

Article

Intelligent Diagnosis of Rolling Element Bearings Under Various Operating Conditions Using an Enhanced Envelope Technique and Transfer Learning

Ali Davoodabadi ¹, Mehdi Behzad ^{1,*}, Hesam Addin Arghand ² , Somaye Mohammadi ¹  and Len Gelman ³

¹ Condition Monitoring Laboratory, Department of Mechanical Engineering, Sharif University of Technology, Tehran 14588-89694, Iran; alidavoodabadi@mech.sharif.edu (A.D.); somaye.mohammadi@sharif.edu (S.M.)

² Department of Mechanical Engineering, Faculty of Engineering, University of Zanjan, Zanjan 45371-38791, Iran; arghand@znu.ac.ir

³ Department of Engineering, School of Computing and Engineering, University of Huddersfield, Huddersfield HD1 3DH, UK; l.gelman@hud.ac.uk

* Correspondence: m_behzad@sharif.edu

Abstract: Rolling element bearings (REBs) are vital in rotating machinery, making fault detection essential for optimal performance and system reliability. This study assesses the effectiveness of a simple convolutional neural network (SCNN) and a transfer learning-based convolutional neural network (TL-CNN) for diagnosing REB faults using time-domain signals, frequency-domain spectra, and envelope frequency spectrum analysis. The study uses diverse datasets, including laboratory and industrial data under various operating conditions, covering fault types like inner race fault (IRF), outer race fault (ORF), rolling element fault (REF), and healthy (H) states. The main innovation is applying Transfer Learning (TL) with fine-tuning to improve model accuracy in identifying REB conditions by leveraging features learned from diverse datasets. An innovative algorithm is also introduced to identify resonance regions for optimal filter selection in envelope analysis, improving fault-related feature extraction and reducing noise. A preprocessing step that removes speed-related variations further enhances model accuracy by isolating fault features and minimizing the impact of rotational speed. The results show that transfer learning with fine-tuning, combined with the resonance region identification algorithm, significantly enhances fault detection accuracy. The TL-CNN model with envelope signal input achieves the highest accuracy across all scenarios, especially under variable operating conditions, and performs reliably on industrial data.

Keywords: condition monitoring; intelligent fault diagnosis; rolling element bearings; convolutional neural network; transfer learning; vibration analysis



Academic Editors: Yancai Xiao, Tianyang Wang and Shaodan Zhi

Received: 16 March 2025

Revised: 15 April 2025

Accepted: 22 April 2025

Published: 23 April 2025

Citation: Davoodabadi, A.; Behzad, M.; Arghand, H.A.; Mohammadi, S.; Gelman, L. Intelligent Diagnosis of Rolling Element Bearings Under Various Operating Conditions Using an Enhanced Envelope Technique and Transfer Learning. *Machines* **2025**, *13*, 351. <https://doi.org/10.3390/machines13050351>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

In industrial rotating machinery, mechanical faults can arise from various critical components, with shaft-related faults and rolling element bearings (REBs) being the most prominent sources of failure. Shaft-related faults, such as imbalance, misalignment, bending, cracks, eccentricity, and torsional deformation, are often caused by operational stresses, fatigue, or manufacturing defects. These faults typically account for 40–50% of mechanical failures in rotating equipment [1,2], highlighting their significant impact on system performance. REBs, which are vital for minimizing friction and facilitating rotational motion, are responsible for approximately 40% of rotating machinery failures, with common defects including damage to the inner or outer race, rolling elements, or the cage [3]. Each

type of fault generates distinct vibration signatures, which, when monitored effectively, can serve as early indicators of impending failures. Given the high failure rates and the associated economic costs, developing intelligent diagnostic models for the early detection of REB faults is crucial for improving the reliability and operational efficiency of industrial systems [4].

In recent years, significant progress has been made in intelligent fault diagnosis of these components through advancements in machine learning (ML) and deep learning (DL) algorithms [5–23]. These algorithms primarily rely on labeled data, which can be easily collected under controlled laboratory conditions. However, in industrial environments, gathering sufficient labeled data is a challenge, since REBs typically operate in healthy states, with failures occurring rarely. Furthermore, continually halting machinery to collect failure-related data is neither practical nor cost-effective. Additionally, manually labeling data is a difficult task due to the large volume of data and the need for expert knowledge. These challenges have resulted in a shortage of labeled data for training accurate models.

Fault simulation under laboratory conditions offers a temporary solution and can provide valuable insights into fault patterns. However, due to differences in operational and environmental conditions, laboratory data and real-world data often exhibit different distributions. As a result, the trained models on laboratory data tend to perform poorly when applied to real-world scenarios. To address this issue, transfer learning (TL) has been introduced in deep learning, enabling the transfer of knowledge from one domain (e.g., laboratory data) to another (e.g., industrial data). TL techniques enhance model performance across diverse conditions by minimizing the discrepancies between the source and target data distributions. These techniques can be broadly categorized into statistical methods and deep neural network (DNN)-based approaches. By adapting the domain and transferring learned features, these methods improve the accuracy of fault diagnosis in REBs in industrial settings, offering an effective solution to overcome existing limitations.

In recent years, numerous studies have been conducted on the fault diagnosis of REBs using TL and DL techniques. Shen et al. [24] proposed an innovative method for detecting faults in REBs by combining singular value decomposition (SVD) with TL. In their approach, vibration signals from REBs were processed using SVD to extract relevant features, and the Transfer AdaBoost (TrAdaBoost) algorithm was employed for classification. The TrAdaBoost algorithm enhanced detection accuracy in scenarios with limited target data by dynamically weighting the target and auxiliary data. Additionally, similarity measurement between the target and auxiliary data, based on label and feature distributions, was performed to mitigate the risk of negative transfer.

Zhang et al. [25] introduced the Wide Deep Convolutional Neural Network (WDCNN) model for fault detection in REBs within mechanical systems. This model utilized raw vibration signals as input and employed wide kernels in the first convolutional layer to capture important features while effectively suppressing high-frequency noise. In subsequent layers, smaller kernels were used to extract more complex and detailed features. To address discrepancies between the source and target data distributions, the Adaptive Batch Normalization (AdaBN) method was integrated, enabling the model to maintain robust performance across varying conditions.

In a related study, Zhang et al. [26] proposed a novel unsupervised domain adaptation (DA) method based on subspace alignment (SA) for fault detection in REBs under varying operating conditions. This method extracted features from vibration signals using fast Fourier transform (FFT) and principal component analysis (PCA), followed by subspace alignment to reduce distribution differences between the source and target domains. This approach demonstrated significant effectiveness in improving fault diagnosis accuracy under diverse operational conditions.

Zhang et al. [27] explored the use of TL with neural networks (NN) for fault diagnosis. They proposed a model that leveraged source data from initial conditions and improved performance with limited data under new conditions. In their approach, the NN was initially trained on source data, and the learned parameters were then transferred to new conditions by adapting the network structure. Qian et al. [28] presented a deep transfer learning network for fault detection in rotating machinery under varying operating conditions. Their proposed method consisted of three stages data preprocessing, network training, and fault detection. The approach utilized stacked autoencoders (SAE) for feature extraction and AdaBN to address domain shifts. Notably, this method did not require labeled target domain data during training, making it highly practical for real-world applications. In a related study, Tong et al. [29] proposed a new method using domain adaptation and transferable features. They transformed vibration data from REBs into the frequency domain using FFT and reduced the data dimensions with PCA. They used maximum mean discrepancy (MMD) to reduce the distribution discrepancy between training and testing data and extracted transferable features for high-accuracy classification. Li et al. [30] introduced the domain adaptive convolutional neural network (DACNN) model, which achieved high accuracy in fault detection by effectively adapting source and target data distributions.

Yang et al. [31] developed a model for reducing feature distribution differences using TL and distribution alignment with MMD. They employed pseudo-labels for unlabeled data in the target domain. Li et al. [32] proposed a deep nonnegativity-constraint sparse autoencoder for REBs diagnosis with limited labeled data. Initially, a novel non-negativity-constraint sparse autoencoder (NSAE) is used to enhance sparsity. A base deep NSAE is then employed to extract features from raw vibration signals. A parameter TL strategy is applied to transfer knowledge from the source model to the target model, addressing the issue of limited labeled data. Shao et al. [33] proposed a DL framework for machine fault diagnosis using TL. In this approach, raw sensor data are converted into time-frequency images using wavelet transformation. A pre-trained network is then used to extract lower-level features, and the higher levels of the NN architecture are fine-tuned using labeled time-frequency images. He et al. [34] proposed an enhanced deep transfer autoencoder for REB diagnosis in different machines. The method improves the quality of vibration data using a scaled exponential linear unit (SELU) and applies a non-negative constraint to enhance the reconstruction effect. The parameter knowledge from a pre-trained source model is transferred to the target model, and the target model is fine-tuned to adapt to the characteristics of new data. Zhao et al. [35] introduced a deep multi-scale convolutional neural network (MSCNN) for intelligent fault diagnosis of REB under varying working conditions and domains. A novel multi-scale module is designed using dilated convolution to extract differential features from various perceptual fields. The architecture and weights of the pre-trained MSCNN are transferred to a target domain and fine-tuned for better adaptation. Li et al. [36] proposed an intelligent cross-machine fault diagnosis method combining deep autoencoders and domain adaptation. The features from different machines are projected into a shared subspace, and a cross-machine adaptation algorithm is employed to minimize distribution discrepancies between data from different machines. This method facilitates knowledge transfer for fault diagnosis across different machines. Wu et al. [37] developed a deep transfer maximum classifier discrepancy (TMCD) method for REB fault diagnosis under conditions with limited labeled data. The method uses a batch-normalized long-short term memory (BNLSTM) model to learn the relationships between different datasets and generate auxiliary samples. The TMCD method applies an adversarial strategy to align the probability distributions between auxiliary samples and unlabeled target domain data.

Dong and Li [38] presented a new framework for intelligent fault detection in REBs, addressing the small sample problem. This framework involved initially training the model with simulated data and then fine-tuning the model with a limited amount of real data. Several parameter transfer strategies, such as freezing and fine-tuning, were explored for knowledge transfer. Finally, Chen et al. [39] reviewed the advancements in deep TL for REB fault detection from 2016 and analyzed the strengths and weaknesses of these methods. Zhong et al. [40] proposed a REB fault diagnosis model based on domain adversarial transfer learning, enhanced with structural adjustment modules. The pre-trained source model is adapted to the target domain using adversarial domain adaptation, while the network architecture is automatically optimized via the Optuna framework. This approach offers strong adaptability to complex and changing industrial conditions without manual tuning.

One of the main weaknesses of previous research is the use of the REB dataset published by Case Western Reserve University (CWRU) [41] for evaluating the proposed algorithms. The vibration of REBs is influenced by factors such as radial load and shaft rotational speed. However, this dataset does not account for variations in radial load, and the range of rotational speed variations is less than 5%. Therefore, this dataset cannot properly simulate the varying operating conditions of the REBs. Due to the uniform distribution of labeled and unlabeled data in this dataset, the algorithms used in previous studies have shown acceptable performance. However, when these algorithms are applied to unlabeled industrial data with a different distribution, their performance drastically decreases. These limitations highlight the necessity of developing TL algorithms for data where changes in operating conditions are more pronounced, and the distribution of labeled and unlabeled data is more different.

In this research, a convolutional neural network (CNN) with an innovative training approach is proposed, which is particularly effective in extracting transferable features from data and accurately identifying faults under variable operating conditions. The proposed model utilizes TL and fine-tuning techniques, where the feature extraction layers remain fixed, and only the classification layers are retrained with real-world data. This training approach enables the model to transfer complex, transferable features learned from simulated or lab data to real-world data, enhancing fault detection accuracy in various operational conditions.

