



## Remaining Useful Life Prediction of Mining Equipment Based on Wavelet Thresholding and ResNet

Zhiguo Ma<sup>1</sup>, Mingjun Tang<sup>1</sup>, Lijun Dai<sup>1</sup>, Xiaoxiong Wu<sup>1</sup>, Tian Wang<sup>1</sup>, Depeng Wang<sup>2\*</sup>

<sup>1</sup> Power China Road Bridge Group Co., Ltd., Beijing 100160, China

<sup>2</sup> Beijing Chongde Construction Engineering Co., Ltd., Beijing 100097, China

Corresponding Author Email: [enjoynow123@163.com](mailto:enjoynow123@163.com)

Copyright: ©2025 The authors. This article is published by IETA and is licensed under the CC BY 4.0 license (<http://creativecommons.org/licenses/by/4.0/>).

<https://doi.org/10.18280/ts.420214>

### ABSTRACT

**Received:** 19 September 2024

**Revised:** 22 January 2025

**Accepted:** 30 January 2025

**Available online:** 30 April 2025

#### Keywords:

*mining equipment, remaining useful life prediction, vibration signal, wavelet threshold denoising, deep residual network*

With the continuous deployment and advancement of mining equipment, the need for effective health management and fault prediction has become critical to ensuring operational safety and enhancing equipment efficiency. Vibration signals, which serve as vital indicators of equipment operating conditions, have been widely employed in fault diagnosis and remaining useful life (RUL) prediction. However, the presence of substantial noise in vibration signals often hinders the effective extraction of meaningful features, thereby compromising the accuracy of RUL prediction. Although deep learning has demonstrated notable progress in signal processing and RUL forecasting, challenges persist in environments with high noise levels. Therefore, how to combine signal processing and deep learning techniques to improve prediction performance has become a current research hotspot. Existing approaches predominantly rely on traditional feature extraction techniques and machine learning algorithms, which tend to underperform under nonlinear or high-noise conditions. To address these limitations, a vibration signal denoising method based on improved wavelet thresholding was proposed in this study to suppress high-frequency noise while preserving key feature information. Furthermore, a method for RUL prediction based on the deep residual network (ResNet) was proposed in this study, wherein the residual learning mechanism enhances the generalization capacity and prediction accuracy of the model. By integrating wavelet threshold denoising with ResNet, a novel framework for predicting the RUL of mining equipment was established. This approach offers improved accuracy, providing both theoretical support and a practical basis for the health management of mining equipment.

## 1. INTRODUCTION

With the continued global development of the mining industry, increasing attention has been directed toward the service life and fault prediction of mining equipment [1, 2]. Failures of such equipment not only result in operational disruptions [3] but may also pose significant safety hazards to personnel [4]. As a result, the accurate prediction of the RUL of mining equipment has emerged as a critical issue in equipment management and maintenance. Among various condition monitoring indicators, vibration signals have been extensively utilized for both fault diagnosis and RUL estimation [5-7]. Enhancing the accuracy of these predictions through advanced signal processing techniques has therefore become a focal point in contemporary research on the health management of mining equipment.

The ability to predict RUL offers several strategic advantages, including early warning of potential equipment failures, optimization of maintenance strategies, and improved operational reliability and safety [8-13]. With the ongoing advancement of sensor technology and data acquisition systems, large volumes of vibration data are being generated, providing a substantial basis for data-driven RUL prediction

[14, 15]. Through in-depth analysis of such data, highly accurate forecasting and decision support can be achieved [16], offering a robust scientific foundation for equipment health management. Accordingly, research on the integration of modern signal processing methods with deep learning algorithms, aiming to improve the performance of RUL prediction, has important theoretical significance and practical application value.

Recent research on the prediction of the RUL of mining equipment has primarily focused on the application of traditional machine learning and deep learning approaches. However, conventional methods often struggle to extract meaningful features when faced with noisy environments and complex signal patterns, thereby limiting the accuracy of RUL predictions [17, 18]. For instance, the feature extraction method based on time-domain or frequency-domain analysis proposed by Tang et al. [19] demonstrated limited capacity in capturing the complete health status of equipment when processing nonlinear and complex vibration signals. Moreover, deep neural networks designed by Sun et al. [20] were shown to be susceptible to noise in training data, resulting in model overfitting or diminished predictive performance. To address these challenges, various signal

denoising techniques—such as wavelet threshold denoising—have been proposed in recent years. Nonetheless, most of these methods have exhibited limited effectiveness in removing complex noise components. Additionally, the training of deep learning models remains highly dependent on the quality and quantity of the input data.

The present study comprises two primary components. First, to mitigate the impact of noise in vibration signals collected from mining equipment, a signal denoising method based on improved wavelet thresholding was introduced. This approach enables the effective suppression of high-frequency noise while retaining informative signal features, thereby providing more accurate data support for subsequent RUL prediction. Second, ResNet was employed for RUL prediction, where the incorporation of a residual learning mechanism facilitates the learning of complex nonlinear relationships and improves prediction accuracy. Taken together, this study is not only innovative in signal processing and machine learning model optimization but also provides a new solution for the health management of mining equipment, with high academic value and practical application prospects.

## 2. VIBRATION SIGNAL DENOISING OF MINING EQUIPMENT BASED ON IMPROVED WAVELET THRESHOLDING

During the operation of mining equipment, vibration signals—serving as a critical indicator of equipment health—are often affected by multiple factors, leading to the existence of a large amount of noise in the signals. Such noise may originate from environmental disturbances, sensor inaccuracies, or mechanical friction, all of which significantly degrade signal quality and, consequently, impair the accuracy of RUL prediction models. Conventional signal processing techniques, such as low-pass or high-pass filtering, often fail to effectively remove noise and retain useful information when applied to complex vibration signals generated by mining equipment. This limitation stems from the highly nonlinear nature of the signals and the wide frequency distribution of the noise components. As a result, linear filtering methods have proven inadequate for practical applications, underscoring the need for a more precise and adaptable denoising approach capable of handling complex signal structures. To address these challenges, an improved wavelet thresholding method was proposed in this study, which combines the advantages of nonlinear filtering with logarithmic wavelet thresholding. This hybrid approach was designed to overcome the limitations of traditional techniques in dealing with vibration signals from mining equipment.

