


Article

A Deep Learning-Based Framework for Bearing RUL Prediction to Optimize Laser Shock Peening Remanufacturing

Yuchen Liang ^{1,*} , Yuqi Wang ², Anping Li ¹, Chengyi Gu ³, Jie Tang ³ and Xianjuan Pang ⁴

- ¹ School of Mechanical Engineering, Jiangsu University, Zhenjiang 212013, China; lianping@stmail.ujs.edu.cn
- ² School of Transportation and Logistics Engineering, Wuhan University of Technology, Wuhan 430070, China; yqwang@whut.edu.cn
- ³ Jiangsu Haiyu Machinery Co., Ltd., Taizhou 225714, China; guchengyi20100703@163.com (C.G.); tangjie20100703@163.com (J.T.)
- ⁴ National United Engineering Laboratory for Advanced Bearing Tribology, Henan University of Science and Technology, Luoyang 471000, China; xjpang2001@haust.edu.cn
- * Correspondence: liangyuchen@ujs.edu.cn

Abstract: Accurate prediction of the remaining useful life (RUL) of bearings is crucial for maintaining the reliability and efficiency of industrial systems. This study introduces a novel methodology integrating advanced machine learning and optimization techniques to address this challenge. (1) A transformer-attention model was developed to process segmented vibration signals, effectively capturing complex patterns. The model showed better performance than traditional approaches, with an RMSE of 0.989. (2) A Deep Neural Network (DNN) was designed to predict the extended RUL of bearings after laser shock peening (LSP) remanufacturing. The fruit fly optimization (FFO) algorithm was employed to optimize the remanufacturing parameters; a 29.33% improvement was achieved in fitness compared to the baseline. (3) The DNN model predictions were validated against Finite Element Analysis (FEA) simulations, with a low relative error of 2.5% to 5.8%; the model showed good accuracy in capturing the effects of optimized LSP parameters on bearing life extension.

Keywords: bearing RUL prediction; laser shock peening; deep learning; remanufacturing; data pre-processing



Citation: Liang, Y.; Wang, Y.; Li, A.; Gu, C.; Tang, J.; Pang, X. A Deep Learning-Based Framework for Bearing RUL Prediction to Optimize Laser Shock Peening Remanufacturing. *Appl. Sci.* **2024**, *14*, 10493. <https://doi.org/10.3390/app142210493>

Academic Editor: Ilaria Cacciotti

Received: 27 September 2024
Revised: 30 October 2024
Accepted: 10 November 2024
Published: 14 November 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Bearings are critical components in different types of industrial machinery, including automotive, aerospace, and manufacturing systems. The performance and lifespan directly influence the overall efficiency and reliability of mechanical systems [1]. Bearings facilitate smooth motion and support loads to have the proper function essential for operational success. Accurate prediction of the remaining useful life (RUL) of bearings is crucial for effective maintenance planning, as it supports organizations to schedule timely interventions and reduce the risk of unexpected failures [2]. Such failures can lead to significant operational disruptions to and costly downtime for equipment [3]. Traditional RUL prediction methods include statistical approaches and simple machine learning models, but the traditional approaches struggle to capture the complex and nonlinear degradation patterns inherent in bearing wear [4,5]. Therefore, it is essential to predict RUL to extend the remaining life of the bearing system.

Based on the remaining life of the bearing system, it is essential to integrate remanufacturing approaches based on RUL prediction. The potential benefits of advanced remanufacturing techniques have not been fully explored in conjunction with RUL assessments [6]. Laser shock peening (LSP) is a surface enhancement technique that can improve the fatigue life of bearings by inducing beneficial compressive residual stresses in the material [7,8]. This process enhances the mechanical properties of bearings, and it also reduces crack initiation and propagation [9,10]. However, the connection between RUL predictions and

LSP remanufacturing techniques is still rarely researched, and the potential for innovations in predictive maintenance and lifecycle management is limited. To fully utilize the advantages of remanufacturing, it is essential to enhance the accuracy of RUL predictions and optimize maintenance strategies. Deep learning is an effective tool for predicting the RUL of equipment and machinery due to its powerful data processing capabilities. It is capable of processing complex, multi-dimensional datasets during industrial processes. By integrating advanced remanufacturing processes with predictive analytics, engineers can achieve more effective utilization of remanufactured components for improved operational efficiency and reduced costs during the machinery process. Figure 1 shows the bearing remanufacturing research background.

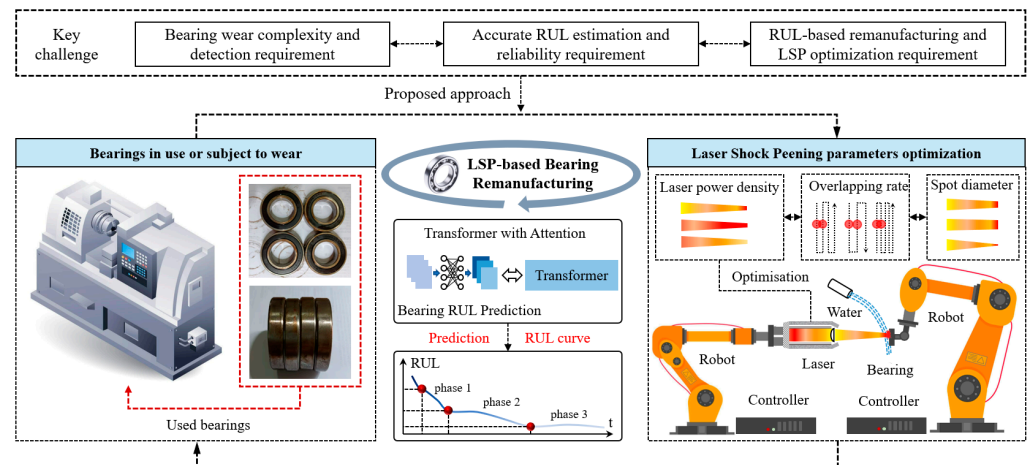


Figure 1. Bearing remanufacturing research background.

In order to address the limitations, this work introduces a novel RUL prediction framework that utilizes a hybrid transformer-attention algorithm developed to support LSP remanufacturing [11]. The transformer-attention algorithm was applied to vibration data collected from bearings operating under various conditions. The model was trained to predict RUL and support LSP remanufacturing process. The predicted RUL was subsequently used to determine the optimal LSP parameters including laser power density, overlapping rate, and spot diameter, and the remanufacturing process was optimized to the specific degradation state of each bearing. This paper proposes a novel solution for predictive maintenance and remanufacturing with the following innovations:

- Hybrid transformer-attention model for enhanced RUL prediction: The integration of transformer and attention mechanisms is proposed to predict bearing RUL. The transformer extracts relevant historical data points with similar vibration characteristics by comparing the features with standard samples. The attention mechanism focuses on samples based on the significance of the degradation trend so that a more accurate prediction of the bearing remaining life can be achieved;
- LSP parameter optimization based on RUL: A deep neural network (DNN) was used to predict the extended RUL of the bearing after LSP remanufacturing process, and the predictions were utilized by an optimization algorithm to determine the ideal LSP parameters, including laser power density, overlapping rate, and spot diameter. The fruit fly optimization (FFO) algorithm was utilized to optimize the remanufacturing performance;
- Experimental validation and case studies: Comprehensive experimental validation and case studies were conducted to validate the effectiveness of the proposed models and optimization strategies. Real-world bearing data and various LSP scenarios were tested to verify the accuracy of the RUL predictions and the efficacy of the optimized LSP parameters.