One of the key innovations of this study is the development of an advanced algorithm for accurately identifying the resonance region and using it in envelope analysis as input to the model. While previous models have used envelope analysis for feature extraction, the resonance region was not precisely and optimally determined. In this research, an algorithm has been designed that accurately identifies the resonance region and uses it as a filtering zone for envelope analysis. This method significantly improves the accuracy and performance of the model in fault detection, especially under variable operational conditions where vibration signals are affected by various factors. Furthermore, a preprocessing procedure has been applied to eliminate the impact of rotational speed variations from the data. Additionally, a set of experimental data under variable conditions and industrial data have been used to evaluate and validate the developed model. In the following sections of the paper, Section 2 introduces the proposed method. Section 3 describes the experimental setup and data collection process. Section 4 discusses the implementation of the developed model on both experimental and industrial data. Section 5 presents the discussion in Section 5.1, followed by the limitations and future research directions in Section 5.2. Finally, Section 6 provides the conclusions.

2. The Proposed Method

In this section, the proposed method for fault detection of REB is outlined. Section 2.1 discusses the architecture and configuration of the simple convolutional neural network (SCNN) and transfer learning-based convolutional neural network (TL-CNN) models. Section 2.2 covers the input data types and the resonance-based filtering technique for fault detection. Section 2.3 describes the approach for mitigating the effects of rotational speed variations on fault detection accuracy.

2.1. Architecture and Configuration of SCNN and TL-CNN Models

In this study, two one-dimensional CNN models are used for fault detection of REBs, which include the SCNN model and the TL-CNN model with fine-tuning. These models are designed to process one-dimensional signal data and extract important features for classifying different types of faults. The SCNN model, as one of the effective tools in DL, is applied to analyze one-dimensional signals from REBs. The architecture of the SCNN model is shown in Figure 1.

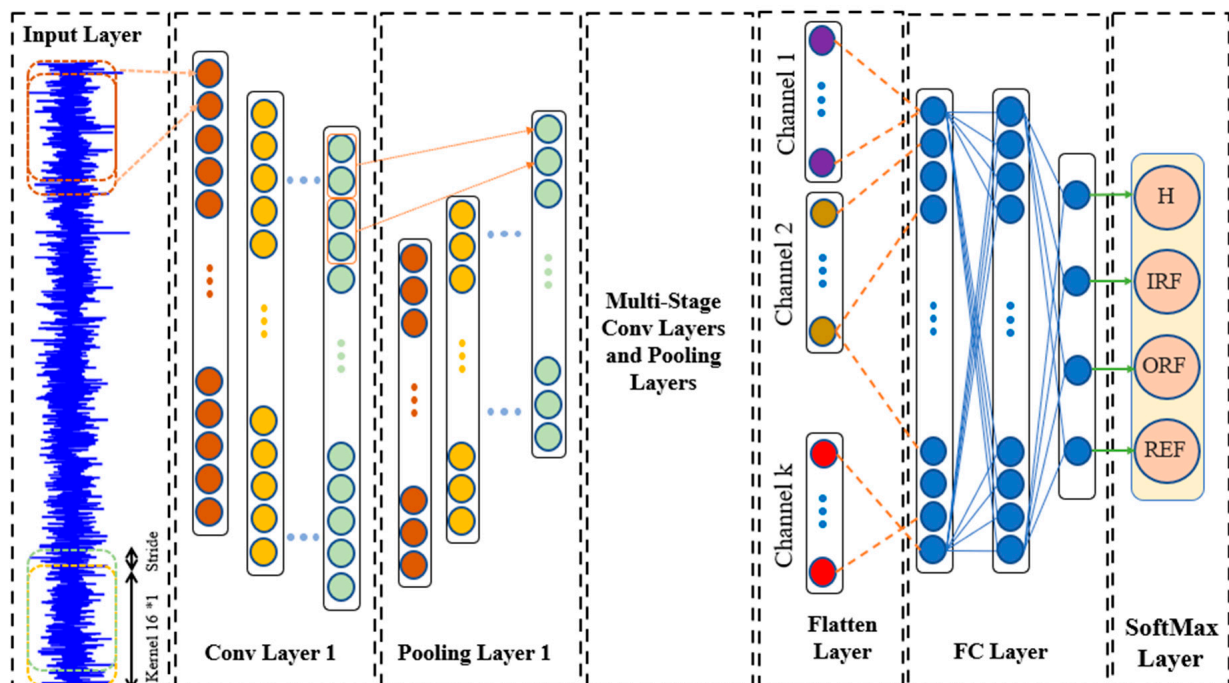


Figure 1. Architecture of SCNN Model to detect healthy (H), inner race fault (IRF), outer race fault (ORF), and rolling element fault (REF) in vibration analysis of REBs.

The TL-CNN model is designed to improve the model's ability to recognize and analyze new data with different distributions. This model is trained in two distinct phases. In the first phase, the model is pre-trained using laboratory data collected under varying operating conditions, including seeded faults. These data are obtained in controlled environments and encompass a wide range of operational conditions for the REB. In the second phase, the pre-trained model is adapted to new data. During this phase, the convolutional layers, which are responsible for feature extraction, remain fixed, while only the fully connected (dense) layers are fine-tuned using the new dataset. This two-phase approach ensures that the model leverages prior knowledge while adapting to new conditions. The architecture of the TL-CNN model is illustrated in Figure 2.

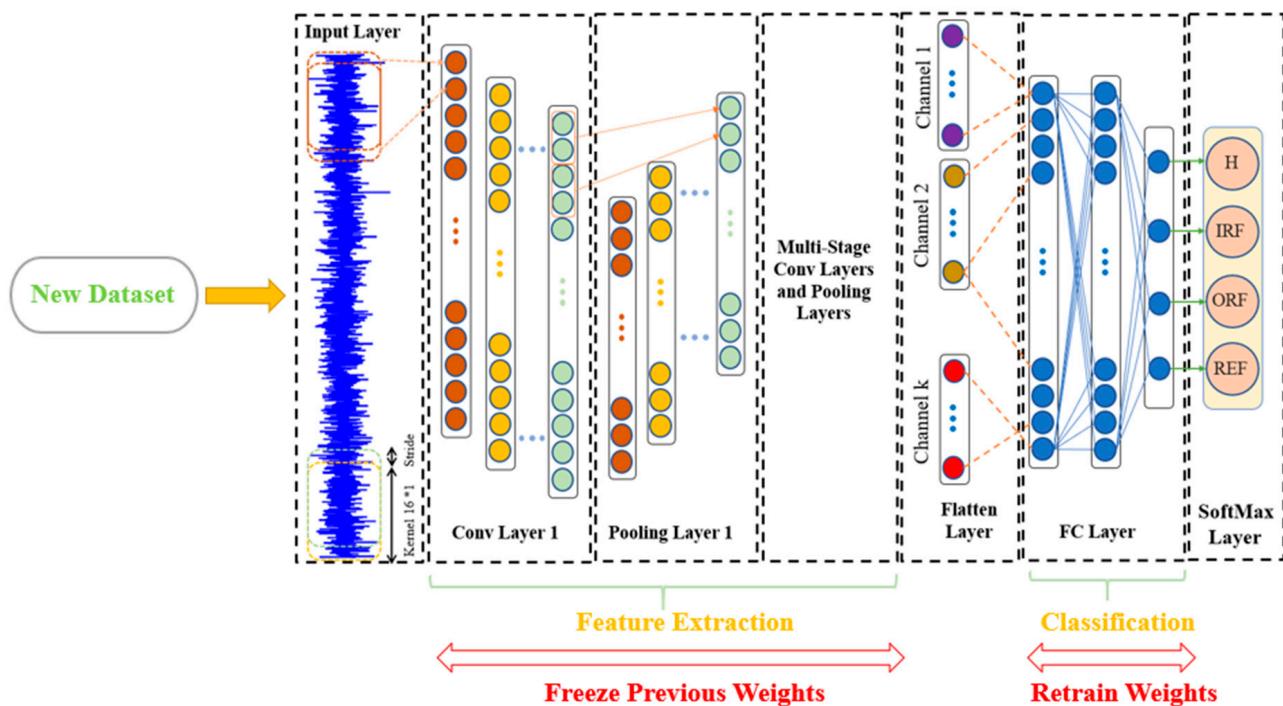


Figure 2. Structure of the proposed TL-CNN model to H, IRF, ORF, and REF in vibration analysis of REBs.

Following the architectural diagrams presented in Figures 1 and 2, which illustrate the SCNN and TL-CNN models, respectively, Table 1 provides a detailed summary of the architecture and configuration of each model. The table outlines the structure of the network, including the type of each layer, the output shape, the number of filters, activation functions, trainable status, and parameter count.

Table 1. Architecture of SCNN and TL-CNN Models.

Layer No.	Layer Type	Number of Filters	Kernel Size	Activation	Output Shape	Params	Trainable Status	Description
1	Conv1D	64	16×1	ReLU	(4081, 64)	1088	Frozen (No Training)	Feature extraction from source data
2	MaxPooling1D	-	2×1	-	(2040, 64)	0	Frozen	Down sampling to reduce dimensionality
3	Conv1D	32	8×1	ReLU	(2033, 32)	16,416	Frozen	Further feature extraction
4	MaxPooling1D	-	2×1	-	(1016, 32)	0	Frozen	Pooling to further reduce dimensionality
5	Conv1D	16	4×1	ReLU	(1013, 16)	2064	Frozen	Final feature extraction step
6	Flatten	-	-	-	-16,208	0	Frozen	Flattening the output for dense layer integration
7	Dense	64	-	ReLU	-64	1,037,376	Trainable	Fully connected layer for feature propagation
8	Dense	64	-	ReLU	-64	4160	Trainable	Fully connected layer, adding complexity to the model

Table 1. Cont.

Layer No.	Layer Type	Number of Filters	Kernel Size	Activation	Output Shape	Params	Trainable Status	Description
9	Dense	4	-	ReLU	−8	260	Trainable	Softmax-based output for four-class classification
12	Total Params					3,184,094		

In the TL-CNN architecture, transfer learning is performed through a two-phase training process. In the first phase, the model is trained using a source dataset with a broad range of operating conditions. In the second phase, the model is adapted to the target dataset by fine-tuning only the fully connected layers, while the convolutional layers remain frozen. This strategy allows the model to retain robust and generalizable feature representations learned from the source data, ensuring that the base knowledge is not overwritten during adaptation to new conditions. Freezing the convolutional layers also significantly reduces computational costs and the risk of overfitting, which is particularly important when working with limited target data.

Furthermore, to ensure optimal training efficiency and generalization performance, Table 2 presents the key optimization hyperparameters used during training and fine-tuning. These include the number of epochs (which is reduced in the TL-CNN phase), batch size, optimizer (Adam), loss function (categorical crossentropy), validation split ratio, and the early stopping mechanism. Specifically, early stopping with a patience value of 3 was employed to terminate training when no further improvement in validation loss was observed, effectively preventing overfitting and reducing training time.

Table 2. Optimization Hyperparameters and Recommended Configurations for Fine-Tuning.