In the context of vibration signal processing, classical hard and soft thresholding methods have achieved a degree of success in denoising tasks; however, both approaches are subject to notable limitations. In situations where high precision is required, hard thresholding functions tend to introduce discontinuities around  $\pm\eta$ , resulting in signal oscillations that reduce smoothness and complicate subsequent feature extraction. On the other hand, soft thresholding functions are known to introduce a consistent bias, especially when handling impulsive noise, and the denoising effect may degrade. Although compromise functions—such as the one expressed in Eq. (1)—have been proposed to balance the characteristics of hard and soft thresholding, the introduction of multiple tuning parameters

has increased algorithmic complexity and failed to fully eliminate bias-related issues.

$$\hat{q}_{k,j} = \begin{cases} SGN(q_{k,j})|q_{k,j} - z\eta|, & |q_{k,j}| \geq \eta \\ 0, & |q_{k,j}| < \eta \end{cases} \quad (1)$$

where,  $z$  denotes the adjustment factor. In addition, new improved thresholding functions have been proposed based on the conventional hard and soft thresholding functions and existing ones. One such function is expressed in Eq. (2):

$$\hat{q}_{k,j} = \begin{cases} SGN(q_{k,j}) \left( |q_{k,j}| - \left( \frac{i\eta}{X} \cdot \frac{1}{Y} \right) \right), & |q_{k,j}| \geq \eta \\ 0, & |q_{k,j}| < \eta \end{cases} \quad (2)$$

where,

$$X = \exp\left(o \frac{|q_{k,j}| - \eta}{\eta}\right) + n \quad (3)$$

$$Y = \sqrt[w]{|q_{k,j}|^w - |\eta|^w} + 1 \quad (4)$$

where,  $i$ ,  $n$ ,  $o$ , and  $w$  represent the adjustable parameters, satisfying the condition  $i=n+1$ . Those parameters of the function are complex. To address the aforementioned limitations, an improved logarithmic thresholding function was proposed in this study. This new formulation was designed to overcome the inherent drawbacks of traditional methods and to provide a more effective algorithm for the denoising of vibration signals from mining equipment. Let  $\eta$  denote the predefined threshold,  $q_{k,j}$  the original wavelet coefficient, and  $\hat{q}_{k,j}$  the thresholded wavelet coefficient. The structure of the improved logarithmic thresholding function is presented in Eq. (5):

$$\hat{q}_{k,j} = \begin{cases} q_{k,j} - \frac{\eta^2}{q_{k,j} \ln(r + |q_{k,j}| - \eta)}, & |q_{k,j}| \geq \eta \\ 0, & |q_{k,j}| < \eta \end{cases} \quad (5)$$

The improved logarithmic thresholding function exhibits significant advantages in the denoising of vibration signals from mining equipment. First, the function demonstrates favorable continuity, with no discontinuities or oscillatory behavior at  $\pm\eta$ . This ensures smoothness in the signal reconstruction process, which is particularly critical for the high-frequency and complex nonlinear characteristics typical of equipment vibration signals.

$$\begin{aligned} \lim_{|q_{k,j}| \rightarrow \eta^+} \hat{q}_{k,j} &= \lim_{|q_{k,j}| \rightarrow \eta^+} \left( \hat{q}_{k,j} - \frac{\eta^2}{q_{k,j} \ln(r + |q_{k,j}| - \eta)} \right) \\ &= \lim_{|q_{k,j}| \rightarrow \eta^+} \left( q_{k,j} - \frac{\eta^2}{q_{k,j}} \right) = 0 \end{aligned} \quad (6)$$

Similarly, it can be shown that  $\lim_{|q_{k,j}| \rightarrow \eta^-} \hat{q}_{k,j} = 0$ ,

demonstrating that the improved function is continuous at  $+\eta$ . Second, in comparison to soft thresholding and traditional compromise threshold approaches, the improved logarithmic thresholding function exhibits superior performance in eliminating constant bias. Owing to the mathematical properties of the logarithmic function, the denoising effect is further enhanced while avoiding the over-compression commonly observed with soft thresholding functions. Moreover, as no additional parameters are introduced, the computational complexity of the proposed method is significantly lower than that of approaches reported in prior literature. This reduction in complexity facilitates improved computational efficiency, rendering the method particularly suitable for signal processing tasks in mining equipment real-time monitoring systems.

$$\begin{aligned}
& \lim_{|q_{k,j}| \rightarrow +\infty} (\hat{q}_{k,j} - q_{k,j}) \\
&= \lim_{|q_{k,j}| \rightarrow +\infty} \left( q_{k,j} - \frac{\eta^2}{q_{k,j} \ln(r + |q_{k,j}| - \eta)} - q_{k,j} \right) \\
&= \lim_{|q_{k,j}| \rightarrow +\infty} \left( -\frac{\eta^2}{q_{k,j} \ln(r + |q_{k,j}| - \eta)} \right) = 0
\end{aligned} \quad (7)$$

Furthermore, it can be shown that  $\lim_{q_{k,j} \rightarrow \infty} (\hat{q}_{k,j} - q_{k,j}) = 0$ , indicating that the difference between  $\hat{q}_{k,j}$  and  $q_{k,j}$  diminishes as the wavelet coefficient  $q_{k,j}$  increases.

The selection of an appropriate threshold plays a critical role in the effectiveness of wavelet thresholding for the denoising of vibration signals from mining equipment. Given the pronounced nonlinear characteristics of such signals and the coexistence of both valuable frequency components and various types of noise, conventional fixed-threshold methods often fail to accommodate this complexity. An excessively low threshold may result in residual noise remaining in the reconstructed signal, which compromises subsequent fault diagnosis and RUL prediction. Conversely, an overly high threshold may mistakenly eliminate significant signal components by treating them as noise, thereby resulting in the loss of essential information. To address the specific characteristics of vibration signals from mining equipment, a threshold selection method based on the properties of wavelet coefficients at different decomposition levels was proposed in this study. According to the observed behavior of wavelet coefficients across multiple scales, noise components tend to diminish with increasing decomposition depth, whereas useful signal components exhibit an increasing trend. Accordingly, a layered threshold selection strategy was employed: the first-level wavelet decomposition is processed using a traditional thresholding rule, and the thresholds for subsequent layers are progressively increased based on the value from the preceding level. This enables the threshold of each layer to better adapt to the signal and noise distribution in different frequency ranges, thereby improving the denoising effect.

Since wavelet coefficients at different decomposition levels exhibit varying sensitivity to noise, it is essential to adjust the threshold according to the noise amplitude and signal strength at each level. To achieve this, a method estimating the Lipschitz exponent based on wavelet transform modulus maxima was adopted in this study. By calculating the Lipschitz exponent of the noise, the relationship among the wavelet coefficients at different levels was evaluated. This

approach allows for dynamic adjustment of the threshold across layers: high-frequency levels are primarily used for suppressing fine-grained noise, while low-frequency levels retain a greater proportion of informative signal content. Additionally, a layer-wise threshold computation rule was introduced. Specifically, the threshold for each level is set to  $2\sigma$  times the threshold at the preceding level. This ensures that during multilevel wavelet decomposition, noise is effectively filtered out while preserving key signal features to the greatest extent possible, thereby improving both the precision and efficiency of the denoising process. Let  $v$  denote the length of the signal, and  $\delta$  the standard deviation of the noise, calculated as  $\delta = (ME[q_{k,j}]) / 0.68$ , where  $ME[q_{k,j}]$  denotes the median of the wavelet decomposition coefficients at level  $k$ . This leads to:

$$\eta = \delta \sqrt{2LN(v)} \quad (8)$$

Let  $k$  denote the number of wavelet decomposition levels, and let the wavelet coefficients at the  $k$ -th level be represented by  $qd_{k,j}$ . A constant is denoted by  $J$ , and the Lipschitz exponent of white noise is represented by  $\sigma$ . Under these assumptions, the wavelet coefficients at level  $k$  satisfy the following condition:

$$|qd_k a(s)| \leq J 2^{k\sigma} \quad (9)$$

Accordingly, the wavelet coefficients at the  $(k+1)$ -th level can be expressed as:

$$|qd_{k+1} a(s)| \leq J 2^{(k+1)\sigma} = J 2^{k\sigma} \cdot 2^\sigma = 2^\sigma \cdot |qd_k a(s)| \quad (10)$$