2. Related Works

2.1. Bearing Remaining Useful Life Prediction

The prediction of the RUL of bearings has been extensively explored to improve the reliability and maintenance of mechanical systems. Zhang et al. [12] proposed a weighted time-embedding transformer network that enhances the extraction of time-correlation features, and prediction error was significantly reduced in industrial applications. Lu et al. [13] introduced a cross-domain RUL prediction model with dynamic hybrid domain adaptation and attention contrastive learning to process domain shifts, and high accuracy and generalization were achieved. Song et al. [14] developed a comprehensive framework utilizing a fractional generalized Pareto degradation model; its performance was enhanced in early fault detection and adaptive failure threshold determination.

Several studies have emphasized the integration of innovative machine learning techniques with domain-specific adaptations. Cui et al. [15] presented a digital twin-driven graph domain adaptation model in which a multi-layered graph neural network was utilized for robust RUL prediction with limited data. Wang et al. [16] addressed the uncertainty in RUL prediction using a multi-task learning mixture density network; model-driven and data-driven strategies were effectively combined for improved prediction accuracy. Similarly, Lu et al. [17] integrated physical principles into an LSTM-based network to align the prediction results with physical laws, and consistency and accuracy were enhanced. Sun and Wang [18] explored a novel approach by combining entropy-based feature extraction with Elman neural networks; the proposed method showed good prediction accuracy over traditional time-domain features. Pan et al. [19] developed a meta-weighted network that quantifies uncertainty and mitigates negative transfer in limited data scenarios. Wei and Wu [20] proposed an attention-aware graph convolutional network that simultaneously handles temporal and feature correlations and is better than traditional models in prediction accuracy. Kumar et al. [21] developed an intelligent framework combining dynamic analysis-assisted wavelet filtering with graph convolution networks, providing a holistic approach to degradation monitoring and RUL estimation.

Bearing RUL prediction from previous studies has been widely researched with good results [15–17], but there is still a research gap in exploring bearing RUL prediction to support LSP remanufacturing optimization. The current studies mainly focus on improving RUL prediction accuracy [20,21], so there is a lack in the integration of RUL prediction results into LSP process optimization, which is essential to enhancing remanufacturing outcomes and extending component life in a more targeted manner.

2.2. Laser Shock Peening and Remanufacturing

LSP is a surface modification technology that enhances the mechanical properties and durability of metallic components. He et al. [22] validated the effectiveness of LSP in improving the high-temperature oxidation resistance of Ti₂AlNb alloys; microhardness and compressive residual stress were significantly improved, which contributed to the formation of protective Al-rich oxidation layers. Wen et al. [23] explored the impact of LSP on the high cycle fatigue properties of laser-welded 2A60 aluminum alloy joints, and LSP introduced compressive residual stresses and grain refinement, by which fatigue strength and lifespan were enhanced. Deng et al. [24] provided a comprehensive review of LSP applications in various metallic materials, emphasizing its potential to generate deep compressive residual stress fields and refine microstructures, which are critical for improving fatigue resistance in extreme environments. Ge et al. [25] combined LSP with laser cladding to enhance the fatigue properties of remanufactured Ti-6Al-4V alloy, and there was a 2.64-fold increase in fatigue life due to the beneficial microstructural modifications and removal of tensile residual stress. Huang et al. [26] studied the effects of femtosecond LSP on the fretting wear resistance of Ti6Al4V, in which the optimal laser energy parameters improved surface hardness and wear resistance. Chi et al. [27] attempted a combination of direct energy deposition (DED) and LSP on titanium alloys, which resulted in refined microstructures and a 12.46% increase in ultimate tensile strength. LSP was proven to be

effective in repairing and enhancing the mechanical properties of DEDed components. Bae et al. [28] explored LSP performance on silicon nitride ceramics with different sintering additive ratios, and LSP improved both surface hardness and residual compressive stress. Sheng et al. [29] introduced a laser non-uniform shock peening (LNUSP) technique for SUS 304 stainless steel welded joints; more uniform residual stress distribution and improved fatigue performance were achieved compared to traditional LSP methods. Lu et al. [30] investigated a hybrid approach combining laser-directed energy deposition with LSP for remanufacturing Ti6Al4V alloy. Interlayer LSP significantly improved microstructural features and mechanical properties, including tensile strength and hardness.

LSP has shown good effectiveness in improving the surface and mechanical properties of different materials, but the application in bearing remanufacturing is still rarely researched. There is a gap in extending the life of LSP remanufactured bearings; therefore, the relationship between LSP parameters and remanufacturing needs to be further researched so that optimization can be applied to find the optimal remanufacturing parameters.

3. Overview of the Research Framework

The proposed framework consists of three main stages. (i) Data acquisition and preprocessing: Vibration signals are collected from bearings under different operational conditions as the input data for the prediction model. The data are preprocessed through steps such as noise reduction and feature extraction to extract relevant features for subsequent analysis. (ii) Transformer-attention based RUL prediction: The preprocessed data are fed into the hybrid transformer-attention model. Transformer is employed to identify historical data points with similar degradation patterns with the current bearing condition, the attention mechanism assigns varying importance to these reference data points, and the most essential features can be extracted for the bearing's remaining useful life prediction. The combined approach enables a more refined and accurate RUL prediction. (iii) LSP parameter optimization: Based on the predicted RUL, the optimal laser shock peening parameters are determined based on the specific wear characteristics of each bearing. The selected parameters are applied to the bearing, which is also validated with FEA results. Figure 2 shows the overview of research framework.

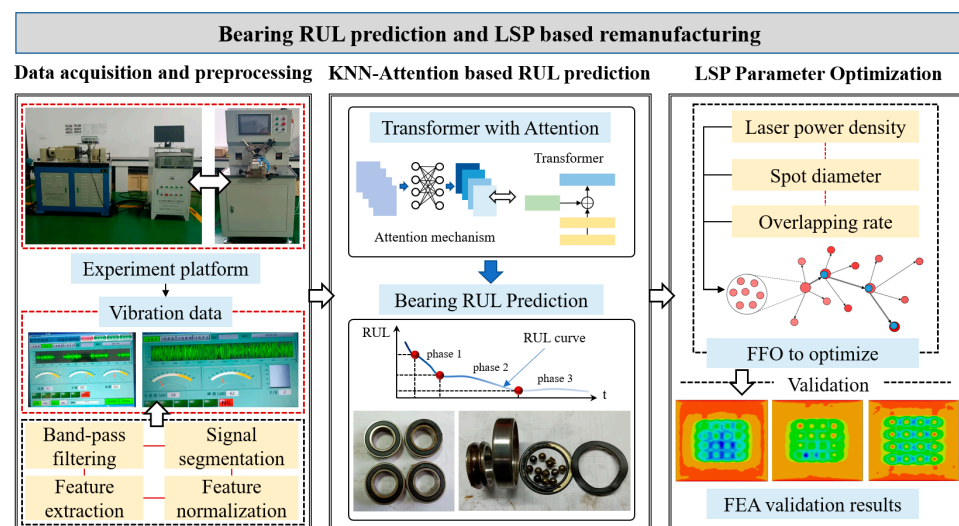


Figure 2. The overview of research framework.

4. Methodology

4.1. Data Collection and Preprocessing

Vibration signals were collected from bearings under different operational conditions using high-precision sensors, and the data were collected from Zhejiang Nulun Bearing Co., Ltd., Xinchang, China and University of Shanghai for Science and Technology in July 2024.

The setup includes accelerometers installed on the bearing housing to record the vibration signals. The time-series raw data $x(t)$ were preprocessed to remove noise and extract relevant features. In order to minimize the influence of ambient noise and interference, a band-pass filter was applied to the raw vibration signal [31]. The band-pass filter can be represented as follows:

$$y(t) = x(t) * h(t). \tag{1}$$

where $y(t)$ is the filtered signal, $x(t)$ is the raw signal, and $h(t)$ is the impulse response of the band-pass filter. The filter parameters were chosen based on the typical frequency range of bearing faults to retain only the most informative frequency components.