Parameter	Value Used	Description and Suggested Settings
Number of Epochs	50 (Base Model), 20 (Transfer Learning)	Fine-tuning epochs can be reduced during transfer learning to avoid overfitting
Batch Size	32	A common value that provides stable learning performance
Optimizer	Adam	Recommended optimizer; the learning rate can be adjusted if needed
Loss Function	Categorical Crossentropy	Suitable for multi-class classification tasks. No need to change unless data-specific adjustments are required
Validation Split	0.1	Typically 10–20% is used for validation data
Early Stopping	Patience = 3	Prevents overfitting by stopping training early if validation loss does not improve

2.2. Input Data Types and Resonance-Based Filtering for Fault Detection

For a comprehensive performance evaluation, both models use diverse data types, including time-domain signals, frequency spectra, and envelope signals as inputs. Time-domain signals are the raw data directly recorded from sensors, representing the instantaneous state of the REB. The frequency spectrum, extracted via FFT of the time-domain signals, reveals key features in the frequency domain. Envelope signal analysis is an effective method for analyzing vibration data of REBs, simulating modulated variations in vibration signals. When a fault occurs in one of the REB components (such as the inner race, outer race, or rolling element), the interaction of other components with the fault generates periodic, short-duration pulses with a distinct frequency. These pulses cause the vibration to be modulated at the characteristic fault frequency. McFadden and Smith [42] demonstrated that envelope analysis can accurately identify these characteristic frequencies. The envelope analysis process involves several stages. First, a bandpass filter is applied

to the signal to remove extra noise. Then, rectification is performed on the signal, and the envelope curve is extracted. Finally, the frequency spectrum of the envelope is obtained using FFT. This method is one of the most efficient techniques for fault detection in REBs due to its high capability in extracting fault-related features.

The selection of the filter region in vibration signal analysis is crucial for accurate fault detection. The resonance region is an optimal choice as it contains the most relevant fault-related information. Vibration signals from REBs with faults typically show high-amplitude frequency components due to resonance, appearing as short-duration pulses modulated at the fault's characteristic frequency. By choosing the resonance region as the filter, fault-related energy is enhanced while reducing noise, improving feature extraction and algorithm performance. In this study, an algorithm has been developed to automatically identify the resonance region in frequency signals, distinguishing relevant peaks from other oscillations and noise, step by step, enhancing feature extraction accuracy.

To identify resonance regions in frequency signals, an algorithm has been developed to automatically distinguish peaks associated with resonance regions from other oscillations and noise in the signal. The algorithm operates in sequential steps, with each step enhancing accuracy and improving the extraction of signal features:

1. In the first step, a smoothing process is applied to reduce the effects of random noise and short-term fluctuations in the signal. This process employs a Gaussian filter, which uses weighting functions based on the Gaussian distribution. Values closer to the center of the smoothing window receive higher weights. By suppressing unwanted oscillations and amplifying primary patterns in the signal, this method highlights the desired features.
2. In the second step, a threshold is established to identify significant peaks in the signal. This threshold is dynamically adjusted based on the mean and standard deviation of the signal's amplitude, enabling the algorithm to adapt to variations in signal conditions. Peaks with amplitudes below this threshold are considered noise and are excluded.
3. In the third step, the width of each identified peak is evaluated. Peak width is defined as the distance between two points where the amplitude equals one-quarter of the peak's maximum amplitude. The mean width of the identified peaks is calculated and used as the minimum acceptable width. Peaks with widths less than this threshold are removed. This stage helps to differentiate genuine peaks from noise-induced ones, leaving only peaks associated with resonance regions.
4. In the fourth step, the valid peaks are grouped into clusters representing resonance regions. This grouping is based on the temporal distances between consecutive peaks. For this purpose, an adaptive threshold is calculated, reflecting the characteristics of peak-to-peak distances. If the temporal distance between two peaks is less than this threshold, they are assigned to the same group; otherwise, a new group is formed.
5. Finally, to select the primary resonance region, the energy of each group is calculated. The energy of a region is defined as the sum of the squared amplitudes of its constituent peaks. The region with the highest energy is identified and reported as the primary resonance region. The flowchart of the proposed algorithm for detecting resonance regions is illustrated in Figure 3. This algorithm effectively distinguishes resonance regions from noise and irrelevant oscillations, accurately identifying regions associated with resonant behavior.

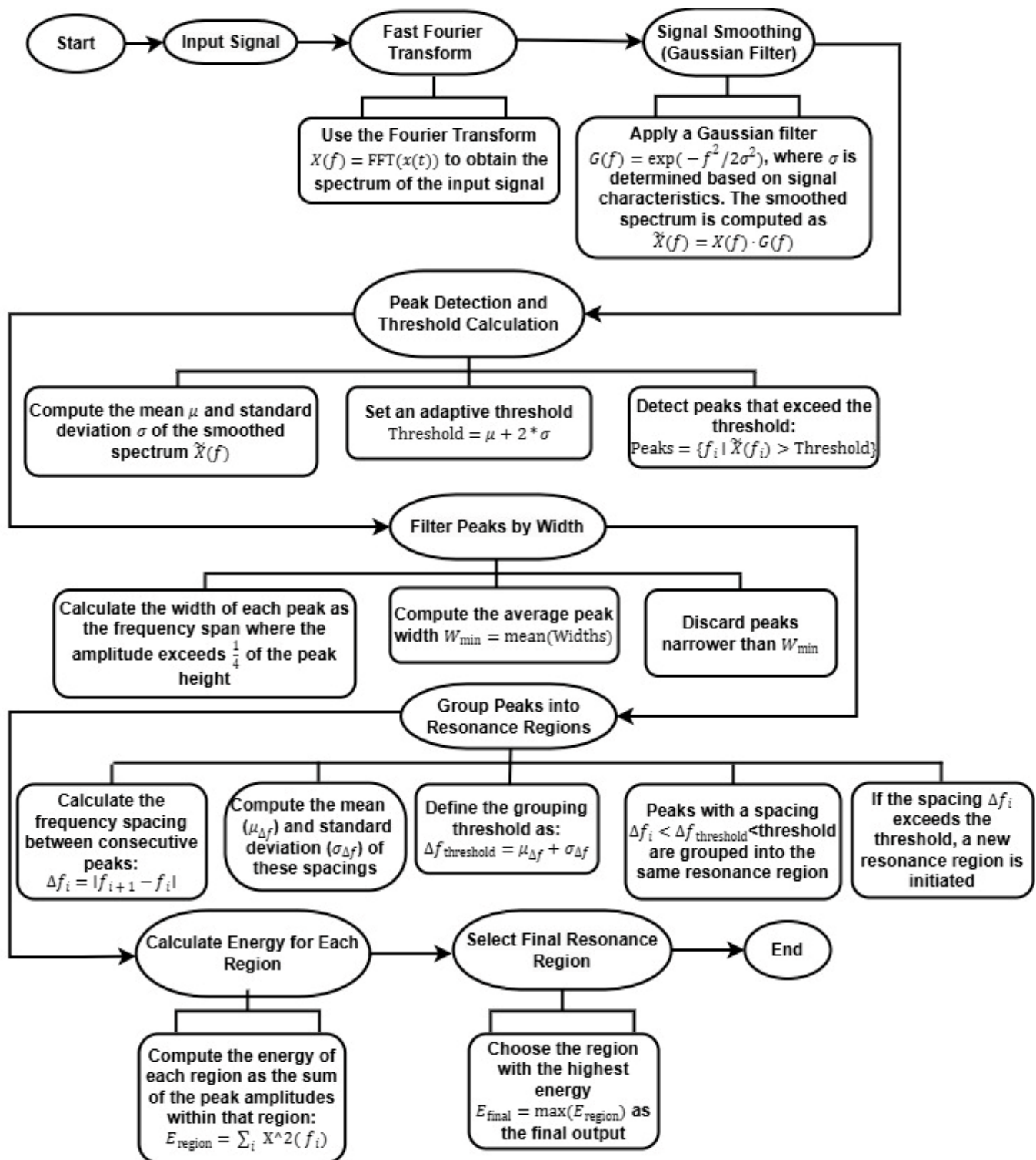


Figure 3. Flowchart of the proposed algorithm for automatically resonance zone detection.

Figure 4 illustrates the detection of the resonance region for four REB health conditions using the developed algorithm. This algorithm effectively distinguishes the resonance regions from noise and irrelevant oscillations.

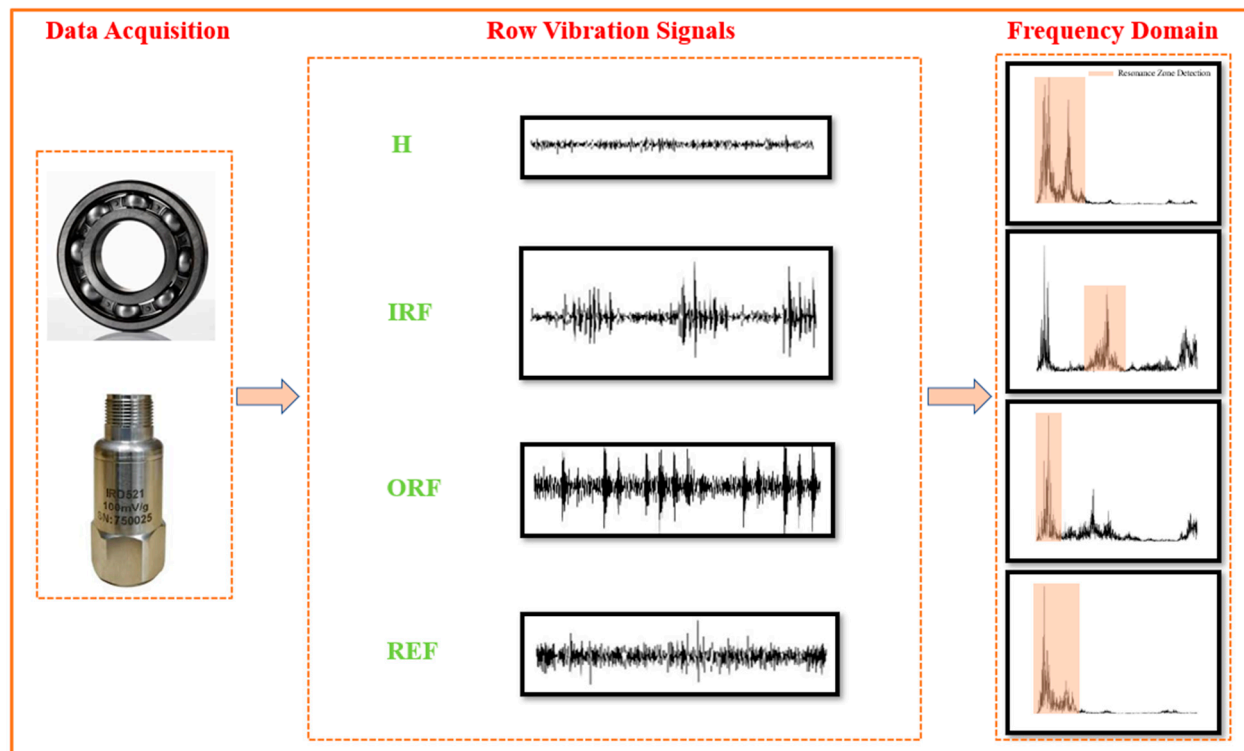


Figure 4. Resonance region detection for four REB health conditions using the developed algorithm.

2.3. Mitigating the Effects of Rotational Speed Variations in Vibration-Based Fault Detection

In vibration data analysis for REB fault detection, one of the key challenges is the impact of rotational speed variations on the collected data. The frequency spectra extracted from the vibration signals are directly dependent on the system speed, and variations in speed can result in changes to both the frequency scale and amplitude of these spectra. These changes make direct comparison or analysis of data collected at different speeds prone to errors due to misalignment of features. Particularly in DL models, such as CNNs, these speed-induced differences can cause the model to focus on speed-related variations rather than fault-relevant features, ultimately reducing fault detection accuracy.

To address the mentioned challenge, it is necessary to map the frequency spectra of data collected at different speeds to a common reference frame. This process must ensure that the fault-related feature information is fully preserved, while the impact of speed variations is eliminated. The goal is to standardize the data to improve the accuracy of DL models and to facilitate the analysis and comparison of vibration data under varying operational conditions. Typically, in industrial and laboratory settings, data are collected over frequency ranges between 3 Hz and 50 Hz. Since data are gathered at different speeds, selecting a fixed reference speed, particularly one centrally located within these frequency ranges, can significantly aid in data standardization. In this study, a reference speed of 25 Hz (Ref. Speed) is selected. This choice has been made because 25 Hz lies centrally within the common frequency range, allowing for efficient mapping of frequency spectra from varying speeds to this reference. A two-stage approach is employed to achieve this goal:

1. In the first stage, all frequencies in the spectra are normalized by dividing by the respective speed (Rot. Speed). This step ensures that frequencies corresponding to different speeds are transferred to a uniform scale.
2. In the second stage, the amplitudes of the frequency spectra are interpolated to the corresponding amplitudes at the reference speed of 25 Hz using linear interpolation

(LI). This amplitude adjustment process preserves critical fault-related information while aligning the frequency spectra to a common reference frame.