Because the amplitude of the wavelet coefficients at the  $(k+1)$ -th level for the noise component is smaller than  $2^\sigma$  times the amplitude of those at level  $k$ , a condition can be formulated below. Let the maximum value of the wavelet decomposition coefficients at the  $k$ -th level be denoted by  $|q_{k,j}|_{MAX}$ , and let the decomposition scale be denoted by  $t$ , satisfying  $t = 2^k$ . Based on the computation of the Lipschitz exponent of the noise, the following expression can be obtained:

$$\sigma = \frac{\log_2 |q_{k,j}|_{MAX}}{\log_2 t} - \frac{1}{2} \quad (11)$$

Based on the above analysis, the threshold can be selected for noise suppression as follows:

$$\eta_{k+1} = 2^\sigma \eta_k = 2^{\left( \frac{\log_2 |q_{k,j}|_{MAX}}{\log_2 t} \right)} \eta_k \quad (12)$$

During the denoising of vibration signals from mining equipment, interference from impulsive noise with pronounced spiking characteristics, such as alpha-stable distributed noise, is frequently encountered. This type of noise significantly compromises signal accuracy. In particular, at locations of strong impulse interference, conventional wavelet threshold denoising methods often fail to effectively suppress such noise spikes, resulting in suboptimal denoising performance. To address this challenge, a combined denoising approach was proposed, integrating nonlinear filtering with the improved wavelet thresholding method. This hybrid

strategy was designed to attenuate strong impulsive noise while preserving critical signal components. Initially, a median-mean filtering algorithm was applied to the raw signal as a preprocessing step. This filtering technique was shown to be highly effective in weakening signals containing strong pulse noise, thereby reducing the influence of such noise on the subsequent wavelet threshold denoising process. Median-mean filtering, owing to its nonlinear properties, offers strong performance in suppressing spike-like noise. Then the logarithmic wavelet thresholding function was applied to further denoise the filtered signal. This approach preserves signal features while minimizing noise, especially for complex noise sources in mining equipment vibration signals, thereby more accurately restoring the health status of the equipment.

The specific steps of the proposed denoising algorithm, which combines nonlinear filtering with logarithmic wavelet thresholding for impulsive interference suppression in mining equipment vibration signals, were taken as follows:

a) The process began with median-mean filtering applied to the raw signal in order to suppress the interference caused by strong impulsive noise, thereby providing a cleaner signal for subsequent processing. Let  $M$  denote the window length,  $L$  a positive integer,  $a(e)$  the current signal point, and  $\hat{a}(e)$  the filtered signal point after median-mean processing. Given that  $2L+1 \leq v$ , where  $v$  denotes the signal length, it leads to:

$$\hat{a}(e) = \begin{cases} \frac{1}{M-2} \sum_{l=-L}^L a(e+l), & M = 2L+1 \\ \frac{1}{M-2} \sum_{l=-L}^{L-1} a(e+l), & M = 2L \end{cases} \quad (13)$$

b) The denoised signal was then subjected to wavelet decomposition based on a suitably determined number of decomposition levels. At each level, the threshold was computed based on the wavelet coefficients from the preceding layer. This ensures effective noise suppression across multiple frequency scales.

c) Logarithmic wavelet thresholding was subsequently applied. This step allows for a more accurate delineation between noise and signal, overcoming the constant bias commonly introduced by traditional methods.

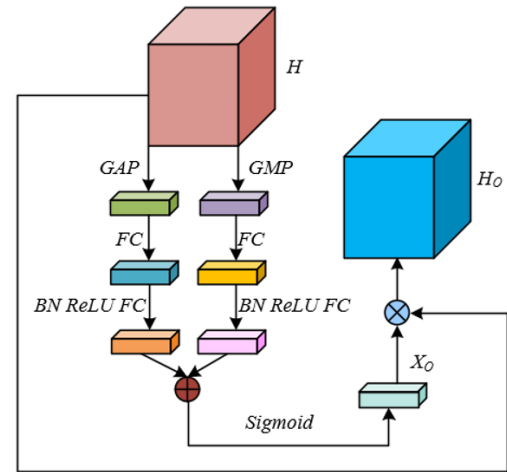
d) Wavelet reconstruction was finally performed to regenerate the denoised signal, ensuring that key features of the original vibration signal are preserved.

By following this structured procedure, the proposed algorithm effectively removed impulse interference from mining equipment vibration signals through the integrated use of nonlinear filtering and wavelet threshold denoising. The result was a substantial improvement in signal quality, providing a more reliable data foundation for subsequent equipment condition monitoring and fault diagnosis.

### 3. RUL PREDICTION OF MINING EQUIPMENT BASED ON RESNET

In the task of predicting the RUL of mining equipment, both temporal and spatial features of the signal exert a significant influence on predictive performance. Traditional deep residual shrinkage networks typically generate feature vectors by applying absolute value transformation and global average pooling (GAP) to the input signal, followed by element-wise multiplication with channel attention weights to determine

thresholds. However, such an approach neglects spatial domain information, which causes the models to fail to fully consider the influence of spatial position on signal features when processing grayscale images of mining equipment vibration signals. These signals often exhibit complex spatial patterns and structural features that are strongly correlated with the equipment's health status and RUL prediction. Variations in spatial features at different time points and under differing operational conditions directly affect the precision of the RUL prediction. As a result, traditional models relying solely on channel-domain information are insufficient to meet the practical requirements of RUL forecasting in mining environments, and often result in degraded prediction performance. To address this limitation, a novel residual shrinkage model was proposed in this study, designed to simultaneously capture both channel-domain and spatial-domain feature information.



