OXFORD CDE

Journal of Computational Design and Engineering, 2025, **12**, 1–20

DOI: 10.1093/jcde/qwaf040
Advance access publication date: 9 April 2025
Research Article

Frequency-enhanced neural networks with a hybrid spall-size estimator for bearing fault diagnosis

Mikyung Hwang, Minseok Choi and Hyunseok Oh *

Department of Mechanical and Robotics Engineering, Gwangju Institute of Science and Technology, 123 Cheomdangwagi-ro, Buk-gu, Gwangju 61005, Republic of Korea
*Correspondence: hsoh@gist.ac.kr

Abstract

In the fault diagnosis of rolling element bearings (REBs), spall size is a typical indicator of fault severity. Conventionally, spall size estimation relies on expert-knowledge-based or data-driven approaches. Expert-knowledge-based approaches require accurate assumptions about spall-excited events, making them challenging to apply in field environments. In contrast, data-driven approaches often struggle with insufficient training data and limited generalization across various operating conditions. To address this challenge, this paper proposes a frequency-enhanced neural network (FENN) with a hybrid spall-size estimator (HSSE). The proposed FENN employs both one-dimensional convolution in the time domain and Fourier convolution on the frequency magnitude, while preserving phase information in the frequency domain to enhance frequency components that are associated with spall in REBs. The novel HSSE proposed here integrates a data-driven spall-size estimator and an expert-knowledge-guided spall-size estimator to capture spall entry and exit events between rolling elements and raceways. Model validation results, which analyzed both simulation and experimental data from roller and ball bearings, demonstrate that the proposed approach provides accurate predictions of spall size, even with limited training data. Additionally, it is confirmed that the proposed model identifies the mechanical frequencies associated with spall events, providing interpretable results from raw vibration signals without requiring further processing.

Keywords: spall-size estimation, expert knowledge, deep learning, domain generalization, regression

1. Introduction

Rolling element bearings (REBs) are critical components in various industries, including manufacturing, transportation, petrochemical, and power generation (Tao *et al.*, 2007; Tyagi & Panigrahi, 2017). Despite proper maintenance, REBs are prone to spalling due to fatigue failure. Failure of a REB can result in significant damage to the system in which it is installed. Thus, accurate spall-size estimation is crucial for condition-based maintenance, as it can accurately indicate the severity of damage. However, the complex nature of spalling—affected by factors such as residual stress and material properties (Branch *et al.*, 2013; Toumi *et al.*, 2018; Gazizulin *et al.*, 2020)—poses a significant challenge to the development of accurate spall-size estimation methods.

The research field of spall-size estimation has evolved through two distinct but complementary methodological approaches, each offering unique insights into bearing fault diagnosis. The first approach, based on expert knowledge, uses various preprocessing strategies to identify the entry and exit points of spall events (Cui *et al.*, 2016; Kogan *et al.*, 2018; Larizza *et al.*, 2020; Moazen Ahmadi *et al.*, 2016; Sawalhi & Randall, 2011; Zhao *et al.*, 2013). This conventional method provides strong theoretical foundations and clear interpretability. However, its effectiveness can vary when applied to diverse field conditions where bearing wear patterns can deviate from theoretical models. The second approach emerged with the advancement of machine learning technologies. These data-driven methods train algorithms to identify patterns and predict outcomes using large datasets (Qin *et al.*, 2022; Xiyang *et al.*, 2023; Kumar *et al.*, 2024). While these approaches have demonstrated remarkable adaptability to complex patterns, their performance depends heavily on training data availability and can face challenges in generalizing across different operating conditions (Ko *et al.*, 2023; Li *et al.*, 2024; Park *et al.*, 2023, 2025; Wang *et al.*, 2023; Xu *et al.*, 2024; Zio, 2022).

Recent developments in prognostics and health management suggest that combining these approaches could leverage their complementary strengths. Several studies have successfully integrated expert knowledge with deep-learning models to enhance system performance and interpretability (Hu *et al.*, 2023; Kan *et al.*, 2024, 2024; Kim & Lee, 2023; Kim *et al.*, 2024). For instance, the machining tool-wear prediction method was enhanced by incorporating physics-guided neural networks with empirical cutting equations (Wang *et al.*, 2020). Similarly, an interpretable neural network was developed for identifying resonant frequency bands in bearing defects (Wang *et al.*, 2022). The knowledge-informed deep network (KIDN) demonstrated how fusing domain expertise with deep learning can improve fault detection accuracy across various experimental datasets (Su *et al.*, 2024). Zhou et al. further showed how combining state-space models with neural networks could enhance bearing condition monitoring under varying operating conditions (Zhou *et al.*, 2023). Nonetheless, the advances in hybrid methodologies have not yet been fully

Received: December 9, 2024. **Revised:** April 5, 2025. **Accepted:** April 5, 2025

explored for the specific challenge of spall-size estimation, where precise defect measurement remains crucial for maintenance decisions.

The primary objective of this research is to enhance the performance of REB fault diagnosis by integrating expert knowledge into a data-driven approach for estimating spall size. This research addresses two key challenges: (1) preserving phase information during frequency analysis and (2) enabling the direct integration of expert knowledge into model training. To this end, this paper introduces a new spall-size estimation model that emphasizes the frequency components linked to raceway defects in REBs.

The proposed model features a multi-Fourier convolution layer, referred to as the frequency-enhanced neural network (FENN) and a hybrid spall-size estimator (HSSE). The Fourier convolution layer combines one-dimensional (1D) convolution in the time domain and Fourier convolution in the frequency domain. HSSE merges a data-driven spall-size estimator (dd-SSE) with an expert-knowledge-guided spall-size estimator (ekg-SSE), connecting rolling element (RE) and raceway interactions. By integrating these complementary approaches, our method overcomes the limitations of both expert-knowledge-based and data-driven methods, while maintaining interpretability through its physics-based components. The proposed model provides robust spall-size estimations using limited training data by employing a raw signal input, eliminating the need for additional processing. The main contributions of this research are as follows:

(1) FENN is proposed to leverage information from the frequency domain that is not easily discernible in the time domain alone. This network extracts spectral features from vibration impulses attributed to the entry and exit of REs into the spall. The novel phase-preserving mechanism in the Fourier convolution layer enables superior capture of both magnitude and phase characteristics of spall-excited events, unlike conventional approaches that either operate purely in the time domain and lose phase information during frequency analysis.
(2) HSSE is proposed by combining an expert-knowledge-based approach with a data-driven approach. A novel algorithm is introduced to enable backpropagation through expert-knowledge formulas during model training, allowing seamless integration of domain expertise with data-driven learning. The cumulative autocorrelation and clamping mechanisms specifically address the challenge of maintaining differentiability while incorporating physics-based constraints.
(3) The proposed FENN model with HSSE demonstrates robust performance in both intra-domain prediction and domain generalization tasks. Through extensive validation using both simulation and experimental data, it is demonstrated that the phase-preserving frequency enhancement and hybrid estimation framework significantly enhances diagnostic accuracy and reduces domain-generalization error compared to conventional approaches.
(4) The proposed hybrid approach provides interpretable results by highlighting specific frequency components associated with spall events, making it more reliable for practical applications. The model identifies mechanical frequencies associated with spall events, ranging from several tens to hundreds of Hz, providing transparent insights into the fault diagnosis process without requiring additional signal processing steps.

The remainder of this paper is organized as follows. Section 2 provides an overview of existing spall-size estimation methods and neural network architectures with Fourier convolution. Section 3 describes FENN and HSSE that estimate the spall size of REBs. Sections 4 and 5 discuss two case studies examined to evaluate the effectiveness of the proposed model. Section 6 presents the conclusions of this study and offers suggestions for future research.

2. Previous Studies

2.1. Spall-size estimation

Spall-size estimation methods developed for REBs can be categorized into two approaches: expert-knowledge-based and data-driven. As shown in Table 1, expert-knowledge-based approaches can be further divided into three groups. In previous studies, spall-excited events during the entry and exit of REs were defined in different ways. One group (Cui *et al.*, 2016; Sawalhi & Randall, 2011; Zhao *et al.*, 2013) assumed two distinct events, including a low-frequency entry event and a high-frequency impact exit event. Under this assumption, for example, the spall size (l) can be estimated as

$$l = \frac{\pi f_r (D_p^2 - d^2)}{D_p f_s} T_i, \tag{1}$$

where T_i is the time-to-impact (samples), f_r is the shaft speed of the rotor, f_s is the sampling frequency (Hz), D_p is the pitch diameter, and d is the ball diameter. Another group defined spall-excited events as low-frequency entry, transient high-frequency impact, and low-frequency exit (or rattling after impact) (Kogan *et al.*, 2018; Moazen Ahmadi *et al.*, 2016). These terms describe the characteristic transitions in the vibration signal as RE interacts with the spall: the low-frequency entry corresponds to the initial contact with the spall, the high-frequency impact represents the direct collision at the spall, and the low-frequency exit indicates the REs departure from the spall. These transitions are used to determine T_i by measuring the timing of these events, which delineates the spall size. The last group defined the spall-excited events into four distinct events: initial unloading, complete loading, high-frequency impact, and partial reloading to estimate spall size (Larizza *et al.*, 2020; Larizza, Howard, Grainger *et al.*, 2020). Prior approaches based on expert knowledge have offered accurate spall-size estimation results when the assumptions about the spall-excited events exactly match the wearout mechanism of the REBs under study. However, it is sometimes challenging to predict the wearout mechanism of REBs that is to be expected in in-use field conditions. Incorrect assumptions about spall-excited events can lead to considerable deviations in spall-size estimation results. Thus, it is difficult to ensure accuracy when applying these methods in real-world settings.

Data-driven approaches do not require expert knowledge about spall-excited events. Deep-learning models can be trained with vibration signals collected from computational simulations and testbeds. For example, Kumar et al. (2024) employed an ensemble tree to predict the spall size. Xiyang *et al.* (2023) proposed a combination of the convolutional neural network (CNN) and the bidirectional long-short-term memory (BiLSTM). This method robustly predicts spall size while accounting for the changes in shaft speeds. Furthermore, the back propagation neural network (BPNN) was developed to correlate the root mean square values of

Table 1: Existing spall-size estimation methods for REBs.

Approach	Spall-excited events	Pre-processing	Entry/exit time point selection	Spall-size estimation formula or model	References
Expert-knowledge-based	Low-frequency entry and high-frequency impact	Pre-whitening and filtered MED	Max envelope Cepstrum	$l = \frac{\pi f_r (D_p^2 - d^2)}{D_p f_s} T_i$	(Sawalhi & Randall, 2011)
		EMD and ApEn	Max envelope Cepstrum		(Zhao *et al.*, 2013)
		–	Matching pursuit		(Cui *et al.*, 2016)
	Low frequency entry, high-frequency impact and low-frequency exit	Pre-whitening and low-pass filter	Max envelope	$l = \frac{D_p + d}{2} \sin(2\beta_2 + \beta_1)$	(Moazen Ahmadi *et al.*, 2016)
		Band-pass and high-pass filter	Mode value from first impact; maximum acceleration	$l = l_{disconnect} + l_{imp} + l_{TE}$	(Kogan *et al.*, 2018)
	Initial unloading, complete loading, high-frequency impact and Partial reloading	Power spectral density	Spectrogram	$l = \frac{D_p}{2}(1.5\alpha_{1b-2b} + \alpha_{2b-3b} + 1.5\alpha_{3b-4b})$	(Larizza *et al.*, 2020)
Data-driven	Not applicable	Wavelet envelope		Ensemble tree	(Kumar *et al.*, 2024)
		Matrix diagram		CNN and BiLSTM	(Xiyang *et al.*, 2023)
		RMS feature		BPNN	(Qin *et al.*, 2022)

measured vibration signals with spall size, representing the evolution law over the bearing's life (Qin *et al.*, 2022). It is evident that advanced architecture can be employed to enhance the performance of this method. However, the model's accuracy relies on the quantity and quality of the training data, limiting its generalizability under different operating conditions. In summary, both the expert-knowledge-based and data-driven approaches have strengths and limitations. Thus, it is natural to devise a hybrid approach that leverages the strengths of these two approaches while minimizing their weaknesses.

2.2. Neural network architecture with Fourier convolution

Neural network architectures often incorporate Fourier transforms to improve efficiency. Previous studies have demonstrated that the neural architecture training process can be accelerated. For example, Mathieu et al. employed fast Fourier transforms (FFT) by computing convolutions as pointwise products in the Fourier domain (Mathieu *et al.*, 2013). Furthermore, this study demonstrated that, with numerical experiments, the simple Fourier transform-based algorithm can accelerate the training and inference processes by a significant factor. A Fourier layer architecture was devised to processes features through parallel local and global paths effectively, where the global path selectively preserves lower frequency components for computational efficiency (Li *et al.*, 2020). In that study, it was demonstrated that the Fourier layer was highly efficient, as the multilayer architecture solved partial differential equations (PDEs) up to three orders of magnitude faster compared to conventional PDE solvers.

While these studies focused on computational efficiency, conventional convolution operations have demonstrated effectiveness in bearing fault diagnosis (Liang & Zhao, 2021, 2022; Liu *et al.*, 2020). The convolution operation of the l^{th} layer is defined as

$$h_i^{l_{Conv.1D}} = \sum_{j=1}^{M} w_{i,j}^{l} * v_j^{l-1} + b_i^{l_{Conv.1D}}, \quad (2)$$

where $h_i^{l_{Conv.1D}}$ is the i^{th} output vector of the convolution of the layer l; $w_{i,j}^{l}$ is the weight of the i^{th} filter at layer l; v_j^{l-1} denotes the j^{th} feature vector with M channels generated from layer l–1; $b_i^{l_{Conv.1D}}$ is the i^{th} bias vector for the convolution at the layer l; and $w_{i,j}^{l}$ and $b_i^{l_{Conv.1D}}$ are the parameters that are optimized through training. Despite the proven effectiveness of convolution operations in CNN models (Kim *et al.*, 2020), conventional convolutions in modern deep networks are limited by their local operation within the receptive field, making them ineffective for learning non-local features (Wang *et al.*, 2018). Although these limitations can be partially addressed through enlarged receptive fields (Han *et al.*, 2019) and multi-receptive fields (Li *et al.*, 2020), additional improvements have been sought through Fourier-based approaches.