The filtered vibration signal $y(t)$ can be partitioned into smaller windows to capture local temporal features. Each segment is defined as $y_i(t)$ over a fixed time window T_w with an overlap of ΔT to improve continuity and smooth transitions between segments:

$$y_i(t) = y(t + i \times \Delta T), i = 0, 1, 2, \dots, N. \tag{2}$$

where N is the total number of segments; t ranges from 0 to T_w . From each segmented window $y_i(t)$, essential features are extracted to characterize the condition of bearing in transformer-attention model.

4.2. Development of the Transformer-Attention RUL Prediction Model

The transformer-attention RUL prediction model was designed to directly utilize the segmented vibration signals $y_i(t)$ as input, and the inherent intricate temporal dynamics can be captured in the bearing’s operational conditions [32,33]. Each segment $y_i(t)$ represents the filtered vibration data over a fixed time window T_w with an overlap of ΔT ; the capture of local temporal features is essential for accurate RUL estimation. Figure 3 shows the structure of transformer-attention algorithm.

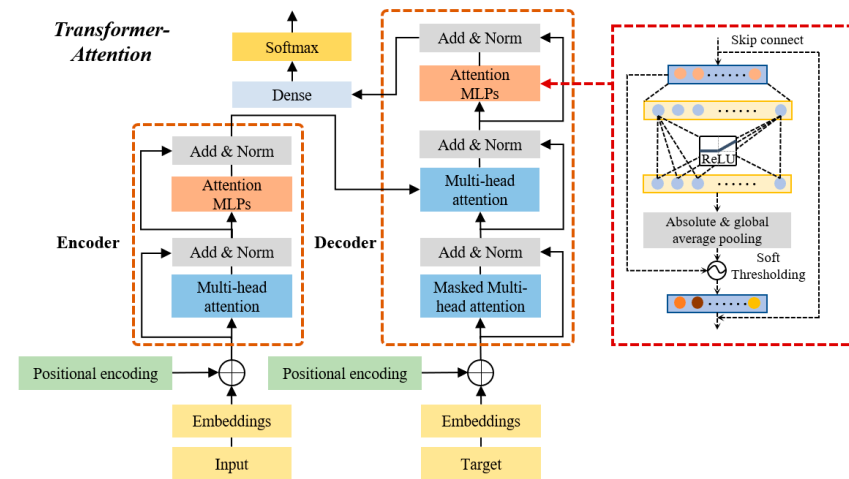


Figure 3. The structure of transformer-attention algorithm.

The segmented signals $\{y_1(t), y_2(t), \dots, y_N(t)\}$ are fed into the transformer encoder, which processes these raw time-series segments using multi-head self-attention mechanisms and positional encodings. This approach enables the model to learn the temporal dependencies and patterns directly from the vibration data, and the potential loss of information can be avoided. In the transformer encoder, each vibration segment $y_i(t)$ is first embedded into a higher-dimensional space through a linear transformation, and the addition of positional encodings is further used to preserve the order of the segments. The embedded sequence is passed through multiple self-attention layers, the model learns to attend to different parts of the sequence, and the key segments reflect significant degradation patterns in the bearing.

The self-attention mechanism calculates the attention scores α_i for each segment, and the importance relative to the current condition of the bearing is calculated [34]. The attention score for each segment is calculated:

$$To\alpha_i = \frac{\exp(Q_i \cdot K_i^\top / \sqrt{d_k})}{\sum_{j=1}^N \exp(Q_j \cdot K_j^\top / \sqrt{d_k})} \quad (3)$$

where Q_i and K_i are the query and key vectors derived from the segment $y_i(t)$, and d_k is the dimensionality of these vectors. α_i is the attention score. Then, the transformer function for each segment is represented as follows:

$$Z = \text{softmax}\left(\frac{QK^\top}{\sqrt{d_k}}\right)V. \quad (4)$$

where Q, K , and V are the query, key, and value matrices derived from the vibration segments $y_i(t)$. The function outputs a context-aware representation Z , which captures the temporal interactions among all segments. In addition, the weighted sum of the segments $y_{\text{weighted}}(t)$ is computed as follows:

$$y_{\text{weighted}}(t) = \sum_{i=1}^N \alpha_i \cdot y_i(t) \quad (5)$$

This aggregated representation $y_{\text{weighted}}(t)$ captures the most informative aspects of the bearing vibration data, and the temporal patterns are most indicative of its degradation state. The results are fed into a regression layer to predict the RUL of the bearing. The regression function is defined as follows:

$$\text{RUL}_{\text{pred}} = f(y_{\text{weighted}}(t)) \quad (6)$$

4.3. Model Training and Validation

The model parameters are optimized during training using the Mean Squared Error (MSE) loss function:

$$\text{MSE} = \frac{1}{N} \sum_{i=1}^N \left(\text{RUL}_{\text{pred},i} - \text{RUL}_{\text{true},i} \right)^2. \quad (7)$$

where $\text{RUL}_{\text{true},i}$ is the actual remaining useful life of the i -th bearing in the training set. The optimization is carried out using gradient descent techniques; the Adam optimizer is utilized, and the Adam optimizer has the ability to adapt the learning rate during training, so faster convergence and improved performance can be achieved. The Adam optimizer updates the model parameters iteratively based on the computed gradients of the loss function:

$$\theta_{t+1} = \theta_t - \eta \frac{\hat{m}_t}{\sqrt{\hat{v}_t + \epsilon}}. \quad (8)$$

where θ_t represents the model parameters at iteration t , η is the learning rate, \hat{m}_t and \hat{v}_t are the bias-corrected first and second moment estimates of the gradients, and ϵ is a small constant to ensure numerical stability.

The final model is validated on an independent test set, which contains vibration data that are not used during training. The training performance is evaluated using metrics such as root mean squared error (RMSE). RMSE provides a measure of the predictive accuracy by quantifying the difference between the predicted and actual RUL values:

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N \left(\text{RUL}_{\text{pred},i} - \text{RUL}_{\text{true},i} \right)^2}. \quad (9)$$

4.4. Laser Shock Peening Parameter Optimization

The DNN for predicting the extended RUL after the LSP remanufacturing process uses a set of input parameters that include laser power density P , overlapping rate O , spot diameter D , and the current remaining life $RUL_{current}$ [35]. The output of the network is the predicted remaining life after LSP remanufacturing $RUL_{extended}$.

The DNN can be represented by the following equation:

$$RUL_{extended} = f_{DNN}(P, O, D, RUL_{current}). \tag{10}$$

where f_{DNN} represents the DNN function mapping the inputs to the output. The network consists of multiple hidden layers. For a given hidden layer l , the transformation is defined as follows:

$$h^{(l)} = \sigma(W^{(l)}h^{(l-1)} + b^{(l)}). \tag{11}$$

where $h^{(l)}$ is the output of the l -th hidden layer, $h^{(l-1)}$ is the input to the l -th layer (or the input features for $l = 1$, $W^{(l)}$ and $b^{(l)}$ are the weights and biases of the l -th layer, and σ is the activation function (ReLU). The output layer of the DNN generates the predicted extended RUL:

$$RUL_{extended} = W^{(L)}h^{(L-1)} + b^{(L)}. \tag{12}$$

where L is the total number of layers in the DNN. The network is trained using a dataset of different LSP parameter combinations and corresponding extended RUL values. The training process minimizes the MSE between the predicted $RUL_{extended}$ and the actual RUL values.

Based on the trained DNN, the fruit fly optimization (FFO) algorithm is employed to optimize the LSP parameters for maximizing the extended RUL of bearing [36,37]. The FFO algorithm begins with an initial population of fruit flies, where each individual represents a candidate solution comprising a unique set of LSP parameters: laser power density, overlapping rate, and spot diameter. The fitness function is calculated as follows:

$$Fitness = RUL_{extended} = f_{DNN}(P, O, D, RUL_{current}). \tag{13}$$

where P , O , and D are the laser power density, overlapping rate, and spot diameter, respectively. Each fruit fly computes its fitness value, representing the predicted improvement in RUL based on its LSP parameter set. Figure 4 shows the structure of DNN and optimization function.

