requirement of big data is usually unachievable for structural materials, in particular for Ti alloys, due to both time and cost. Hence, it is of great importance to establish a rational machine-learning approach to fatigue life prediction based on sparse and scattering datasets. Indeed, the idea of employing the machine learning approach is to benefit from its power on processing data, i.e., identifying data features and relationships between features, so as to guide the discovery of influencing factors. Previous studies have somewhat demonstrated the adverse effect of the location, size, and morphology of defects on the fatigue

properties of additive manufactured Ti-6Al-4V[32,34-38]. However, due to the characteristics of rotating-bending fatigue, the relation between defect parameters, in particular pore location, is completely different from ordinary fatigue tests.

Therefore, in the present paper, employing forging/additive hybrid manufactured Ti-17 alloys, we focus on their rotating-bending fatigue life and further examine the methodology of dealing with small and sparse datasets. Incorporating the back propagation neural network, generative adversarial network (GAN), and physical knowledge, we aim to provide a systematic platform for the prediction of rotating-bending fatigue properties of hybrid fabricated titanium alloys. As such, we intend to stress the scientific importance of (1) the appropriate combination of various machine learning techniques on processing small datasets; (2) the use of physical knowledge to supervise the model; and (3) the application of the obtained physical model on evaluating the performance of forged/additively manufactured Ti-17 alloys.

MATERIALS AND METHODS

Specimen and fatigue test

Ti-17 segments were produced with WAAMed layers on a forged base [Figure 1A]. Standard bar samples of ϕ4 mm [Figure 1B] were taken from the hybrid regions and the rotating-bending fatigue tests were performed at ambient temperature in line with the GB/T 4337-2008. The effective stress at the narrowest point of the specimen was set as 500 MPa. The samples rotated at a speed of 5,000 r/min with a stress ratio R = -1.

Theoretical models

Back propagation neural network

The accuracy of the back propagation neural network prediction results is strongly influenced by the training process, as they are based on maximum likelihood algorithms inspired by neural networks in the brain. In engineering, back propagation neural networks are widely used to solve complex problems because of their faster and more accurate predictive power. In these networks, the mathematical equations for how inputs and outputs map together can be learned through training to get the best possible result. The schematic diagram of the back propagation neural network used is shown in Figure 2A.

The first and last layers of a neural network are called the input and output layers, respectively, while the middle layer is called the hidden layer. A neural network is trained by passing signals between the layers. The operation of a single neuron can be expressed as follows:

$$y_i = f_i(\sum w_{ij} x_j - t_i) \quad (1)$$

where y_i represents the output, w_{ij} is the weights, x_j is the inputs, t_i is the inputs and f_i is the activation function.

Sigmoid, Relu and Tanh functions are some of the commonly used activation functions[39,40]. With gradient descent, the back propagation neural network uses error back propagation to train its Multilayer feed-forward network. It uses mathematical chain rules to calculate the gradient of each layer, and then returns to the input layer to adjust the weights. This algorithm calculates the error rate of the predicted data[41]. As the network's actual output value compares with its desired output value, this step is repeated until the mean square error (MSE) is as small as possible. The coefficient of determination (R^2) and MSE, which are used to assess the prediction accuracy, are expressed as

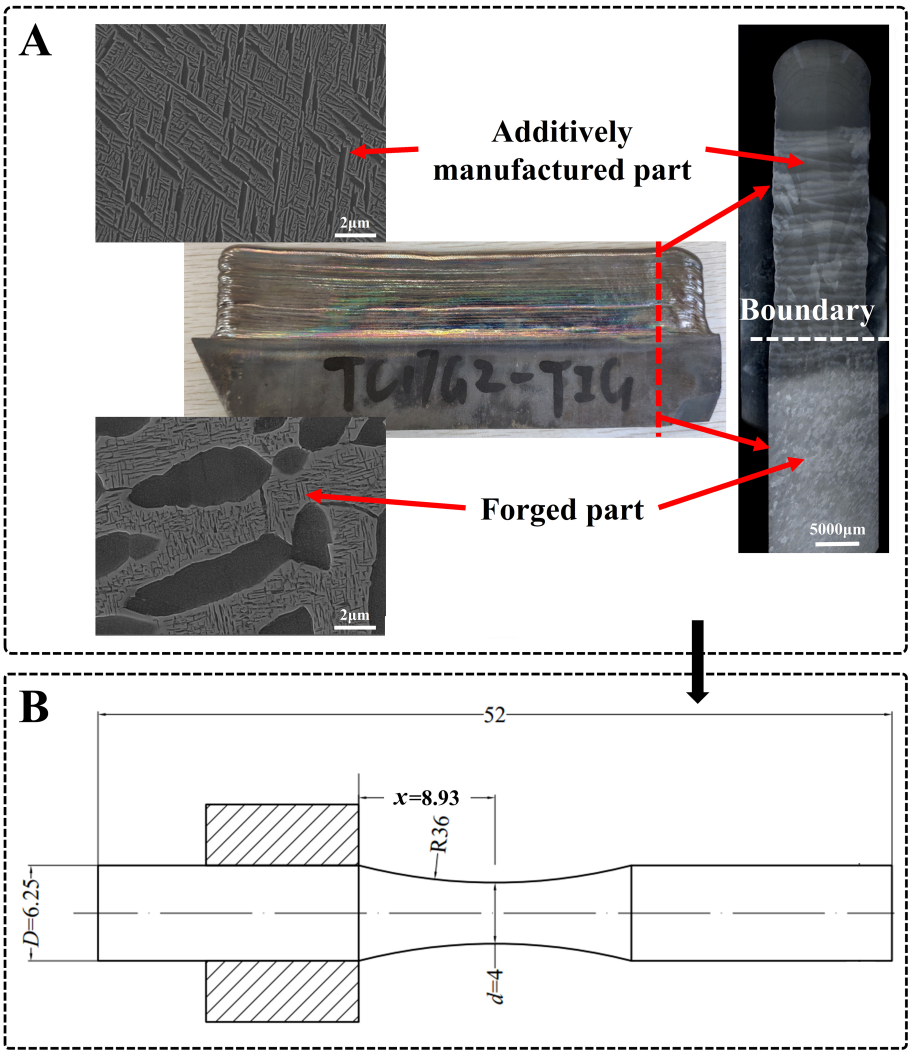

Figure 1. (A) Forging/additive hybrid manufactured Ti-17 part and the microstructures of the additively manufactured and forged parts; (B) Schematic of the rotating bar bending fatigue sample (unit in mm). Ti-17: Ti-5Al-2Sn-2Zr-4Mo-4Cr.

$$R^2 = \frac{\sum_n^{i=1} (\hat{y}_i - \hat{y}_i)^2}{\sum_n^{i=1} (y_i - \hat{y}_i)^2} \tag{2}$$

$$\text{MSE} = \frac{1}{n} \sum_n^{i=1} (y_i - \hat{y}_i)^2 \tag{3}$$

where $\hat{y}_i$ is the predicted value, and the rest of the parameters are the same as above.

The input variables are the defect size (pore size) and the distance to surface (pore location), while the output variable is the fatigue life. Using a random distribution, 70% of the dataset was employed for training, while 15% was allocated for validation and 15% for testing. The reference or hidden layer of a back propagation neural network model is generally determined without a uniform rule. This work uses back propagation neural networks with a topology of three, four and five neurons in the first hidden layer [Table 1]. According to the results, when the number of neurons in the first hidden layer is 4, the model has

Table 1. Predicted performance of the back propagation neural network models with different numbers of neurons in the first hidden layer

Back propagation neural network	Number of neurons 1st layer	Metrics					
		R^2			MSE		
		Training	Validation	Testing	Training	Validation	Testing
Model 1	3	0.835997	0.829817	0.837631	0.088284	0.657599	0.249147
Model 2	4	0.804974	0.098125	0.944086	0.161155	0.095284	0.275916
Model 3	5	0.807096	0.920224	0.921055	0.201762	0.210129	0.155929

R^2: The coefficient of determination; MSE: mean square error.

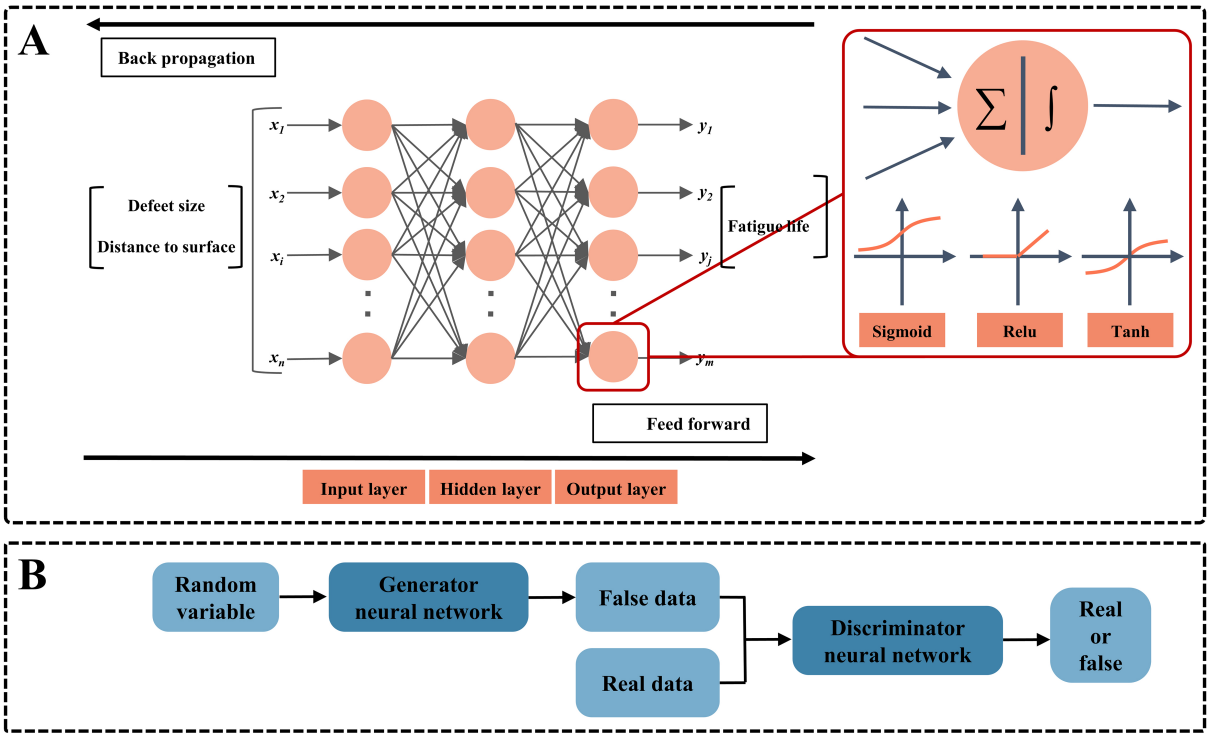

Figure 2. (A) Back propagation neural network model structure diagram; (B) GAN model structure diagram. GAN: Generative adversarial network.

the best overall performance.

GAN

The primary idea of the GAN is the concept of a zero-sum game as applied to deep learning neural networks, where the generative network (generator) and the discriminator network (discriminator) play a continuous game while the generator learns the data distribution. Traditional neural network topological structures are constrained and typically are only able to predict numerical or categorical outcomes based on the input data. GANs, on the other hand, are an unsupervised learning technique, and trained GANs are capable of producing totally new data on their own with a wide range of possible applications.

The two neural networks that make up the GAN, a generator and a discriminator, are built to compete with one another [Figure 2B]. After training, the generator can produce data that can be faked to trick the discriminator, as the discriminator is trained to classify the data in the training set as real data and the data

generated by the generator as fake data. The standard GAN training loop contains three steps:
(1) Use genuine training data to train the discriminator.
(2) Train the discriminator using the created data.
(3) Train the generator to generate data and trick the discriminator.