**Figure 1.** Architecture of the improved channel attention network

In the task of RUL prediction for mining equipment, extreme data values embedded within the signal often reflect critical information regarding the onset or progression of mechanical faults. These extreme signals are typically indicative of failure modes or abnormal operational behaviors and serve as valuable cues for accurate RUL predication. However, conventional GAP, while effective in capturing overall signal information, tends to obscure such extreme values by computing only the mean level of the signal. In the context of mining equipment vibration data, this limitation can result in reduced sensitivity to key, extreme outliers associated with early fault symptoms, thereby diminishing the model's diagnostic effectiveness. To address this issue, global max pooling (GMP) was introduced into the channel attention network within the residual shrinkage module. The incorporation of GMP enhances the model's capacity to focus on critical outliers, thereby improving sensitivity to failure-related features. Figure 1 illustrates the architecture of the improved channel attention network. Specifically, for a feature map  $H$  with  $Z$  channels obtained through an operation  $h(\cdot)$ , both GAP and GMP were applied in parallel. Let  $Q$ ,  $G$ , and  $Z$  denote the height, width, and number of channels of the feature map, respectively. The resulting compressed global feature vectors are denoted as  $D_o^{AVG}$  and  $D_o^{MAX}$ , respectively, and are defined as follows:

$$D_o^{AVG} = GAP(H) \quad (14)$$

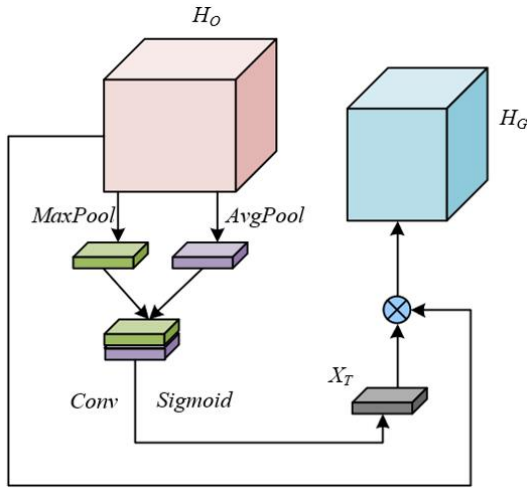
$$D_o^{MAX} = GMP(H) \quad (15)$$

Let  $d(\cdot)$  denote the composite operation applied to both  $D_o^{AVG}$  and  $D_o^{MAX}$ , and let  $\delta(\cdot)$  represent the Sigmoid activation function. The channel attention weight coefficients  $X_o$  can then be computed as:

$$X_o = \delta\left(d\left(D_o^{AVG}\right) + d\left(D_o^{MAX}\right)\right) \quad (16)$$

Finally, element-wise multiplication was performed between  $X_o$  and  $H$  to yield the feature map filtered by the channel attention network, expressed as:

$$H_o = H \otimes X_o \quad (17)$$



**Figure 2.** Architecture of the spatial attention network

In the context of RUL prediction for mining equipment, vibration signal grayscale images contain not only rich temporal features but also significant spatial information. Spatial information often reflects the operational state and

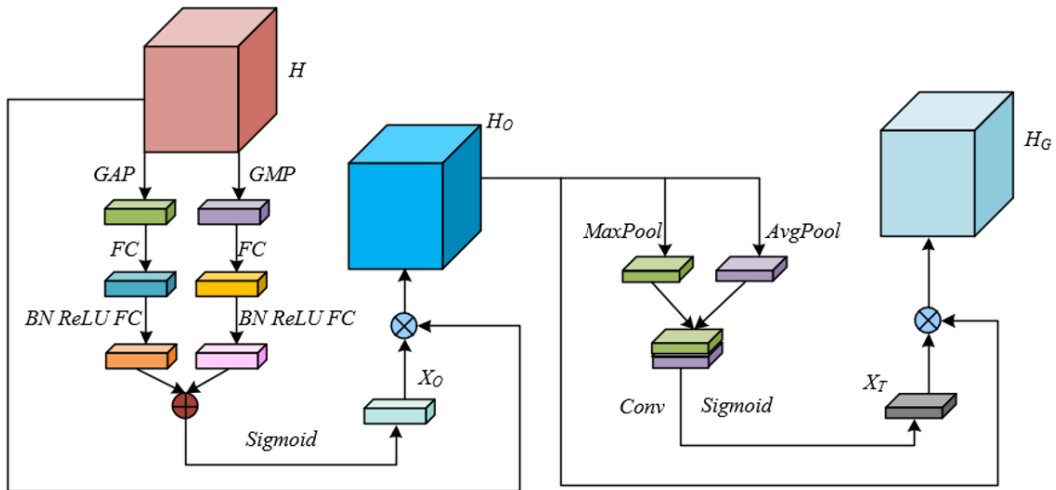
interrelationships among different components of the equipment, which are critical for accurate RUL prediction. Traditional channel attention networks primarily focus on inter-channel feature dependencies. Although such networks assist in capturing important global features, they overlook the spatial positional relationships within feature maps. In many cases, mechanical faults in mining equipment manifest through significant vibration signals concentrated in specific regions. Consequently, relying solely on channel attention mechanisms may result in the omission of critical spatially localized fault signatures. To address this limitation, a spatial attention network was integrated into the proposed model. This addition enhances the model's capacity to learn spatial relationships within input feature maps and to focus on local regions that play a decisive role in fault classification, thereby significantly improving the prediction accuracy for the RUL of mining equipment. Figure 2 presents the architecture of the spatial attention network.

Let the spatial attention weight coefficients be denoted as  $X_T$ , and the input to the spatial attention network as  $H_o$ . Specifically, GAP and GMP were applied to  $H_o$ , resulting in two spatially compressed feature maps:  $D_o^{AVG}$  and  $D_o^{MAX}$ . These two feature maps were then concatenated along the channel dimension and passed through a convolutional layer. The output was subsequently activated by the sigmoid function to obtain the spatial attention weight coefficients. Let the convolution operation be represented by  $CONV(\cdot)$ , and the channel-wise concatenation operation by  $CAT(\cdot)$ . The computation is expressed as:

$$X_T = \delta\left(CONV\left(CAT\left(D_o^{AVG}, D_o^{MAX}\right)\right)\right) \quad (18)$$

Finally, element-wise multiplication was performed between  $X_T$  and  $H_o$ , yielding the feature map filtered by the spatial attention network:

$$H_G = H_o \otimes X_T \quad (19)$$



**Figure 3.** Architecture of the hybrid attention network

Subsequently,  $H_G$  can be obtained as:

$$H_G = H \otimes X_o \otimes X_T \quad (20)$$

The implementation of the hybrid attention mechanism was

realized by serially connecting the channel attention network and the spatial attention network. Figure 3 illustrates the architecture of the hybrid attention network. Specifically, the input feature map was first subjected to an absolute value operation, resulting in a feature map  $H$  with  $Z$  channels. The



feature map  $H$  was then passed through the channel attention network to compute the channel attention weight coefficients  $X_O$ . These weight coefficients were element-wise multiplied with the feature map  $G$ , producing a feature map  $H_O$  filtered by the channel attention. This operation allows the model to automatically emphasize the most informative channel-specific features in the vibration signal, thereby filtering redundant information and preserving those features most relevant to RUL prediction. This step enhances the model's ability to extract robust and comprehensive channel-domain information, improving representational capacity across channels. Next, the refined feature map  $H_O$  was input into the spatial attention network to generate the spatial attention weight coefficients  $X_T$ , which were then element-wise multiplied with the feature map  $H$  to yield the final threshold representation. This process not only strengthens the extraction of channel-domain information but also enriches the feature representation through spatial attention, ensuring that the spatial distribution of vibration signal features is adequately captured. The integration of spatial-domain information enables the model to recognize subtle spatial variations in vibration patterns, thereby enhancing sensitivity