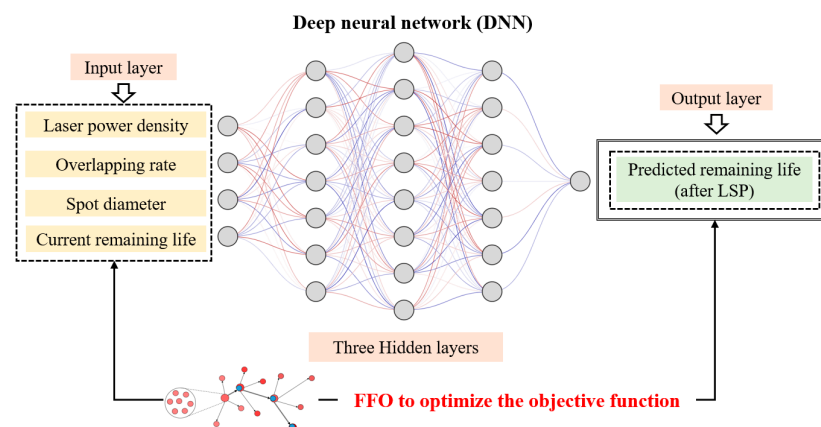


Figure 4. The structure of DNN and optimization function.

During the optimization process, the fruit flies adjust their positions in the search space by moving towards the fruit fly with the highest fitness value; the swarm behavior is

simulated in locating the optimal food source [38]. The new position $(P, O, D)_{\text{new}}$ for each fruit fly is updated:

$$(P, O, D)_{\text{new}} = (P, O, D)_{\text{best}} + \lambda \cdot \text{random}(0, 1). \quad (14)$$

where $(P, O, D)_{\text{best}}$ is the position of the best-performing fruit fly, λ is a scaling factor, and $\text{random}(0, 1)$ is a random Gaussian noise to introduce variability in the search. The final set of parameters is expected to maximize the bearing's extended RUL.

4.5. Validation and Case Studies

The validation of the optimized LSP parameters and the predicted extended RUL is conducted using Finite Element Analysis (FEA). The primary objective is to compare the extended RUL obtained from the DNN model with the RUL calculated from FEA simulations after LSP remanufacturing. The comparison validates the effectiveness of the optimization process and the accuracy of the predictive models.

The FEA model simulates the bearing's response under cyclic loading conditions with the effects of the optimized LSP parameters on the residual stress distribution and fatigue life. The methodology for calculating the extended remaining life using FEA involves the following steps:

The residual stress profile $\sigma_{\text{residual}}(d)$ as a function of depth d is obtained from the simulation. The dynamic response of the material is governed by the following wave equation:

$$\rho \frac{\partial^2 \mathbf{u}}{\partial t^2} = \nabla \cdot \boldsymbol{\sigma} + \mathbf{F}_{\text{LSP}}. \quad (15)$$

where ρ is the material density, \mathbf{u} is the displacement vector, $\boldsymbol{\sigma}$ is the stress tensor, and \mathbf{F}_{LSP} represents the LSP-induced force. The effective stress amplitude σ_{eff} experienced by the material under cyclic loading is modified due to the presence of residual stresses:

$$\sigma_{\text{eff}} = \sigma_{\text{applied}} - \sigma_{\text{residual}}. \quad (16)$$

where σ_{applied} is the applied stress amplitude from service loading; σ_{residual} is the compressive residual stress at the critical location obtained from FEA. The extended remaining life N_{FEA} is calculated using the Basquin equation, which relates the effective stress amplitude to the fatigue life:

$$N_{\text{FEA}} = \left(\frac{\sigma'_f}{\sigma_{\text{eff}}} \right)^{1/b} \quad (17)$$

where N_{FEA} is the number of cycles to failure after LSP, σ'_f is the fatigue strength coefficient, and b is the fatigue strength exponent. The DNN model predicts the extended remaining life N_{DNN} of the bearing after LSP remanufacturing based on the optimized LSP parameters and the current remaining life N_{current} :

$$N_{\text{DNN}} = f_{\text{DNN}}(P, O, D, N_{\text{current}}) \quad (18)$$

where f_{DNN} represents the trained DNN function mapping inputs to the extended remaining life. The accuracy of the DNN model is validated by comparing N_{DNN} with the FEA-calculated extended life N_{FEA} . A low relative error indicates strong agreement between the DNN predictions and the FEA results.

The relative error is computed as follows:

$$\text{Relative Error (\%)} = \left(\frac{|N_{\text{DNN}} - N_{\text{FEA}}|}{N_{\text{FEA}}} \right) \times 100\%. \quad (19)$$

5. Results

5.1. Data Collection and Preprocessing

The collected raw vibration data from the bearings were preprocessed to reduce noise using a band-pass filter. The filtering step minimized the impact of ambient noise and other interferences. The filtered signal was divided into overlapping segments to capture local temporal features. Each segment represents a specific time window with partial overlap. This segmentation strategy can be used for detailed analysis of the vibration characteristics, patterns, and anomalies that can be identified in the specific time frame. Essential features were extracted from each segment for further use as inputs for the transformer-attention model. Figure 5 shows the data collection and preprocessing example.

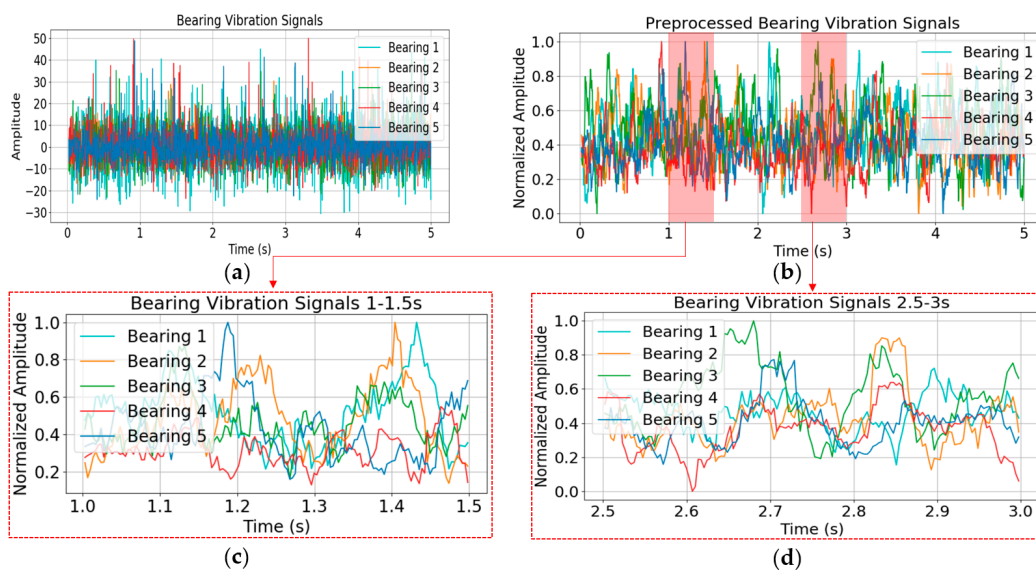


Figure 5. The preprocessed bearing vibration data collection and preprocessing. (a) Bearing raw data. (b) Pre-processed data. (c) Partitioned data sample 1. (d) Partitioned data sample 2.