Physical knowledge model
The general relationship between the fatigue stress and the fatigue life[42-46], i.e., the S-N curve, is written as

$$\sigma_{eff} * N_f = C \tag{4}$$

In additive manufactured materials[47-50], the existence of pores produces additional stress concentration, which largely influence the fatigue performance[51-53]. Regarding the relationship between fatigue life and defect parameters[5,54,55], previous investigations have derived

$$\{S\sigma_a(area_{eff})^{1/12}L^{\beta}\}^{\alpha}N_f = C \tag{5}$$

$$L = 1 - \frac{d_1 + d_2}{2r} \tag{6}$$

where N_f refers to the fatigue life, with C as a constant, S the defect shape factor, σ_a the unified stress amplitude, $area_{eff}$ the effective size of the defects, α and β the material dependent constants, r the radius of the fatigue fracture surface, and d_1 and d_2 the distances from defect to fatigue fracture surface. However, in rotating-bending fatigue tests, since the overall stress is related to the distance to sample surface, the above equations do not apply and the parameter of the distance of defects to the sample surface needs to be considered. In this study, the pore size ranges between 10 to 250 μm, and under the rotating-bending condition, stress concentration in the whole sample needs to be considered. We thus employ finite element analysis to evaluate stress concentration and reemploy the common fatigue model of Equation (4) in the subsequent model development.

We first evaluate the influence of pore size and location on stress distribution in the samples by simulating the bending of pore-free and pore-containing samples with different pore diameter and locations with the finite element method (FEM) employing the software ANSYS. Tetrahedral mesh partitioning method is used to build a mesh model for the three-dimensional solid model and mesh size is 0.05 mm for all parts. The minimum time step was set to 10^{-5} s. The left end of the sample was fixed and a 500 N torsional force was applied along the Y direction at the right end [Figure 3].

The equivalent stress distribution on the sample cross sections is systematically studied and Figure 4A-D exemplifies the cases for near-surface [Figure 4A], half-radius [Figure 4B], center [Figure 4C] pores of 100 μm in radius and a pore-free case [Figure 4D]. To evaluate the stress concentration, the overall stress ratio at a fixed point P (0.05 mm from the surface) with respect to pore size and location is shown in Figure 4E and F. In general, by fitting the FEM results, the stress ratio increases with the pore size and decreases with pore distance to surface with the following relationship.

$$\frac{\sigma_{E-pore-contain}}{\sigma_{E-pore-free}} = \frac{S*1.005}{1+x^{1.1291}} * 1.2295^{1.0172} + 0.8868 \tag{7}$$

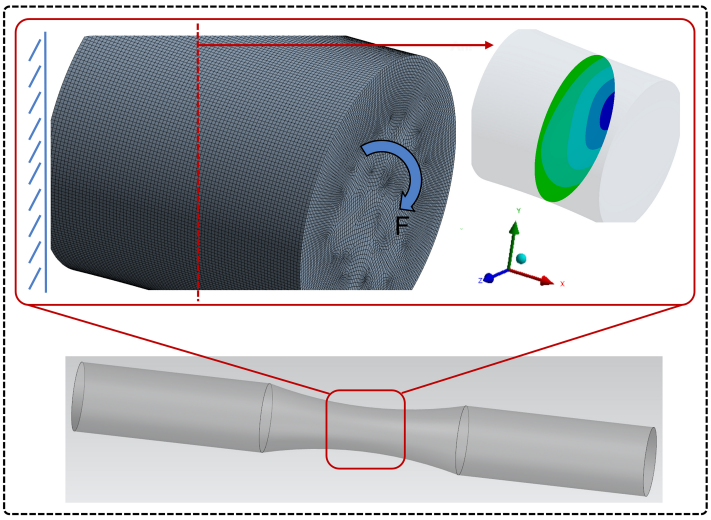

Figure 3. Boundary conditions for finite element simulation.

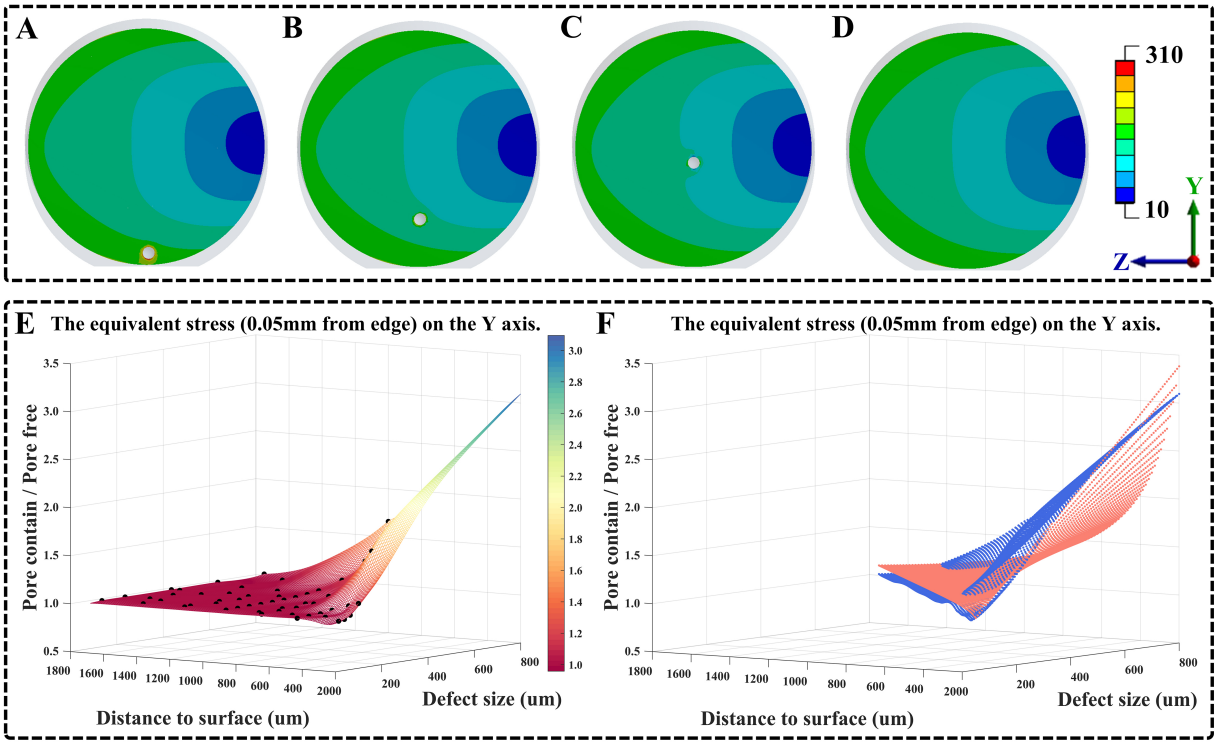

Figure 4. Equivalent stress distribution on the cross sections of the pore-containing (A-C) and pore-free (D) samples, pores are 100 um in radius. Equivalent stress distribution of pore size and location on the Y axis negative direction from the edge to the center of the circle at 0.05 mm from the fixed point; (E) black dots are equivalent stress and (F) data were fitted using Equation (7). Blue dots are finite element simulation data, red dots are equation fitting data.

where S is the pore size and x is the pore location.

Regarding the above stress concentration ratio and considering the potential size and location limit in this work, the effective stress after can be written as:

$$\sigma_{eff} = \sigma_0 * \left[1 + \frac{a*(\frac{S}{S_0})^b}{(1+\frac{x}{x_0})^c}\right]^m \tag{8}$$

where σ_0 is the fatigue stress during the test (500 MPa in the present work), with a, b, c, m, S_0, and x_0 being constants.

The relationship between σ_{eff} and fatigue life (N_f) is:

$$\sigma_{eff} * N_f = \sigma_0 * \left[1 + \frac{a*\left(\frac{S}{S_0}\right)^b}{\left(1+\frac{x}{x_0}\right)^c}\right]^m * N_f = I \tag{9}$$

$$\log N_f = \log I - \log \sigma_0 - \log\left[1 + \frac{a*(\frac{S}{S_0})^b}{(1+\frac{x}{x_0})^c}\right]^m = d - m * \log\left[1 + \frac{a*(\frac{S}{S_0})^b}{(1+\frac{x}{x_0})^c}\right] \tag{10}$$

where d is the material constants, and the rest of the parameters are the same as above.

Overall strategy

In this work [Figure 5], 34 sets of defect size, defect to distance and fatigue life data were obtained from fracture photos after forging/additive components experienced rotating-bending fatigue fracture. The sparse and limited data set then prevented further analysis of the work. In order to overcome the influence of data quality, the corresponding fatigue life equation is established on the basis of physical knowledge, and it is applied to data processing and model interpretation. In order to overcome the influence of data quantity, a GAN based on Pytorch deep learning framework is constructed that conforms to the laws of defect size, distance to surface and fatigue life. Finally, a machine learning framework with physical knowledge predicts the fatigue life to be smooth by defect size and defect to distance.

RESULTS AND DISCUSSION

Fracture morphology

The fracture morphology was observed under scanning electron microscopy for spherical crack sources. Typical fractographs are shown in Figure 6 with the fatigue cycles to failure and defect characters. Generally, samples with larger defects and shorter distance to surface exhibit shorter fatigue life. The sample with a pore size of Φ127.5 μm closest to the surface displays the shortest life of 29,700 cycles [Figure 6A]. Samples with intermediate pore size and distance to surface demonstrate intermediate lives of 254,000 [Figure 6B] and 545,000 cycles [Figure 6C]. The sample with the smallest pore size and largest distance to surface presents the longest fatigue life of 6,410,000 cycles [Figure 6D].

Data preprocessing

Owing to the existence of pores, a majority of the fatigue samples fail at the hybrid regions. Due to the lack of data for arbitrarily shaped pores, only that from spherical pores is collected. The original 34 sets of experimental data with the fatigue life against the pore size and location are shown in Figure 7A and B. For data preprocessing, we have operations such as logging the fatigue life. We find that the fatigue life processed by log can better find the relationship with the defect parameters [Figure 7C and D]. However, remarkable sparsity is readily observed, which hinders the immediate employment of machine learning algorithms. Therefore, preprocessing of the raw data is necessary for the successful and accurate prediction