Building upon these developments, fast Fourier convolution has been studied to further enhance the accuracy of neural network performance. The fast Fourier convolution architecture was developed to leverage the complementary nature of local and global information processing effectively (Chi *et al.*, 2020). This approach allows for internal information sharing between the local and global paths, enabling each path to benefit from the features extracted by the other. This architecture demonstrated significant performance improvements in computer vision tasks. More recently, Sinha et al. proposed the non-local attention-aided Fourier convolution architecture to widen the receptive field and learn long-range dependencies (Sinha *et al.*, 2022). For image super resolution tasks, this Fourier-convolution-based architecture achieved performance comparable to that of other approaches, while using fewer model parameters.

In summary, the application of Fourier transforms in existing neural architectures has demonstrated promise in accelerating model training and enhancing performance for various tasks, including image recognition, human key-point detection, and image super-resolution. For spall-size estimation in bearing fault diagnostics, prior work has reported that the frequency components between several tens and hundreds of Hz are closely associated with the spall size, when analysing the vibration impulses measured from bearings with spalls (Moazen Ahmadi *et al.*, 2016; Sawalhi & Randall, 2011). Consequently, it is anticipated that a Fourier-convolution-based method has more potential than a conventional-convolution-based method for enhancing specific frequency components in vibration impulses. However, to the best of our knowledge, the potential of a Fourier-convolution-based neural architecture has not yet been exploited for spall-size estimation of REBs.

3. Proposed Method

As stated earlier, it is desirable to leverage the benefits of both the expert-knowledge-based and data-driven approaches.

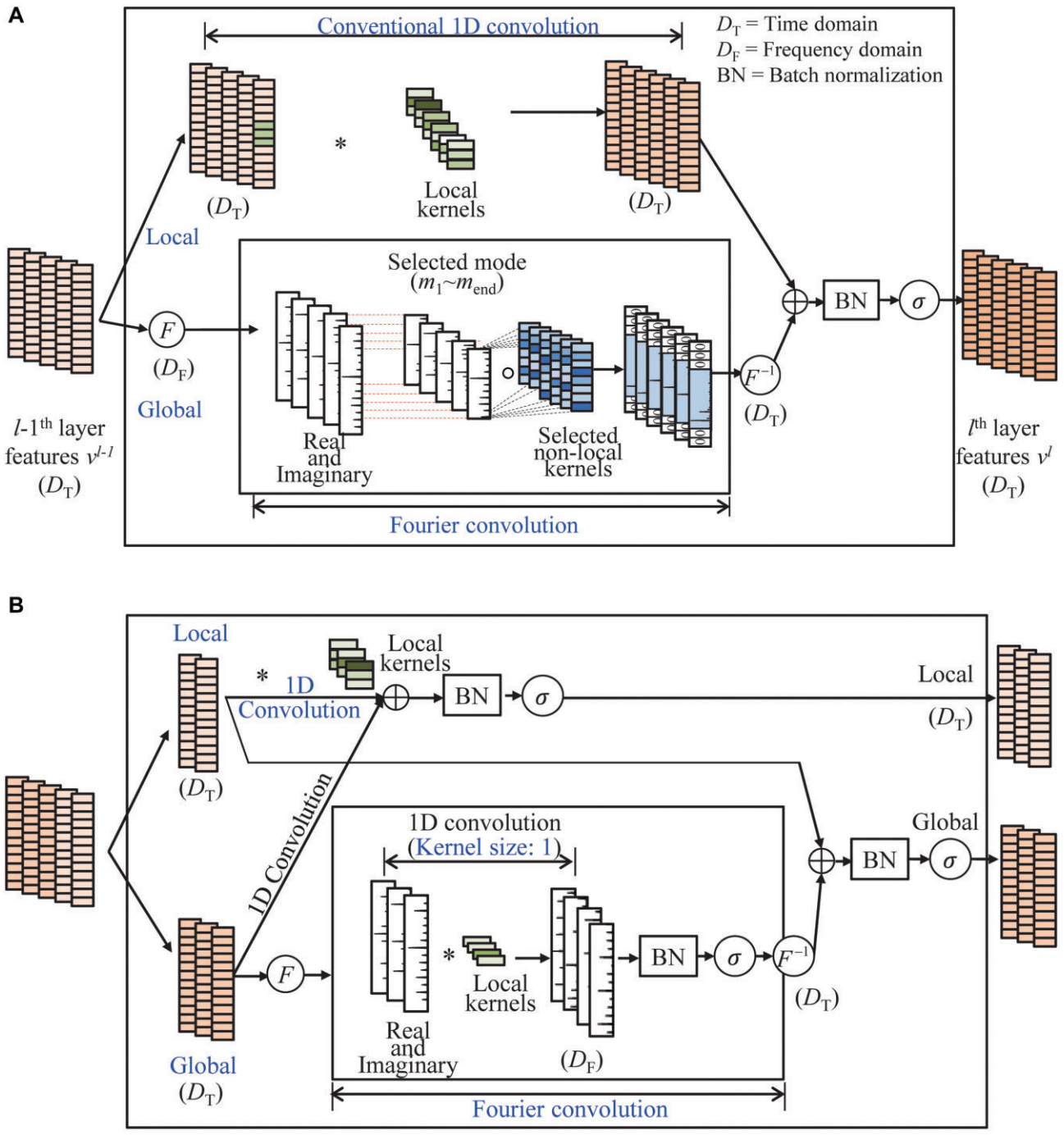

Figure 1: Existing Fourier-layer architecture. (A) Fourier layer (Li *et al.*, 2020). (B) Fast Fourier convolution layer (Chi *et al.*, 2020).

Moreover, incorporating Fourier convolution into a neural network architecture has great potential for spall-size estimation tasks. The key questions are (1) how to incorporate the expert knowledge (i.e., spall-size estimation formula) into the neural network and (2) how to design the Fourier convolution network architecture that can enhance the frequency components associated with the vibration of spall-like defects. To address these challenges, Section 3.1 proposes a modified Fourier convolution layer that can extract the spectral features associated with spall-excited events without any assumptions about the spall-excited events. Section 3.2 presents ekg-SSE. The section outlines the novel algorithm that was devised to enable backpropagation through the expert-knowledge-based formula during model training. Section 3.3 describes the proposed HSSE model for estimating the spall size of REBs.

3.1. Modified Fourier convolution layer

As depicted in Figure 1A (Li *et al.*, 2020), the local path performs conventional convolutions on a subset of the input feature channels, while the global path operates in the frequency domain. The Fourier convolution part is specially designed to selectively preserve the lower frequency components (m_1 to m_{end}) while

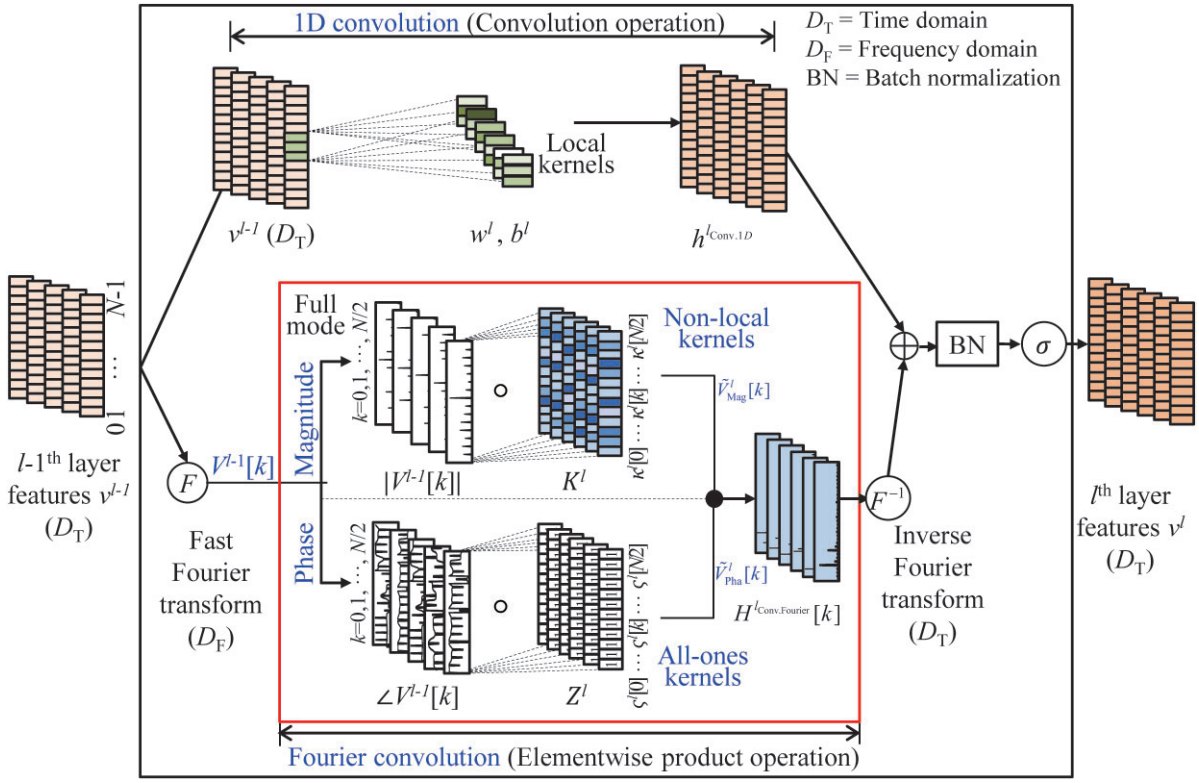

Figure 2: Modified Fourier convolution layer.

filtering out the higher components. This selective preservation enhances computational efficiency by focusing on the most significant frequency components, ensuring robust performance without compromising accuracy. As presented in Figure 1B (Chi *et al.*, 2020), the fast Fourier convolution layer also conducts conventional and Fourier convolution, where the information extracted from the two paths is internally shared, complementing one another.

Inspired by the existing Fourier neural network architecture in Chi *et al.* (2020) and Li *et al.*, (2020), the proposed frequency layer incorporates fast Fourier convolutions that have local and Fourier convolution within the convolution unit, as shown in Figure 2. In particular, the proposed Fourier convolution has a novel architecture to accommodate the characteristics of the vibration impulses observed from defective REBs with spalls. The Fourier convolution process starts with the Fourier transform of the input feature channels, providing the magnitude and phase for each channel:

$$V_j^{l-1}[k] = \sum_{n=0}^{N-1} v_j^{n-1}[n] \cdot \exp^{-j(2\pi/N)kn} = \left|V_j^{l-1}[k]\right| \cdot \exp^{j\angle V_j^{l-1}[k]},$$
$$\left|V_j^{l-1}[k]\right| = \sqrt{R\left(V_j^{l-1}[k]\right)^2 + I\left(V_j^{l-1}[k]\right)^2},$$
$$\angle V_j^{l-1}[k] = \tan^{-1}\left(I\left(V_j^{l-1}[k]\right) \Big/ R\left(V_j^{l-1}[k]\right)\right), \quad (3)$$

where $V_j^{l-1}[k]$ is the Fourier transform of $v_j^{l-1}[n]$ with an even length N as the j^{th} feature vector at the l-1 layer; k is the frequency index, which ranges from 0 to $N/2$ for a positive frequency; $|V_j^{l-1}[k]|$ is the magnitude of $V_j^{l-1}[k]$ at the frequency index k; and $\angle V_j^{l-1}[k]$ denotes the phase angle of $V_j^{l-1}[k]$.

Then, the magnitude ($|V_j^{l-1}[k]|$) of each channel's frequency index k is element-wise multiplied with the filter, κ_{ij}^l, that consists of learnable parameters with $(N/2) + 1$ length. Meanwhile, the phase ($\angle V_j^{l-1}[k]$) of the feature spectrum undergoes element-wise multiplication with a filter ς_{ij}^l where all k index values are set to one:

$$\tilde{V}^l_{i_{Mag.}}[k] = \sum_{j=1}^{M} \kappa_{ij}^l[k] \cdot \left|V_j^{l-1}[k]\right|,$$
$$\tilde{V}^l_{i_{Pha.}}[k] = \sum_{j=1}^{M} \varsigma_{ij}^l[k] \cdot \angle V_j^{l-1}[k]. \quad (4)$$

Finally, the magnitude of the feature spectrum, along with the retained phase information, is transformed back to the time domain via inverse Fourier transform:

$$h_i^{l_{Conv.Fourier}} = F^{-1}\left(\tilde{V}^l_{i_{Mag.}}[k] \cdot \exp^{j\tilde{V}^l_{i_{Pha.}}[k]}\right)$$
$$= \frac{1}{N}\sum_{k=0}^{N/2}\left(\tilde{V}^l_{i_{Mag.}}[k] \cdot \exp^{j\tilde{V}^l_{i_{Pha.}}[k]}\right) \cdot \exp^{j(2\pi/N)kn}, \quad (5)$$

where $h_i^{l_{Conv.Fourier}}$ is the i^{th} output vector found by the Fourier convolution at the Fourier layer l.

The magnitude of the complex spectrum contains information about the energy content of the signal across each frequency component. The role of the filter κ_{ij}^l is to identify the frequencies associated with the entry and exit events appearing in the signal when the RE passes through the spall. As the phase represents the relative position of each frequency component with respect to the time reference point (Oppenheim & Lim, 1981),

it is necessary to maintain the phase information of the entry and exit events. Therefore, by element-wise multiplication of ς_{ij}^{l} filters, in which all values are one, and phases, the phases of each channel of the feature spectrum can be consistently maintained.