5.2. Development of the Transformer-Attention RUL Prediction Model

The proposed transformer-attention model demonstrated good performance in predicting the RUL of bearings compared to other benchmark models. It achieved the lowest RMSE of 0.989 and a corresponding loss function value of 0.173; it showed the highest accuracy in capturing the complex temporal dependencies and degradation patterns present in the vibration signals. The CNN-LSTM model showed a higher RMSE of 1.356 and a loss value of 0.276, which was less accurate in extraction capabilities. The transformer model without the attention mechanism yielded an RMSE of 2.111; the importance of attention mechanisms is validated in the proposed approach. TCN and MLP showed higher RMSE values of 1.477 and 1.283, respectively, and the transformer-attention model was the most accurate in predicting bearing RUL. Figure 6 and Table 1 show the training accuracy of different models.

Table 1. Comparison of different models.

Model	RMSE	Loss Function
Transformer-Attention	0.989	0.173
CNN-LSTM	1.356	0.276
Transformer	2.111	0.268
TCN	1.477	0.481
2 layer MLP	1.283	0.399

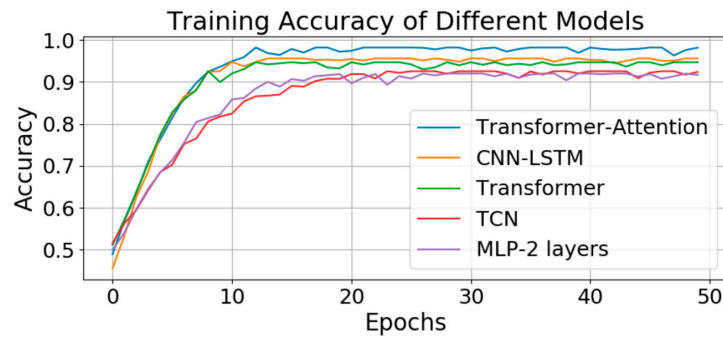


Figure 6. Training accuracy of different models.

5.3. Model Training and Validation

The transformer-attention model was optimized during training using the Mean Squared Error (MSE) loss function, with the Adam optimizer effectively adjusting model parameters. The performance was evaluated using RMSE, validation accuracy, and training accuracy metrics on the collected dataset. Across different parameter settings, the RMSE values ranged from 0.989 to 1.213; the model is thus sensitive to variations in attention heads and learning rates. Validation accuracy varied between 0.912 and 0.963, while training accuracy changed from 0.944 to 0.982.

Configurations with a higher number of attention heads generally have improved performance, with RMSE values falling consistently below 1.2 and validation accuracy above 0.92. For instance, settings with 8 to 16 attention heads and lower learning rates exhibit better predictive accuracy, with RMSE values ranging from 0.989 to 1.109. Configurations with fewer attention heads or higher learning rates tend to show slightly higher RMSE values and lower accuracy metrics.

The learning rate is essential for model performance. Lower learning rates tend to have better results by providing more stable parameter updates, which lead to lower RMSE and higher accuracy. The best result was achieved with eight attention heads and a learning rate of 0.002, with an RMSE of 0.989, validation accuracy of 0.963, and training accuracy of 0.982. Therefore, a moderate number of attention heads combined with a well-tuned lower learning rate is crucial for optimal model performance. Conversely, higher learning rates (e.g., 0.01 or above) tend to have slightly increased RMSE and decreased accuracy, which indicates instability in training. Figure 7 shows the testing results under different parameters and Table 2 shows the training and validation accuracy under different parameter settings.

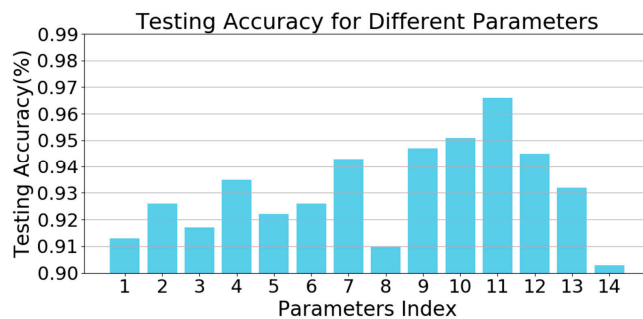


Figure 7. The testing results under different parameters.

Table 2. The training and validation accuracy under different parameter settings.

Index	Parameters Settings	RMSE	Validation Accuracy	Training Accuracy
1	heads = 2, lr = 0.001	1.213	0.912	0.944
2	heads = 4, lr = 0.001	1.152	0.921	0.954
3	heads = 8, lr = 0.001	1.134	0.925	0.957
4	heads = 16, lr = 0.001	1.112	0.935	0.962
5	heads = 4, lr = 0.01	1.135	0.928	0.955
6	heads = 4, lr = 0.05	1.124	0.929	0.958
7	heads = 8, lr = 0.0008	1.109	0.937	0.963
8	heads = 8, lr = 0.005	1.204	0.916	0.947
9	heads = 8, lr = 0.01	1.098	0.946	0.968
10	heads = 16, lr = 0.005	1.073	0.956	0.974
11	heads = 8, lr = 0.002	0.989	0.963	0.982
12	heads = 4, lr = 0.0008	1.105	0.932	0.961
13	heads = 4, lr = 0.005	1.198	0.919	0.948
14	heads = 2, lr = 0.002	1.125	0.925	0.955

5.4. Laser Shock Peening Parameter Optimization

The DNN model was trained to predict the extended RUL of bearings after LSP remanufacturing, using input parameters such as laser power density, overlapping rate, spot diameter, and the current RUL. The model provided accurate predictions of extended RUL, which were utilized as the fitness function in various optimization algorithms to find the optimal LSP parameters that maximize bearing life. Table 3 shows the effectiveness of different optimization algorithms for determining the optimal LSP parameters. The fruit fly optimization (FFO) algorithm performed better than other methods; a fitness value of 0.97 with a convergence in 23 iterations was achieved, and there was a 29.33% improvement in fitness over the fitness without optimization. FFO algorithm showed efficient optimization performance for maximizing extended RUL of remanufactured bearing.

Table 3. Comparison of different optimization algorithms for LSP parameters.

Algorithm	Fitness (0–1)	Convergence Iterations	Improvement in Fitness (%)
Without Optimization	0.75	-	-
FFO Algorithm	0.97	23	29.33
ACO Optimization	0.92	25	22.67
MOPSO Optimization	0.95	26	26.67
QPSO Optimization	0.90	31	20.00
SA Optimization	0.91	18	21.33
WOA Optimization	0.96	21	28.00
GWO Optimization	0.94	32	25.33
NSGA-II Optimization	0.89	27	18.67
GA Optimization	0.93	30	24.00

Ant colony optimization (ACO), multi-objective particle swarm optimization (MOPSO), quantum particle swarm optimization (QPSO), simulated annealing (SA), whale optimization algorithm (WOA), grey wolf optimizer (GWO), nondominated sorting genetic algorithm II (NSGA-II), and genetic algorithm (GA) also showed improvements in fitness values ranging from 0.89 to 0.96; the improvements were between 18.67% and 28.00%. However, the algorithms converged at a slightly slower pace or achieved lower fitness values compared to FFO. Therefore, the FFO algorithm was deemed most suitable for balancing convergence speed and maximizing the improvement in bearing life through optimal LSP parameter selection (See Figure 8).

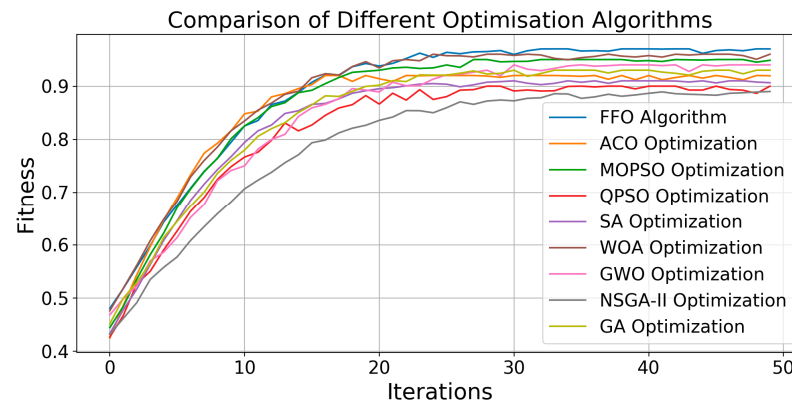


Figure 8. Comparison of different optimization algorithms.