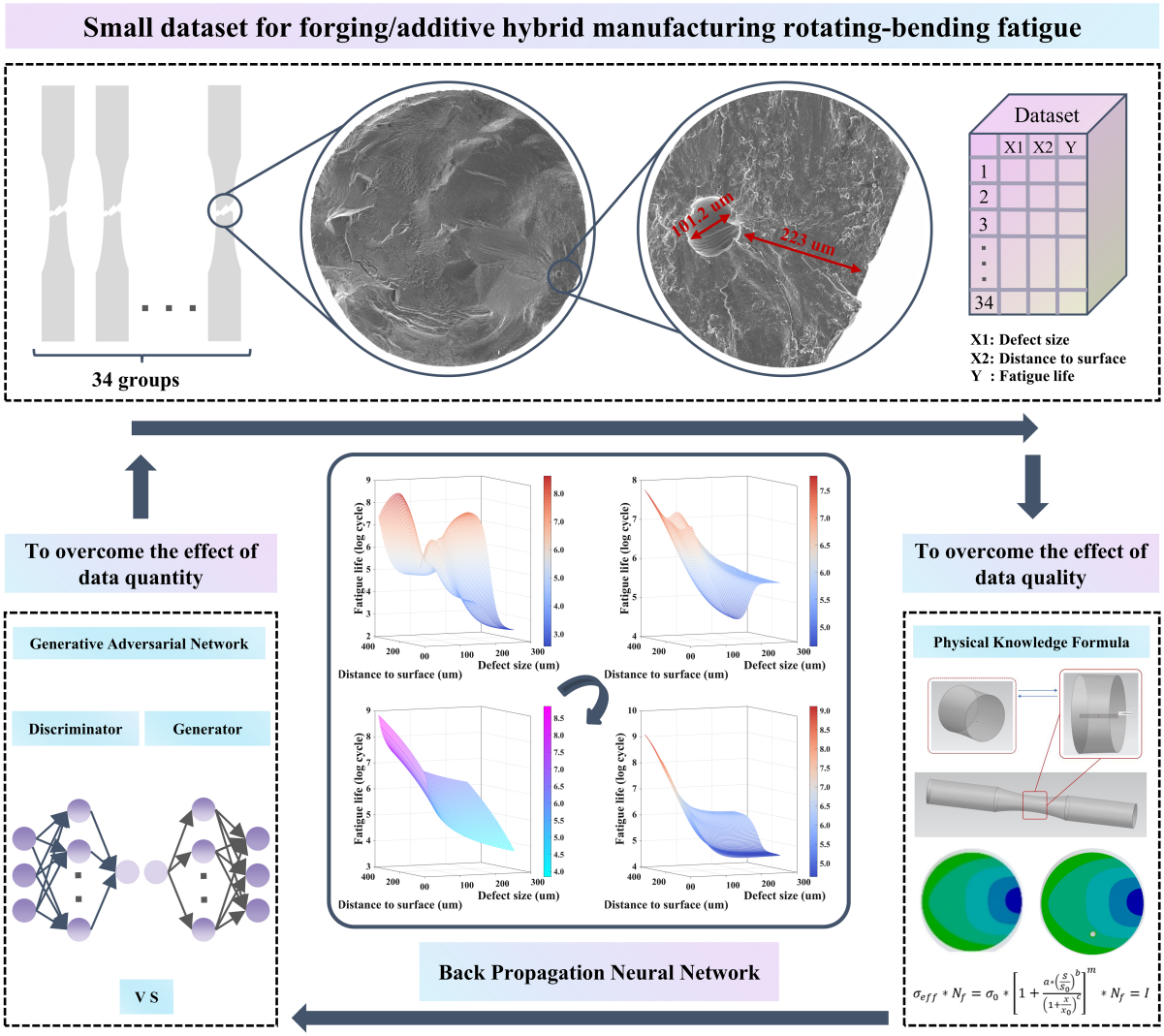

Figure 5. The overall flow chart of the present work. A total of 34 sets of defect size, defect to distance and fatigue life data were obtained from fracture photos after forging/additive components experienced rotating-bending fatigue fracture. The sparse and limited data set then prevented further analysis of the work. In order to overcome the influence of data quality, the corresponding fatigue life equation is established on the basis of physical knowledge, and it is applied to data processing and model interpretation. In order to overcome the influence of data quantity, a GAN based on Pytorch deep learning framework is constructed that conforms to the laws of defect size, distance to surface and fatigue life. Finally, a machine learning framework with physical knowledge predicts the fatigue life to be smooth by defect size and defect to distance.

of fatigue life.

Then, outlier rejection was performed using statistical methods, employing two distinct approaches. Values that do not fall within the interval, i.e., mean ± 2× standard deviation, are considered outliers. Outliers are defined as values that are not within the interval, i.e., lower four medians -1.5× median deviation, upper four medians +1.5× median deviation. Figure 8A and B shows the box line plots of defect size, distance to surface and fatigue life. After discrete value rejection of the raw data, the data quality has been greatly improved. Finally, the data of defect size, distance to surface and fatigue life are in the vicinity of 50, 100 and 5.8 intervals, respectively. Figure 8C and D shows the heatmaps of defect size, distance to surface and fatigue life from the raw data and the data after discrete value rejection. With such data processing, the relationship

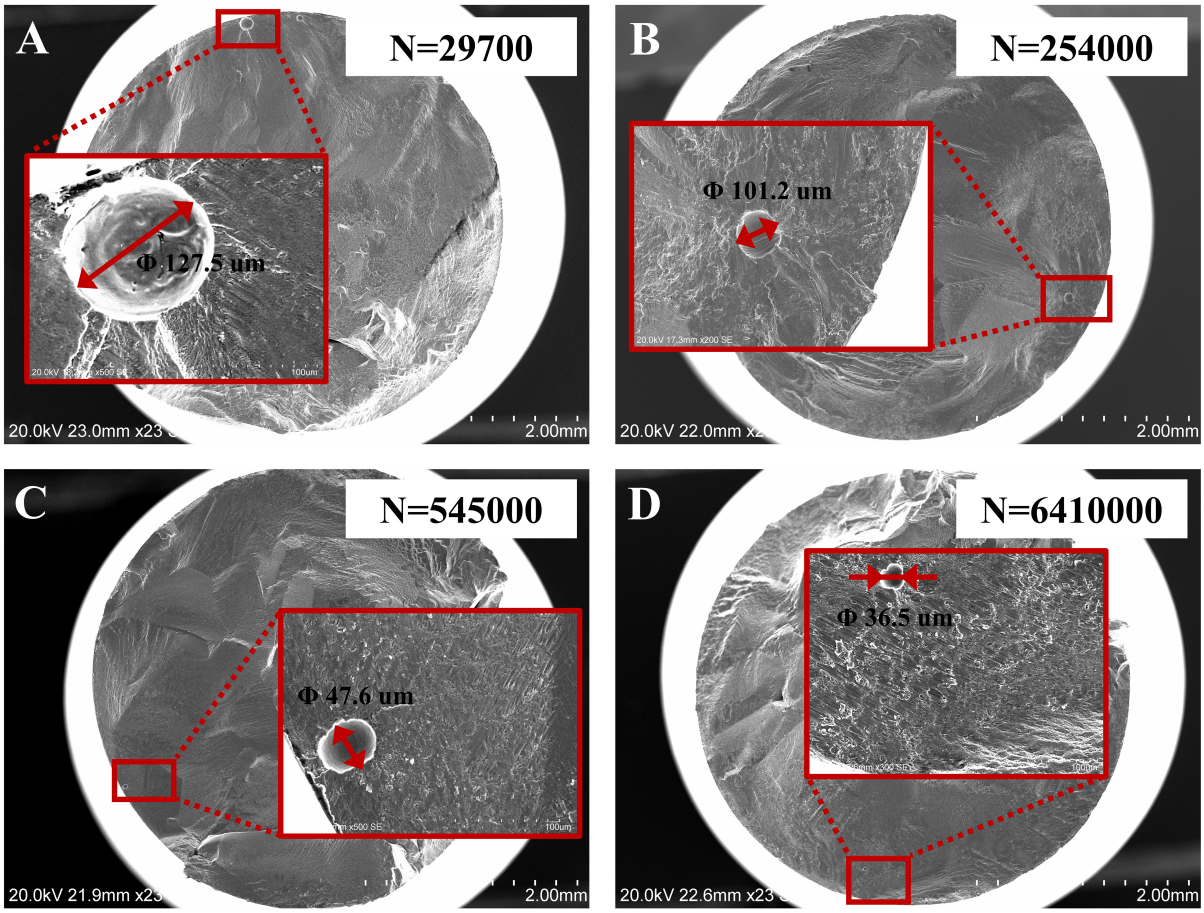

Figure 6. Fatigue fractographs with the size and location of the spherical pores after cycles of (A) 29,700, (B) 254,000, (C) 561,000 and (D) 6,410,000.

between the independent and dependent variables gradually improves and is more conducive to subsequent model training. After preprocessing, there are 30 groups of data.

Of course, operations such as feature selection, data cleaning, cross-validation and splitting and their specific effects on the model are described in turn in Section "Model performance and remarks", in Section "Data filtering" and in Section "Fatigue life model". It is important to note that these operations are formally different from data processing in regular machine learning models, and we have incorporated them into every step of the process in this work.

Data filtering

When Equation (10) was fitted to the initial datasets with S_0 = 100 and x_0 = 2,000 (the sample radius), the resulting R^2 was found to be 0.5527 with d = 6.7106, m = 0.8087, a = 185.1749, b = 3.3880 and c = 15.2355. The unacceptably small R^2 indicates the deviation of certain data points from the physical knowledge model, which leads to catastrophic results during the fitting and implies the necessity of data filtering. To avoid arbitrariness, we combine d, m, a, b and c parameters in the range of 0.1, so as to reasonably select the data with the largest deviation to all possible models. The parameter combination of d, m, a, b, c and corresponding R^2 for the 243 physical knowledge models are listed in Table 2.

Table 2. The parameter combination of *d*, *m*, *a*, *b*, *c* and corresponding R^2 for the 243 physical knowledge models

Model	*d*	*m*	*a*	*b*	*c*	R^2	Model	*d*	*m*	*a*	*b*	*c*	R^2
1	6.6106	0.7087	185.0749	3.2880	15.1355	0.5532	123	6.7106	0.8087	185.1749	3.3880	15.3355	0.5525
2	6.6106	0.7087	185.0749	3.2880	15.2355	0.5530	124	6.7106	0.8087	185.1749	3.4880	15.1355	0.5522
3	6.6106	0.7087	185.0749	3.2880	15.3355	0.5528	125	6.7106	0.8087	185.1749	3.4880	15.2355	0.5520
4	6.6106	0.7087	185.0749	3.3880	15.1355	0.5529	126	6.7106	0.8087	185.1749	3.4880	15.3355	0.5518
5	6.6106	0.7087	185.0749	3.3880	15.2355	0.5527	127	6.7106	0.8087	185.2749	3.2880	15.1355	0.5532
6	6.6106	0.7087	185.0749	3.3880	15.3355	0.5525	128	6.7106	0.8087	185.2749	3.2880	15.2355	0.5530
7	6.6106	0.7087	185.0749	3.4880	15.1355	0.5522	129	6.7106	0.8087	185.2749	3.2880	15.3355	0.5528
8	6.6106	0.7087	185.0749	3.4880	15.2355	0.5520	130	6.7106	0.8087	185.2749	3.3880	15.1355	0.5529
9	6.6106	0.7087	185.0749	3.4880	15.3355	0.5518	131	6.7106	0.8087	185.2749	3.3880	15.2355	0.5527
10	6.6106	0.7087	185.1749	3.2880	15.1355	0.5532	132	6.7106	0.8087	185.2749	3.3880	15.3355	0.5525
11	6.6106	0.7087	185.1749	3.2880	15.2355	0.5530	133	6.7106	0.8087	185.2749	3.4880	15.1355	0.5522
12	6.6106	0.7087	185.1749	3.2880	15.3355	0.5528	134	6.7106	0.8087	185.2749	3.4880	15.2355	0.5520
13	6.6106	0.7087	185.1749	3.3880	15.1355	0.5529	135	6.7106	0.8087	185.2749	3.4880	15.3355	0.5518
14	6.6106	0.7087	185.1749	3.3880	15.2355	0.5527	136	6.7106	0.9087	185.0749	3.2880	15.1355	0.5532
15	6.6106	0.7087	185.1749	3.3880	15.3355	0.5525	137	6.7106	0.9087	185.0749	3.2880	15.2355	0.5530
16	6.6106	0.7087	185.1749	3.4880	15.1355	0.5522	138	6.7106	0.9087	185.0749	3.2880	15.3355	0.5528
17	6.6106	0.7087	185.1749	3.4880	15.2355	0.5520	139	6.7106	0.9087	185.0749	3.3880	15.1355	0.5529
18	6.6106	0.7087	185.1749	3.4880	15.3355	0.5518	140	6.7106	0.9087	185.0749	3.3880	15.2355	0.5527
19	6.6106	0.7087	185.2749	3.2880	15.1355	0.5532	141	6.7106	0.9087	185.0749	3.3880	15.3355	0.5525
20	6.6106	0.7087	185.2749	3.2880	15.2355	0.5530	142	6.7106	0.9087	185.0749	3.4880	15.1355	0.5522
21	6.6106	0.7087	185.2749	3.2880	15.3355	0.5528	143	6.7106	0.9087	185.0749	3.4880	15.2355	0.5520
22	6.6106	0.7087	185.2749	3.3880	15.1355	0.5529	144	6.7106	0.9087	185.0749	3.4880	15.3355	0.5518
23	6.6106	0.7087	185.2749	3.3880	15.2355	0.5527	145	6.7106	0.9087	185.1749	3.2880	15.1355	0.5532
24	6.6106	0.7087	185.2749	3.3880	15.3355	0.5525	146	6.7106	0.9087	185.1749	3.2880	15.2355	0.5530
25	6.6106	0.7087	185.2749	3.4880	15.1355	0.5522	147	6.7106	0.9087	185.1749	3.2880	15.3355	0.5528
26	6.6106	0.7087	185.2749	3.4880	15.2355	0.5520	148	6.7106	0.9087	185.1749	3.3880	15.1355	0.5529
27	6.6106	0.7087	185.2749	3.4880	15.3355	0.5518	149	6.7106	0.9087	185.1749	3.3880	15.2355	0.5527
28	6.6106	0.8087	185.0749	3.2880	15.1355	0.5532	150	6.7106	0.9087	185.1749	3.3880	15.3355	0.5525
29	6.6106	0.8087	185.0749	3.2880	15.2355	0.5530	151	6.7106	0.9087	185.1749	3.4880	15.1355	0.5522
30	6.6106	0.8087	185.0749	3.2880	15.3355	0.5528	152	6.7106	0.9087	185.1749	3.4880	15.2355	0.5520
31	6.6106	0.8087	185.0749	3.3880	15.1355	0.5529	153	6.7106	0.9087	185.1749	3.4880	15.3355	0.5518
32	6.6106	0.8087	185.0749	3.3880	15.2355	0.5527	154	6.7106	0.9087	185.2749	3.2880	15.1355	0.5532