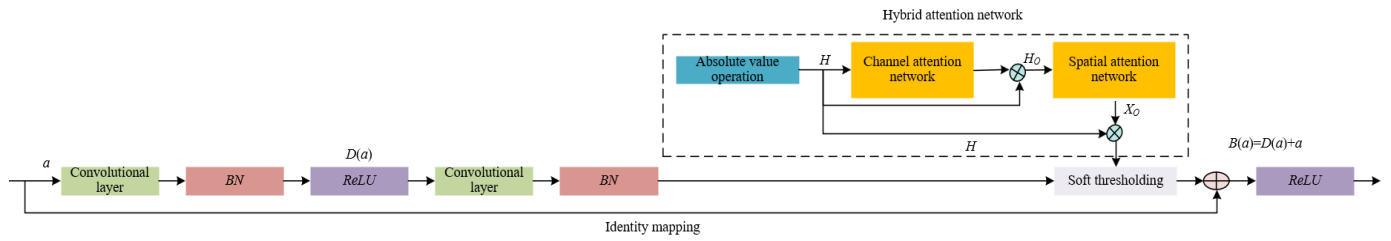
to changes in equipment health status. Let the feature map  $D$  be treated as a generalized function and the feature map  $H$  as the network input, and let the absolute value operation be denoted by  $ADS(.)$ . Then, it leads to:

$$H = ADS(D) \quad (21)$$

The final threshold representation obtained after processing through the hybrid attention network is expressed as:

$$\pi = H \otimes X_T \quad (22)$$

The improved residual shrinkage module enables comprehensive consideration of vibration signal characteristics across multiple dimensions, thereby optimizing feature extraction and information integration in the prediction of RUL for mining equipment. The introduction of the hybrid attention mechanism allows the model not only to assess the relative importance of each channel but also to effectively model spatial details within the vibration signal. As a result, the model exhibits enhanced robustness and prediction accuracy.



**Figure 4.** Architecture of the RUL prediction model for mining equipment

Based on the previously described improved residual shrinkage module, a ResNet model was further proposed for predicting the RUL of mining equipment. Figure 4 illustrates the architecture of the RUL prediction model. By employing multi-level feature extraction and a progressive refinement learning process, the model significantly improves the accuracy of RUL prediction. The input to the model consists of grayscale images representing vibration signals from mining equipment. Initially, the input was processed through a convolutional layer, after which the resulting feature map was passed into residual shrinkage layers and a multi-scale feature extraction module. The residual shrinkage layers adopt the architecture described in the earlier sections. The model is composed of four serially connected residual shrinkage layers, each containing two residual shrinkage modules, following a structure similar to that of ResNet-18. These layers employ progressive downsampling operations to effectively capture multi-level features ranging from fine-grained patterns to high-level semantic representations, while simultaneously avoiding feature loss commonly caused by large strides in traditional convolutional networks. To further enhance representational capacity, dilated convolution operations were incorporated within each residual shrinkage layer. By adjusting the dilation rate, the receptive field was expanded across multiple scales, allowing the model to extract rich, multi-scale features from vibration signals while preserving global contextual information. This mechanism enhances the model's ability to identify early indicators of equipment faults.

In mining equipment RUL prediction tasks, vibration signals often contain critical information distributed across various scales, which is essential for accurate estimation. To

address the limitations of traditional convolutional networks—particularly the potential for information loss during receptive field expansion—a network architecture based on dilated convolution feature expansion was proposed in this study. By introducing varying dilation rates within convolutional kernels, the receptive field was effectively expanded without the need for downsampling. As a result, broader feature representations can be captured. The primary advantage of dilated convolution lies in its ability to enlarge the receptive field while avoiding the information degradation typically associated with pooling or downsampling operations. This is especially critical for preserving the detailed characteristics of vibration signals. Within the proposed model, different dilation rates were strategically applied to extract global features at multiple scales. These features were then fused with the output from the residual shrinkage layers, further enriching the network's understanding of the multi-scale structure of the vibration signals, thereby enabling the model to identify faults at varying depths and granularity.

Specifically, in the proposed model, the output from each dilated convolutional layer was concatenated with the corresponding output from the residual shrinkage layer. A  $1 \times 1$  convolutional layer was subsequently applied to adjust the number of channels, thereby aligning the concatenated feature map with the processing requirements of the subsequent network layers. For each selected dilation rate, the stride and padding of the dilated convolutional layer were matched to the dimensions of the output feature map from the corresponding residual shrinkage layer. Let  $f$  denote the dilation rate and  $j$  the size of a standard convolutional kernel, then the size of the dilated convolutional kernel is defined as:

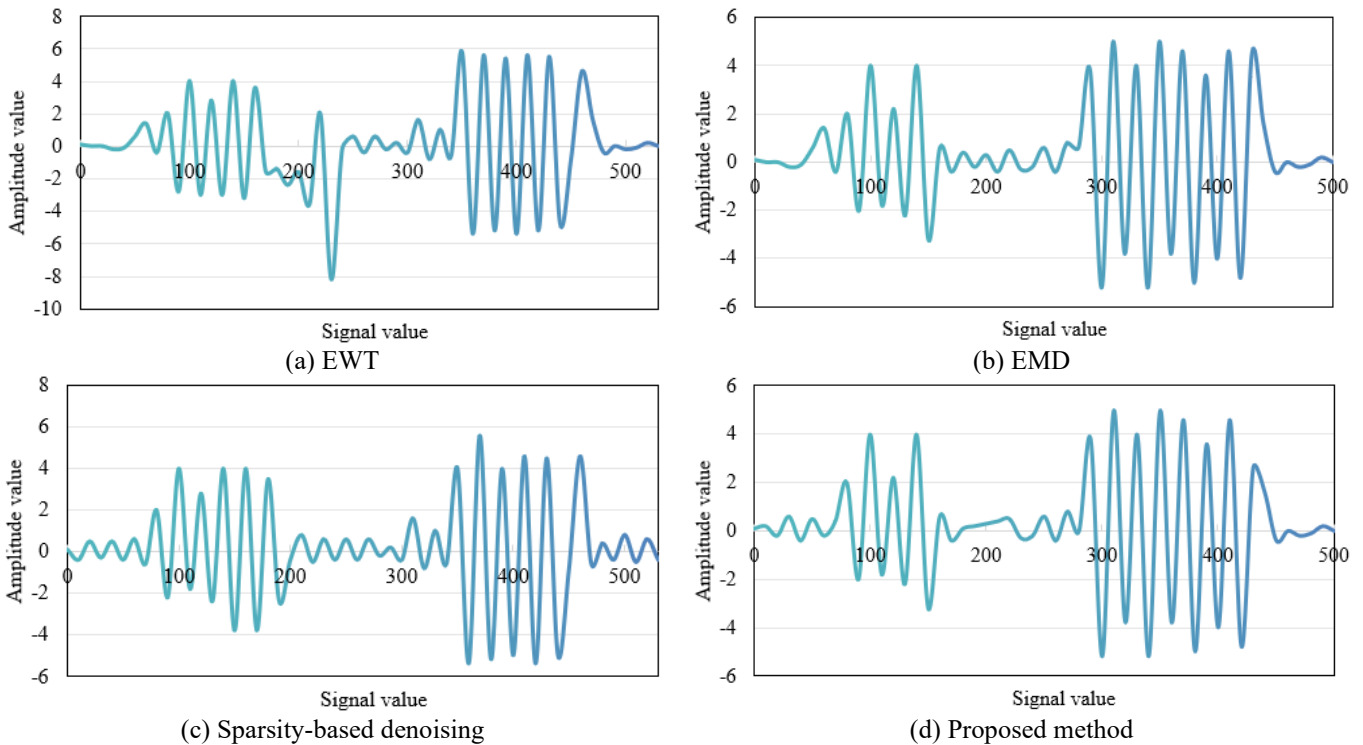
$$j' = j + (j-1) \times (f-1) \quad (23)$$