5.5. Validation and Case Studies

The validation results showed a strong correlation between the extended remaining life predicted by the DNN model and the life calculated through FEA simulations. For various LSP parameter settings, the relative error between the DNN-predicted extended RUL N_{DNN} and the FEA-calculated extended RUL N_{FEA} remained consistently low, ranging from 2.5% to 5.8%. This indicates that the DNN model accurately captured the effects of different LSP parameters on the bearing’s fatigue life. When the optimized LSP parameters were applied, the FEA results showed an increase in the bearing RUL; there was an increase in N_{FEA} of 30.31%. The DNN predictions closely matched these results; The N_{DNN} value was increased by 29.33% compared to the DNN outcomes. Figures 9 and 10 show the FEA simulation under different LSP parameters.

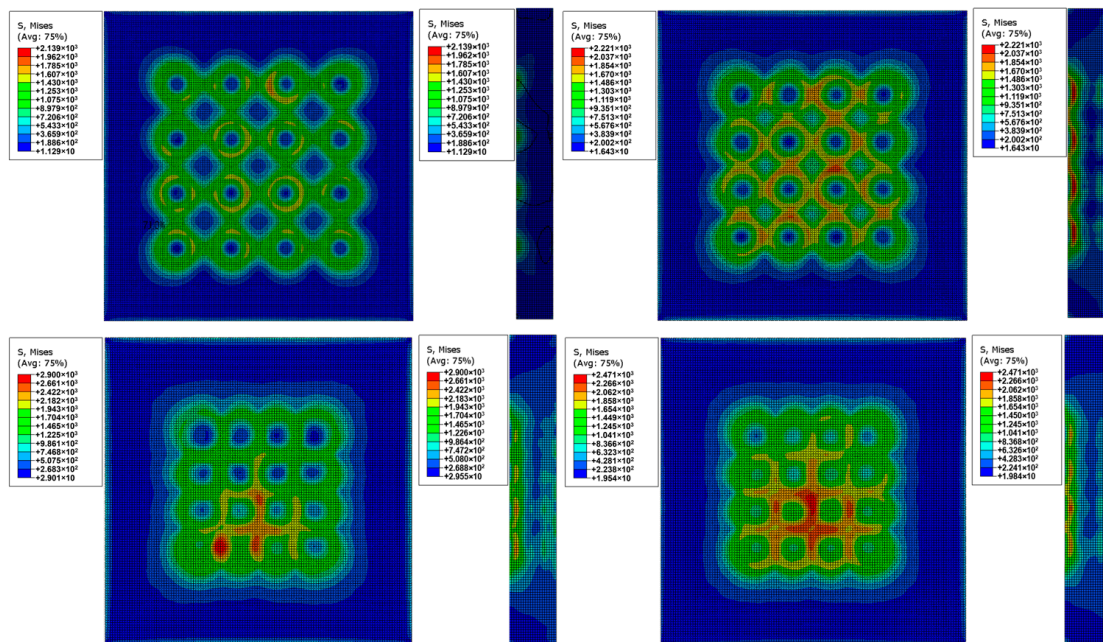


Figure 9. FEA simulation under different parameters.

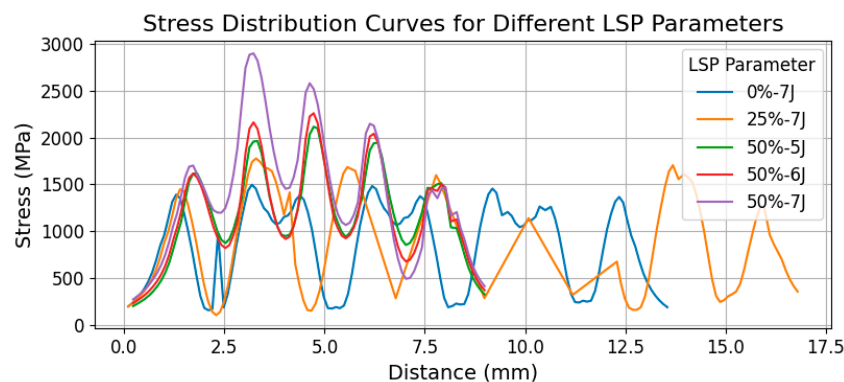


Figure 10. Stress distribution under different LSP parameters example.

6. Discussion

6.1. Data Collection and Preprocessing

The preprocessing of the raw vibration data using a band-pass filter effectively reduces ambient noise and external interferences, and a cleaner signal can show the relevant frequency components associated with bearing faults. The segmentation of this filtered signal into overlapping windows captures local temporal features, and a detailed analysis of the vibration characteristics over time can be implemented. This approach enables the identification of specific patterns and anomalies within defined time frames, and a comprehensive dataset can be provided to keep essential features for model training.

6.2. Development of the Transformer-Attention RUL Prediction Model

The performance of the transformer-attention model in predicting the RUL of bearings is primarily due to its ability to capture complex patterns within the vibration signals. The multi-head self-attention mechanism enables the model to focus on the most relevant segments of the data. This focused attention achieved more precise feature extraction and improved learning from the sequential data, lower RMSE, and better prediction accuracy. Other models without this mechanism could have higher errors and less reliable predictions. The advantage of the vanilla method is its simplicity and ease of implementation, as it is efficient and less prone to overfitting. It can be quickly deployed with a clear baseline for performance comparison, which can further simplify the model when iterating on improvements in future applications [39].

6.3. Model Training and Validation

The optimization of the transformer-attention model using the MSE loss function and the Adam optimizer demonstrated significant sensitivity to variations in attention heads and learning rates. Configurations with 8 to 16 attention heads and lower learning rates achieved the best performance, with RMSE values consistently below 1.2 and validation accuracy above 0.92; the model benefits from a higher number of attention heads, which enhances its ability to capture detailed patterns in the data. Configurations with fewer attention heads or higher learning rates showed slightly higher RMSE values and reduced accuracy, and these settings may not provide sufficient granularity or may lead to unstable training. In order to improve the training performance further, incorporating advanced generative deep learning methods (including manifold learning flows and diffusion models) can provide a more advanced approach to predicting the RUL by capturing more complex degradation patterns and nonlinear dynamics in bearing systems [40,41].

6.4. Laser Shock Peening Parameter Optimization

The DNN model predicts the extended RUL of bearings after LSP remanufacturing, and the key input parameters were analyzed. This accurate prediction capability is crucial for using the output as a fitness function in optimization algorithms to determine the optimal LSP parameters. Among the tested methods, the FFO algorithm showed the best

performance: A fitness value of 0.97 and a 29.33% improvement over the baseline without optimization were achieved, and convergence was achieved in 23 iterations. The FFO algorithm is effective in rapidly exploring the parameter space and identifying optimal solutions. Other algorithms have relatively slower convergence and slightly lower performance, with improvements ranging from 21.33% to 28.00%. Therefore, the FFO was found most suitable for applications to achieve bearing life extension.

6.5. Validation and Case Studies

The low relative error ranging from 2.5% to 5.8% between the DNN-predicted extended RUL and the FEA-calculated values indicates that the DNN model reflects the effects of different LSP parameters on bearing fatigue life. It shows that the proposed deep learning-based remanufacturing approach is validated. Additionally, there were similar increases in RUL observed in both DNN predictions and FEA simulations (29.33% and 30.31%). Therefore, the DNN model is reliable in predicting LSP remanufacturing performance.