33	6.6106	0.8087	185.0749	3.3880	15.3355	0.5525	155	6.7106	0.9087	185.2749	3.2880	15.2355	0.5530
34	6.6106	0.8087	185.0749	3.4880	15.1355	0.5522	156	6.7106	0.9087	185.2749	3.2880	15.3355	0.5528
35	6.6106	0.8087	185.0749	3.4880	15.2355	0.5520	157	6.7106	0.9087	185.2749	3.3880	15.1355	0.5529
36	6.6106	0.8087	185.0749	3.4880	15.3355	0.5518	158	6.7106	0.9087	185.2749	3.3880	15.2355	0.5527
37	6.6106	0.8087	185.1749	3.2880	15.1355	0.5532	159	6.7106	0.9087	185.2749	3.3880	15.3355	0.5525
38	6.6106	0.8087	185.1749	3.2880	15.2355	0.5530	160	6.7106	0.9087	185.2749	3.4880	15.1355	0.5522
39	6.6106	0.8087	185.1749	3.2880	15.3355	0.5528	161	6.7106	0.9087	185.2749	3.4880	15.2355	0.5520
40	6.6106	0.8087	185.1749	3.3880	15.1355	0.5529	162	6.7106	0.9087	185.2749	3.4880	15.3355	0.5518
41	6.6106	0.8087	185.1749	3.3880	15.2355	0.5527	163	6.8106	0.7087	185.0749	3.2880	15.1355	0.5532
42	6.6106	0.8087	185.1749	3.3880	15.3355	0.5525	164	6.8106	0.7087	185.0749	3.2880	15.2355	0.5530
43	6.6106	0.8087	185.1749	3.4880	15.1355	0.5522	165	6.8106	0.7087	185.0749	3.2880	15.3355	0.5528
44	6.6106	0.8087	185.1749	3.4880	15.2355	0.5520	166	6.8106	0.7087	185.0749	3.3880	15.1355	0.5529
45	6.6106	0.8087	185.1749	3.4880	15.3355	0.5518	167	6.8106	0.7087	185.0749	3.3880	15.2355	0.5527
46	6.6106	0.8087	185.2749	3.2880	15.1355	0.5532	168	6.8106	0.7087	185.0749	3.3880	15.3355	0.5525
47	6.6106	0.8087	185.2749	3.2880	15.2355	0.5530	169	6.8106	0.7087	185.0749	3.4880	15.1355	0.5522
48	6.6106	0.8087	185.2749	3.2880	15.3355	0.5528	170	6.8106	0.7087	185.0749	3.4880	15.2355	0.5520
49	6.6106	0.8087	185.2749	3.3880	15.1355	0.5529	171	6.8106	0.7087	185.0749	3.4880	15.3355	0.5518
50	6.6106	0.8087	185.2749	3.3880	15.2355	0.5527	172	6.8106	0.7087	185.1749	3.2880	15.1355	0.5532
51	6.6106	0.8087	185.2749	3.3880	15.3355	0.5525	173	6.8106	0.7087	185.1749	3.2880	15.2355	0.5530
52	6.6106	0.8087	185.2749	3.4880	15.1355	0.5522	174	6.8106	0.7087	185.1749	3.2880	15.3355	0.5528
53	6.6106	0.8087	185.2749	3.4880	15.2355	0.5520	175	6.8106	0.7087	185.1749	3.3880	15.1355	0.5529
54	6.6106	0.8087	185.2749	3.4880	15.3355	0.5518	176	6.8106	0.7087	185.1749	3.3880	15.2355	0.5527
55	6.6106	0.9087	185.0749	3.2880	15.1355	0.5532	177	6.8106	0.7087	185.1749	3.3880	15.3355	0.5525
56	6.6106	0.9087	185.0749	3.2880	15.2355	0.5530	178	6.8106	0.7087	185.1749	3.4880	15.1355	0.5522
57	6.6106	0.9087	185.0749	3.2880	15.3355	0.5528	179	6.8106	0.7087	185.1749	3.4880	15.2355	0.5520
58	6.6106	0.9087	185.0749	3.3880	15.1355	0.5529	180	6.8106	0.7087	185.1749	3.4880	15.3355	0.5518
59	6.6106	0.9087	185.0749	3.3880	15.2355	0.5527	181	6.8106	0.7087	185.2749	3.2880	15.1355	0.5532
60	6.6106	0.9087	185.0749	3.3880	15.3355	0.5525	182	6.8106	0.7087	185.2749	3.2880	15.2355	0.5530
61	6.6106	0.9087	185.0749	3.4880	15.1355	0.5522	183	6.8106	0.7087	185.2749	3.2880	15.3355	0.5528
62	6.6106	0.9087	185.0749	3.4880	15.2355	0.5520	184	6.8106	0.7087	185.2749	3.3880	15.1355	0.5529
63	6.6106	0.9087	185.0749	3.4880	15.3355	0.5518	185	6.8106	0.7087	185.2749	3.3880	15.2355	0.5527
64	6.6106	0.9087	185.1749	3.2880	15.1355	0.5532	186	6.8106	0.7087	185.2749	3.3880	15.3355	0.5525
65	6.6106	0.9087	185.1749	3.2880	15.2355	0.5530	187	6.8106	0.7087	185.2749	3.4880	15.1355	0.5522
66	6.6106	0.9087	185.1749	3.2880	15.3355	0.5528	188	6.8106	0.7087	185.2749	3.4880	15.2355	0.5520

67	6.6106	0.9087	185.1749	3.3880	15.1355	0.5529	189	6.8106	0.7087	185.2749	3.4880	15.3355	0.5518
68	6.6106	0.9087	185.1749	3.3880	15.2355	0.5527	190	6.8106	0.8087	185.0749	3.2880	15.1355	0.5532
69	6.6106	0.9087	185.1749	3.3880	15.3355	0.5525	191	6.8106	0.8087	185.0749	3.2880	15.2355	0.5530
70	6.6106	0.9087	185.1749	3.4880	15.1355	0.5522	192	6.8106	0.8087	185.0749	3.2880	15.3355	0.5528
71	6.6106	0.9087	185.1749	3.4880	15.2355	0.5520	193	6.8106	0.8087	185.0749	3.3880	15.1355	0.5529
72	6.6106	0.9087	185.1749	3.4880	15.3355	0.5518	194	6.8106	0.8087	185.0749	3.3880	15.2355	0.5527
73	6.6106	0.9087	185.2749	3.2880	15.1355	0.5532	195	6.8106	0.8087	185.0749	3.3880	15.3355	0.5525
74	6.6106	0.9087	185.2749	3.2880	15.2355	0.5530	196	6.8106	0.8087	185.0749	3.4880	15.1355	0.5522
75	6.6106	0.9087	185.2749	3.2880	15.3355	0.5528	197	6.8106	0.8087	185.0749	3.4880	15.2355	0.5520
76	6.6106	0.9087	185.2749	3.3880	15.1355	0.5529	198	6.8106	0.8087	185.0749	3.4880	15.3355	0.5518
77	6.6106	0.9087	185.2749	3.3880	15.2355	0.5527	199	6.8106	0.8087	185.1749	3.2880	15.1355	0.5532
78	6.6106	0.9087	185.2749	3.3880	15.3355	0.5525	200	6.8106	0.8087	185.1749	3.2880	15.2355	0.5530
79	6.6106	0.9087	185.2749	3.4880	15.1355	0.5522	201	6.8106	0.8087	185.1749	3.2880	15.3355	0.5528
80	6.6106	0.9087	185.2749	3.4880	15.2355	0.5520	202	6.8106	0.8087	185.1749	3.3880	15.1355	0.5529
81	6.6106	0.9087	185.2749	3.4880	15.3355	0.5518	203	6.8106	0.8087	185.1749	3.3880	15.2355	0.5527
82	6.7106	0.7087	185.0749	3.2880	15.1355	0.5532	204	6.8106	0.8087	185.1749	3.3880	15.3355	0.5525
83	6.7106	0.7087	185.0749	3.2880	15.2355	0.5530	205	6.8106	0.8087	185.1749	3.4880	15.1355	0.5522
84	6.7106	0.7087	185.0749	3.2880	15.3355	0.5528	206	6.8106	0.8087	185.1749	3.4880	15.2355	0.5520
85	6.7106	0.7087	185.0749	3.3880	15.1355	0.5529	207	6.8106	0.8087	185.1749	3.4880	15.3355	0.5518
86	6.7106	0.7087	185.0749	3.3880	15.2355	0.5527	208	6.8106	0.8087	185.2749	3.2880	15.1355	0.5532
87	6.7106	0.7087	185.0749	3.3880	15.3355	0.5525	209	6.8106	0.8087	185.2749	3.2880	15.2355	0.5530
88	6.7106	0.7087	185.0749	3.4880	15.1355	0.5522	210	6.8106	0.8087	185.2749	3.2880	15.3355	0.5528
89	6.7106	0.7087	185.0749	3.4880	15.2355	0.5520	211	6.8106	0.8087	185.2749	3.3880	15.1355	0.5529
90	6.7106	0.7087	185.0749	3.4880	15.3355	0.5518	212	6.8106	0.8087	185.2749	3.3880	15.2355	0.5527
91	6.7106	0.7087	185.1749	3.2880	15.1355	0.5532	213	6.8106	0.8087	185.2749	3.3880	15.3355	0.5525
92	6.7106	0.7087	185.1749	3.2880	15.2355	0.5530	214	6.8106	0.8087	185.2749	3.4880	15.1355	0.5522
93	6.7106	0.7087	185.1749	3.2880	15.3355	0.5528	215	6.8106	0.8087	185.2749	3.4880	15.2355	0.5520
94	6.7106	0.7087	185.1749	3.3880	15.1355	0.5529	216	6.8106	0.8087	185.2749	3.4880	15.3355	0.5518
95	6.7106	0.7087	185.1749	3.3880	15.2355	0.5527	217	6.8106	0.9087	185.0749	3.2880	15.1355	0.5532
96	6.7106	0.7087	185.1749	3.3880	15.3355	0.5525	218	6.8106	0.9087	185.0749	3.2880	15.2355	0.5530
97	6.7106	0.7087	185.1749	3.4880	15.1355	0.5522	219	6.8106	0.9087	185.0749	3.2880	15.3355	0.5528
98	6.7106	0.7087	185.1749	3.4880	15.2355	0.5520	220	6.8106	0.9087	185.0749	3.3880	15.1355	0.5529
99	6.7106	0.7087	185.1749	3.4880	15.3355	0.5518	221	6.8106	0.9087	185.0749	3.3880	15.2355	0.5527
100	6.7106	0.7087	185.2749	3.2880	15.1355	0.5532	222	6.8106	0.9087	185.0749	3.3880	15.3355	0.5525