The core objective of this study's task is to predict the RUL of mining equipment by analyzing historical vibration signals. This prediction problem is typically formulated as a multi-class classification task. To estimate the RUL at different life stages, the Softmax function was used as the activation function in the final fully connected output layer. The Softmax function can convert multiple neuron values output by the model into a probability distribution, ensuring that the output of each class falls within the interval (0, 1), and that the sum of all class probabilities equals one. Each output corresponds to the probability that the equipment belongs to a specific RUL interval. The final prediction results are represented by the probability values at different life stages. To optimize the model's performance in RUL prediction, the cross-entropy loss function was employed. This loss function quantifies the discrepancy between the predicted probability distribution and the target distribution, and is widely used in classification tasks due to its effectiveness in measuring divergence between probabilistic outputs and ground truth labels. In the context of RUL prediction for mining equipment, the target probability distribution is generated based on historical condition data and known failure modes, while the predicted distribution represents the model's estimation of the equipment's remaining life. By minimizing the cross-entropy loss, the model was iteratively refined to produce predictions that more accurately reflect actual conditions. Let  $o(a)$  denote the target probability distribution and  $w(a)$  the predicted distribution. The objective function is defined as:

$$G(o, w) = -\sum_a o(a) \log w(a) \quad (24)$$

#### 4. EXPERIMENTAL RESULTS AND ANALYSIS

Upon examining the denoising results presented in Figure 5, it can be observed that residual noise remains in the signals processed using Empirical Wavelet Transform (EWT), Empirical Mode Decomposition (EMD), and sparsity-based denoising. The waveforms exhibit limited smoothness in finer details, and irregular fluctuations persist—particularly around the signal value of 200 and in other regions of notable variation. In contrast, the signal processed by the proposed method displays a significantly smoother overall trend, with noise effects markedly attenuated. The resulting waveform more closely approximates the characteristics of an ideal clean signal, thus visually reinforcing the denoising advantages achieved by the proposed approach. By employing an optimized wavelet thresholding strategy, the proposed method enables more precise identification and elimination of noise components while maximizing the retention of informative signal features. In terms of signal fidelity, the waveform reconstructed using the proposed method exhibits clearer structure and reduced noise interference. This indicates a superior capability in noise filtering, resulting in improved signal purity and more accurate representation of underlying characteristics, thereby providing a more accurate data foundation for subsequent RUL prediction of mining equipment.

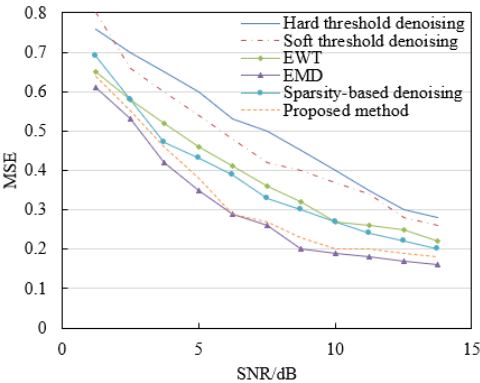


**Figure 5.** Comparison of denoising outcomes using different methods

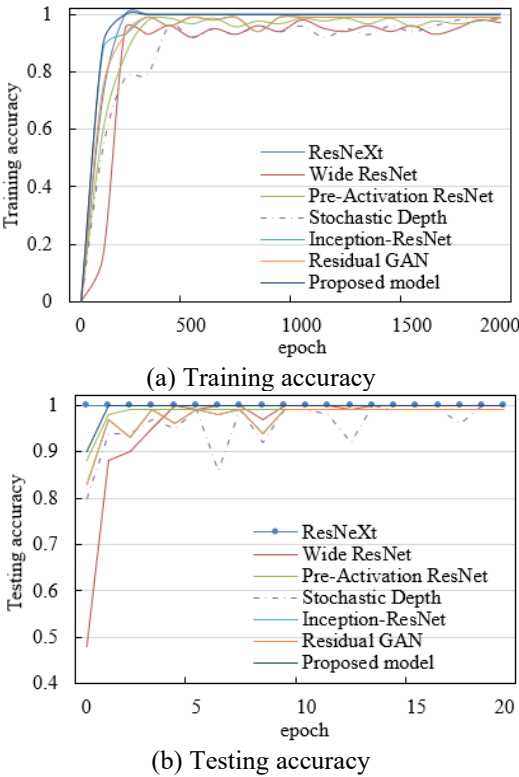
As observed in Figure 6, a decreasing trend in mean squared error (MSE) is exhibited by the proposed method with increasing input signal-to-noise ratio (SNR). When the SNR is 0 dB, the MSE achieved by the proposed method is approximately 0.65, which is notably lower than the values obtained by hard threshold denoising, soft threshold

denoising, and EWT. As the SNR increases to 15 dB, the MSE of the proposed method is further reduced to approximately 0.2, significantly outperforming the comparison algorithms under the same noise conditions. For example, the MSE for hard threshold denoising remains around 0.3, while that of EMD is approximately 0.18 but exhibits considerable

fluctuation. These results demonstrate that across a range of SNR conditions, the MSE of the proposed method consistently decreases with rising SNR and remains lower than those of the baseline algorithms at all evaluated points. This performance advantage is attributed to the optimized wavelet thresholding strategy, which enables more accurate noise identification and suppression while preserving the informative components of the signal. By contrast, other algorithms present distinct limitations: hard and soft threshold denoising methods exhibit insufficient flexibility in noise suppression, often resulting in either loss of signal features or retention of residual noise. Additionally, methods such as EWT and EMD display limited adaptability when processing signals with varying SNRs, leading to less pronounced MSE reduction or unstable error performance. Therefore, the proposed method demonstrates superior robustness and adaptability across a wide range of noise conditions. Its trend of decreasing MSE with increasing SNR is more significant.