The result of this method is the generation of a set of uniform frequency spectra that can be used as inputs for DL models. These data uniformity ensure that the CNN focuses solely on fault-relevant features, eliminating the detrimental effects of rotational speed variations. This approach not only enhances fault detection accuracy but also improves the model's generalization capability for new data, representing a significant step forward in the advancement of condition monitoring systems. The selection of 25 Hz as the reference speed not only facilitates better alignment between datasets but also ensures that fault-related spectral features, which are often more prominent in this frequency range, are fully preserved. This method reduces the impact of speed variations and significantly improves the performance of CNNs in REB fault detection.

For example, in Figures 5–7, the process of synchronizing the frequency spectrum of vibration signals to a reference framework is illustrated. Figure 5 presents the frequency spectrum of a signal collected at a rotational speed of 1x, where the frequency components and corresponding amplitudes are clearly observable. Figure 6 depicts the frequency spectrum of a signal at 2x speed, showing significant changes in the position and intensity of frequency components due to the increase in rotational speed. Figure 7 demonstrates the transformation of the frequency spectrum from 1x to a reference speed of 2x. This transformation is achieved through frequency normalization relative to rotational speed and linear interpolation of amplitudes, ensuring the alignment of data within a common reference framework. For instance, the 5x rotational speed frequency in this case corresponds to the outer race fault of the bearing. This approach enhances the accuracy of vibration data comparison and eliminates the adverse effects of rotational speed variations on DL-based fault analysis. Consequently, standardizing vibration data not only improves the generalization capability of signal processing models but also enhances the accuracy of fault detection.

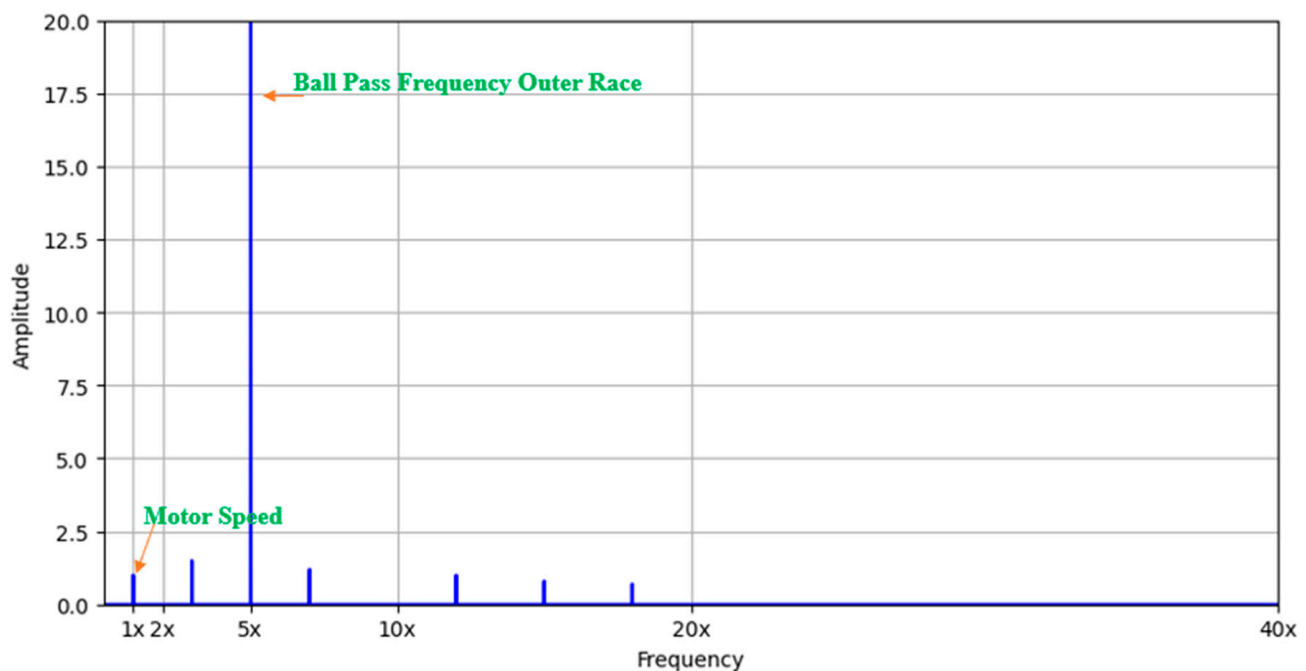


Figure 5. Example of a frequency spectrum for a vibration signal at 1x rotational speed.

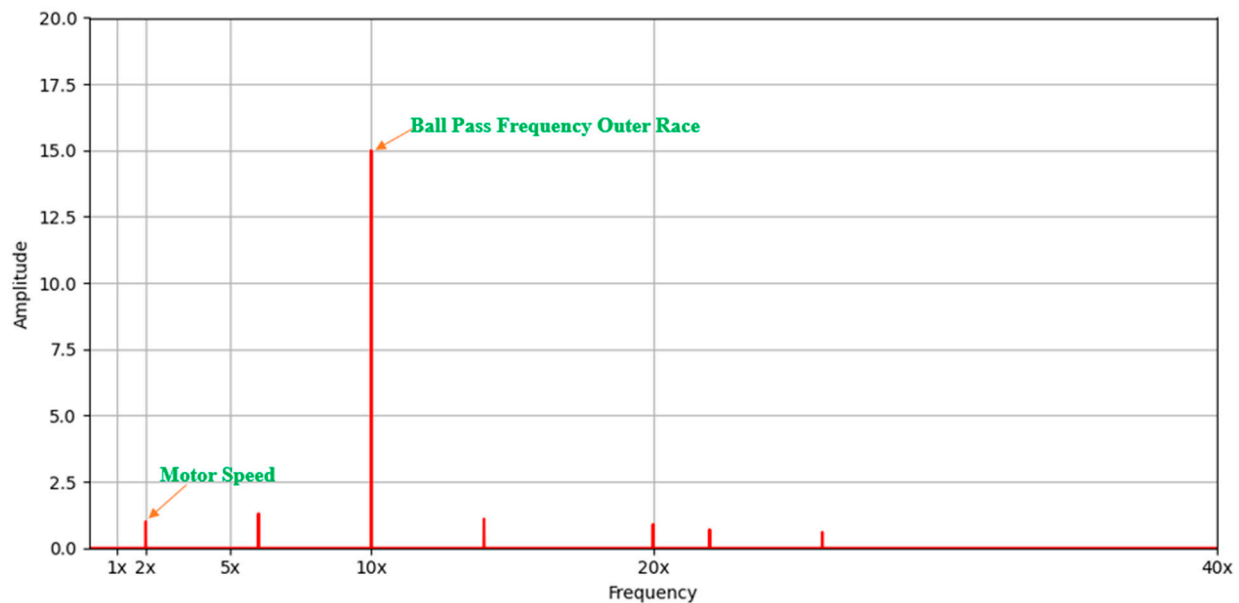


Figure 6. Example of a frequency spectrum for a vibration signal at 2x rotational speed.

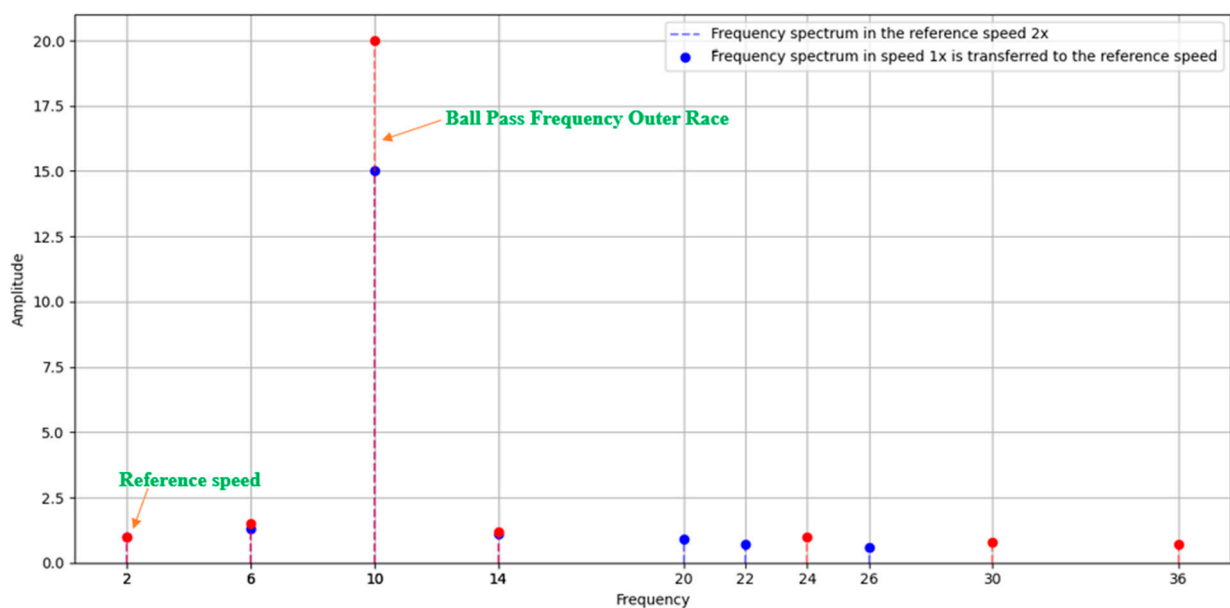


Figure 7. Example of frequency spectrum transformation from 1x rotational speed to the reference speed of 2x.

3. Experimental Setup

This section describes the experimental setups used to evaluate the performance of the proposed models. Section 3.1 introduces the first laboratory test setup and outlines the initial experiments conducted. Section 3.2 provides an overview of the different data acquisition scenarios and their application in evaluating the models.

3.1. Introduction of the First Laboratory Test Setup

To evaluate the proposed model, an REB test setup was required. In this regard, a setup with the capability to vary load and speed during the REB degradation process was developed in the Condition Monitoring Laboratory of Sharif University of Technology. A CAD model of the setup is shown in Figure 8. For emergency load control, a hydraulic power pack load control system is used, which is managed through a computer. In this

design, the target REB is placed between two REBs that act as supports. Another feature of this setup is the controllability of the rotational speed through its motor drive. The speed value and the vertical and horizontal vibrations of the load applied to the REB can be measured in real-time by this setup. In this study, data from a horizontal accelerometer sensor, which is aligned with the applied load, are utilized.

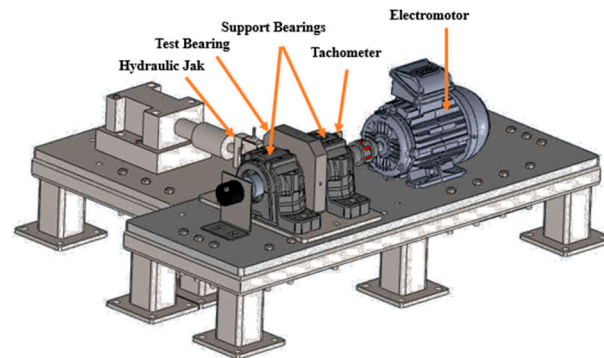


Figure 8. A CAD model of accelerated life tests on REBs experiments test.