101	6.7106	0.7087	185.2749	3.2880	15.2355	0.5530	223	6.8106	0.9087	185.0749	3.4880	15.1355	0.5522
102	6.7106	0.7087	185.2749	3.2880	15.3355	0.5528	224	6.8106	0.9087	185.0749	3.4880	15.2355	0.5520
103	6.7106	0.7087	185.2749	3.3880	15.1355	0.5529	225	6.8106	0.9087	185.0749	3.4880	15.3355	0.5518
104	6.7106	0.7087	185.2749	3.3880	15.2355	0.5527	226	6.8106	0.9087	185.1749	3.2880	15.1355	0.5532
105	6.7106	0.7087	185.2749	3.3880	15.3355	0.5525	227	6.8106	0.9087	185.1749	3.2880	15.2355	0.5530
106	6.7106	0.7087	185.2749	3.4880	15.1355	0.5522	228	6.8106	0.9087	185.1749	3.2880	15.3355	0.5528
107	6.7106	0.7087	185.2749	3.4880	15.2355	0.5520	229	6.8106	0.9087	185.1749	3.3880	15.1355	0.5529
108	6.7106	0.7087	185.2749	3.4880	15.3355	0.5518	230	6.8106	0.9087	185.1749	3.3880	15.2355	0.5527
109	6.7106	0.8087	185.0749	3.2880	15.1355	0.5532	231	6.8106	0.9087	185.1749	3.3880	15.3355	0.5525
110	6.7106	0.8087	185.0749	3.2880	15.2355	0.5530	232	6.8106	0.9087	185.1749	3.4880	15.1355	0.5522
111	6.7106	0.8087	185.0749	3.2880	15.3355	0.5528	233	6.8106	0.9087	185.1749	3.4880	15.2355	0.5520
112	6.7106	0.8087	185.0749	3.3880	15.1355	0.5529	234	6.8106	0.9087	185.1749	3.4880	15.3355	0.5518
113	6.7106	0.8087	185.0749	3.3880	15.2355	0.5527	235	6.8106	0.9087	185.2749	3.2880	15.1355	0.5532
114	6.7106	0.8087	185.0749	3.3880	15.3355	0.5525	236	6.8106	0.9087	185.2749	3.2880	15.2355	0.5530
115	6.7106	0.8087	185.0749	3.4880	15.1355	0.5522	237	6.8106	0.9087	185.2749	3.2880	15.3355	0.5528
116	6.7106	0.8087	185.0749	3.4880	15.2355	0.5520	238	6.8106	0.9087	185.2749	3.3880	15.1355	0.5529
117	6.7106	0.8087	185.0749	3.4880	15.3355	0.5518	239	6.8106	0.9087	185.2749	3.3880	15.2355	0.5527
118	6.7106	0.8087	185.1749	3.2880	15.1355	0.5532	240	6.8106	0.9087	185.2749	3.3880	15.3355	0.5525
119	6.7106	0.8087	185.1749	3.2880	15.2355	0.5530	241	6.8106	0.9087	185.2749	3.4880	15.1355	0.5522
120	6.7106	0.8087	185.1749	3.2880	15.3355	0.5528	242	6.8106	0.9087	185.2749	3.4880	15.2355	0.5520
121	6.7106	0.8087	185.1749	3.3880	15.1355	0.5529	243	6.8106	0.9087	185.2749	3.4880	15.3355	0.5518
122	6.7106	0.8087	185.1749	3.3880	15.2355	0.5527							

R^2: The coefficient of determination.

For example, the parameter combination is performed with the step interval of 0.1. With d = 6.6106/6.7106/6.8106, m = 0.7087/0.8087/0.9087, a = 185.0749/185.1749/185.2749, b = 3.2880/3.3880/3.4880 and c = 15.1355/15.2355/15.3355, a total of 243 combinations of d, m, a, b, c and corresponding R^2 are obtained.

Generally, R^2 of 0.5-0.6 is reached after the 243 physical knowledge models [Table 2], but still far from satisfactory. Figure 9 shows the experimental data and one of the fitted surfaces with d = 6.7106, m = 0.8087, a = 185.1749, b = 3.3880 and c = 15.2355 with R^2 = 0.5527. Remarkable deviation from the fitted models exists for a number of data points. We thus gathered the mean absolute percentage deviation (MAPE) of the 30 groups of experimental data to all 243 physics knowledge models. The MAPE is defined according to

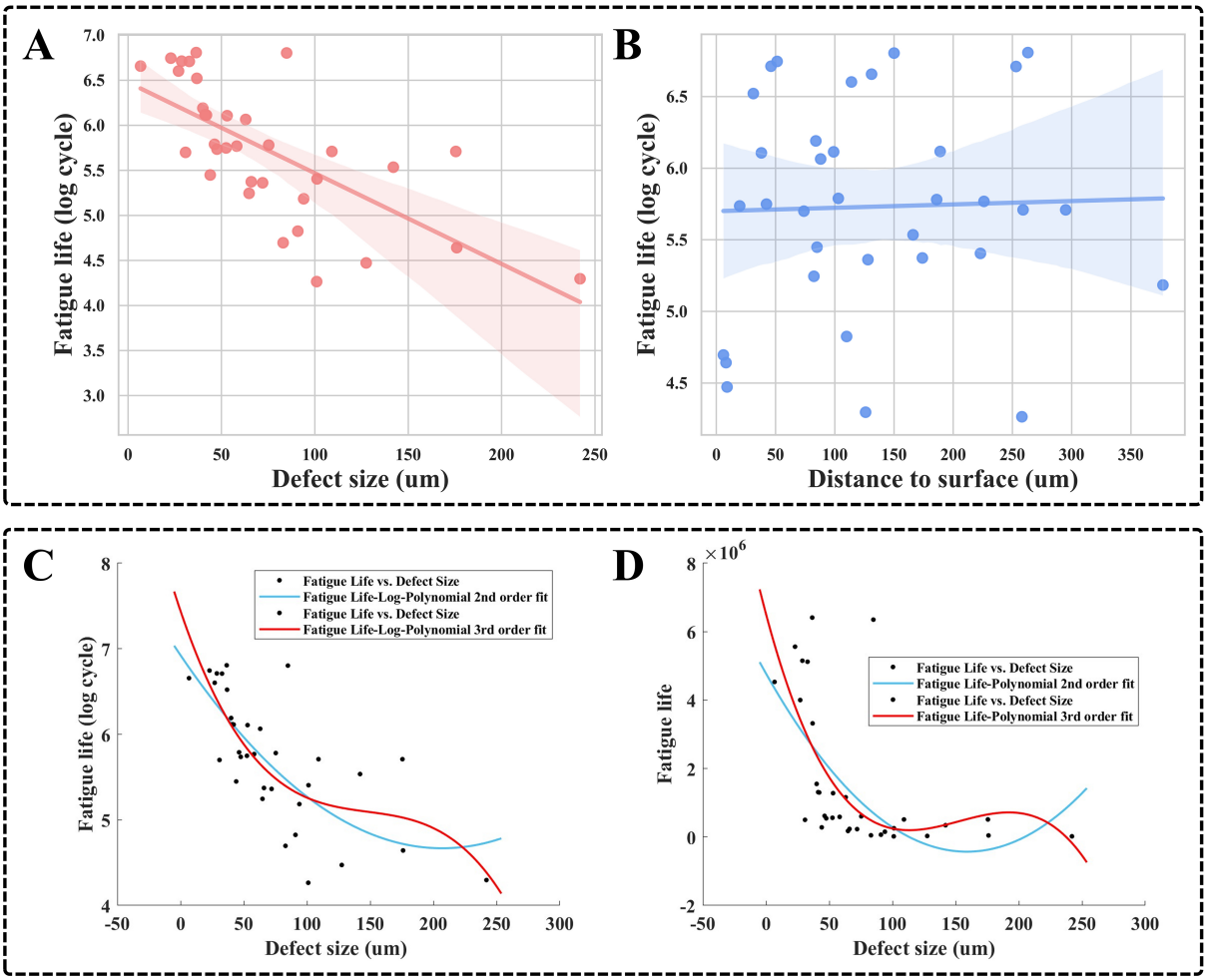

Figure 7. (A and B) Relationship between fatigue life with and without log treatment and defect parameters. The experimental data with the fatigue life of the bending samples against (C) the size and (D) location of the spherical pores.

$$\mathrm{MAPE} = 100 \times \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \tag{11}$$

where n is the total number of data, and y_i and $\hat{y}_i$ the actual and fitted values of the ith data, respectively. The MAPE for each group of data is shown in Figure 10. A general criterion of 10% is set to filter the data and groups 11, 24, 25, 30 are thus filtered. However, despite the relatively large deviation of data group 2 (10.5%), it is retained due to relatively small deviation on certain models. After filtering, there are 26 groups of data.

Data augmentation

Subsequently, a GAN is built. The GAN learns to generate values that match patterns in the defect size, the distance to surface, and the fatigue life. The generator, a neural network with three output values, is trained to provide data that matches the defect size, the distance to surface, and the fatigue life pattern. On the other hand, the discriminator tries to determine whether it is from the real data source or from the generator, based on these three values. The steps for constructing and training the GAN in this paper are as follows:

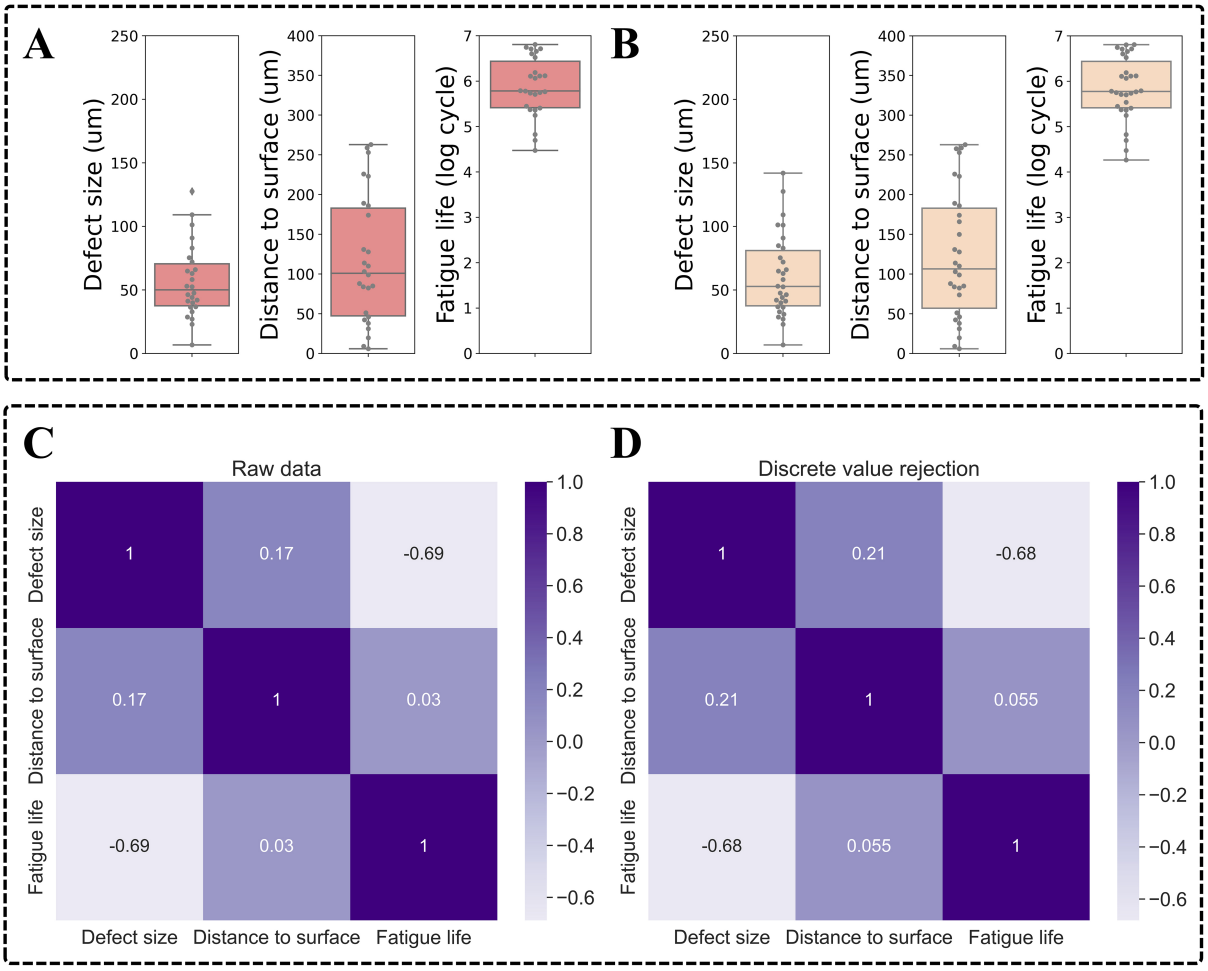

Figure 8. Box line plots of the independent and dependent variables. (A) Raw data (B) Data after discrete value rejection. Heatmaps of defect size, distance to surface and fatigue life. (C) Raw data (D) Data after discrete value rejection.

(1) Previewing data from a real dataset; (2) Determining whether the discriminator can at least distinguish between actual data and random noise; (3) Determining whether an untrained generator can produce data in the required format; (4) Displaying the observed loss values to understand the training progress; (5) Producing fresh data; (6) Repeat the data filtering process.

The loss values of the GAN, the discriminator and the generator are shown in Figure 11A-C, respectively. When the discriminator is ineffective at telling the authentic data from the fake data, it cannot decide whether to output 0 or 1, and instead outputs 0.5, and the loss value is $0.5^2 = 0.25$. The GAN advances gradually as the training goes on since the loss value reduces gradually and perceivably. The loss value increases again to 0.25 later in the training. This indicates that the generator has mastered the art of producing data in a predictable fashion, rendering the discriminator ineffective. In other words, the discriminator's output is 0.5 and the loss value jumps to 0.25. Note that due to the initially limited number of experiment data with large sparsity, GAN was able to produce only two qualified groups of data. After augmentation, there are 28 groups of data.

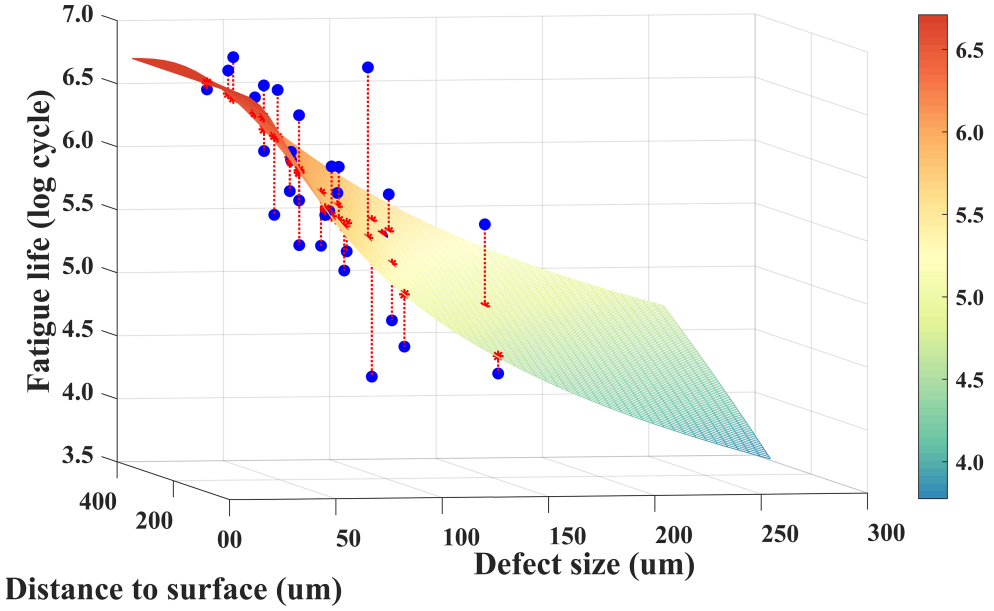

Figure 9. The fatigue life against the defect size and distance to surface for the experimental data (red dots) and one of the fitted surfaces with d = 6.7106, m = 0.8087, a = 185.1749, b = 3.3880 and c = 15.2355 with R^2 = 0.5527. R^2: The coefficient of determination.

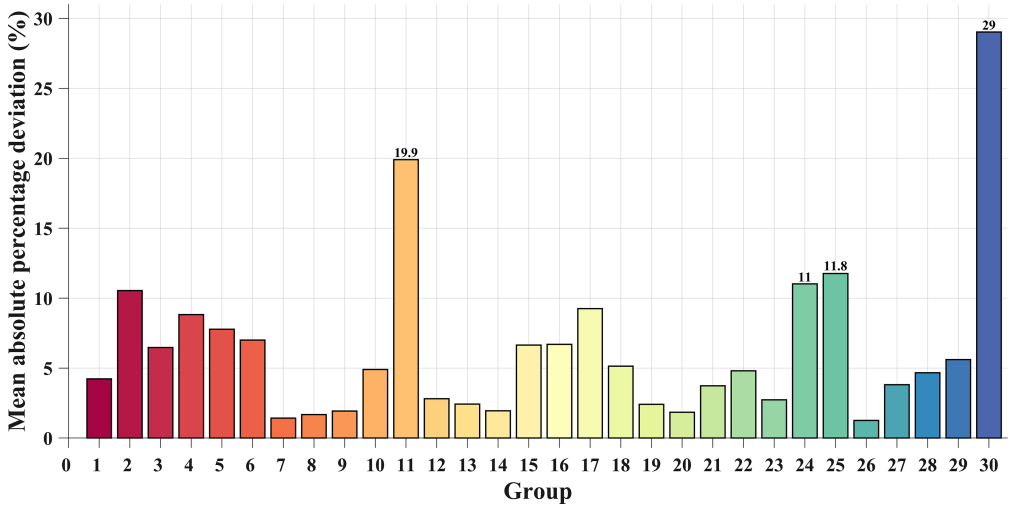

Figure 10. The overall deviation of the 30 groups of experimental data to all 243 physics knowledge models.

Fatigue life model

Finally, after data filtering and augmentation [Figure 12], the prediction surface becomes increasingly smooth. A total of 28 groups of data were employed to predict the fatigue life with only the back propagation neural network, as shown in Figure 13A and B, with a confidence interval of nearly 90% in Figure 13C. The predicted surface is close to the physical knowledge model (e.g., Figure 9) with the fatigue life decreasing with the defect size and increasing with the distance to surface. However, without physical knowledge, the resulting model is overfitted without interpretability.

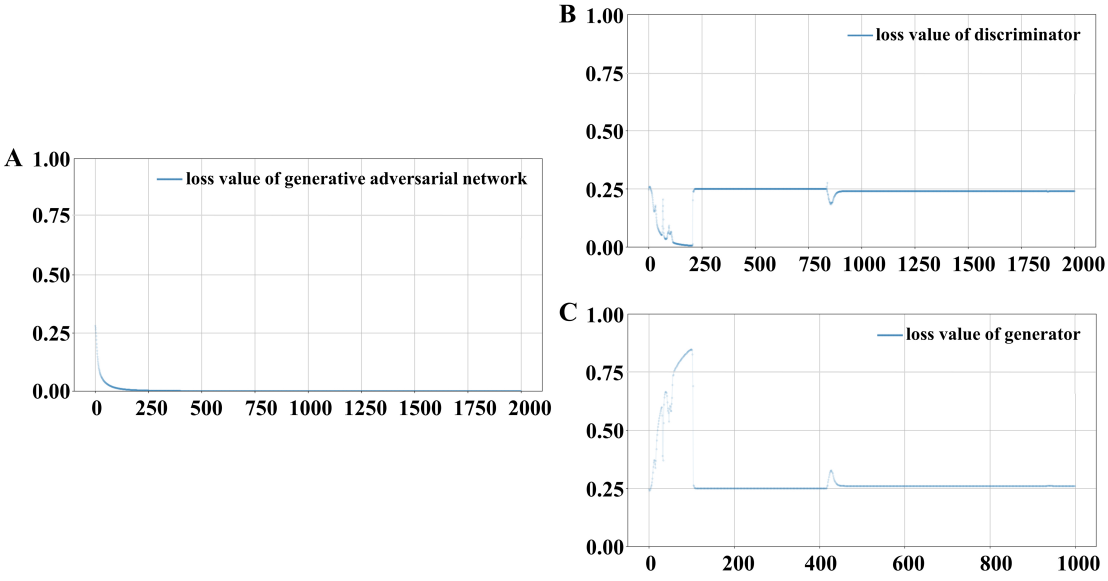

Figure 11. Loss value. (A) GAN; (B) Discriminator; (C) Generator. GAN: Generative adversarial network.

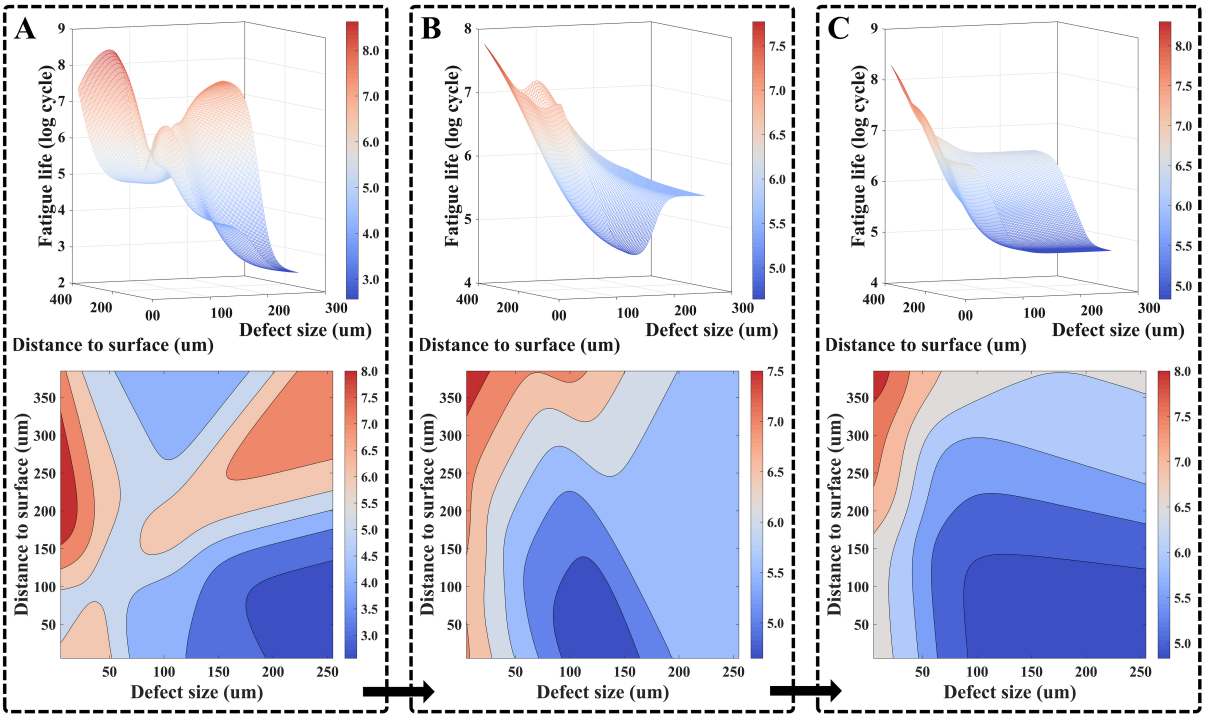

Figure 12. In the process of model selection, the progress of the model is reflected in an increasingly smooth prediction surface.