7. Conclusions

This study addressed the critical need for accurate prediction of bearing RUL and optimization of LSP parameters for enhanced bearing performance. Predictive maintenance is essential for reducing downtime and maintenance costs in industrial applications, and the existing methods often lack the capability to effectively capture complex degradation patterns in bearing data. To overcome these challenges, the present research integrates advanced machine learning models with optimization techniques, providing a comprehensive and reliable solution. (1) The study developed a transformer-attention model to effectively capture complex patterns in vibration data, and the prediction accuracy RMSE was 0.989. This accuracy is better than that of CNN-LSTM and standard transformer models. (2) A DNN was implemented to predict the extended RUL after LSP remanufacturing, using key parameters such as laser power density, overlapping rate, spot diameter and current RUL. The FFO algorithm was employed to optimize these parameters, and there was a 29.33% improvement in RUL, which is better than other optimization methods. (3) The model's accuracy was validated through FEA, showing a low relative error ranging from 2.5% to 5.8% between predicted and simulated extended RUL.

This study shows the key advantages of the proposed approach compared to traditional methods in predicting bearing RUL and optimizing LSP parameters. The transformer-attention model captures complex degradation patterns in vibration data more effectively than conventional models, which enhances prediction accuracy and reliability. Additionally, the DNN-based RUL prediction (optimized with the FFO algorithm) adjusts LSP parameters, including laser power density and overlap rate, to improve bearing life, which achieves more effective parameter tuning than standard strategies. Validation through FEA further supported the model's precision, showing good alignment between the predicted and simulated RUL values.

This study has certain limitations for future exploration. (1) The transformer-attention model achieved a high prediction accuracy (RMSE 0.989). However, its effectiveness may be constrained by the current dataset's limited diversity. Broader testing across varied bearing types and degradation patterns in real-world settings could improve its generalizability and robustness. Future research should focus on data diversification and robustness testing to strengthen model adaptability to diverse industrial environments. (2) The DNN model for extended RUL prediction relies heavily on precise input parameters, such as laser power density and overlap rate, which are often challenging to maintain consistently in industrial applications. Variations in these parameters could impact prediction reliability. Future work could address this limitation by incorporating uncertainty modeling to manage parameter variability, which enhances prediction accuracy under fluctuating conditions.

Author Contributions: Writing—original draft and editing and supervision, Y.L.; writing—original draft and visualizations, Y.W.; data analysis, A.L.; project administration, C.G. and J.T.; investigation X.P. All authors have read and agreed to the published version of the manuscript.

Funding: This research was sponsored by the China Postdoctoral Science Foundation (Project No. 2023M741426) and the open fund project of the National United Engineering Laboratory for Advanced Bearing Tribology (Project No. 202405).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The raw data supporting the conclusions of this article will be made available by the first authors on request.

Acknowledgments: This research particularly thanks Zhejiang Nulun Bearing Co., Ltd. for providing experimental support.

Conflicts of Interest: Authors Chengyi Gu and Jie Tang were employed by the company Jiangsu Haiyu Machinery Co., Ltd. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

- Xiao, Y.; Liu, D.; Cui, L.; Wang, H. Heterogeneous Graph Representation-Driven Multiplex Aggregation Graph Neural Network for Remaining Useful Life Prediction of Bearings. *Mech. Syst. Signal Process.* **2024**, *220*, 111679. [[CrossRef](#)]
- Zuo, T.; Zhang, K.; Zheng, Q.; Li, X.; Li, Z.; Ding, G.; Zhao, M. A Hybrid Attention-Based Multi-Wavelet Coefficient Fusion Method in RUL Prognosis of Rolling Bearings. *Reliab. Eng. Syst. Saf.* **2023**, *237*, 109337. [[CrossRef](#)]
- Wu, J.; He, D.; Li, J.; Miao, J.; Li, X.; Li, H.; Shan, S. Temporal Multi-Resolution Hypergraph Attention Network for Remaining Useful Life Prediction of Rolling Bearings. *Reliab. Eng. Syst. Saf.* **2024**, *247*, 110143. [[CrossRef](#)]
- Pei, X.; Li, X.; Gao, L. A Novel Machinery RUL Prediction Method Based on Exponential Model and Cross-Domain Health Indicator Considering First-to-End Prediction Time. *Mech. Syst. Signal Process.* **2024**, *209*, 111122. [[CrossRef](#)]
- Magadán, L.; Granda, J.C.; Suárez, F.J. Robust Prediction of Remaining Useful Lifetime of Bearings Using Deep Learning. *Eng. Appl. Artif. Intell.* **2024**, *130*, 107690. [[CrossRef](#)]
- Sun, R.; Chen, Z.; Li, Q.; Zhang, X.; Li, R. Atomistic Insight into Laser Shock Peening Response of α -Titanium: An Experimentally Verified Simulation Study. *Acta Mater.* **2024**, *281*, 120359. [[CrossRef](#)]
- Jiang, L.; Lan, L.; Wang, H.; Yan, X.; Gao, S.; He, B.; Chen, C. Laser Shock Peening of Laser Melting Deposited TiAl Alloy for Enhancing Its Corrosion Resistance. *Surf. Coat. Technol.* **2024**, *483*, 130745. [[CrossRef](#)]
- Guan, Z.; Chen, S.; Li, Z.; Zou, S.; Hua, Y.; Cai, J. Tailoring Surface Properties and Corrosion Resistance of Laser Shock Peened Ti6Al4V Alloy by Low-Temperature Annealing. *J. Manuf. Process.* **2024**, *129*, 24–34. [[CrossRef](#)]
- Zhu, J.; Wang, Z.; Wang, C.; Luo, K.; Lu, J. Significant Improvement in Hot Corrosion Resistance of Inconel 690 by Laser Shock Peening. *Surf. Coat. Technol.* **2024**, *492*, 131202. [[CrossRef](#)]
- Dai, W.; Guo, W.; Li, Q.; Xiao, J.; Li, W.; Zhang, H. Homogenization of Local Microstructure and Mechanical Properties in Friction Stir Welded Al-Cu Alloy Joint Achieved through Laser Shock Peening. *J. Mater. Process. Technol.* **2024**, *333*, 118579. [[CrossRef](#)]
- Ebrahimi, M.; Basiri, A. RACEKNN: A Hybrid Approach for Improving the Effectiveness of the K-Nearest Neighbor Algorithm. *Knowl.-Based Syst.* **2024**, *301*, 112357. [[CrossRef](#)]
- Zhang, M.; He, C.; Huang, C.; Yang, J. A Weighted Time Embedding Transformer Network for Remaining Useful Life Prediction of Rolling Bearing. *Reliab. Eng. Syst. Saf.* **2024**, *251*, 110399. [[CrossRef](#)]
- Lu, X.; Yao, X.; Jiang, Q.; Shen, Y.; Xu, F.; Zhu, Q. Remaining Useful Life Prediction Model of Cross-Domain Rolling Bearing via Dynamic Hybrid Domain Adaptation and Attention Contrastive Learning. *Comput. Ind.* **2025**, *164*, 104172. [[CrossRef](#)]
- Song, W.; Wang, Z.; Kudreyko, A. Predictive Framework for Remaining Useful Life of Roller Bearings: Utilizing Fractional Generalized Pareto Degradation Model in Performance Evaluation. *Measurement* **2025**, *241*, 115772. [[CrossRef](#)]
- Cui, L.; Xiao, Y.; Liu, D.; Han, H. Digital Twin-Driven Graph Domain Adaptation Neural Network for Remaining Useful Life Prediction of Rolling Bearing. *Reliab. Eng. Syst. Saf.* **2024**, *245*, 109991. [[CrossRef](#)]
- Wang, X.; Li, Y.; Noman, K.; Nandi, A.K. Multi-Task Learning Mixture Density Network for Interval Estimation of the Remaining Useful Life of Rolling Element Bearings. *Reliab. Eng. Syst. Saf.* **2024**, *251*, 110348. [[CrossRef](#)]
- Lu, W.; Wang, Y.; Zhang, M.; Gu, J. Physics Guided Neural Network: Remaining Useful Life Prediction of Rolling Bearings Using Long Short-Term Memory Network through Dynamic Weighting of Degradation Process. *Eng. Appl. Artif. Intell.* **2024**, *127 Pt B*, 107350. [[CrossRef](#)]
- Sun, Y.; Wang, Z. Remaining Useful Life Prediction of Rolling Bearing via Composite Multiscale Permutation Entropy and Elman Neural Network. *Eng. Appl. Artif. Intell.* **2024**, *135*, 108852. [[CrossRef](#)]
- Pan, T.; Chen, J.; Liu, Z. A Meta-Weighted Network Equipped with Uncertainty Estimations for Remaining Useful Life Prediction of Turbopump Bearings. *Expert Syst. Appl.* **2024**, *252 Pt B*, 124161. [[CrossRef](#)]
- Wei, Y.; Wu, D. Remaining Useful Life Prediction of Bearings with Attention-Aware Graph Convolutional Network. *Adv. Eng. Inform.* **2023**, *58*, 102143. [[CrossRef](#)]