Subsequently, the modified Fourier convolution layer outputs a feature map by summing the result vectors of the 1D convolution and the Fourier convolution in an element-wise manner. The summation of the two output feature vectors can be input into the nonlinear activation function:

$$v_i^l = \sigma\left(h_i^{l_{\text{Conv.1D}}} + h_i^{l_{\text{Conv.Fourier}}}\right), \tag{6}$$

where v_i^l is the i^{th} output feature vector that sums the results of each convolution operation at the Fourier layer l; σ is the nonlinear activation function. Pooling layers were excluded to preserve frequency information. Consequently, the modified Fourier convolution layer leverages both local convolutions through the 1D convolution and global convolutions in the Fourier convolution to extract spall-related signal characteristics.

3.2. Expert-knowledge-guided spall-size estimator

In this work, ekg-SSE is devised to effectively incorporate an empirical equation developed by domain experts. Unlike conventional methods that directly extract T_i from raw signals through signal processing techniques, the proposed approach uses the modified Fourier convolution-based neural network to learn and enhance spall-related features. The feature map output from the modified Fourier convolution layer contains spatial information regarding spall-excited events, specifically, the moments when RE passes the spall's entry (T_{entry}) and exit (T_{exit}) points. For example, as shown in the channel-averaged feature map in Figure 3A, the neural network processes vibration signals to effectively capture impact signals associated with spalling. Using the feature map, ekg-SSE determines the time difference ($T_i = T_{\text{exit}} - T_{\text{entry}}$), which correlates with the physical size of the spall as shown in Equation 1 (or another equivalent formula), given the time difference, bearing geometry, and the shaft speed.

Locating the time difference between the entry and exit moments within the feature map vector appears to be straightforward, as it merely involves indexing that exceeds a threshold within the feature map vector. However, simple indexing can impede backpropagation. To address this challenge, a novel algorithm is proposed in this section to calculate the time difference while allowing backpropagation. This algorithm incorporates cumulative autocorrelation, clamping, and an element-wise product with linear functions, which is analogous to the reparameterization trick (Kingma & Welling, 2013) that enables backpropagation. The time difference calculation module consists of eight steps, as illustrated in Figure 3A.

The first step involves computing the channel-wise average of the feature maps at the modified Fourier convolution layer.

$$x_t = \sum_{j=1}^{M} X_{j,t}^{l_F} \tag{7}$$

where x_t represents the t^{th} value of the averaged feature map vector; $X_{j,t}^{l_F}$ is the feature vector of the t^{th} value at the j^{th} channel of the layer l_F; and M is the total number of channels.

In the second step, the autocorrelation of the averaged feature map is computed, allowing the identification of the periodically repeating T_i pattern. The autocorrelation coefficient vector (R) is expressed below:

$$\begin{aligned} R &= \{r_0, r_1, \cdots, r_L\}, \\ r_t &= \sum_{t=l}^{L} (x_t - \bar{x})(x_{(t-l)} - \bar{x}) \Big/ \sum_{t=0}^{L} (x_t - \bar{x})^2, \end{aligned} \tag{8}$$

where r_0, r_1, …, r_L are the autocorrelation coefficient values; $x_{(t\text{-}l)}$ represents the $(t\text{-}l)^{\text{th}}$ value of x_t; $\bar{x}$ denotes the average value of x_t; and L is the length of x_t.

The third step is to transform the autocorrelation coefficient vector into a new vector (C) by sequentially accumulating the sum of its individual components:

$$C = \{c_0,\ c_1,\ \cdots c_L\} = \left\{\sum_{k=0}^{0} r_k,\ \sum_{k=0}^{1} r_k,\ \cdots,\ \sum_{k=0}^{L} r_k\right\}, \tag{9}$$

where c_0, c_1,…, c_L are the cumulative sum of the autocorrelation coefficient values.

The fourth step is to isolate the single T_i pattern. This is achieved by comparing the components of the new vector with a trainable threshold parameter (c_{th}). In this step, the individual components of the C vector are set to the threshold value if they are smaller than the threshold:

$$C_{\text{clamp}} = \begin{cases} c_i & \text{if } c_i \geq c_{\text{th}} \\ c_{\text{th}} & \text{if } c_i < c_{\text{th}} \end{cases}, \tag{10}$$

The initial value of the threshold parameter is set to the average of the cumulative autocorrelation values corresponding to the spall sizes of the training dataset for each shaft speed. The threshold parameter is updated during model training.

In the fifth step, the discrete second difference (S) is computed to identify the corner point in C_{clamp}.

$$S = \Delta^2 C_{\text{clamp}} = C_{\text{clamp},i+1} - 2C_{\text{clamp},i} + C_{\text{clamp},i-1}. \tag{11}$$

The sixth step is to calculate an element-wise product of the normalized S and the linearly increasing vector (I_{idx}).

$$\begin{aligned} T &= {}^{S}\!/_{\max(S)} \cdot I_{idx} = {}^{S}\!/_{\max(S)} \cdot \{1, 2, ..., L-2\} \\ &= \{T_1, T_2, \cdots, T_i, \cdots, T_{L-2}\} \end{aligned} \tag{12}$$

In the seventh step, the sample point corresponding to T_i is determined by extracting the maximum value from the feature vector T:

$$T_i = \max(T_1, T_2, \cdots, T_i, \cdots, T_{L-2}). \tag{13}$$

Finally, the spall size in ekg-SSE is calculated by inputting the estimated T_i into Equation 1. The pseudo code used to implement the steps is presented in Algorithm 1. During the backpropagation of model training, the gradient is passed through the element associated with the maximum value, ensuring that the backpropagation operation is possible.

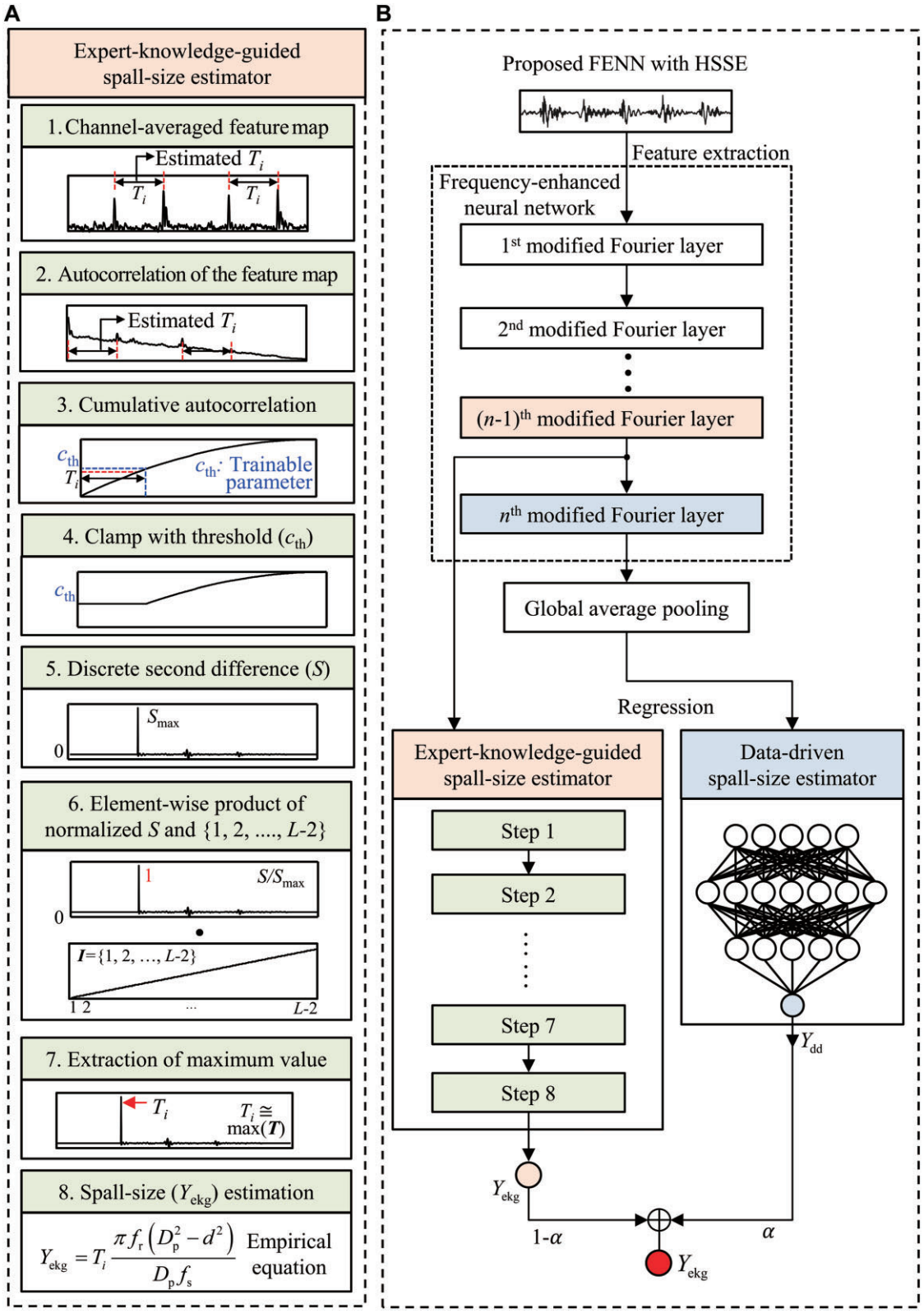

Figure 3: Proposed method. (A) Step-by-step algorithmic flow of ekg-SSE. (B) Overall architecture of the proposed FENN model with HSSE.

3.3. Proposed frequency-enhanced neural network model with the hybrid spall-size estimator

The proposed model, as shown in Figure 3B, has two parts: feature extraction and regression. The feature-extraction part consists of a sequence of modified Fourier convolution layer blocks, while the spall-size estimation part integrates ekg-SSE and dd-SSE in parallel.

The modified Fourier convolution layer is stacked to construct FENN that works as the feature extractor. The channels of the

Algorithm 1: Expert-knowledge-guided spall-size estimation

Input: Feature map ($\boldsymbol{X}=[x_{i,j,k}]_{\text{Batch}\times\text{Channel}\times\text{Length}}$), rotation speed ($\boldsymbol{RS}=[r_i]_{\text{Batch}}$), and trainable parameter of threshold ($\boldsymbol{C}_{\text{th}}=[c_{\text{th},l}]_{\#\text{of\_Rotation\_Speed}}$)

Output: Expert-knowledge-guided spall size ($\boldsymbol{Y}_{\text{ekg}}=[y_{\text{ekg},i}]_{\text{Batch}}$)

1. Initialize $\boldsymbol{C}_{\text{th}}$ as the average target threshold for each RPM based on the training dataset
2. Assign the threshold $[c_{\text{th},l}]$ corresponding to $[r_i]$ of $[x_{i,k}]$ to $\boldsymbol{C}_{\text{th}}=[c_{\text{th},i}]_{\text{Batch}}$
3. Function: ekg-SSE($\mathbf{X}$, $\boldsymbol{RS}$, $\boldsymbol{T}$)
4. Average the channels of feature map from $\boldsymbol{X}=[x_{i,j,k}]$ to $\boldsymbol{X}=[x_{i,k}]$
5. Calculate the autocorrelation of $\boldsymbol{X}$ to $\boldsymbol{R}=[r_{i,k}]_{\text{Batch}\times(\text{Length}+1)}$
6. Calculate the cumulative sum of $\boldsymbol{R}$ to $\boldsymbol{C}=[c_{i,k}]_{\text{Batch}\times(\text{Length}+1)}$
7. Clamp ($\boldsymbol{C}_{\text{clamp}}$) all elements $[c_{i,k}]$ of $\boldsymbol{C}$ into $[c_{\text{th},i}]$ if $[c_{i,k}]<[c_{\text{th},i}]$; otherwise $[c_{i,k}]$
8. Calculate the second difference($\Delta^2\boldsymbol{C}_{\text{clamp}}$) of $\boldsymbol{C}_{\text{clamp}}$ as $\boldsymbol{S}$
9. Calculate dot product of $\boldsymbol{S}_{\text{norm}}=\boldsymbol{S}/\max(\boldsymbol{S})$ and $\boldsymbol{I}=[idx_{k=1\text{ to }L\text{-}2}]_{(\text{Length-}2)}$
10. Calculate the samples of $\boldsymbol{T}_i=[t_i]_{\text{Batch}}$ from $\max(\boldsymbol{T})$
11. Calculate the $\boldsymbol{Y}_{\text{ekg}}$ by substituting $\boldsymbol{T}_i$ into the spall-size estimation formulation
12. Return the $\boldsymbol{Y}_{\text{ekg}}$

last layer from the FENN structure are compressed through the global average pooling (GAP) layer. Compressed channel information is then transmitted to dd-SSE. As previously discussed, ekg-SSE computes the spall size (Y_{ekg}) through the expert-knowledge-guided process, incorporating a trainable threshold parameter (c_{th}). On the other hand, dd-SSE employs a conventional neural network architecture that includes a GAP layer and fully connected layers to compute the spall size (Y_{dd}). Both estimators are trained simultaneously, and HSSE outputs the estimate by using the weighted sum of the two estimates derived from ekg-SSE and dd-SSE:

$$\hat{Y}_{\text{spall}} = (1-\alpha)Y_{\text{dd}} + \alpha Y_{\text{ekg}}, \tag{14}$$

where α is the weight that controls the strength of Y_{dd} and Y_{ekg}. In this study, α was set to 0.5, assigning an equal weight to the contributions of ekg-SSE and dd-SSE. HSSE is designed to complement the limitations of the data-driven and expert-knowledge-based methods.

The spall size estimated by HSSE is set to the learning target by minimizing the mean absolute error (MAE) for the training sample:

$$\text{MAE} = \frac{1}{N}\sum_{i=1}^{N}\left|\hat{Y}_{\text{spall,i}} - Y_{\text{spall,i}}\right|, \tag{15}$$

where N is the total number of samples in the training data set and $Y_{spall,i}$ is the i^{th} actual spall size.