In order to achieve a precise model with physical knowledge, the 28 groups of data are again employed together with the physical knowledge models in Section "Physical knowledge model". The lsqcurvefit function in MATLAB is utilized and the optimal constants are obtained as: d = 13.1254, m = 8.5685, a = 7.6417, b = 0.2401, c = 1.3565, S_0 = 51.0409 and x_0 = 564.6523 [Equation (12)]. The final fatigue life model after Equation (10) is as follows, which is visualized in Figure 13D-F.

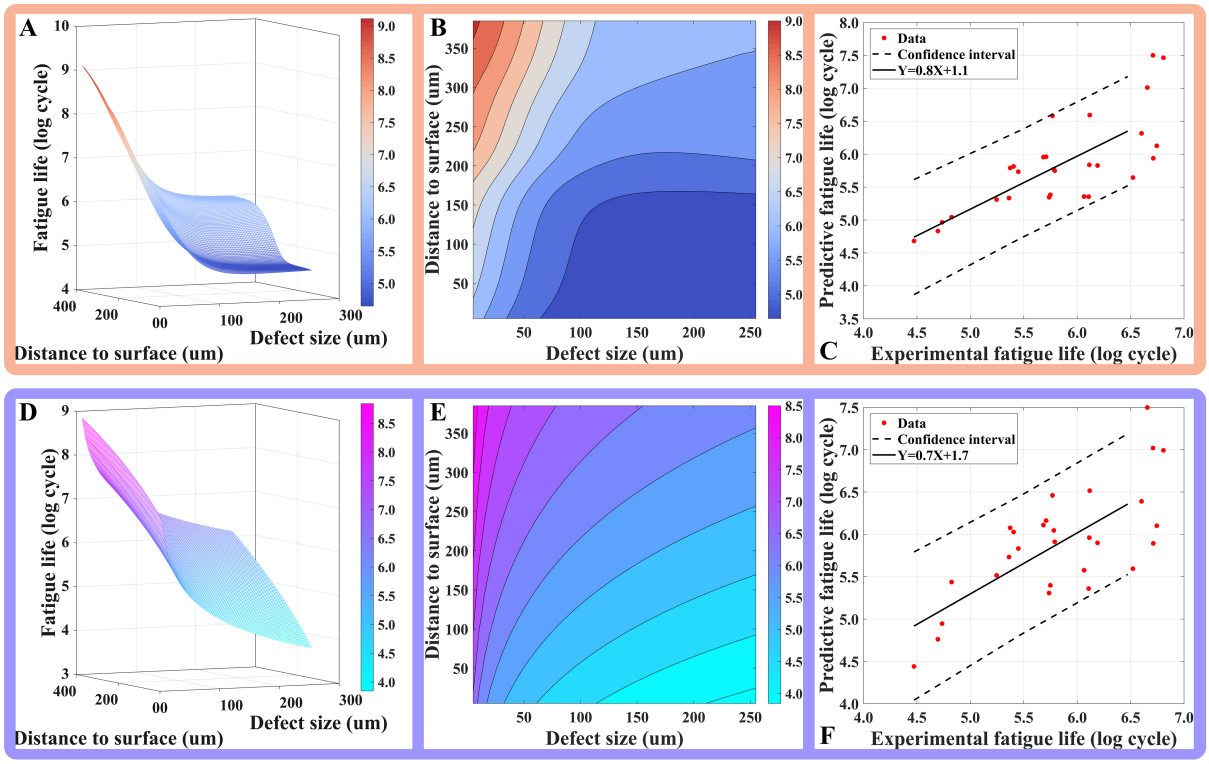

Figure 13. Third prediction of experimental data after augmentation of GAN model using back propagation neural network. (A) 3D surface; (B) Contour map; (C) 2D confidence interval. The final physical model optimized with the lsqcurvefit function in MATLAB. (D) 3D surface; (E) Contour map; (F) 2D confidence interval. GAN: Generative adversarial network.

$$\left[1+\frac{a*\left(\frac{S}{S_0}\right)^b}{\left(1+\frac{x}{x_0}\right)^c}\right]\times N_f = 10^d$$
$$\left[1+\frac{7.6417*\left(\frac{S}{51.0409}\right)^{0.2401}}{\left(1+\frac{x}{564.6523}\right)^{1.3565}}\right]^{8.5685}\times N_f = 10^{13.1254} \quad (12)$$

The present model thus indicates the influence of the two main factors, i.e., pore radius and location, on the fatigue life. At a fixed pore location, the fatigue life decreases with pore radii, while at a fix pore radius, the fatigue life decreases with pore distance to surface. In particular, pores located near the surface will be detrimental to fatigue life even if they are very small, while pores located at the central axis of the sample will be remarkably safer. More subtle relationship between S and x can thus be evaluated under the circumstance of rotating-bending fatigue. With $N_f \geq 10^7$ as a criterion, the following expression holds,

$$S \leq 4.1619 \times \left(1+\frac{x}{564.6523}\right)^{5.6497} \text{ and } x \geq 438.7438 \times S^{0.1770} - 564.6523 \quad (13)$$

which thus indicates the mutual relationship between x and S; i.e., to have a satisfactory fatigue life, the pores should be smaller as their location approaches the surface. Specifically, for near-surface pores with $x \to 0$, only very small pores with S ~ 4 μm are allowed without significant reducing fatigue life, while for typical pores under current additive manufacturing techniques with S ~ 1,000 μm, they can be located at least ~ 1,000 μm from the surface.

Model performance and remarks

We have also compared the methods used and the results obtained with the existing work. Since machine

learning is a data-driven algorithm, the data itself has a significant effect on the model performance. Therefore, researchers have used different methods to prepare data before training a model. Li *et al.* developed a database for machine learning models using the Monte Carlo method[56]. Zhan *et al.*[57,58] proposed a continuous damage mechanics[59,60] model that also incorporates the additive manufacturing effect with theoretical calculations for data augmentation. It was estimated that the R^2 for the training, validation and test sets are 0.9854, 0.9857 and 0.9854, respectively, with the overall R^2 = 0.9855. Nevertheless, it is unclear whether the high dispersion of fatigue data leads to a localized distribution of data, and whether the so-called theoretical models consider all the factors that influence fatigue data. In this paper, we propose a completely new approach for processing data combining GAN and physical knowledge models. The R^2 values for the training, validation and test sets are 0.8338, 0.9975 and 0.8912, respectively, while the total R^2 = 0.8701. A well-developed theoretical system and novel ideas underpin the GAN in deep learning. To overcome the limitations of localization of traditional methods, a continuous game of generative and discriminative networks can be used to learn the characteristics of data distribution.

The most essential activity in the model, i.e., feature engineering, influences the model performance. To estimate the fatigue life, Bao *et al.*[13] have used ten sets of experimental data with three independent variables of defect size, defect location and defect type appearance, while Li *et al.*[56] have used 22 sets of experimental data with five independent variables, i.e., flaw depth, defect size, defect-to-surface distance, maximum stress and construction direction. It can be shown that the geometric parameters of the critical flaws have a considerable influence on the fatigue life. As a result, in this paper, the fatigue life is predicted using two independent variables: defect size and distance to surface. We also provide a physical knowledge method for data processing based on these two variables. Lastly, it is worth mentioning that the present work has the largest data to variable ratio, i.e., 34 to 2, among the published work.

It is worth mentioning that, first, the present model only considers spherical pores, while in the hybrid manufacture samples, pores with other shapes also exist; second, the present model has not explicitly considered microstructure features, while pores located at grain/phase boundaries may have detrimental influence on fatigue life; third, the present model omits other metallurgical defects, e.g., lack of fusion, microcrack, *etc.*, which is supposed to distinguish with appropriate additive manufacturing process and subsequent post-processing. While the above issues are covered in ongoing investigations, the results will be reported in forthcoming publications.

CONCLUSION

A thorough investigation was conducted on the prediction of fatigue life against defect parameters for forging/additive hybrid manufactured Ti-17 alloys, with the combination of machine learning and physical knowledge to deal with sparse and limited data. The following conclusions are drawn.

(1) Strong and reliable physical supervision enhances the power of machine learning to deal with sparse and limited datasets. Specifically for fatigue life prediction, a machine learning framework with a back propagation neural network and GAN can thus achieve a relatively high level of 90% confidence.

(2) Pore-induced stress concentration influences the fatigue life of hybrid manufactured Ti-17 alloys. The resulting physical model indicates that the rotating/bending fatigue life decreases with pore size and increases with its distance to surface.

(3) Specifically for rotating/bending fatigue tests, near-surface pores with size ~ 4 μm or in-depth pores with size ~ 1,000 μm have negligible influence on fatigue life with respect to pore-free samples.

The present model thus aids in evaluating rotating/bending fatigue life of forging/additive hybrid manufactured Ti-17 alloys, and contributes to the general understanding of metal fatigue regarding the influence of pores. In the future, when enough data, including other pore characteristics, other types of defects and defect-microstructure coupling, is available, a comprehensive model for fatigue life prediction can be expected with the help of machine learning.

DECLARATIONS

Authors' contributions

Conceptualization: Wang H, Liu J, Yang R
Methodology: Gao S, Wang H, Liu J
Formal analysis: Yang Y, Li S, Qu S, Chen Z, Wang, H, Liu J
Investigation: Gao S, Li W, Ma Y, Wang B
Data curation: Gao S, Li W, Dong X
Draft preparation: Gao S, Li W, Ma Y, Wang B, Dong X
Review and editing: Yang Y, Li S, Qu S, Chen Z, Wang H, Liu J, Yang R
Validation: Wang H, Liu J, Yang R
Project administration: Wang H, Yang R
All authors have read and agreed to the published version of the manuscript.

Availability of data and materials

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Financial support and sponsorship

The authors acknowledge the financial support of National Natural Science Foundation of China (U2241245 and 91960202), National Science and Technology Major Project (2019-VII-0004-0144), National Key Laboratory Foundation of Science and Technology on Materials under Shock and Impact (6142902220301), the Aeronautical Science Foundation of China (2022Z053092001), Liaoning Science and Technology Major Project, Shanghai Engineering Research Center of High-Performance Medical Device Materials (20DZ2255500) and Natural Science Foundation of Shenyang (23-503-6-05).

Conflicts of interest

Qu S and Chen Z are affiliated with AECC Shenyang Liming Areo-engine Co., Ltd, while the other authors have declared that they have no conflicts of interest.