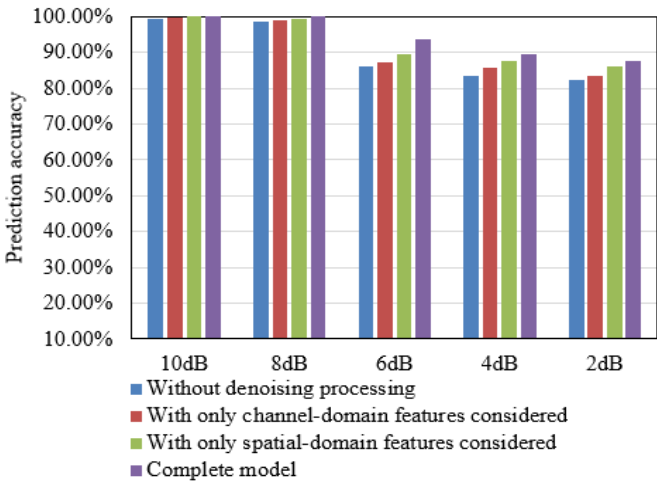
**Figure 6.** SNR vs. MSE curves for different denoising methods



**Figure 7.** Training and testing accuracy of the proposed model

As observed in Figure 7, the training and testing accuracy curves of the proposed model demonstrate exceptionally fast

convergence. The training accuracy stabilizes and reaches 100% after only 280 iterations. In terms of testing accuracy, a value of 99.5% is attained after the very first epoch, with further increases leading to stable convergence at 100% in subsequent epochs. By comparison, the six benchmark models—including Aggregated Residual Transformations for Deep Neural Networks (ResNeXt) and Wide ResNet—exhibit noticeable fluctuations during both training and testing phases. For instance, the training accuracy curve of the Stochastic Depth model shows repeated oscillations, and the testing accuracy of Wide ResNet displays multiple declines across epochs. These variations indicate that the training processes of the baseline models are less stable. The experimental results confirm that the proposed model achieves high testing accuracy at an early stage—99.5% in the first epoch—and rapidly stabilizes at 100%. The instability observed in the benchmark models is largely attributed to the absence of targeted signal preprocessing techniques and suboptimal network architecture design, resulting in reduced performance when handling noisy inputs or learning complex feature relationships. This comparison provides strong evidence of the proposed model’s significant advantages in terms of convergence speed and accuracy stability. Specifically, the proposed method converges faster and reaches 100% training and testing accuracy earlier and more consistently than the other six models evaluated.



**Figure 8.** Ablation study results of the proposed model

As shown in Figure 8, the complete model consistently outperforms the other three configurations under various SNR conditions. At an SNR of 10 dB, the prediction accuracies for the models without denoising, with only channel-domain features, and with only spatial-domain features are approximately 95%, 98%, and 98%, respectively, whereas the complete model achieves 100%. As the SNR decreases to 2 dB, the prediction accuracy of the model without denoising drops to approximately 80%, while the models with only channel-domain and spatial-domain features achieve approximately 85% and 88%, respectively. In contrast, the complete model maintains a higher accuracy of approximately 90%, demonstrating superior stability and robustness across different noise levels. The results of the ablation study validate the critical importance of both the improved wavelet threshold denoising process and multi-domain feature fusion—specifically the integration of channel-domain and spatial-domain features—in enhancing prediction accuracy. When denoising is omitted, the presence of high-frequency noise



significantly degrades performance, confirming the effectiveness of the proposed denoising strategy in preserving useful signal characteristics. When only a single domain of features is considered, the model’s representation of signal characteristics is incomplete, resulting in lower prediction accuracy compared to the full model. The complete model, by incorporating the improved wavelet thresholding method, effectively suppresses noise and provides more accurate data

**Table 1.** Complexity comparison of different models

| Model  | Input Size | FLOPs   | Number of Parameters |
|--|------------|---------|----------------------|
| Without denoising processing                 | 31×31      | 0.54218 | 11.25485             |
| With only channel-domain features considered | 31×31      | 0.64235 | 12.23015             |
| With only spatial-domain features considered | 31×31      | 0.54859 | 11.20325             |
| Complete model                               | 31×31      | 0.64213 | 12.22356             |

As shown in Table 1, the input dimensions of all models are  $31 \times 31$ . The floating-point operations (FLOPs) for the model without denoising amount to 0.54218; for the channel-domain-only and spatial-domain-only models, the values are 0.64235 and 0.54859, respectively; and for the complete model, the value is 0.64213. Regarding the number of parameters, the values are 11.25485 for the model without denoising, 12.23015 for the channel-domain-only model, 11.20325 for the spatial-domain-only model, and 12.22356 for the complete model. These results indicate that although variations exist across the models in terms of complexity, FLOPs and parameter count of the complete model are similar to those of the channel-domain-only configuration and are slightly lower. Importantly, despite integrating both the improved wavelet thresholding for denoising and multi-domain feature fusion, the complete model does not exhibit a significant increase in computational complexity. Both the operation count and parameter size remain within a reasonable range. Compared with configurations that omit denoising or consider only a single feature domain, the complete model achieves higher input quality through denoising and more comprehensive learning of complex nonlinear relationships via ResNet-based feature fusion while ensuring a certain level of computational complexity. This effective balance between complexity and performance enables enhanced prediction performance through the joint application of denoising and feature fusion strategies without incurring excessive computational burden. The overall results validate the feasibility and effectiveness of the proposed method in real-world mining equipment RUL prediction scenarios, offering a practical solution that achieves both accuracy and efficiency.

## 5. CONCLUSION

This study centers around two core components: signal denoising and RUL prediction. In practical mining equipment applications, vibration signals are often contaminated by high-frequency noise, which adversely affects the accuracy of subsequent RUL predictions. To address this challenge, an improved wavelet thresholding-based denoising method was proposed. This approach enables effective suppression of high-frequency noise while preserving essential signal features, thereby providing more accurate input data for RUL prediction tasks. The enhanced signal quality achieved through this innovative denoising method offers a more robust

foundation for subsequent prediction tasks. Furthermore, the use of a ResNet for multi-domain feature integration enables more comprehensive learning of complex nonlinear relationships. As a result, higher and more stable prediction accuracy is achieved across varying SNR conditions, highlighting the synergistic advantage of combining denoising and multi-domain feature fusion in the proposed method.

foundation for health management and early fault detection in mining equipment. In addition, a ResNet framework was introduced for RUL prediction. By incorporating residual learning mechanisms, the network was capable of capturing complex nonlinear relationships embedded in the vibration signals, while also mitigating the vanishing gradient issue commonly encountered in deep neural network training. Leveraging a hybrid attention mechanism based on the improved residual shrinkage module, the model was designed to perform feature selection and enhancement across multiple dimensions, significantly improving prediction accuracy. This study presents innovations in both signal processing and machine learning model optimization. Notably, the integration of the hybrid attention mechanism for feature extraction contributes not only to enhanced model performance but also to the development of a new solution for health management of mining equipment.