Optimal structural parameters for the proposed HSSE model must be determined. A grid search is employed to find structural hyperparameters, including the modified Fourier convolution layers, the channel combinations, and ekg-SSE location. The search spaces for the modified Fourier convolution layers and channel combinations are determined based on findings from a previous study (Kim *et al.*, 2022) and the doubling principle of VGGNet (Simonyan & Zisserman, 2014), respectively. The layer location fed to ekg-SSE is determined through an exhaustive search within the Fourier convolution layer search spaces. Weights and biases are updated using the adaptive moment estimation (Adam) optimizer. The optimal hyperparameter set that yields the best performance is selected using a validation dataset. Finally, the model's accuracy is evaluated using a test dataset.

In this research, all computations were conducted on a computer equipped with an Intel Core i9-10980XE processor, 128 GB of RAM, and an NVIDIA GeForce RTX 3090 graphics card, using the PyTorch 1.7.1 framework.

4. Case Study 1: Cylindrical Roller Bearing

This study evaluates the proposed model using FAG N206E-TVP2 cylindrical roller bearings, as shown in Figure 4 for intra-domain prediction and domain generalization. The intra-domain prediction tests spall size estimation within simulation data, while domain generalization examines the model's transfer from simulation to experimental data. Bearing specifications are provided in Table 2.

4.1. Data description

A simulation dataset was generated using a multi-body nonlinear dynamic model (Moazen Ahmadi *et al.*, 2015), as shown in Figure 5. The parameters of the simulation model were calibrated to maximize the agreement between the computational frequency response of the experimental results in (Larizza *et al.*, 2020). The calibrated simulation model parameters are listed in Table 3. The simulation was carried out using MATLAB Simulink 2019b. The acceleration signal contains 1 024 data points

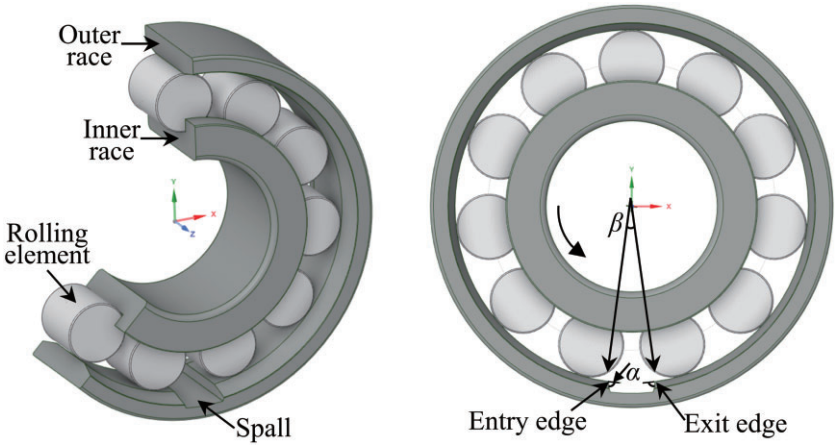

Figure 4: Rolling element bearings with a spall-like defect on the outer raceway.

Table 2: Specifications of the cylindrical roller bearing studied (FAG N206E-TVP2).

Number of rollers	Separation angle (°)	Roller diameter (mm)	Pitch diameter (mm)	Inner raceway diameter (mm)	Outer raceway diameter (mm)	Dynamic load capacity (N)
13	27.6	9.00	46.50	37.50	55.50	44 000

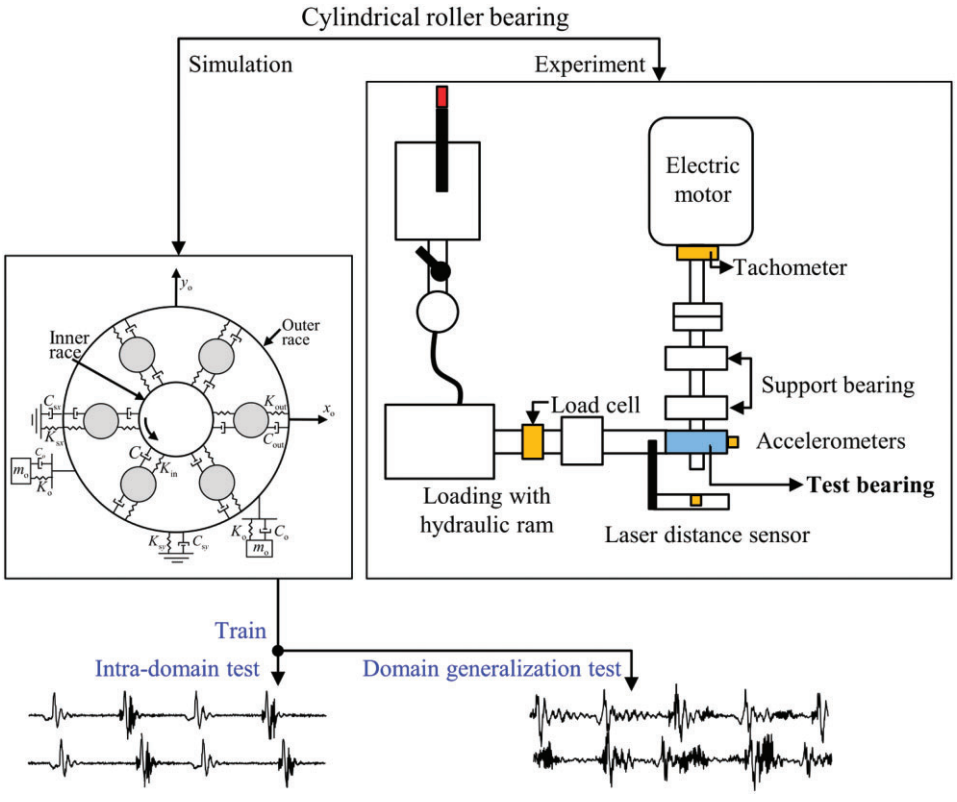

Figure 5: Data preparation including simulation model (Moazen Ahmadi *et al.*, 2015) and test rig (Larizza *et al.*, 2020).

calculated from the bearing cage frequency to capture the spall-like vibration signal as RE passes over the spall at the lowest shaft speed.

In the simulation, the spall size varied from 7.51 to 9.69 mm. The corresponding arc angles (β in Figure 4) varied between 15.5 and 20.5° in 0.5° increments. The bearing defect geometry was modeled with a spall depth of 110 μm. Data were collected with constant shaft speeds of 300, 450, 500, and 750 RPM, and constant loads of 2500, 5000, and 10 000 N. To further emulate real-world conditions, white Gaussian noise was added to the simulated signals, resulting in a signal-to-noise ratio of 20 dB (Sawalhi *et al.*, 2017).

Table 3: Simulation model parameters.

Hertzian contacts	Resonant spring system for outer raceway	Bearing inner raceway characteristics	Pedestal characteristics
$K_{in} = 249$ MN/m	$m_o = 0.5$ kg	$m_i = 0.2$ kg	$M_p = 1.2403$ kg
$K_{out} = 249$ MN/m	$K_o = 4934.8$ MN/m	$K_{sx}, K_{sy} = 1152.7$ MN/m	$K_{px} = 0.0687$ MN/m, $K_{py} = 0.0387$ MN/m
$C = 800$ N s/m	$C_o = 993.5$ N s/m	$C_{sx}, C_{sy} = 1214.7$ N s/m	$C_{px}, C_{py} = 4006$ N s/m

Variables
K: stiffness [MN/m], M: mass [kg], C: damping coefficient [N s/m]

Subscripts
in: inner raceway, out: outer raceway, o: resonant spring system for outer raceway, s_i: support at i-direction, p_i: pedestal at i-direction

The experimental dataset was obtained using a SpectraQuest test rig, as shown in Figure 5. This dataset is publicly available (Larizza, Howard, Grainger *et al.*, 2020). The spall sizes were 8.2 and 9.69 mm, with spall edge angles varying at 5°, 20°, 30°, 45°, and 90° (denoted as α in Figure 4). Further details can be found in Larizza *et al.* (2020).

4.2. Task configuration and model training

The two tasks, along with the domain, spall geometry, and dataset usage, are shown in Table 4. The intra-domain prediction task was designed to evaluate the model's prediction accuracy for different spall sizes within the training domain using simulated data. The domain-generalization task assessed the model's ability to predict spall sizes accurately across unseen domains by utilizing simulation data for training and experimental data for testing.

The results of the hyperparameter study are presented in Table 5. Within the search space, it was observed that MAE decreased as the number of layers and channels increased. It was expected that including additional layers and channels could improve the performance of the approach. However, the maximum allowed GPU memory is limited in practice. Therefore, a tradeoff between performance and computational resources was needed. The optimal layer location fed into ekg-SSE was the fifth layer, directly before the last. This placement was most effective in this case study. The neural network learns more abstract knowledge in deeper layers, making the penultimate layer ideal for interpreting a spall's mechanical properties. Consequently, the proposed model's architecture has 2^5 channels in the first layer. As the layers are stacked, the number of channels expands by 2^1, resulting in a total of six layers. The model architecture was found to be most suitable for ekg-SSE located in the fifth of these six layers. The mini-batch size, learning rate, and stop criterion were 1024, 10^{-4}, and 350 epochs, respectively.

The performance of the proposed model was tested for each condition of shaft speed and load per spall size. The number of training samples was set to 35 for the intra-domain task and 83 for the domain-generalization task. The number of test samples was set to 262, which represents the total cycles of bearing outer race fault frequency measured over 10 seconds. To account for the effect of weight initialization on model performance, prediction results were averaged across five random seeds.

4.3. Results and discussion

4.3.1. *Results*

The spall-size estimation performance is presented in Table 6. To evaluate the model's accuracy, the correct spall size was compared with the predicted spall size, absolute error, and relative error as key metrics. For the intra-domain prediction task, the model achieved an average absolute error of 0.635 mm and a relative error of 7.18% across five different spall sizes. In the domain-generalization task, the model demonstrated an average relative error of 8.57% for two spall sizes. These results indicate that the proposed model performs comparably to previously reported methods in the literature, which estimate spall size using conventional signal-processing techniques with an approximate error of 5% (Larizza *et al.*, 2020).

Table 7 compares spall-size estimation methods across signal processing, data-driven, and hybrid approaches, revealing distinct patterns in their performance characteristics. The conventional signal processing method by Chen and Kurfess (2019) showed reasonable accuracy in intra-domain prediction with a relative error of 7.42 (±7.50)%. However, its performance degraded significantly in domain generalization tasks, reaching 14.13 (±12.44)%. This limitation suggests that pure signal processing approaches may lack robustness when faced with varying operational conditions. The data-driven approach by Jeong (2023) encountered substantial challenges, particularly in the ensemble averaging process, resulting in high relative errors of 22.10 (±4.03)% for intra-domain prediction and 34.48% (±16.63)% for domain generalization. However, Kumar *et al.* (2024) demonstrated that combining data-driven techniques with signal processing could significantly reduce these errors, achieving 8.02 (±5.94)% and 9.58 (±7.19)% for intra-domain and domain generalization tasks, respectively. The proposed model, which integrates physical knowledge with deep learning, achieved the most consistent performance across both tasks. Its relative errors of 7.18 (±5.00)% for intra-domain prediction and 8.57 (±4.13)% for domain generalization demonstrate the effectiveness of incorporating domain expertise into machine learning frameworks. These results highlight how the synthesis of physical understanding and data-driven approaches can enhance both accuracy and generalization capability in bearing fault diagnosis. A noteworthy advantage of the proposed model is its time and cost efficiency, as training can be conducted using simulation data rather than requiring extensive experimental testing. Moreover, its demonstrated ability to generalize across different bearing represents an advancement for practical industrial applications, where diagnostics must remain reliable across varied equipment configurations and operating environments.

The feature visualization process of the proposed model uses raw time-domain vibration signals as input, and multiple feature map channels in each FENN layer are averaged into a single channel for visualization. The averaged feature maps are analyzed in both time and frequency domains, and the frequency analysis is inferred by transforming from the time-domain feature

Table 4: Task configuration.

Task	Domain	Spall size (mm)	Edge angle (°)	Usage
Intra-domain prediction	Simulation	7.51, 7.99, 8.48, 8.96, 9.44, 9.93	90	Training
	Simulation	7.75, 8.23, 8.72, 9.20, 9.69	90	Test
Domain generalization	Simulation	7.51, 7.99, 8.48, 8.96, 9.44, 9.93	90	Training
	Experiment	8.23	5	Test
	Experiment	9.69	20, 30, 45, 90	Test

Table 5: Hyperparameter study results of case study 1.

		Layer fed to ekg-SSE (mean absolute error)					
Number of layers	Channel combination	1	2	3	4	5	6
2	$[2^1, 2^2]$	1.21	1.08	–	–	–	–
	$[2^2, 2^3]$	1.26	1.15	–	–	–	–
	$[2^3, 2^4]$	1.25	1.05	–	–	–	–
	$[2^4, 2^5]$	1.28	1.16	–	–	–	–
	$[2^5, 2^6]$	1.29	1.27	–	–	–	–
4	$[2^1, 2^2, 2^3, 2^4]$	1.21	1.10	1.08	1.07	–	–
	$[2^2, 2^3, 2^4, 2^5]$	1.18	1.04	1.00	1.02	–	–
	$[2^3, 2^4, 2^5, 2^6]$	1.27	1.15	1.12	1.12	–	–
	$[2^4, 2^5, 2^6, 2^7]$	1.09	0.99	0.96	0.87	–	–
	$[2^5, 2^6, 2^7, 2^8]$	1.11	1.10	1.09	1.00	–	–
6	$[2^1, 2^2, 2^3, 2^4, 2^5, 2^6]$	1.20	1.07	1.05	1.03	1.05	1.04
	$[2^2, 2^3, 2^4, 2^5, 2^6, 2^7]$	1.11	0.98	0.99	0.99	0.94	0.90
	$[2^3, 2^4, 2^5, 2^6, 2^7, 2^8]$	1.11	1.03	0.94	0.93	0.92	0.88
	$[2^4, 2^5, 2^6, 2^7, 2^8, 2^9]$	1.22	1.24	1.27	1.26	1.27	1.23
	$[2^5, 2^6, 2^7, 2^8, 2^9, 2^{10}]$	1.07	1.10	1.10	1.08	**0.78**	0.79

Table 6: Performance evaluation of the proposed model.