In Figure 9, the direction of the applied load and the placement of the sensors and tachometer are shown. Figure 10 displays the constructed experimental setup.

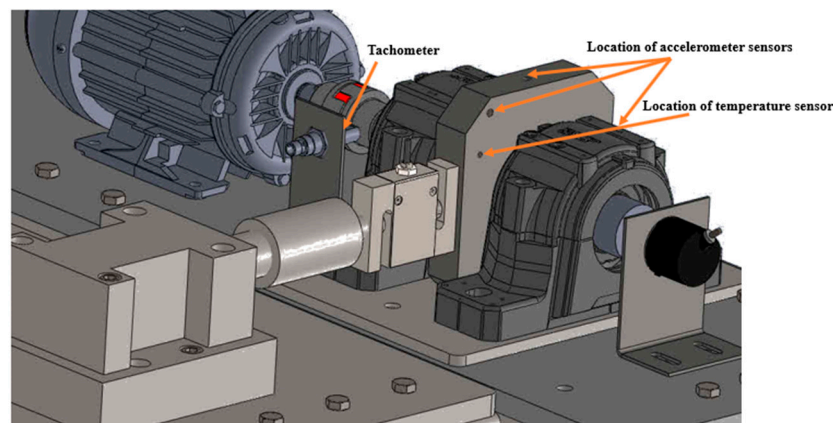


Figure 9. A view of loading on the target REB.



Figure 10. Accelerated life tests on REBs experiments test rig.

Vibrations were measured using the SDT Vigilant data acquisition system [43]. For measuring the vibration signal, the HS1001005008 sensor were used. The sampling frequency of the measured signals is 25,600 Hz. In this study, a self-aligning double-row REB has been used for test. Figure 11 shows the artificial faults created on the elements of REB.

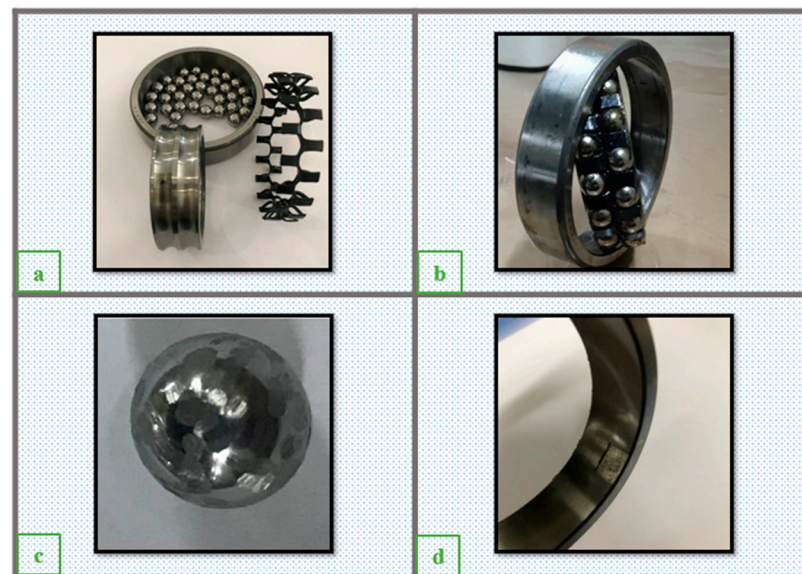


Figure 11. The artificial damages caused in the test REB, including (a) IRF, (b) H, (c) REF and (d) ORF.

Table 3 lists the dimensions of the ETN9 REB, the characteristic frequencies corresponding to IRF, ORF, and REF, as well as the dynamic specifications of the test REB.

Table 3. Specifications of the ETN9 REB 1210 [44].

Parameter		Symbol	Value
Dimensional Specifications of the REB (mm)	Outer Diameter	O	90
	Inner Diameter	I	50
	Width	B	20
	Number of Balls	N	34
Frequency Specifications of the REB (Motor speed = 1 Hz)	Ball Pass Frequency Outer	BPFO	7.26
	Ball Pass Frequency Inner	BPFI	9.28
	Ball Spin Frequency	BSF	6.54
	Fundamental Train Frequency	FTF	0.42
Dynamic Specifications of the REB (kN)	Static Load Capacity	C ₀	9.15
	Dynamic Load Capacity	C	26.5

3.2. Introduction to Data Acquisition Scenarios

For the purposes of this study, artificial faults were introduced into the REBs, and laboratory tests were conducted under four different fault conditions, including H, IRF, ORF, and REF, across 36 different operating conditions (Figure 12). It is worth mentioning that, from a practical point of view, the type of fault (fault existence on the either of elements) has little importance. On the other hand, ensuring degradation on any of the IR, OR, or RE leads to a same recommendation, that is, “Replacing the REB”. However, identifying the fault type enables experts and machine owners to make more informed and confident repair decisions

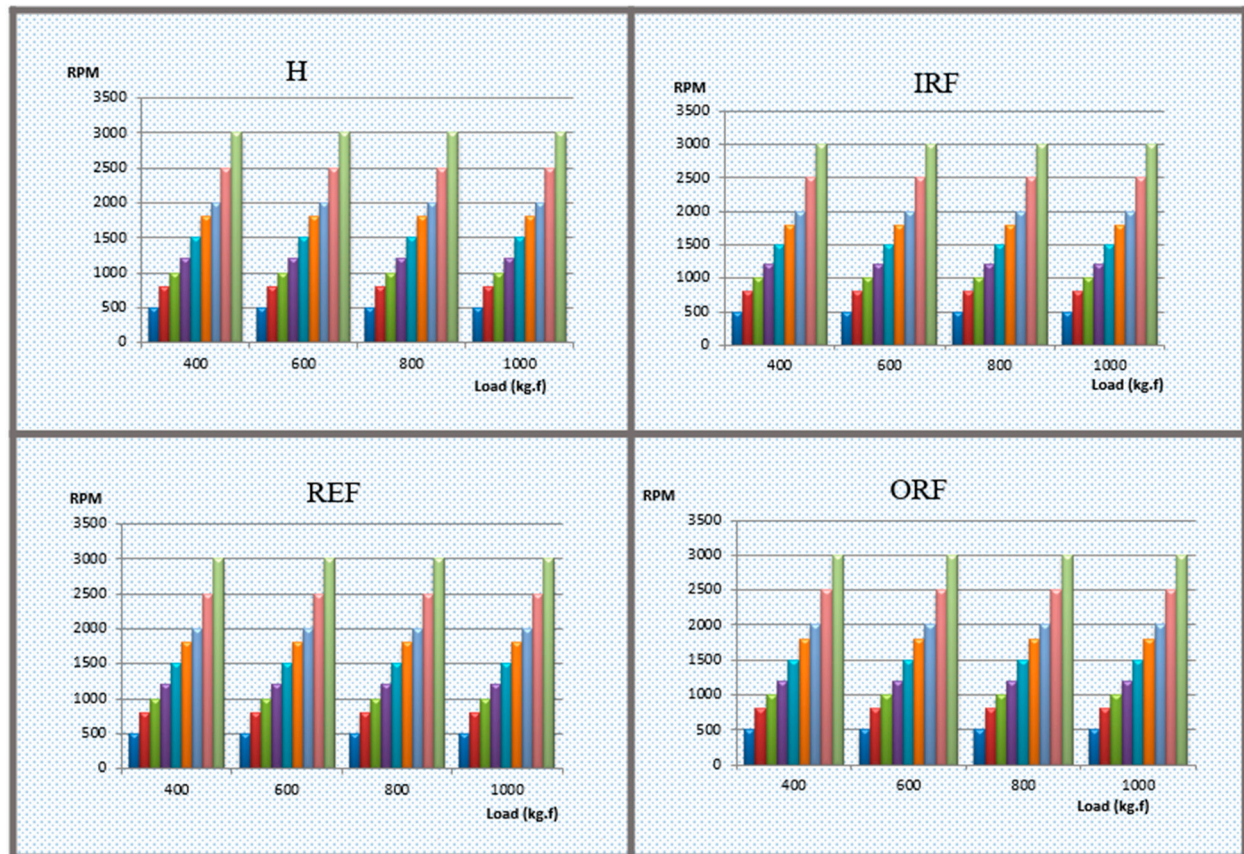


Figure 12. Number of data recorded in four conditions, nine speeds, and five different loads.

The rationale for considering variable operating conditions is to enable the developed model to perform intelligent fault detection in diverse conditions, thereby reducing its dependence on the operating environment. In these experiments, the rotational speed varied from 500 rpm to 3000 rpm and the applied load ranged from 400 kg·f to 1000 kg·f. In contrast to the CWRU dataset, where the rotational speed is limited to a range of 1720 rpm to 1797 rpm with variations of less than 5% and the radial load remains constant, this study incorporates a broader range of speed and load variations to precisely examine their impact on fault detection. This dataset facilitates a more comprehensive analysis of the performance of fault detection models under variable operating conditions. During these experiments, measurements were simultaneously recorded using three vibration sensors aligned in three orthogonal directions: vertical, horizontal, and axial. A total of 4741 datasets were recorded for each of the four fault conditions, encompassing all load and speed combinations. Each time-domain data point was recorded over a duration of 160 ms, with 4096 samples collected per measurement.

Furthermore, this dataset is planned to be publicly released in the future to support further research in fault diagnosis.

To present the results, two commonly used industrial features in fault diagnosis are employed. The first feature is the root mean square (RMS) of acceleration, which is calculated for all operating conditions and the four faults, including H, IRF, ORF, and REF. The results are shown in Figure 13, where, for each fixed radial load, the rotational speed ranges from 500 rpm to 3000 rpm. As seen in the graphs, the vibration amplitude under the ORF condition is the highest compared to the other fault conditions. The vibration amplitudes in the IRF and REF conditions are the second highest, following the ORF.

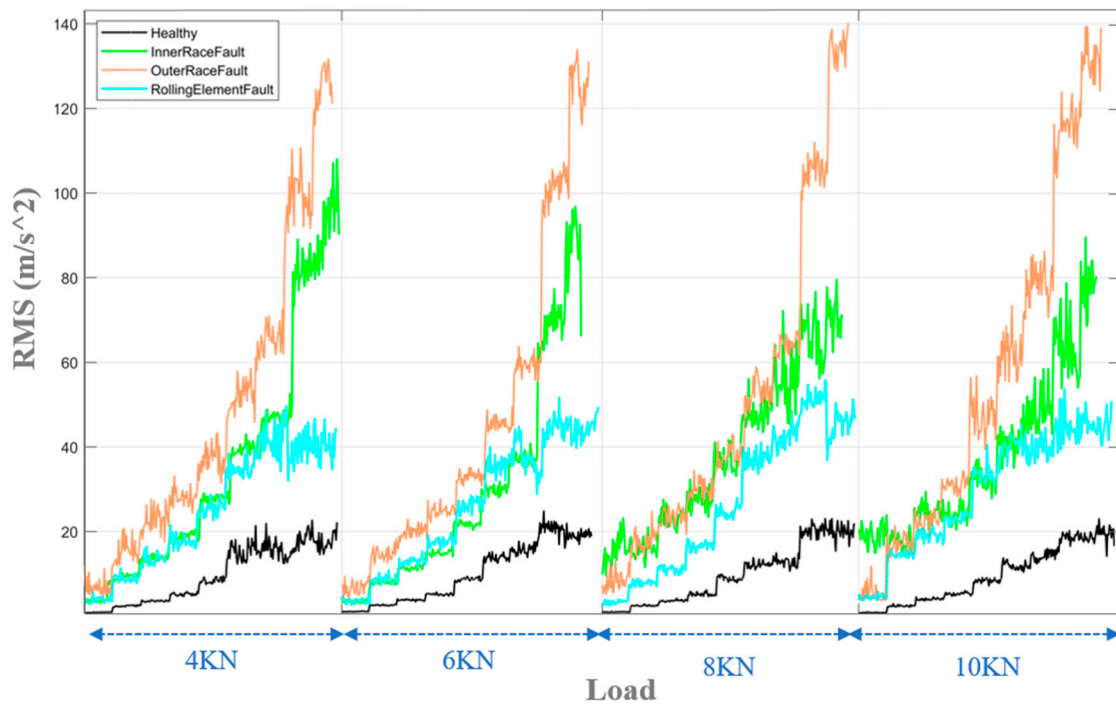


Figure 13. RMS feature representation for four REB conditions as rotational speed varies from 500 rpm to 3000 rpm at fixed radial loads.