Ethical approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

REFERENCES

1. Zhang D, Qiu D, Gibson MA, et al. Additive manufacturing of ultrafine-grained high-strength titanium alloys. *Nature* 2019;576:91-5. DOI PubMed
2. Korkmaz ME, Waqar S, Garcia-collado A, Gupta MK, Krolczyk GM. A technical overview of metallic parts in hybrid additive manufacturing industry. *J Mater Res Technol* 2022;18:384-95. DOI
3. Hirtler M, Jedynak A, Sydow B, Sviridov A, Bambach M. A study on the mechanical properties of hybrid parts manufactured by forging and wire arc additive manufacturing. *Procedia Manuf* 2020;47:1141-8. DOI
4. Murakami Y, Endo M. Effects of defects, inclusions and inhomogeneities on fatigue strength. *Int J Fatigue* 1994;16:163-82. DOI
5. Murakami Y. Effects of small defects and nonmetallic inclusions on the fatigue strength of metals. *JSME Int J Ser* 1989;32:167-80. DOI
6. Piascik RS, Gangloff RP. Environmental fatigue of an Al-Li-Cu alloy: part I. Intrinsic crack propagation kinetics in hydrogenous environments. *Metall Trans A* 1991;22:2415-28. DOI
7. Piascik RS, Gangloff RP. Environmental fatigue of an Al-Li-Cu alloy: Part II. Microscopic hydrogen cracking processes. *Metall Trans A* 1993;24:2751-62. DOI
8. Davidson CJ, Griffiths JR, Machin AS. The effect of solution heat-treatment time on the fatigue properties of an Al-Si-Mg casting alloy. *Fatigue Fract Eng Mat Struct* 2002;25:223-30. DOI
9. Zhu M, Xuan F, Chen J. Influence of microstructure and microdefects on long-term fatigue behavior of a Cr–Mo–V steel. *Mat Sci Eng A* 2012;546:90-6. DOI
10. Zhu M, Xuan F, Du Y, Tu S. Very high cycle fatigue behavior of a low strength welded joint at moderate temperature. *Int J Fatigue* 2012;40:74-83. DOI
11. Zhu M, Jin L, Xuan F. Fatigue life and mechanistic modeling of interior micro-defect induced cracking in high cycle and very high cycle regimes. *Acta Mater* 2018;157:259-75. DOI
12. Jordan MI, Mitchell TM. Machine learning: trends, perspectives, and prospects. *Science* 2015;349:255-60. DOI PubMed
13. Bao H, Wu S, Wu Z, Kang G, Peng X, Withers PJ. A machine-learning fatigue life prediction approach of additively manufactured metals. *Eng Fract Mech* 2021;242:107508. DOI
14. Zhan Z, Hu W, Meng Q. Data-driven fatigue life prediction in additive manufactured titanium alloy: a damage mechanics based machine learning framework. *Eng Fract Mech* 2021;252:107850. DOI
15. Ye S, Huang W, Li M, Feng X. Deep learning method for determining the surface elastic moduli of microstructured solids. *Extreme Mech Lett* 2021;44:101226. DOI
16. Luo C, Zhu S, Keshtegar B, Macek W, Branco R, Meng D. Active Kriging-based conjugate first-order reliability method for highly efficient structural reliability analysis using resample strategy. *Comput Method Appl M* 2024;423:116863. DOI
17. Wu S, Liu J, Wang Y, Lu J, Zhang Z. Multiaxial low-cycle fatigue life model for notched specimens considering small sample characteristics. *IJSI* 2024;15:948-73. DOI
18. Yang S, Meng D, Wang H, Yang C. A novel learning function for adaptive surrogate-model-based reliability evaluation. *Philos Trans A Math Phys Eng Sci* 2024;382:20220395. DOI
19. Tarsha-Kurdi F, Landes T, Grussenmeyer P, Koehl M. Model-driven and da-ta-driven approaches using LIDAR data: analysis and comparison. 2007. Available from: https://www.researchgate.net/publication/32227577_Model-driven_and_data-driven_approaches_using_LIDAR_data_Analysis_and_comparison. [Last accessed on 29 Nov 2024].
20. Wang J, Chang Q, Xiao G, Wang N, Li S. Data driven production modeling and simulation of complex automobile general assembly plant. *Comput Ind* 2011;62:765-75. DOI
21. Pozdnoukhov A, Foresti L, Kanevski M. Data-driven topo-climatic mapping with machine learning methods. *Nat Hazards* 2009;50:497-518. DOI
22. Srinivasan V. Low cycle fatigue and creep–fatigue interaction behavior of 316L(N) stainless steel and life prediction by artificial neural network approach. *Int J Fatigue* 2003;25:1327-38. DOI
23. Nasiri S, Khosravani MR, Weinberg K. Fracture mechanics and mechanical fault detection by artificial intelligence methods: a review. *Eng Fail Anal* 2017;81:270-93. DOI
24. Barbosa JF, Correia JA, Júnior RF, Jesus AMD. Fatigue life prediction of metallic materials considering mean stress effects by means of an artificial neural network. *Int J Fatigue* 2020;135:105527. DOI
25. Romano S, Brandão A, Gumpinger J, Gschweitl M, Beretta S. Qualification of AM parts: extreme value statistics applied to tomographic measurements. *Mater Design* 2017;131:32-48. DOI
26. Plessis A, Yadroitsava I, Yadroitsev I. Effects of defects on mechanical properties in metal additive manufacturing: a review focusing on X-ray tomography insights. *Mater Design* 2020;187:108385. DOI
27. Bhadeshia HKDH. Neural networks in materials science. *ISIJ Int* 1999;39:966-79. DOI
28. Zhou Y, Zhang Z, Wang D, et al. New trends in additive manufacturing of high-entropy alloys and alloy design by machine learning: from single-phase to multiphase systems. *J Mater Inf* 2022;2:18. DOI
29. Wang W, Yin J, Chai Z, et al. Big data-assisted digital twins for the smart design and manufacturing of advanced materials: from atoms to products. *J Mater Inf* 2022;2:1. DOI
30. Xu P, Ji X, Li M, Lu W. Virtual sample generation in machine learning assisted materials design and discovery. *J Mater Inf* 2023;3:16. DOI

31. Zhou Z, Shang Y, Yang Y. A critical review of the machine learning guided design of metallic glasses for superior glass-forming ability. *J Mater Inf* 2022;2:2. DOI
32. Hu Y, Wu S, Withers P, et al. The effect of manufacturing defects on the fatigue life of selective laser melted Ti-6Al-4V structures. *Mater Design* 2020;192:108708. DOI
33. Zhang M, Sun C, Zhang X, et al. High cycle fatigue life prediction of laser additive manufactured stainless steel: a machine learning approach. *Int J Fatigue* 2019;128:105194. DOI
34. Serrano-Munoz I, Buffiere JY, Mokso R, Verdu C, Nadot Y. Location, location &size: defects close to surfaces dominate fatigue crack initiation. *Sci Rep* 2017;7:45239. DOI PubMed PMC
35. Benedetti M, Fontanari V, Bandini M, Zanini F, Carmignato S. Low- and high-cycle fatigue resistance of Ti-6Al-4V ELI additively manufactured via selective laser melting: Mean stress and defect sensitivity. *Int J Fatigue* 2018;107:96-109. DOI
36. Ren S, Chen Y, Liu T, Qu X. Effect of build orientation on mechanical properties and microstructure of Ti-6Al-4V manufactured by selective laser melting. *Metall Mater Trans A* 2019;50:4388-409. DOI
37. Barba D, Alabort C, Tang Y, Viscasillas M, Reed R, Alabort E. On the size and orientation effect in additive manufactured Ti-6Al-4V. *Mater Design* 2020;186:108235. DOI
38. Liu J, Zhang K, Liu J, et al. Investigation of fatigue behavior of laser powder bed fusion Ti-6Al-4V: roles of heat treatment and microstructure. *Int J Fatigue* 2023;176:107839. DOI
39. Sharma S, Sharma S, Athaiya A. Activation functions in neural networks. *Int J Eng Appl Sci Technol* 2020;4:310-6. Available from: https://www.ijeast.com/papers/310. [Last accessed on 29 Nov 2024]
40. Apicella A, Donnarumma F, Isgrò F, Prevete R. A survey on modern trainable activation functions. *Neural Netw* 2021;138:14-32. DOI PubMed
41. Rumelhart DE, Hinton GE, Williams RJ. Learning representations by back-propagating errors. *Nature* 1986;323:533-6. DOI
42. Dieter GE. Mechanical metallurgy. 3rd edition. McGraw Hill, New York, 1986. Available from: https://books.google.com/books?id=4pZRAAAAMAAJ&hl=zh-CN&source=gbs_book_other_versions. [Last accessed on 29 Nov 2024].
43. Murakami Y. Chapter 6 - Effects of nonmetallic inclusions on fatigue strength. In: Metal Fatigue. Elsevier; 2002. pp. 75-127. DOI
44. Stromeyer CE. The determination of fatigue limits under alternating stress conditions. *Proc R Soc Lond A* 1914;90:411-25. DOI
45. Paris P, Erdogan F. A critical analysis of crack propagation laws. *J Basic Eng* 1963;85:528-33. DOI
46. Yukitaka M, Masahiro E. Quantitative evaluation of fatigue strength of metals containing various small defects or cracks. *Eng Fract Mech* 1983;17:1-15. DOI
47. Deng K, Wei H, Liu W, Zhang M, Zhao P, Zhang Y. Probabilistic-based random maximum defect estimation and defect-related fatigue life prediction for laser direct deposited 316L parts. *J Mater Process Tech* 2022;299:117389. DOI
48. Liu F, He C, Chen Y, Zhang H, Wang Q, Liu Y. Effects of defects on tensile and fatigue behaviors of selective laser melted titanium alloy in very high cycle regime. *Int J Fatigue* 2020;140:105795. DOI
49. Tang M, Pistorius PC. Fatigue life prediction for AlSi10Mg components produced by selective laser melting. *Int J Fatigue* 2019;125:479-90. DOI
50. Zhang M, Sun C, Zhang X, Wei J, Hardacre D, Li H. Predictive models for fatigue property of laser powder bed fusion stainless steel 316L. *Mater Design* 2018;145:42-54. DOI
51. Liu A, Xu Z, Liang Z, Wang X. An evaluation on high cycle fatigue fracture characteristics of 2024-T351 Al alloy with different surface defects. *Mech Mater* 2022;164:104133. DOI
52. Chang D, Wang Q, Chen H, Zhang X, Wen J, Tu S. Effects of defect size and location on high cycle fatigue life of a maraging stainless steel at ambient and cryogenic temperatures. *Int J Fatigue* 2022;161:106906. DOI
53. Tiryakioğlu M. Relationship between defect size and fatigue life distributions in Al-7 Pct Si-Mg alloy castings. *Metall Mater Trans A* 2009;40:1623-30. DOI
54. Queiroz H, Araújo J, Silva C, Ferreira J. A coupled critical plane- area methodology to estimate fatigue life for an AISI 1045 steel with small artificial defects. *Theor Appl Fract Mec* 2022;121:103426. DOI
55. Murakami Y, Takagi T, Wada K, Matsunaga H. Essential structure of S-N curve: prediction of fatigue life and fatigue limit of defective materials and nature of scatter. *Int J Fatigue* 2021;146:106138. DOI
56. Li J, Yang Z, Qian G, Berto F. Machine learning based very-high-cycle fatigue life prediction of Ti-6Al-4V alloy fabricated by selective laser melting. *Int J Fatigue* 2022;158:106764. DOI
57. Zhan Z, Li H. A novel approach based on the elastoplastic fatigue damage and machine learning models for life prediction of aerospace alloy parts fabricated by additive manufacturing. *Int J Fatigue* 2021;145:106089. DOI
58. Zhan Z, Li H. Machine learning based fatigue life prediction with effects of additive manufacturing process parameters for printed SS 316L. *Int J Fatigue* 2021;142:105941. DOI
59. Ayoub G, Naït-abdelaziz M, Zaïri F, Gloaguen J, Charrier P. Fatigue life prediction of rubber-like materials under multiaxial loading using a continuum damage mechanics approach: effects of two-blocks loading and R ratio. *Mech Mater* 2012;52:87-102. DOI
60. Darabi MK, Abu Al-rub RK, Little DN. A continuum damage mechanics framework for modeling micro-damage healing. *Int J Solids Struct* 2012;49:492-513. DOI