Task	Answer spall size (mm)	Predicted spall size (mm)	Absolute error (mm)	Relative error (%)
Intra-domain prediction	7.75	7.35	0.619	7.98
	8.23	8.23	0.316	3.84
	8.72	8.40	0.476	5.46
	9.20	8.68	0.725	7.88
	9.69	8.90	1.039	10.72
	Average		0.635	7.18
Domain generalization	8.23	8.64	0.685	8.33
	9.69	9.37	0.836	8.63
	Average		0.859	8.57

maps, assuming a one-second period. The objective was to ascertain whether the frequencies identified in the feature map were associated with the spall entry and exit events of REs. In previous studies (Larizza *et al.*, 2020), it was reported that frequencies of 200 and 1 050 Hz were linked to the spall entry and exit events, respectively. Feature map visualizations were conducted for all experimental conditions listed in Table 6. As a representative example, the results obtained at 5 000 N and 600 RPM, intermediate values of the dynamic conditions of load and shaft speed, are shown in Figure 6.

In the intra-domain prediction task, from the time-domain visualization on the left side of Figure 6A, it is evident that as the layers deepen from the first to the fifth, the spike components associated with spall entry and exit events become progressively more pronounced. This indicates that the feature map from the fifth layer is well-suited for input to ekg-SSE. In contrast, at the sixth and final layer, the previously emphasized components in the time domain are diminished. The frequency-domain visualization on the right side of Figure 6A clarifies this behaviour, showing that the network effectively suppresses non-spall-related frequency components while enhancing spall-related signals, such as the 200 Hz component linked to spall entry. This selective feature extraction ensures that the final-layer feature map retains only the most relevant spectral components. Therefore, it is appropriate to use this feature map as input to dd-SSE, demonstrating the effectiveness of the modified Fourier convolution layers in extracting spall-related features.

In the domain-generalization task, the visual inspection results were similar to those of the intra-domain prediction task. An observation is that, in the feature map of the 6th layer, the frequency components at 200 Hz showed a prominent amplitude, as shown in Figure 6B. Furthermore, the frequency range from 925 to 1 075 Hz was identified in previous research (Larizza *et al.*, 2020) as significant for capturing spall exit events, although this range appears with much lower amplitude than the 200 Hz component in our spectrum analysis. This differential prominence between entry and exit frequency components is consistent with previous observations. This suggests that the proposed model has an enhanced ability to broadly predict spall characteristics.

The stability of the proposed model was evaluated by an ablation study, as presented in Table 8. For example, in the domain-generalization task, the prediction accuracy and standard deviation increased from 72.68% (Hybrid CNN) to 91.66% (Proposed model) and from 20.57% to 4.21%, respectively, when the modified Fourier convolution layers were adopted. As another example, the prediction accuracy and standard deviation increased from 78.81% (FENN) to 91.66% (Proposed model) and from 11.28% to 4.21%, respectively, when ekg-SSE was adopted. Thus, it was corroborated that the proposed model is accurate and robust in the intra-domain prediction and domain-generalization tasks.

4.3.2. *Discussion on different training sample sizes*

The performance of the proposed model was examined across different amounts of training data. The relative error of the spall-size predictions is described in Figure 7. As expected, the relative error for both tasks decreased as the amount of training data increased. The training data intervals are scaled logarithmically. The error bar in the figure indicates the variation between different shaft speeds and different loading conditions.

For the intra-domain prediction task, as presented in Figure 7A, the relative error results of the proposed model quickly converged. This result suggests that spall-size estimation was possible with a small number of training data, when the training and test data were sampled from identical domains (i.e., simulation model). Conversely, in the case of the domain-generalization task, the increased number of training data is required for convergence in the relative error, as shown in Figure 7B, when the training and

Table 7: Comparison of spall size estimation between existing and proposed models.

Method	Signal processing (Chen and Kurfess, 2019)	Data-driven (Jeong, 2023)	Data-driven with signal processing (Kumar *et al.*, 2024)	Proposed FENN model with HSSE
Intra-domain prediction	7.42 (±7.50)%	22.10 (±4.03)%	8.02 (±5.94)%	7.18 (±5.00)%
Domain generalization	14.13 (±12.44)%	34.48 (±16.63)%	9.58 (±7.19)%	8.57 (±4.13)%

test data were sampled from different domains (i.e., simulation model and experimental using a test rig). It is worth noting that, in the case of a dearth of real data for AI model training, the proposed model can provide accurate prediction results of spall sizes in bearings of a real test rig, when the calibrated simulation model can precisely emulate the physical behavior of real bearings.

4.3.3. *Discussion on performance evaluation under dynamic conditions*

The performance of the proposed model was evaluated for predicting spall size under dynamic conditions of load and shaft speed. As a representative example, the spall size of 8.23 mm was selected. The proposed approach demonstrated superior accuracy in interpolation tasks, even with a training sample size as small as 35 samples. Figure 8A shows that the estimated spall size increases with increasing load. The estimate remains close to the true value, with a maximum relative error of 2.49%. Figure 8B presents that the estimated spall sizes also increase as the shaft speed increases. Across all four shaft-speed conditions, the model maintained predictions within a 2.85% relative error of the true spall size, which is considered negligible. Previous research (Mufazzal *et al.*, 2023) has demonstrated that increasing the radial load yields stronger characteristic signals in vibration impulses. Moreover, it was reported that, at low shaft speeds, spall entry-exit impact peaks are sharp, while at higher shaft speeds, the influence on the exit impulse diminishes. The results described above, as observed in this study, are consistent with the findings from previous research. Thus, it was corroborated that the proposed model estimated the mean of the actual defect size with negligible deviations, suggesting that the proposed model is accurate in spall-size estimation across various dynamic operating conditions.

5. Case Study 2: Deep-Groove Ball Bearing

This study examines the proposed model's extrapolation capabilities with SKF 6205–2RS JEM deep-groove ball bearings. It tests the model's prediction accuracy beyond its training range, reflecting scenarios where bearing damage progresses over time. The specifications of the ball bearing studied in this research are shown in Table 9.

5.1. Data description

An experimental dataset was generated using a testbed in the Smart Diagnosis and Design Optimization (SDDO) Laboratory at the Gwangju Institute of Science and Technology (GIST), Korea. As shown in Figure 9, the rotational shaft is supported by two bearings. One bearing included a spall defect. Based on the previous study (Chen & Kurfess, 2019; Zhang *et al.*, 2021), the axial length of the spall was determined to be 5 mm. The width of the spall varied from 0.6 to 1.6 mm, with an interval of 0.1 mm. The spall depth was 100 μm (Sawalhi & Randall, 2011). The defect was precisely fabricated by electrical discharge machining.

The defective bearing was installed on the right-hand side of the testbed. An accelerometer whose sensing direction was the y-axis was mounted on the bearing housing. The sampling frequency was 25.6 kHz. Vibration data were collected at the shaft speeds of 450, 600, and 750 RPM.

5.2. Model training

The input data length was determined based on the bearing's cage frequency at three shaft speeds. The minimum required data points were calculated to be 952 at the lowest speed. Based on this calculation, the input length was set to 1024 data points, a value reasonable to accommodate signal properties across different speeds. The input signal underwent preprocessing, including division without overlap and standardization to zero mean and unit variance. The number of training samples was determined by multiplying the ten-second data acquisition time by the bearing outer-race fault frequencies, resulting in 269 samples per shaft speed. The test dataset size was set to one-third of the training samples.

A study was conducted to determine a set of relevant hyperparameters. The datasets used for the hyperparameter optimization task are shown in Table 10. During the training process, the model with MAE loss function was selected from combinations of hyperparameters, including the number of layers, channel combinations, and ekg-SSE locations, as detailed in Table 11. The batch size, learning rate, and number of epochs were set to 1024, 10^{-4}, and 250, respectively. To account for the influence of weight initialization, a grid search approach was used to determine the model architecture through three random seed tests. The optimal configuration was achieved with six layers, channel combinations from 2^5 to 2^{10}, and the ekg-SSE layer positioned at the fifth layer. This setup resulted in an MAE of 0.108.

The performance of the proposed model was evaluated using the dataset from the extrapolation estimation task, as shown in Table 10. To ensure robustness, the evaluation results were averaged across five unique random seeds.

5.3. Results and discussion

5.3.1. *Results*

The proposed model predicted the spall size with an average relative error of 23.37% when evaluated using the dataset used in the hyperparameter optimization task. This result can also be considered as the performance in the interpolation task.

The extrapolation performance of the HSSE model is summarized in Table 12, where the predicted average spall size is compared with the true spall size and the relative error is analyzed at each rotational speed. The average relative error increased as the spall size deviated from the training domain. The smallest error (22.77%) was observed at 1.2 mm, which was closest to the training domain. A comparison with the average interpolation performance showed similar spall-size estimations within a 1% error margin. For the largest tested spall size of 1.6 mm, which

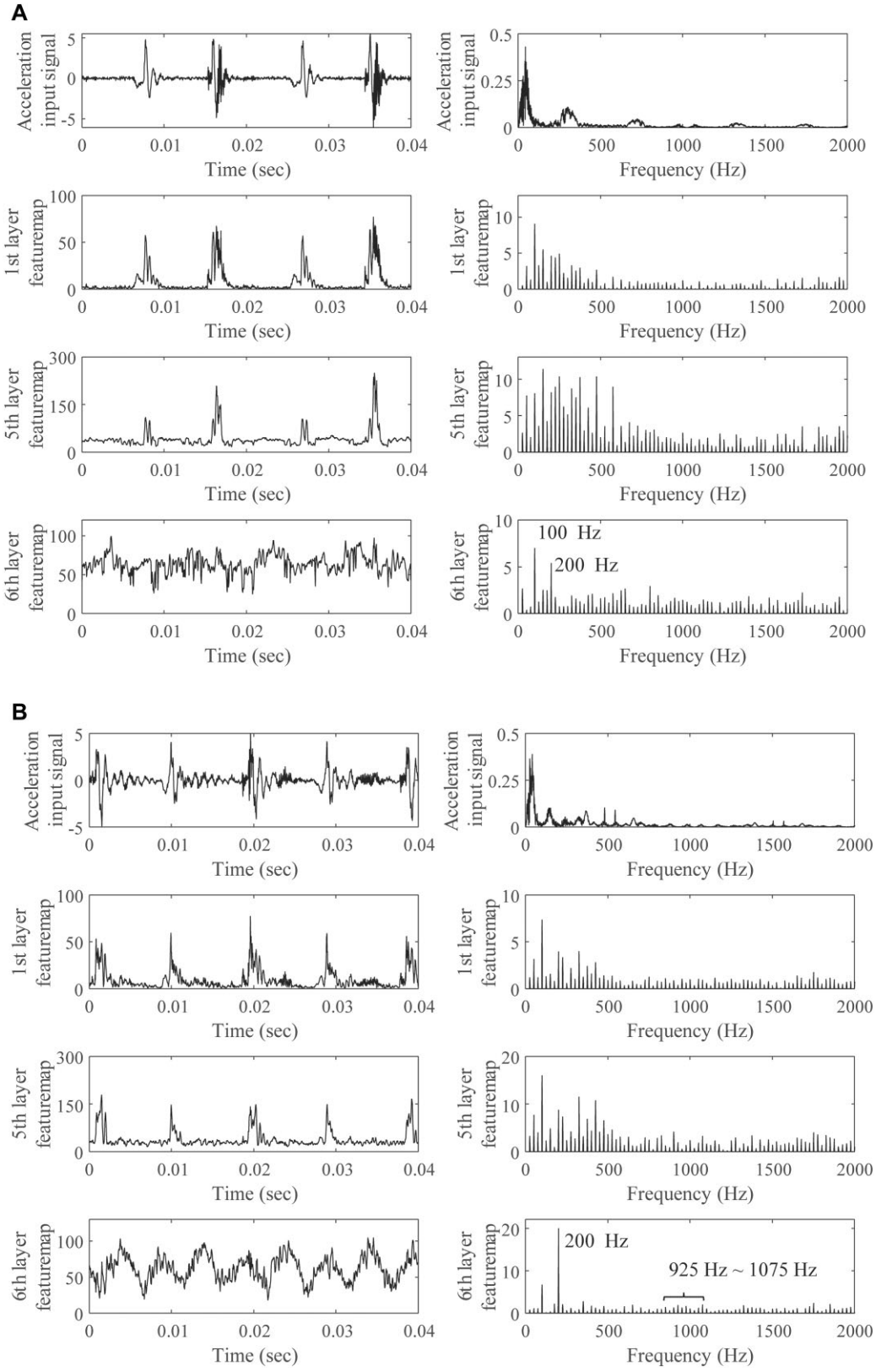

Figure 6: Representative example of feature map visualization of modified Fourier convolution layers. (A) Intra-domain prediction task. (B) Domain-generalization task.

Table 8: Ablation study: spall-size prediction accuracy (%).