The second feature is the peak, which is calculated for all operating conditions and the four fault scenarios. Similar results are shown in Figure 14.

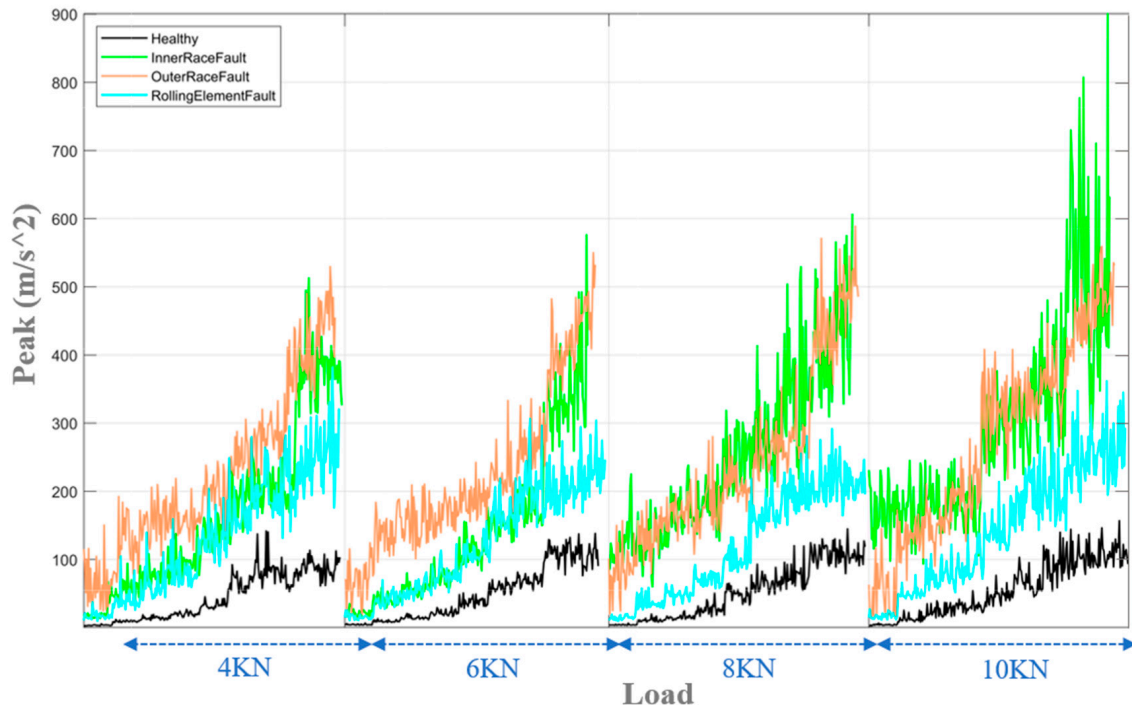


Figure 14. Peak feature representation for four REB conditions as rotational speed varies from 500 rpm to 3000 rpm at fixed radial loads.

4. Model Verification

This section discusses the model verification process. Section 4.1 focuses on the implementation of the developed model on the first laboratory dataset, describing how the

model is evaluated with these data. Section 4.2 addresses the implementation of the model on industrial data, where the model's performance is tested under real-world conditions. Section 4.3 discusses the implementation of the model on the second experimental dataset, providing an evaluation of the model's performance on new data under different conditions.

4.1. Implementation of the Developed Model on the First Laboratory Dataset

As mentioned in previous sections, CNNs extract key features from signals associated with different REB faults and classify the data accordingly. To optimize network performance, various inputs, including time-domain signals, frequency spectra, and envelope spectra, are used across all operating conditions and four REB fault states. To systematically evaluate the performance of the SCNN and TL-CNN models under various operating conditions, ten distinct experimental scenarios were defined, as illustrated in Figure 15. Each scenario involves specific combinations of rotational speeds and loading conditions across four REB health conditions (H, IRF, ORF, and REF). The goal is to simulate domain shifts and evaluate the generalization ability of the models in unseen conditions. For example, in Scenario 1, the SCNN model was trained using data collected under two loading conditions (4 kN and 8 kN) across all available rotational speeds. This includes data from all four REB health conditions. The trained SCNN model was then tested using data from two unseen loading conditions (6 kN and 10 kN), using rotational speeds of 500, 800, 1000, 2000, 2500, and 3000 RPM. In contrast, for the TL-CNN model under the same scenario, the initial convolutional layers were transferred from the pre-trained SCNN model, and only the fully connected layers were fine-tuned. Fine-tuning was performed using a small subset of data collected under the same target domain conditions (i.e., 6 kN and 10 kN) but at intermediate rotational speeds (1200, 1500, and 1800 RPM, specifically). After fine-tuning, the TL-CNN model was evaluated on the same test set used for the SCNN model. This structure was repeated consistently across all ten scenarios, ensuring that each scenario simulates a realistic domain shift to assess the effectiveness of transfer learning. The diversity in operating conditions across scenarios helps demonstrate the model's ability to generalize across different mechanical and environmental variations.

In each of the ten scenarios, the dataset was systematically divided into training and testing subsets based on varying operational conditions. As illustrated in Figure 15, the distribution of training and testing data is defined for each scenario, with varying proportions across different cases. Additionally, to ensure robust model evaluation, 10% of the training data in each scenario was allocated for validation.

The performance of the SCNN and TL-CNN models for REB fault diagnosis was evaluated using three types of input data, including time-domain signals, FFT, and signal envelopes, across 10 different scenarios. The results are presented in Tables 4 and 5. According to the findings in Table 4, the performance of the SCNN model is input-dependent. Among the inputs, the signal envelope achieved the highest accuracy of 79.8%, followed by FFT with 76.7% and time-domain signals with 65.1%. The SCNN model faced challenges in processing time-domain signals, particularly in complex scenarios. In contrast, the TL-CNN model demonstrated higher accuracy, especially with the signal envelope input, reaching 90.8%. The FFT input also performed better in TL-CNN (88.9%) compared to SCNN. Even with time-domain signals, TL-CNN outperformed SCNN with an accuracy of 79.2%. These improvements are attributed to TL-CNN's adaptability and its effective utilization of both training and new data. Overall, the TL-CNN model, particularly with signal envelope and FFT inputs, outperformed the SCNN model. The results indicate that TL enhances fault detection performance by extracting significant features and optimizing learning across diverse datasets.

Training Data	
Retraining Data	
Test Data	

Scenario 1					Scenario 2					Scenario 3					Scenario 4					Scenario 5				
Motor Speed (RPM)	Load (KN)				Motor Speed (RPM)	Load (KN)				Motor Speed (RPM)	Load (KN)				Motor Speed (RPM)	Load (KN)				Motor Speed (RPM)	Load (KN)			
	4	6	8	10		4	6	8	10		4	6	8	10		4	6	8	10		4	6	8	10
500					500					500					500					500				
800					800					800					800					800				
1000					1000					1000					1000					1000				
1200					1200					1200					1200					1200				
1500					1500					1500					1500					1500				
1800					1800					1800					1800					1800				
2000					2000					2000					2000					2000				
2500					2500					2500					2500					2500				
3000					3000					3000					3000					3000				
Scenario 6					Scenario 7					Scenario 8					Scenario 9					Scenario 10				
Motor Speed (RPM)	Load (KN)				Motor Speed (RPM)	Load (KN)				Motor Speed (RPM)	Load (KN)				Motor Speed (RPM)	Load (KN)				Motor Speed (RPM)	Load (KN)			
	4	6	8	10		4	6	8	10		4	6	8	10		4	6	8	10		4	6	8	10
500					500					500					500					500				
800					800					800					800					800				
1000					1000					1000					1000					1000				
1200					1200					1200					1200					1200				
1500					1500					1500					1500					1500				
1800					1800					1800					1800					1800				
2000					2000					2000					2000					2000				
2500					2500					2500					2500					2500				
3000					3000					3000					3000					3000				

Figure 15. Ten different scenarios have been defined for the partitioning of training, testing, and retraining datasets.

Table 4. Accuracy of SCNN with various inputs, including time signal, FFT and envelope on first laboratory data.

Accuracy of Test Data for 10 Scenarios											
Input	Scenario 1	Scenario 2	Scenario 3	Scenario 4	Scenario 5	Scenario 6	Scenario 7	Scenario 8	Scenario 9	Scenario 10	Average
Time wave	75.1	73.2	76.2	51.2	57.2	67.1	62.5	48.9	59.7	80	65.1
FFT	82.2	81.4	84.1	73.2	74.2	75.2	69.4	64.5	73.2	90.3	76.7
Envelope	85.3	84.1	86.1	78.1	75.3	80.3	72.6	67.7	76.1	92.5	79.8

Table 5. Accuracy of TL-CNN with various inputs, including time signal, FFT, and envelope on first laboratory data.

Accuracy of Test Data for 10 Scenarios											
Input	Scenario 1	Scenario 2	Scenario 3	Scenario 4	Scenario 5	Scenario 6	Scenario 7	Scenario 8	Scenario 9	Scenario 10	Average
Time wave	82.1	81.2	84.2	75.2	77.3	79.6	78.2	72.3	74.3	87.5	79.2
FFT	89.2	88.6	91.3	88.2	89.7	88.9	87.3	85	88.1	93.3	88.9
Envelope	91.3	90.2	92.1	90.8	89.4	92.5	90.4	86.9	89.1	95.9	90.8

Figure 16 presents the confusion matrix results for the SCNN and TL-CNN models in REB fault detection using time-domain signals, FFT, and signal envelopes. The SCNN model shows lower accuracy with time-domain signals due to more off-diagonal errors. Accuracy improves with FFT, while the signal envelope achieves the best performance. The TL-CNN model significantly enhances classification accuracy across all inputs. Errors decrease notably with time-domain signals, and FFT input yields better feature extraction. The signal envelope input in TL-CNN shows the highest accuracy, demonstrating

strong generalization. Overall, Figure 16 confirms that TL-CNN outperforms SCNN, especially with the signal envelope, which provides key diagnostic information and minimizes prediction errors.