21. Kumar, A.; Parkash, C.; Tang, H.; Xiang, J. Intelligent Framework for Degradation Monitoring, Defect Identification and Estimation of Remaining Useful Life (RUL) of Bearing. *Adv. Eng. Inform.* **2023**, *58*, 102206. [[CrossRef](#)]
22. He, D.; Li, L.; Guo, W.; He, G.; Peng, P.; Shao, T.; Huan, H.; Zhang, G.; Han, G.; Yan, J. Improvement in oxidation resistance of Ti2AlNb alloys at high temperatures by laser shock peening. *Corros. Sci.* **2021**, *184*, 109364. [[CrossRef](#)]
23. Wen, F.; Long, Z.; Xing, Z.; Liu, X.; Huang, X.; Zhou, L. The effect of laser shock peening on very high cycle fatigue properties of laser welded 2A60 aluminum alloy joints. *Eng. Fract. Mech.* **2023**, *290*, 109537. [[CrossRef](#)]
24. Deng, W.; Wang, C.; Lu, H.; Meng, X.; Wang, Z.; Lv, J.; Luo, K.; Lu, J. Progressive developments, challenges and future trends in laser shock peening of metallic materials and alloys: A comprehensive review. *Int. J. Mach. Tools Manuf.* **2023**, *191*, 104061. [[CrossRef](#)]
25. Ge, M.Z.; Tang, Y.; Zhang, Y.K.; Wang, Y. Enhancement in fatigue property of Ti-6Al-4V alloy remanufactured by combined laser cladding and laser shock peening processes. *Surf. Coat. Technol.* **2022**, *444*, 128671. [[CrossRef](#)]
26. Huang, X.; Chen, K.; Zhou, L.; Lakshmi Narayan, R.; Ramamurty, U. Effect of laser energy on the fretting wear resistance of femtosecond laser shock peened Ti6Al4V. *Surf. Coat. Technol.* **2024**, *494*, 131353. [[CrossRef](#)]
27. Chi, J.; Cai, Z.; Zhang, H.; Zhang, W.; Wan, Z.; Han, G.; Peng, P.; Zeng, Z. Combining manufacturing of titanium alloy through direct energy deposition and laser shock peening processes. *Mater. Des.* **2021**, *203*, 109626. [[CrossRef](#)]
28. Bae, S.; Kim, H.; Lee, J.; Jeong, S. Effects of laser shock peening on silicon nitride ceramic with varying sintering additive ratios. *Ceram. Int.* **2024**; *in press*. [[CrossRef](#)]
29. Sheng, Y.; Feng, A.; Cao, X.; Lin, G.; Ngankam Ronaldo, C. Effects of laser non-uniform shock peening on the microstructure and fatigue performance of SUS 304 stainless steel welded joints. *Mater. Sci. Eng. A* **2024**, *914*, 147186. [[CrossRef](#)]
30. Lu, H.; Wu, L.; Wei, H.; Cai, J.; Luo, K.; Xu, X.; Lu, J. Microstructural evolution and tensile property enhancement of remanufactured Ti6Al4V using hybrid manufacturing of laser directed energy deposition with laser shock peening. *Addit. Manuf.* **2022**, *55*, 102877. [[CrossRef](#)]
31. Ben, Y.; Gao, Q.; Wei, T.; Gong, S.; Li, Q. Real-Time Heave Motion Measurement by Adaptive Band-Pass Filter Based on Strapdown INS. *Ocean Eng.* **2022**, *262*, 112278. [[CrossRef](#)]
32. Zhao, Z.-H.; Zhou, Y.; Zhang, S.; Chen, S.; Wen, C.-B.; Xu, Q.; Li, W.-H. PulseNet: Multi-Task Learning-Based Non-Contact Pulse Condition Diagnosis Using Multi-Scale Fusion and Transformer. *Knowl.-Based Syst.* **2024**, *302*, 112333. [[CrossRef](#)]
33. Hou, Y.; Li, T.; Wang, J.; Ma, J.; Chen, Z. A Lightweight Transformer Based on Feature Fusion and Global-Local Parallel Stacked Self-Activation Unit for Bearing Fault Diagnosis. *Measurement* **2024**, *236*, 115068. [[CrossRef](#)]
34. Wang, R.; Liu, Q.; You, W.; Chen, Y. A Multi-Task Deep Learning Model Based on Comprehensive Feature Integration and Self-Attention Mechanism for Predicting Response to Anti-PD1/PD-L1. *Int. Immunopharmacol.* **2024**, *142*, 113099. [[CrossRef](#)]
35. Singh, A.; Sinha, R.K. SS-DNN: A Hybrid Strang Splitting Deep Neural Network Approach for Solving the Allen-Cahn Equation. *Eng. Anal. Bound. Elem.* **2024**, *169*, 105944. [[CrossRef](#)]
36. Ren, X.; Huang, Z.; He, Y. Financial Warning for Coal Mining Investments: Evidence from the Fruit Fly Optimisation Algorithm with Backpropagation Neural Networks. *Energy Econ.* **2024**, *134*, 107594. [[CrossRef](#)]
37. Sutradhar, S.; Karforma, S.; Bose, R.; Roy, S. A Dynamic Step-Wise Tiny Encryption Algorithm with Fruit Fly Optimization for Quality of Service Improvement in Healthcare. *Healthc. Anal.* **2023**, *3*, 100177. [[CrossRef](#)]
38. Ibrahim, I.A.; Hossain, M.J.; Duck, B.C. A Hybrid Wind Driven-Based Fruit Fly Optimization Algorithm for Identifying the Parameters of a Double-Diode Photovoltaic Cell Model Considering Degradation Effects. *Sustain. Energy Technol. Assess.* **2022**, *50*, 101685. [[CrossRef](#)]
39. Kharga, R.; Nikoukar, A.; Hwang, I.-S.; Goudarzi, H.; Jafarippanah, S. Energy Savings in 25G-NGEPON for Spatial Computing Multimedia Services Based on Vanilla-RNN. *Opt. Fiber Technol.* **2024**, *87*, 103914. [[CrossRef](#)]
40. Flouris, K.; Konukoglu, E. Canonical Normalizing Flows for Manifold Learning. *Adv. Neural Inf. Process. Syst.* **2023**, *36*, 27294–27314.
41. Lu, C.; Zhou, Y.; Bao, F.; Chen, J.; Li, C.; Zhu, J. DPM-Solver: A Fast ODE Solver for Diffusion Probabilistic Model Sampling in Around 10 Steps. *Adv. Neural Inf. Process. Syst.* **2022**, *35*, 5775–5787. [[CrossRef](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.