		Conventional CNN	Hybrid CNN	FENN	Proposed FENN model with HSSE
	Including Fourier convolution layers	No	No	Yes	Yes
Task	Including ekg-SSE	No	Yes	No	Yes
Intra-domain prediction		53.54 (±44.01)	83.03 (±20.26)	88.33 (±8.69)	92.82 (±5.88)
Domain generalization		17.96 (±35.97)	72.68 (±20.57)	78.81 (±11.28)	91.66 (±4.21)

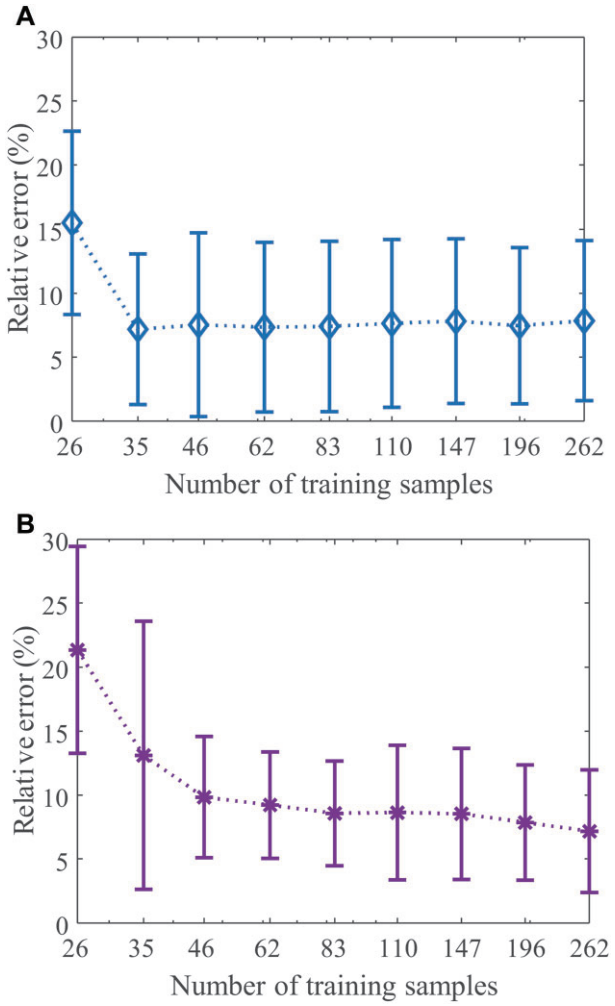

Figure 7: Relative error of the average spall-size estimation with respect to the number of training data. (A) Intra-domain prediction task. (B) Domain-generalization task.

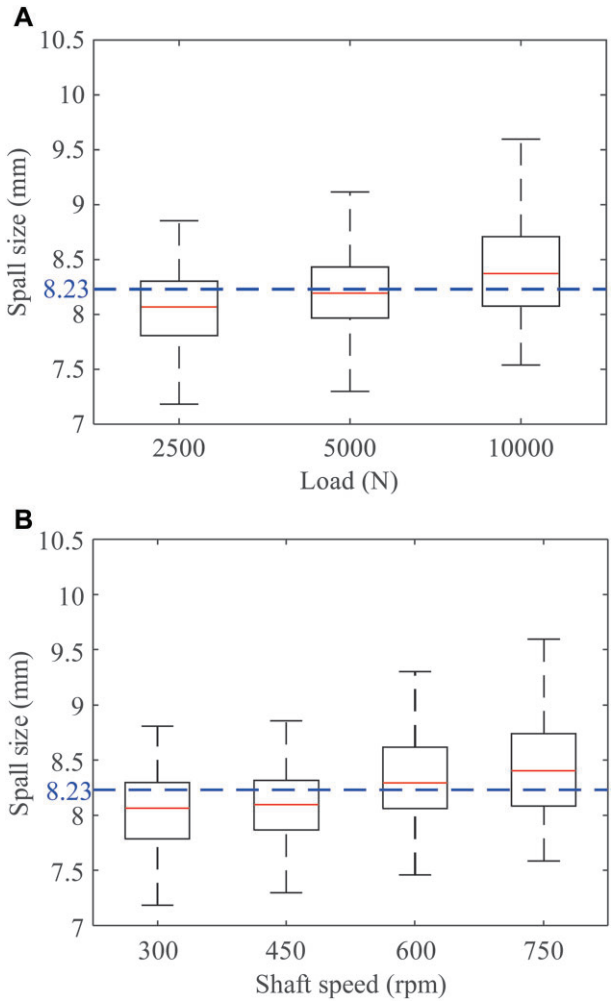

Figure 8: A representative example of spall-size estimation results. (A) Varying load conditions. (B) Varying shaft-speed conditions.

represented the furthest extrapolation from the training domain, the model showed the highest relative error of 37.21%.

However, this prediction pattern showed inconsistent deviations; notably, the 1.4 mm spall size exhibited a higher error (29.98%) than the 1.5 mm size (27.11%), despite being closer to the training domain. Previous studies have indicated that spall depth can affect the contact conditions and vibration response of REs, potentially leading to size-estimation errors (Mufazzal *et al.*, 2023). This study also found that deeper spalls can lead to contact loss with the raceway, resulting in stronger impact forces that may distort spall entry- and exit-based size estimation. Consequently, further analysis was conducted to identify whether the observed prediction error was due to machining effects of spall size or spall depth influence; this analysis is presented in Section 5.3.2.

5.3.2 *Model performance assessment based on preciously measured spall size and depth*

The spall-size estimation results were further analyzed by measuring the actual machined dimensions, including spall size and depth. Spall dimensions were measured using Bruker's DektakXT and Micro-vu's EXCEL 501UC instruments, providing detailed surface and three-dimensional measurements. As specified in Table 13, the eleven spall widths, representing spall sizes, were manufactured with an average error of 1.17% (0.01 mm). For the

Table 9: Specifications of the deep-groove ball bearing studied (SKF 6205–2RS JEM).

Number of balls	Separation angle (°)	Ball diameter (mm)	Pitch diameter (mm)	Inner raceway diameter (mm)	Outer raceway diameter (mm)	Dynamic load capacity (N)
9	40	7.938	39.04	25.00	52.00	14 000

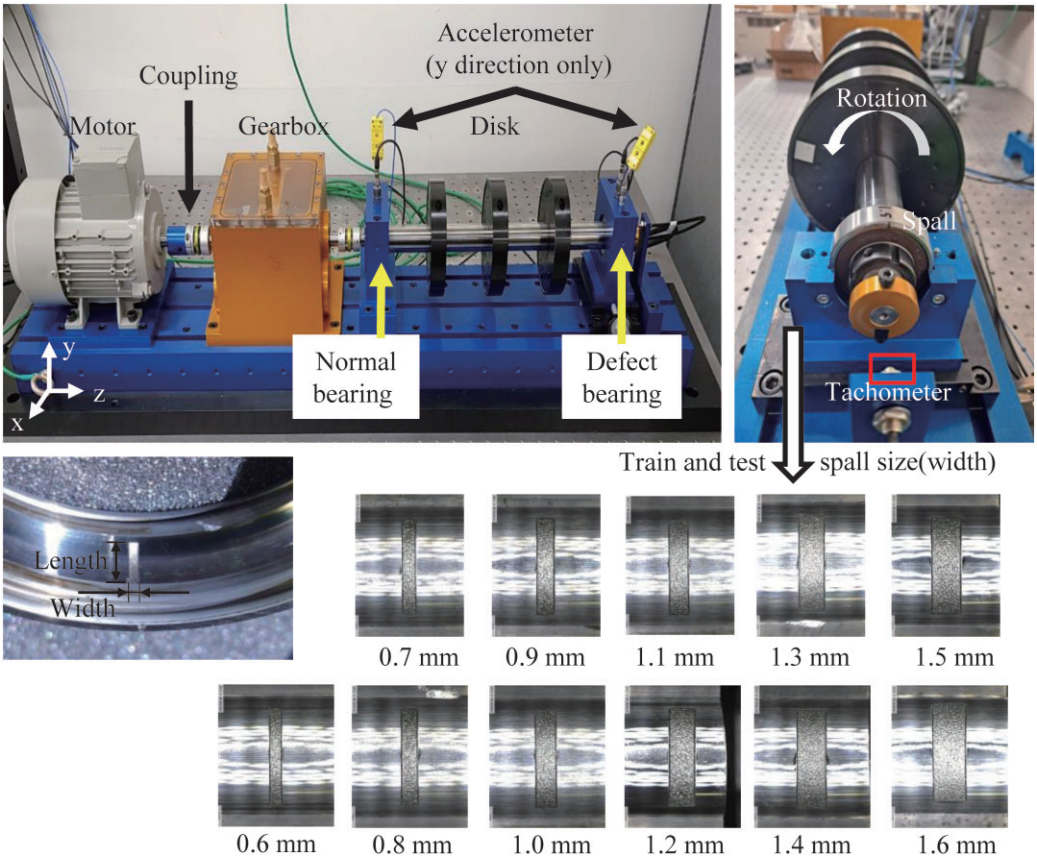

Figure 9: Rotor testbed and ball bearings with different defect sizes.

Table 10: Bearing datasets with different spall sizes.

Task	Purpose	Spall size (mm)					
Hyperparameter optimization	Training	0.6	0.8	1.0	1.2	1.4	1.6
	Validation	0.7	0.9	1.1	1.3	1.5	
Extrapolation	Training	0.6	0.7	0.8	0.9	1.0	1.1
	Test	1.2	1.3	1.4	1.5	1.6	

spall sizes of interest, 1.4 and 1.5 mm, the manufacturing error was evaluated to be within 1%, indicating a reliable level of production accuracy. For the 1.6 mm spall size, which represented the furthest extrapolation test point, the manufacturing error was similarly well-controlled at 1.25%, indicating that the increased prediction error at this size was not attributable to manufacturing variations.

To further investigate the cause of the higher error in the 1.4 mm spall size, a detailed analysis of the defect depth was conducted using 3D laser profilometry. This analysis aimed to determine whether the depth characteristics of the spall could contribute to the observed prediction errors, as the manufacturing process had already been confirmed to be accurate (Table 13). The three-dimensional profile of the processed spall depth according to the spall size is shown in Figure 10, which illustrates the depth characteristics through high-resolution surface measurements. For a comprehensive understanding, Figures 10A–10D present the surface topography across different spall widths (0.6, 1.3, 1.4, and 1.5 mm), while Figure 10E defines the key surface profile parameters used for quantitative analysis. The depth variations are quantified using three critical parameters defined in Figure 10E and measured in Table 14: R_p (Maximum profile peak height), R_v (Maximum profile valley depth), and R_z (Maximum height of profile, calculated as the difference between R_p and R_v).

Table 11: Hyperparameter study results of case study 2.

Number of layers	Channel combination	Location of ekg-SSE (mean absolute error) 1	2	3	4	5	6
2	$[2^1, 2^2]$	0.144	0.138	–	–	–	–
	$[2^2, 2^3]$	0.168	0.136	–	–	–	–
	$[2^3, 2^4]$	0.167	0.195	–	–	–	–
	$[2^4, 2^5]$	0.158	0.143	–	–	–	–
	$[2^5, 2^6]$	0.153	0.142	–	–	–	–
4	$[2^1, 2^2, 2^3, 2^4]$	0.143	0.132	0.142	0.131	–	–
	$[2^2, 2^3, 2^4, 2^5]$	0.158	0.140	0.142	0.138	–	–
	$[2^3, 2^4, 2^5, 2^6]$	0.151	0.137	0.124	0.125	–	–
	$[2^4, 2^5, 2^6, 2^7]$	0.148	0.149	0.148	0.161	–	–
	$[2^5, 2^6, 2^7, 2^8]$	0.145	0.136	0.128	0.152	–	–
6	$[2^1, 2^2, 2^3, 2^4, 2^5, 2^6]$	0.185	0.164	0.174	0.155	0.143	0.145
	$[2^2, 2^3, 2^4, 2^5, 2^6, 2^7]$	0.203	0.174	0.166	0.160	0.175	0.174
	$[2^3, 2^4, 2^5, 2^6, 2^7, 2^8]$	0.203	0.179	0.172	0.178	0.185	0.194
	$[2^4, 2^5, 2^6, 2^7, 2^8, 2^9]$	0.195	0.199	0.189	0.171	0.210	0.218
	$[2^5, 2^6, 2^7, 2^8, 2^9, 2^{10}]$	0.119	0.108	0.111	0.112	**0.108**	0.112

Unlike when the spall width is 0.6 mm in Figure 10A, a difference in depth flatness can be visually confirmed as the spall width increases from 1.3 to 1.5 mm, as shown in Figures 10B–D. In particular, when quantifying the depth roughness in the case of the 1.4 mm spall width in Figure 10C, it was confirmed that R_z was 155.0 μm, as shown in Table 14. On the other hand, the depths R_z of the 1.3 and 1.5 mm spall sizes were 122.5 and 124.2 μm, respectively. It was confirmed that the R_z of the 1.4 mm spall had a greater height deviation than the R_z of the 1.3 and 1.5 mm spall sizes by approximately 20.9 and 19.6%, respectively. Therefore, it was found that the 1.4 mm spall size causes a difference in the peak times of the entry and exit events of the measurement signal due to the depth deviation when the RE passes through the raceway, leading to an error deviation in the model prediction. For the 1.6 mm spall size, despite having well-controlled manufacturing precision (1.25% error) and relatively uniform depth characteristics ($R_z = 110.5$ μm), the model showed reduced prediction accuracy. This observation, combined with the previous results, suggests that while the proposed model maintains reliable predictions for spall sizes near its training domain, its performance degrades when estimating sizes that substantially exceed its training experience, even under ideal manufacturing conditions.

5.3.3. *Discussion on frequency enhancement*

The proposed model was analyzed by comparing feature maps with conventional CNN, Hybrid CNN, and FENN models. A spall size of 1.2 mm at 750 RPM was selected as a representative example, as shown in Figure 11. The conventional CNN model's feature maps (Figure 11A) show spall-like impulses at 0.01 and 0.03 s, with time-domain patterns differing only in amplitude between the first and last layers. No specific frequency components were emphasized in its frequency-domain feature maps. The hybrid CNN model (Figure 11B) showed similar limitations, with a decrease in feature amplification at higher frequencies. The FENN model (Figure 11C) demonstrated reduced amplification in its last layer's time-domain feature map, while showing a slight deviation from the monotonic decrease pattern in the 2.4–3.8 kHz range of the corresponding frequency-domain feature map. This difference arises because Figures 11A and B illustrate the spectral bias tendency of AI networks. In these networks, learning of lower frequencies is prioritized leading to a continuous decline in emphasis as frequency increases. While the difference is subtle, the FENN model shows less pronounced spectral bias in this frequency range, suggesting a potential preservation of higher-frequency features that may be relevant to fault diagnosis. The proposed model (Figure 11D) exhibited time-domain feature patterns similar to the FENN model, with its last layer spectrum emphasizing frequencies at 500 Hz, 1.2 kHz, and 2.4 kHz. While the FENN model exhibits moderate frequency component amplification related to spall-like defect characteristics, the proposed model, which integrates both FENN and HSSE, demonstrates enhanced capability in capturing spectral features associated with spall-like defects through more pronounced frequency-component amplification.