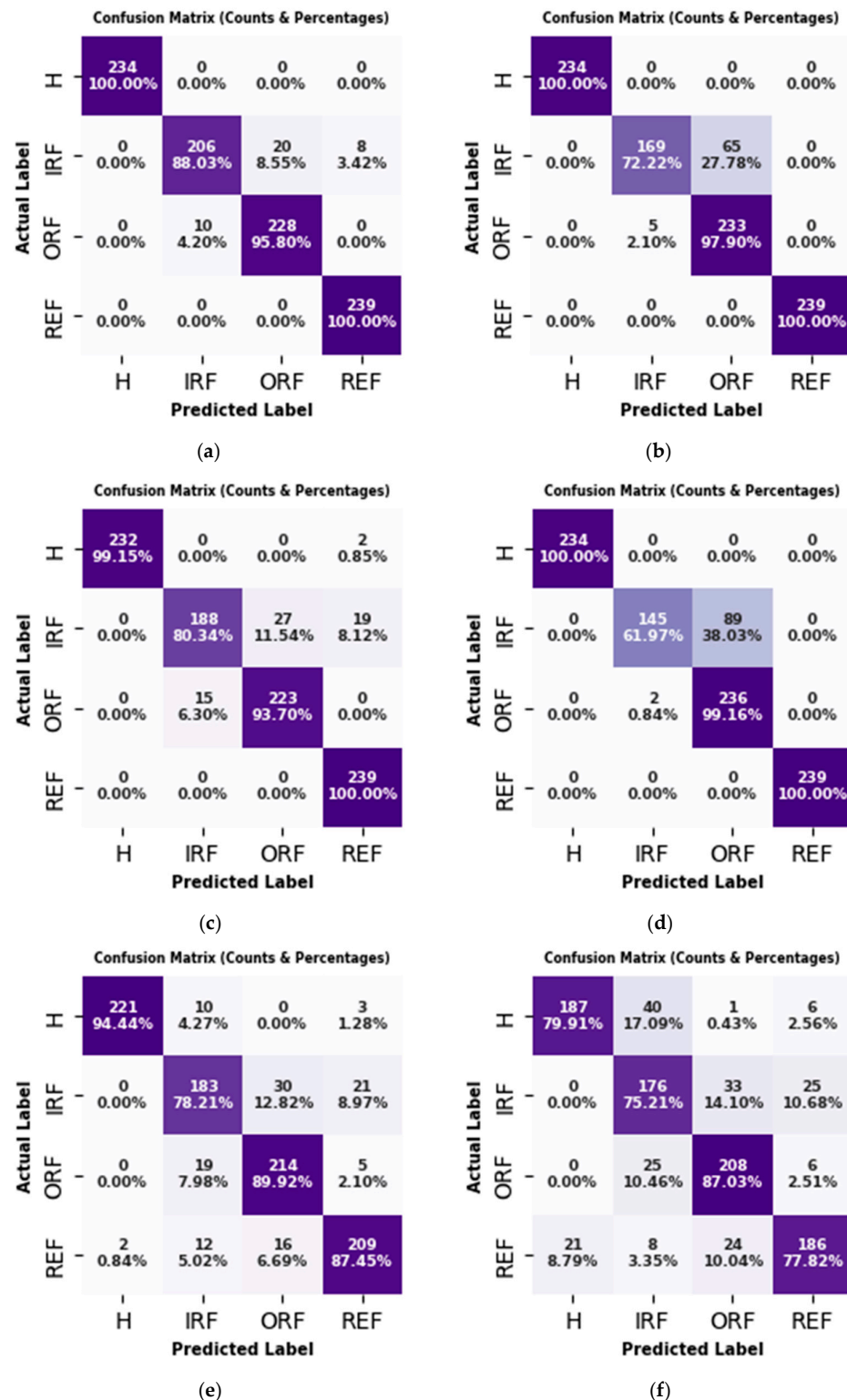


Figure 16. The results correspond to Scenario 10 and include confusion matrices for the test data obtained from both the SCNN and TL-CNN models using first laboratory data, with the following input types: (a) TL-CNN model with envelope input, (b) SCNN model with envelope input, (c) TL-CNN model with FFT input, (d) SCNN model with FFT input, (e) TL-CNN model with time-domain wave input, and (f) SCNN model with time-domain wave input.

As mentioned in previous sections, the performance of the SCNN and TL-CNN models in detecting REB faults was evaluated using various types of input data under diverse operational conditions. To investigate the phenomenon of overfitting, changes in the loss values (error) for training and validation data were closely monitored throughout the training process. For example, as shown in Figure 17, this plot corresponds to the tenth scenario with envelope input for the SCNN model. In this plot, it can be observed that the training loss decreases steadily, while the validation loss either plateaus or slightly increases after a certain point. These changes indicate that the model is learning features more accurately from the training data, but its ability to generalize to the validation data is diminishing, which is commonly recognized as a sign of overfitting. To address this issue and improve the model's generalization ability, the early stopping technique was employed. In this method, the “patience” parameter was set to 3 (i.e., stopping the training process after three consecutive epochs without improvement in validation loss), thus preventing further training when no improvement was observed in the validation data. The value of “3” for patience was chosen after various evaluations and tests, considering different parameter configurations. This value allows the model to account for the natural fluctuations in the data and the changes in the validation loss, thereby preventing premature stopping of the training process. At the same time, this value helps to prevent overfitting and ensures that the model stops training before it overfits the training data. Consequently, the model can effectively learn the features and maintain its performance on unseen data. As shown in Figure 17, this strategy led to stopping the training process before the model overfit the training data, thus preventing overfitting.

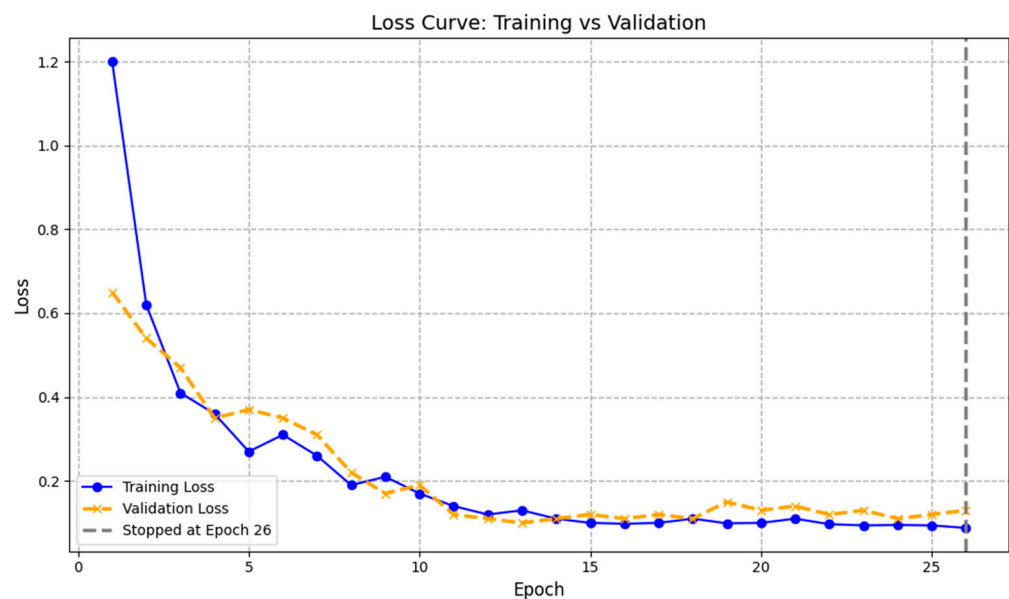


Figure 17. Training and validation loss curve for SCNN model in Scenario 10 using envelope input (with early stopping applied).

This section evaluates the performance of the TL-CNN and SCNN models in diagnosing REB faults using three key metrics, including sensitivity, specificity, and false negative rate (FNR). The analysis is based on average results from ten laboratory scenarios, summarized in Table 6. The TL-CNN model outperforms SCNN across all metrics. It achieves higher sensitivity, correctly identifying faulty data with the highest value of 98.9% for envelope input compared to 93.2% for SCNN. Similar trends are observed for FFT (94.1% vs. 88.6%) and time-domain signals (87.2% vs. 79.4%). For specificity, TL-CNN performs better in classifying healthy data, achieving 98.1% for envelope input, 97.2% for FFT, and

91.5% for time-domain signals, higher than SCNN in all cases. The FNR, reflecting the misclassification of faulty data as healthy, is significantly lower in TL-CNN. The lowest FNR of 1.1% is seen for envelope input, compared to 6.8% for SCNN. Similar reductions are observed for FFT (5.9% vs. 11.4%) and time-domain signals (12.8% vs. 20.6%). Overall, Table 6 confirms that TL-CNN, especially with envelope input, delivers superior fault detection performance due to its enhanced feature extraction and classification accuracy.

Table 6. Performance evaluation of TL-CNN and SCNN models using sensitivity, specificity, and FNR metrics based on average results from ten laboratory scenarios.

Metric	Sensitivity of Faulty Data ($\frac{TP}{TP+FN}$)		Specificity of Healthy Data ($\frac{TN}{TN+FP}$)		False Negative Rate of Faulty Data ($\frac{FN}{TP+FN}$)	
	TL_CNN	SCNN	TL_CNN	SCNN	TL_CNN	SCNN
Time wave	87.2	79.4	91.5	81.1	12.8	20.6
FFT	94.1	88.6	97.2	89.2	5.9	11.4
Envelope	98.9	93.2	98.1	92.1	1.1	6.8

4.2. Implementation of the Developed Model on Industrial Data

For evaluating the performance of the developed model, industrial data with a completely different distribution from laboratory data were used. In this regard, industrial data related to REBs used in various industries such as mining, refining, wood, and paper, among others, were collected. These data encompass a wide range of REBs with different models, including ball, cylindrical, barrel, single-row, double-row, fixed and self-aligning, one-piece and two-piece, and lubricated with grease or oil, etc. These data also include different operating conditions such as speed, loading, mounting orientation (horizontal or vertical), and so on. Figure 18 shows an example of the faults in these industrial REBs, which were visually observable after replacement and cutting. A total of 88 data points were collected for four conditions, including healthy, ORF, IRF, and REF (22 signals for each condition).

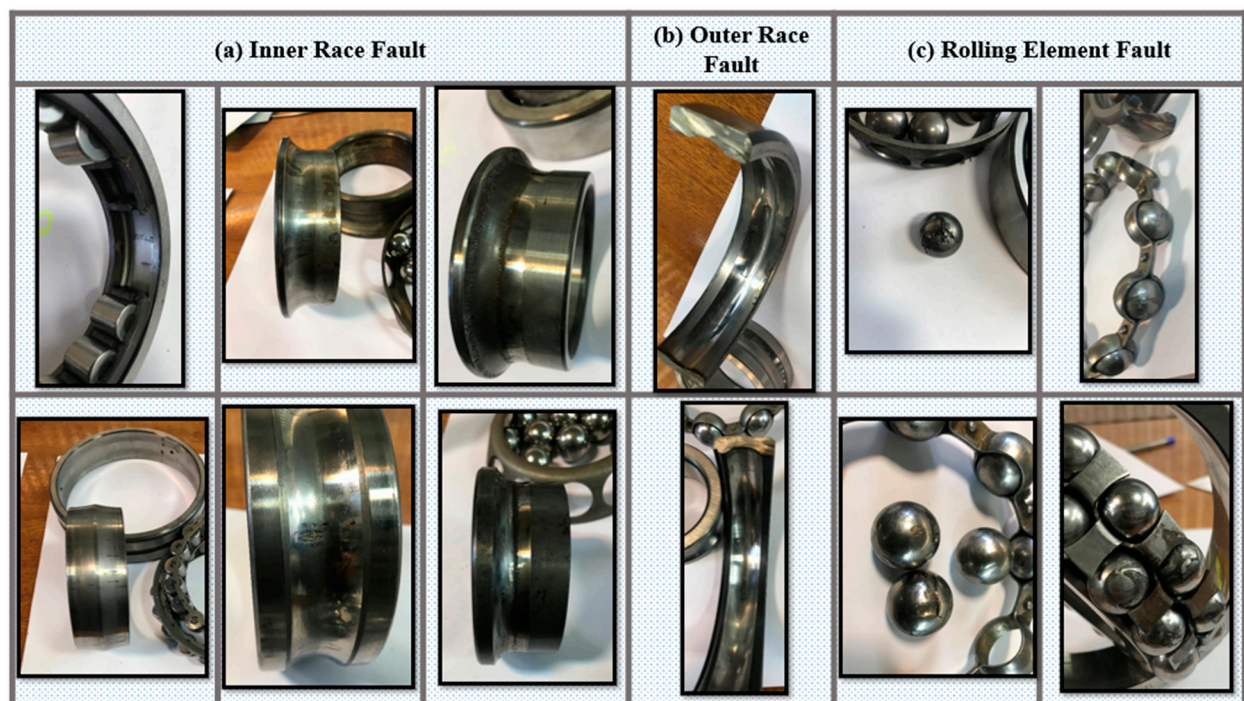


Figure 18. Visual inspection results of some industrial REBs, including (a) IRF, (b) ORF, and (c) REF.