However, several limitations remain in the present study. Although the proposed signal denoising method effectively removes high-frequency noise, residual low-frequency noise may still persist under certain complex operating conditions. Further optimization of the denoising algorithm to accommodate more diverse noise patterns represents an important direction for future research. Additionally, while the ResNet demonstrates a strong capacity to learn complex nonlinear relationships, its performance under extreme conditions may still be constrained by the diversity of the training data. Therefore, enhancing the generalization capability of the model through the construction of more diverse and representative training datasets remains a key challenge. Future research may proceed in several directions. First, the signal denoising algorithm could be further refined through the integration of multi-scale processing techniques and advanced wavelet-based transformations to improve adaptability to complex noise environments. Second, the interpretability of ResNets could be enhanced by leveraging explainable machine learning approaches, thereby increasing the transparency of the model’s predictive behavior and facilitating practical deployment. Third, to address multiple fault patterns in mining equipment, the development of more robust multi-task learning models could be pursued, enabling more efficient and accurate RUL prediction. Collectively, these future directions are expected to further advance the field of mining equipment health management with broad application prospects by improving maintenance efficiency and reducing fault incidence.

## REFERENCES

- [1] Sitko, J., Farhad, Z. (2020). Analysis of mechanical equipment failure at the hard coal mine processing plant. *Acta Montanistica Slovaca*, 25(3): 350-360. <https://doi.org/10.46544/AMS.v25i3.8>
- [2] Kohani, M., Pecht, M. (2017). Malfunctions of medical devices due to electrostatic occurrences big data analysis of 10 years of the FDA's reports. *IEEE access*, 6: 5805-5811. <https://doi.org/10.1109/ACCESS.2017.2782088>
- [3] Kong, Z., Yue, C., Shi, Y., Yu, J., Xie, C., Xie, L. (2021). Entity extraction of electrical equipment malfunction text by a hybrid natural language processing algorithm. *IEEE Access*, 9: 40216-40226. <https://doi.org/10.1109/ACCESS.2021.3063354>
- [4] Ryan, A. (2017). Heat stress management in underground mines. *International Journal of Mining Science and Technology*, 27(4): 651-655. <https://doi.org/10.1016/j.ijmst.2017.05.020>
- [5] Antunović, R., Halep, A., Bučko, M., Perić, S., Vučetić, N. (2018). Vibration and temperature measurement based indicator of journal bearing malfunction. *Tehnički Vjesnik*, 25(4): 991-996. <https://doi.org/10.17559/TV-20160530090714>
- [6] Aliev, T., Babayev, T., Musaeva, N., Mammadova, A., Alibayli, E. (2022). Technologies and intelligent systems for adaptive vibration control in rail transport. *Transport Problems*, 17(3): 31-38. <https://doi.org/10.20858/tp.2022.17.3.03>
- [7] Bachschmid, N., Pennacchi, P. (2006). Faults identification and corrective actions in rotating machinery at rated speed. *Shock and Vibration*, 13(4-5): 485-503. <https://doi.org/10.1155/2006/204098>
- [8] Zou, K., Ren, D.S., Ou-Yang, Q., Li, H., Zheng, J. (2017). Using microfluidic devices to measure lifespan and cellular phenotypes in single budding yeast cells. *Journal of Visualized Experiments: JoVE*, (121): 55412. <https://doi.org/10.3791/55412>
- [9] Wang, T.C., Teng, Q.J. (2025). A Multi-scale temporal convolutional network approach for remaining useful life prediction of rolling bearings. *Precision Mechanics & Digital Fabrication*, 2(1): 31-43. <https://doi.org/10.56578/pmdf020103>
- [10] Gultekin, E., Aktaş, M. (2024). Toward proactive maintenance: A multi-tiered architecture for industrial equipment health monitoring and remaining useful life prediction. *International Journal of Software Engineering and Knowledge Engineering*, 34(12): 1831-1856. <http://doi.org/10.1142/s0218194024500396>
- [11] Aicart, J., Tallobre, L., Surrey, A., Gervasoni, B., et al. (2024). Lifespan evaluation of two 30-cell electrolyte-supported stacks for hydrogen production by high temperature electrolysis. *International Journal of Hydrogen Energy*, 60: 531-539. <https://doi.org/10.1016/j.ijhydene.2024.02.239>
- [12] Karaoğlu, U., Mbah, O., Zeeshan, Q. (2023). Applications of machine learning in aircraft maintenance. *Journal of Engineering Management and Systems Engineering*, 2(1): 76-95. <https://doi.org/10.56578/jemse020105>
- [13] Wang, T.C., Teng, Q.J., Jin, G.H. (2024). A remaining useful life prediction method for rolling bearings based on broad learning system-multi-scale temporal convolutional network. *Precision Mechanics & Digital Fabrication*, 1(3): 145-157. <https://doi.org/10.56578/pmdf010303>
- [14] Hrinchenko, H., Antonenko, N., Khomenko, V., Artiukh, S. (2024). Assessment of power equipment operational safety in the sustainable management of residual lifespan. *Economics Ecology Socium*, 8(3): 78-91. <https://doi.org/10.61954/2616-7107/2024.8.3-7>
- [15] Mihigo, I.N., Zennaro, M., Uwitonze, A., Rwigema, J., Rovai, M. (2022). On-device IoT-based predictive maintenance analytics model: Comparing tinytstm and tinymodel from edge impulse. *Sensors*, 22(14): 5174. <https://doi.org/10.3390/s22145174>
- [16] Liang, Z., Fang, Y., Cheng, H., Sun, Y., et al. (2024). Innovative transformer life assessment considering moisture and oil circulation. *Energies*, 17(2): 429. <https://doi.org/10.3390/en17020429>
- [17] Makhutov, N.A., Gadenin, M.M., Ivanov, V.V., Miodushevskii, P.V. (2016). Methodological basics of nondestructive testing, diagnostics, and monitoring of conditions for materials and engineering systems. *Inorganic Materials*, 52: 1532-1540. <https://doi.org/10.1134/S0020168516150103>
- [18] Mikalauskas, R., Volkovas, V., Uldinskas, E. (2020). Investigation of vibroactivity of the mechanical equipment aiming to reduce emitting noise. *Mechanics*, 26(1): 51-54. <https://doi.org/10.5755/j01.mech.26.1.24115>
- [19] Tang, J., Sun, X., Yan, L., Qu, Y., Wang, T., Yue, Y. (2023). Sound source localization method based time-domain signal feature using deep learning. *Applied Acoustics*, 213: 109626. <https://doi.org/10.1016/j.apacoust.2023.109626>
- [20] Sun, C., Ma, M., Zhao, Z., Tian, S., Yan, R., Chen, X. (2018). Deep transfer learning based on sparse autoencoder for remaining useful life prediction of tool in manufacturing. *IEEE Transactions on Industrial Informatics*, 15(4): 2416-2425. <https://doi.org/10.1109/TII.2018.2881543>