6. Conclusions

The proposed HSSE model integrates the interpretability of expert-knowledge-based methods with the adaptability of data-driven approaches. The HSSE model consists of two main modules: (1) a feature extractor, FENN, and (2) a hybrid spall-size estimator, HSSE. The feature-extraction step combines 1D convolution in the time domain with both phase-preservation and frequency-magnitude Fourier convolution in the frequency domain, enhancing local and global feature maps to extract spall-related features. In the HSSE step, the spall size is estimated in a data-driven manner using the output of the hidden layers from the FENN network. Additionally, an expert-knowledge-guided estimator aligns the feature map output with the physical principles of spall-size estimation, enabling end-to-end learning through backpropagation.

The proposed FENN model, integrated with HSSE, exhibited effective spall size prediction performance in two case studies. The first case study analyzed the relative errors for different spall sizes using both simulated and experimental signals. For intra-domain prediction tasks, the model achieved an average relative error of 7.18%. For domain-generalization tasks, it maintained accuracy with an 8.57% relative error when applying simulation-trained models to experimental data. The proposed model's feature maps also revealed prominent spall-related entry and exit

Table 12: Estimated relative error in the extrapolation task.

Spall size (mm)	Predicted spall size (mm)	Relative error (%) 450 RPM	600 RPM	750 RPM	Average
1.2	1.01 ± 0.09	27.87 ± 8.72	25.11 ± 9.26	15.32 ± 7.78	22.77 ± 10.15
1.3	1.03 ± 0.07	24.02 ± 6.71	22.25 ± 6.09	23.47 ± 6.27	23.24 ± 6.40
1.4	1.01 ± 0.06	29.77 ± 4.93	31.11 ± 5.31	29.06 ± 4.80	29.98 ± 5.09
1.5	1.36 ± 0.11	27.84 ± 7.13	29.49 ± 5.83	24.00 ± 6.33	27.11 ± 6.85
1.6	1.40 ± 0.11	37.53 ± 6.38	35.32 ± 6.72	38.78 ± 5.50	37.21 ± 6.38

Table 13: Bearing samples' defect dimensions, measured by 3D laser profilometry.

	Spall size (mm)										
Dimension	**0.60**	**0.70**	**0.80**	**0.90**	**1.00**	**1.10**	**1.20**	**1.30**	**1.40**	**1.50**	**1.60**
Width (mm)	0.59	0.70	0.80	0.91	1.01	1.10	1.25	1.33	1.41	1.51	1.58
Measurement error (%)	1.67	0.00	0.00	1.11	1.00	0.00	4.17	2.31	0.71	0.67	1.25

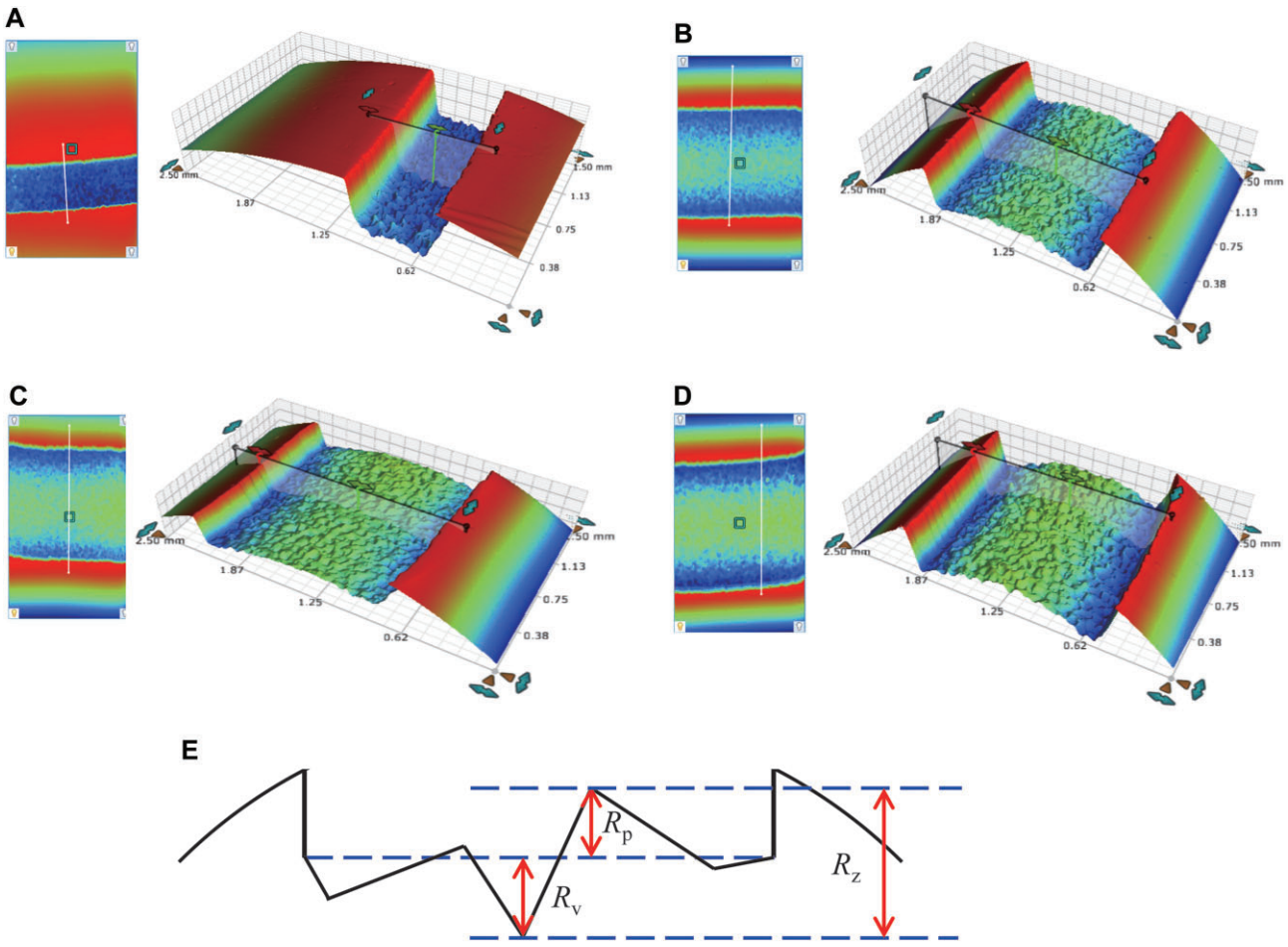

Figure 10: Depth profile of spall-like defective deep-groove ball bearings. (A) 0.6 mm spall size. (B) 1.3 mm spall size. (C) 1.4 mm spall size. (D) 1.5 mm spall size. (E) Depth profile parameters.

Table 14: Depth surface profile parameters of three bearing spall sizes.

	Spall size (mm)				
Parameters	**1.2**	**1.3**	**1.4**	**1.5**	**1.6**
R_p (Maximum profile peak height) (μm)	65.7	60.0	69.8	67.6	54.6
R_v (Maximum profile valley depth) (μm)	−66.1	−54.5	−85.1	−56.5	−55.8
R_z (Maximum height of profile) (μm)	131.7	122.5	155.0	124.2	110.5

frequency components at 200 Hz and in the frequency band of 925 to 1075 Hz, consistent with the existing signal processing literature. In the second case study, the extrapolation task demonstrated the model's ability to estimate spall sizes beyond the training range, with an average error of 22.77%, which is similar to its interpolation performance. The feature map highlighted spall-related specific frequencies as it passed through the spall. This indicates that the HSSE model provides interpretable results through its hybrid architecture.

The experimental design incorporated several real-world complexities: dynamic operating conditions with varying speeds, multiple fault severities, and environmental variations such as measurement noise. Under these conditions, the proposed model maintained robust and consistent performance across different operating scenarios. However, industrial applications may present additional challenges not fully captured in our current validation, such as varying load distributions, complex system interactions, and more severe environmental disturbances. These limitations suggest the need for further validation in actual industrial settings where bearing operations face more diverse operational conditions and environmental factors.

Future research directions include testing with diverse datasets to improve applicability to real-world spalls, which often differ

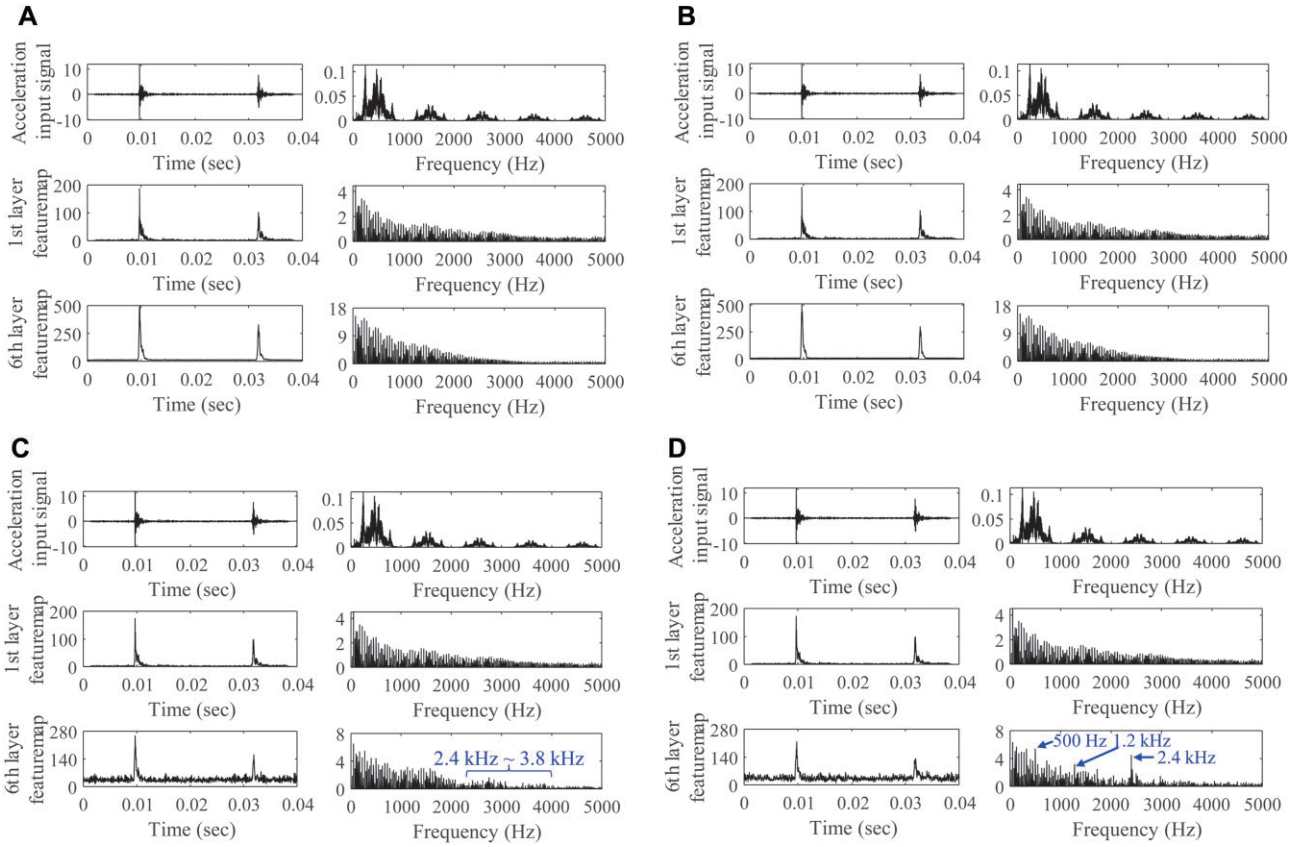

Figure 11: SDDO testbed feature-map analysis (first layer and the last layer feature map from the input signal) for each model in the extrapolation task (1.20 mm target) with an operating condition of 750 RPM. (A) Conventional CNN. (B) Hybrid CNN. (C) FENN. (D) Proposed FENN model with HSSE.

from artificially induced defects in terms of shape, depth, roughness, and irregular wear patterns. Additionally, future research will focus on enhancing the model's extrapolation capabilities under noisy industrial environments through advanced noise-robust feature extraction methods and transfer learning techniques. Future research will also examine whether the observed relationship between network layer depth and specific mechanical frequency capture is consistent across different bearing configurations. To further enhance model accuracy and generalization, efforts will focus on incorporating applied load variations and extended simulation-based testing, ultimately validating the robustness and effectiveness of the model for field applications under realistic conditions. Additionally, validations in industrial environments will be crucial to assess the model's performance under more complex operating conditions, including varying load distributions, system interactions, and environmental factors. These improvements will significantly enhance the accuracy and applicability of the proposed spall-size estimation method in real-world scenarios.

Conflicts of Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Author Contributions

Mikyung Hwang: Conceptualization, Formal analysis, Investigation, Methodology, Software, Validation, Visualization, Writing-original draft. **Minseok Choi**: Data curation, Investigation, Software, Resources. **Hyunseok Oh**: Funding acquisition, Project administration, Supervision, Writing-review & editing.

Funding

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. RS-2025-00517566, No. RS-2022-00144441).

Data Availability

Data available on request.

References

Branch, N. A., Arakere, N. K., Forster, N., & Svendsen, V. (2013). Critical stresses and strains at the spall edge of a case hardened bearing due to ball impact. *International Journal of Fatigue*, **47**, 268–278. https://doi.org/10.1016/j.ijfatigue.2012.09.008.

Chen, A., & Kurfess, T. R. (2019). Signal processing techniques for rolling element bearing spall size estimation. *Mechanical Systems and Signal Processing*, **117**, 16–32. https://doi.org/10.1016/j.ymssp.2018.03.006.

Chi, L., Jiang, B., & Mu, Y. (2020). Fast Fourier convolution. *Advances in Neural Information Processing Systems*, **33**, 4479–4488.

Cui, L., Wu, N., Ma, C., & Wang, H. (2016). Quantitative fault analysis of roller bearings based on a novel matching pursuit method with a

new step-impulse dictionary. *Mechanical Systems and Signal Processing*, **68-69**, 34–43. https://doi.org/10.1016/j.ymssp.2015.05.032.

Gazizulin, D., Rosado, L., Schneck, R., Klein, R., & Bortman, J. (2020). A new efficient rolling element—Spall edge interaction model. *International Journal of Fatigue*, **131**, 105330. https://doi.org/10.1016/j.ijfatigue.2019.105330.

Han, Y., Tang, B., & Deng, L. (2019). An enhanced convolutional neural network with enlarged receptive fields for fault diagnosis of planetary gearboxes. *Computers in Industry*, **107**, 50–58. https://doi.org/10.1016/j.compind.2019.01.012.

Hu, C., Goebel, K., Howey, D., Peng, Z., Wang, D., Wang, P., & Youn, B. D. (2023). Editorial: Special issue on physics-informed machine learning enabling fault feature extraction and robust failure prognosis. *Mechanical Systems and Signal Processing*, **192**, 110219. https://doi.org/10.1016/j.ymssp.2023.110219.

Jeong, H. (2023). *A Study of Deep Learning-based Spall Size Distribution Estimation for Rolling Element Bearing Diagnosis*. Master's thesis, Seoul National University,.

Kan, K., Zhou, J., Feng, J., Xu, H., Zheng, Y., Chen, H., & Chen, J. (2024). Intelligent optimization of axial-flow pump using physics-considering machine learning. *Journal of Computational Design and Engineering*, **11**, 325–342. https://doi.org/10.1093/jcde/qwae013.

Kim, I., Wook Kim, S., Kim, J., Huh, H., Jeong, I., Choi, T., Kim, J., & Lee, S. (2024). Single domain generalizable and physically interpretable bearing fault diagnosis for unseen working conditions. *Expert Systems with Applications*, **241**, 122455. https://doi.org/https://doi.org/10.1016/j.eswa.2023.122455.

Kim, M. S., Yun, J. P., & Park, P. (2020). An explainable convolutional neural network for fault diagnosis in linear motion guide. *IEEE Transactions on Industrial Informatics*, **17**, 4036–4045. https://doi.org/10.1109/TII.2020.3012989.

Kim, T., & Lee, S. (2023). A novel unsupervised clustering and domain adaptation framework for rotating machinery fault diagnosis. *IEEE Transactions on Industrial Informatics*, **19**, 9404–9412. https://doi.org/10.1109/TII.2022.3228395.

Kim, T., & Lee, S. (2024). Deep learning integrated bayesian health indicator for cross-machine health monitoring. *Structural Health Monitoring*, **23**, 3416–3429. https://doi.org/10.1177/14759217241227599.

Kim, Y., Na, K., & Youn, B. D. (2022). A health-adaptive time-scale representation (HTSR) embedded convolutional neural network for gearbox fault diagnostics. *Mechanical Systems and Signal Processing*, **167**, 108575. https://doi.org/10.1016/j.ymssp.2021.108575.

Kingma, D. P., & Welling, M. (2013). Auto-encoding variational bayes. In *Proceedings of the International Conference on Learning Representations (ICLR)*.

Ko, J. U., Lee, J., Kim, T., Kim, Y. C., & Youn, B. D. (2023). Deep-learning-based fault detection and recipe optimization for a plastic injection molding process under the class-imbalance problem. *Journal of Computational Design and Engineering*, **10**, 694–710. https://doi.org/10.1093/jcde/qwad011.

Kogan, G., Klein, R., & Bortman, J. (2018). A physics-based algorithm for the estimation of bearing spall width using vibrations. *Mechanical Systems and Signal Processing*, **104**, 398–414. https://doi.org/10.1016/j.ymssp.2017.11.011.

Kumar, A., Kumar, R., Tang, H., & Xiang, J. (2024). A comprehensive study on developing an intelligent framework for identification and quantitative evaluation of the bearing defect size. *Reliability Engineering & System Safety*, **242**, 109768. https://doi.org/10.1016/j.ress.2023.109768.

Larizza, F., Howard, C. Q., & Grainger, S. (2020). Defect size estimation in rolling element bearings with angled leading and trailing edges, *Structural Health Monitoring*, **20**, 1102–1116. https://doi.org/10.1177/1475921720938296.

Larizza, F., Howard, C. Q., Grainger, S., & Wang, W. (2020). A nonlinear dynamic vibration model of a defective bearing: The importance of modelling the angle of the leading and trailing edges of a defect. *Structural Health Monitoring*, **20**, 2604–2625. https://doi.org/10.1177/1475921720963950.

Li, S., Peng, Y., Shen, Y., Zhao, S., Shao, H., Bin, G., Guo, Y., Yang, X., & Fan, C. (2024). Rolling bearing fault diagnosis under data imbalance and variable speed based on adaptive clustering weighted oversampling. *Reliability Engineering & System Safety*, **244**, 109938. https://doi.org/10.1016/j.ress.2024.109938.

Li, T., Zhao, Z., Sun, C., Yan, R., & Chen, X. (2020). Multireceptive field graph convolutional networks for machine fault diagnosis. *IEEE Transactions on Industrial Electronics*, **68**, 12739–12749. https://doi.org/10.1109/TIE.2020.3040669.

Li, Z., Kovachki, N., Azizzadenesheli, K., Liu, B., Bhattacharya, K., Stuart, A., & Anandkumar, A. (2020). *Fourier Neural Operator for Parametric Partial Differential Equations*, arXiv preprint arXiv:2010.08895. https://doi.org/10.48550/arXiv.2010.08895.

Liang, H., & Zhao, X. (2021). Rolling bearing fault diagnosis based on one-dimensional dilated convolution network with residual connection. *IEEE Access*, **9**, 31078–31091. https://doi.org/10.1109/ACCESS.2021.3059761.

Liu, X., Centeno, J., Alvarado, J., & Tan, L. (2022). One dimensional convolutional neural networks using sparse wavelet decomposition for bearing fault diagnosis. *IEEE Access*, **10**, 86998–87007. https://doi.org/10.1109/ACCESS.2022.3199381.

Liu, Z., Wang, H., Liu, J., Qin, Y., & Peng, D. (2020). Multitask learning based on lightweight 1DCNN for fault diagnosis of wheelset bearings. *IEEE Transactions on Instrumentation and Measurement*, **70**, 1–11. https://doi.org/10.1109/TIM.2020.3017900.

Mathieu, M., Henaff, M., & LeCun, Y. (2013). Fast training of convolutional networks through FFTs. arXiv preprint arXiv:1312.5851.

Moazen Ahmadi, A., Howard, C. Q., & Petersen, D. (2016). The path of rolling elements in defective bearings: Observations, analysis and methods to estimate spall size. *Journal of Sound and Vibration*, **366**, 277–292. https://doi.org/10.1016/j.jsv.2015.12.011.

Moazen Ahmadi, A., Petersen, D., & Howard, C. (2015). A nonlinear dynamic vibration model of defective bearings—The importance of modelling the finite size of rolling elements. *Mechanical Systems and Signal Processing*, **52–53**, 309–326. https://doi.org/10.1016/j.ymssp.2014.06.006.

Mufazzal, S., Muzakkir, S., & Khanam, S. (2023). A dynamic model for simulating vibration response of ball bearings with extended outer race defects for precise spall size estimation. *Journal of Sound and Vibration*, **552**, 117606. https://doi.org/10.1016/j.jsv.2023.117606.

Oppenheim, A. V., & Lim, J. S. (1981). The importance of phase in signals. *Proceedings of the IEEE*, **69**, 529–541. https://doi.org/10.1109/PROC.1981.12022.

Park, J., Yoo, J., Kim, T., Ha, J. M., & Youn, B. D. (2023). Multi-head denoising autoencoder-based multi-task model for fault diagnosis of rolling element bearings under various speed conditions. *Journal of Computational Design and Engineering*, **10**, 1804–1820. https://doi.org/10.1093/jcde/qwad076.

Park, J., Yoo, J., Kim, T., Kim, M., Park, J., Ha, J. M., & Youn, B. D. (2025). Fault frequency band segmentation and domain adaptation with fault simulated signal for fault diagnosis of rolling element bearings. *Journal of Computational Design and Engineering*, **12**, 34–52. https://doi.org/10.1093/jcde/qwae105.

Qin, Y., Wu, X., & Luo, J. (2022). Data-model combined driven digital twin of life-cycle rolling bearing. *IEEE Transactions on Industrial*

Informatics, **18**, 1530–1540. https://doi.org/10.1109/TII.2021.3089340.

Sawalhi, N., & Randall, R. B. (2011). Vibration response of spalled rolling element bearings: Observations, simulations and signal processing techniques to track the spall size. *Mechanical Systems and Signal Processing*, **25**, 846–870. https://doi.org/10.1016/j.ymssp.2010.09.009.

Sawalhi, N., Wang, W., & Becker, A. (2017). Vibration signal processing for spall size estimation in rolling element bearings using autoregressive inverse filtration combined with bearing signal synchronous averaging. *Advances in Mechanical Engineering*, **9**, 1687814017703007. https://doi.org/10.1177/1687814017703007.

Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556.

Sinha, A. K., Moorthi, S. M., & Dhar, D. (2022). Nl-FFC: Non-local fast Fourier convolution for image super resolution. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. IEEE.

Su, Y., Shi, L., Zhou, K., Bai, G., & Wang, Z. (2024). Knowledge-informed deep networks for robust fault diagnosis of rolling bearings. *Reliability Engineering & System Safety*, **244**, 109863. https://doi.org/10.1016/j.ress.2023.109863.

Tao, B., Zhu, L., Ding, H., & Xiong, Y. (2007). An alternative time-domain index for condition monitoring of rolling element bearings—A comparison study. *Reliability Engineering & System Safety*, **92**, 660–670. https://doi.org/10.1016/j.ress.2006.03.005.

Toumi, M. Y., Murer, S., Bogard, F., & Bolaers, F. (2018). Numerical simulation and experimental comparison of flaw evolution on a bearing raceway: Case of thrust ball bearing. *Journal of Computational Design and Engineering*, **5**, 427–434. https://doi.org/10.1016/j.jcde.2018.01.004.

Tyagi, S., & Panigrahi, S. K. (2017). An improved envelope detection method using particle swarm optimisation for rolling element bearing fault diagnosis. *Journal of Computational Design and Engineering*, **4**, 305–317. https://doi.org/10.1016/j.jcde.2017.05.002.

Wang, D., Chen, Y., Shen, C., Zhong, J., Peng, Z., & Li, C. (2022). Fully interpretable neural network for locating resonance frequency bands for machine condition monitoring. *Mechanical Systems and Signal Processing*, **168**, 108673. https://doi.org/10.1016/j.ymssp.2021.108673.

Wang, J., Li, Y., Zhao, R., & Gao, R. X. (2020). Physics guided neural network for machining tool wear prediction. *Journal of Manufacturing Systems*, **57**, 298–310. https://doi.org/10.1016/j.jmsy.2020.09.005.

Wang, X., Girshick, R., Gupta, A., & He, K. (2018). Non-local neural networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. IEEE.

Wang, Z., Liu, T., Wu, X., & Liu, C. (2023). A diagnosis method for imbalanced bearing data based on improved SMOTE model combined with CNN-AM. *Journal of Computational Design and Engineering*, **10**, 1930–1940. https://doi.org/10.1093/jcde/qwad081.

Xiyang, L., Guo, C., Tengfei, H., & Wenping, P. (2023). A combined deep learning model for damage size estimation of rolling bearing. *International Journal of Engine Research*, **24**, 1362–1373. https://doi.org/10.1177/14680874221086601.

Xu, Y., Xia, T., Jiang, Y., Wang, Y., Wang, D., Pan, E., & Xi, L. (2024). A temporal partial domain adaptation network for transferable prognostics across working conditions with insufficient data. *Reliability Engineering & System Safety*, **250**, 110273. https://doi.org/https://doi.org/10.1016/j.ress.2024.110273.

Zhang, H., Borghesani, P., Smith, W. A., Randall, R. B., Shahriar, M. R., & Peng, Z. (2021). Tracking the natural evolution of bearing spall size using cyclic natural frequency perturbations in vibration signals. *Mechanical Systems and Signal Processing*, **151**, 107376. https://doi.org/10.1016/j.ymssp.2020.107376.

Zhao, S., Liang, L., Xu, G., Wang, J., & Zhang, W. (2013). Quantitative diagnosis of a spall-like fault of a rolling element bearing by empirical mode decomposition and the approximate entropy method. *Mechanical Systems and Signal Processing*, **40**, 154–177. https://doi.org/10.1016/j.ymssp.2013.04.006.

Zhou, H., Wang, B., Zio, E., Wen, G., Liu, Z., Su, Y., & Chen, X. (2023). Hybrid system response model for condition monitoring of bearings under time-varying operating conditions. *Reliability Engineering & System Safety*, **239**, 109528. https://doi.org/https://doi.org/10.1016/j.ress.2023.109528.

Zio, E. (2022). Prognostics and health management (PHM): Where are we and where do we (need to) go in theory and practice. *Reliability Engineering & System Safety*, **218**, 108119. https://doi.org/https://doi.org/10.1016/j.ress.2021.108119.

Received: December 9, 2024. **Revised:** April 5, 2025. **Accepted:** April 5, 2025