In this study, the performance of two models, the SCNN and the proposed TL-CNN, were evaluated for fault detection in REBs using three different inputs, namely time-domain signals, FFT, and signal envelope, across the 10 scenarios mentioned previously with industrial data. The industrial dataset consists of 88 samples, divided into four conditions, including H, ORF, IRF, and REF (22 samples for each condition). For model fine-tuning, five samples from each class were used, and the remaining data were utilized for evaluation in the testing step.

In industrial data, the sampling frequencies significantly differ from those used in laboratory data. To address this issue and align the data, the resampling technique is employed to adjust the number of samples in time-domain signals. Resampling refers to the process of altering the sampling frequency of a signal. In this process, the signal is first transformed into the frequency domain, and then the number of samples is adjusted using appropriate methods. After these adjustments, the signal is transformed back to the time domain to preserve its overall structure. This process is carried out using anti-aliasing filters to minimize aliasing effects and ensure that the signal's information remains undistorted.

According to Table 7, the SCNN model with the signal envelope input achieved the best performance, with an average accuracy of 77.3%, while the FFT and time-domain signal inputs provided accuracies of 73.8% and 60.1%, respectively. The model demonstrated the weakest performance with time-domain signals. In contrast, as shown in Table 8, the TL-CNN model with industrial data and the signal envelope input achieved a higher accuracy of 86.4%, with the FFT and time-domain signal inputs showing accuracies of 84.3% and 65.7%, respectively. These results emphasize the effectiveness of the TL-CNN model, particularly in extracting key features and its ability to generalize better across diverse data with different distribution.

Table 7. Accuracy of SCNN with various inputs, including time signal, FFT and envelope on industrial data.

Input	Accuracy of Test Data for 10 Scenarios										Average
	Scenario 1	Scenario 2	Scenario 3	Scenario 4	Scenario 5	Scenario 6	Scenario 7	Scenario 8	Scenario 9	Scenario 10	
Time wave	66.7	68.4	67.1	50.1	61.2	62.3	64.5	41.4	49.2	70.5	60.1
FFT	78.1	76.2	76.9	77.2	72.1	70	65.4	67.1	74.2	80.8	73.8
Envelope	79.3	80.1	82.4	76.1	75.3	78.5	68.4	69.2	77.1	86.7	77.3

Table 8. Accuracy of TL-CNN with various inputs, including time signal, FFT, and envelope on industrial data.

Input	Accuracy of Test Data for 10 Scenarios										Average
	Scenario 1	Scenario 2	Scenario 3	Scenario 4	Scenario 5	Scenario 6	Scenario 7	Scenario 8	Scenario 9	Scenario 10	
Time wave	72.2	70.4	71.3	59.2	61.3	63.1	61.8	62.2	61.2	74.1	65.7
FFT	86.2	87.3	89.1	83.2	82.6	83.9	80.1	80.4	77.3	93.1	84.3
Envelope	89.1	88.4	90	84.8	85.1	86.3	82.1	81.4	83.2	94.1	86.4

Figure 19 shows the performance of the SCNN and TL-CNN models for REB fault detection using industrial data. With time-domain signals, SCNN had many incorrect predictions, indicating low accuracy, while TL-CNN performed better, though its accuracy was still lower than other inputs. The FFT input improved results for both models, with

more predictions along the main diagonal and fewer errors. The TL-CNN model excelled with this input by extracting key frequency features. Finally, the signal envelope input provided the best performance, with TL-CNN showing the highest accuracy and fewest errors, highlighting its superiority in REB fault detection.

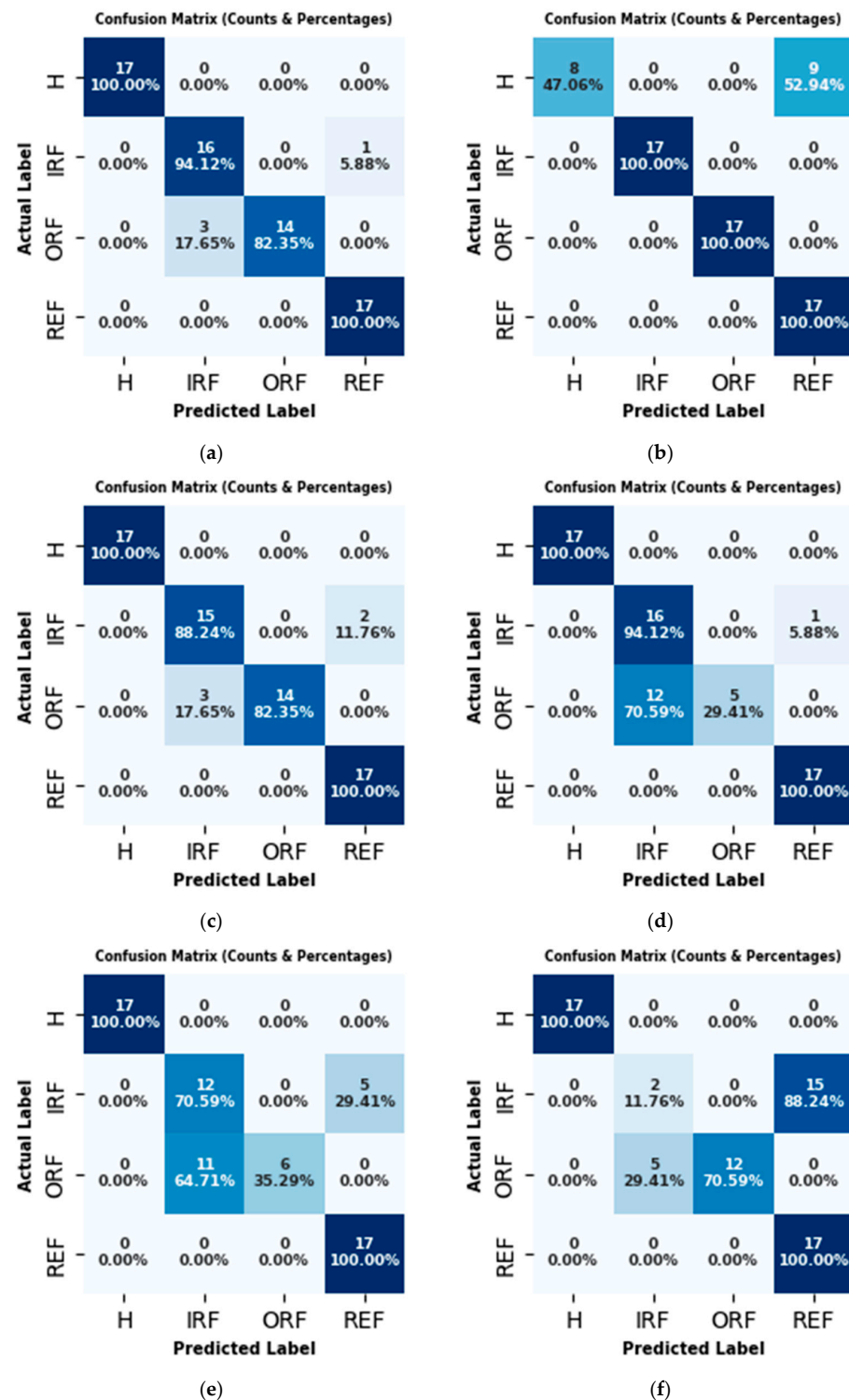


Figure 19. The results correspond to Scenario 10 and include confusion matrices for the test data obtained from both SCNN and TL-CNN models using industrial data, with the following input types: (a) TL-CNN model with envelope input, (b) SCNN model with envelope input, (c) TL-CNN model with FFT input, (d) SCNN model with FFT input, (e) TL-CNN model with time-domain wave input, and (f) SCNN model with time-domain wave input.

In Table 9, the performance of the TL-CNN and SCNN models for REB fault detection is evaluated based on industrial data. The results show that the TL-CNN model outperforms the SCNN model across all input types. Specifically, TL-CNN demonstrates superior sensitivity, specificity, and lower false negative rates, particularly when using the envelope input. Overall, TL-CNN proves to be more effective in detecting REB faults in industrial data.

Table 9. Performance evaluation of TL-CNN and SCNN models using sensitivity, specificity, and FNR metrics based on average results from ten scenarios on industrial data.

Metric	Sensitivity of Faulty Data ($\frac{TP}{TP+FN}$)		Specificity of Healthy Data ($\frac{TN}{TN+FP}$)		False Negative Rate of Faulty Data ($\frac{FN}{TP+FN}$)	
	TL_CNN	SCNN	TL_CNN	SCNN	TL_CNN	SCNN
Input						
Time wave	86.4	77.4	82.2	78.9	13.6	22.6
FFT	92.8	85.1	89.1	84.4	7.2	14.9
Envelope	98.1	93.2	91.9	88.1	1.9	6.8

4.3. Implementation of the Developed Model on the Second Experimental Dataset

For the re-evaluation of the developed model's performance, data from another set of laboratory experiments, which have a completely different distribution compared to the initial laboratory data, were used. These data were collected at the Vibration Laboratory of Sharif University of Technology for the purpose of studying the remaining useful life of REBs. To achieve the objectives of the study, a series of accelerated life tests for REBs were designed on a laboratory platform (Figure 20). To this aim, eight accelerated life tests for REBs were conducted on this platform. In these tests, measurements were made using a vibration sensor on the REBs. The REBs under study were subjected to severe loading, and their vibration data were recorded from the moment of installation to the final failure.

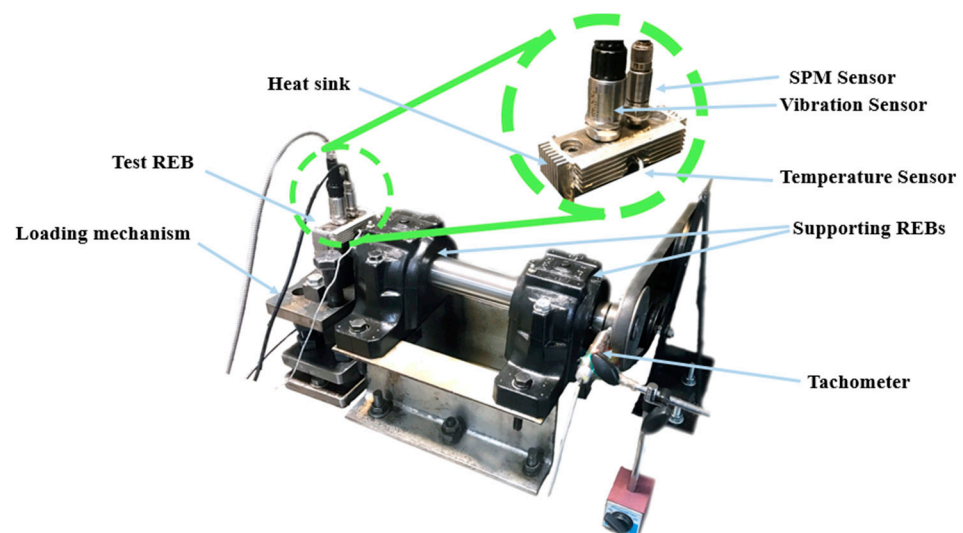


Figure 20. Accelerated life testing conducted on the REBs experimental test rig.

In this platform, shown in Figure 20, a vibration sensor and a temperature sensor were mounted on the test REB. The REB used in these experiments is deep groove with the code 6907. The accelerated life tests were conducted under constant operating conditions with a speed of 2000 rpm and a radial load of 9000 N. The sampling frequency of the accelerometer sensor was 25.6 kHz. Figure 21 shows the observed damage in the components of each REB after disassembly at the end of each test.

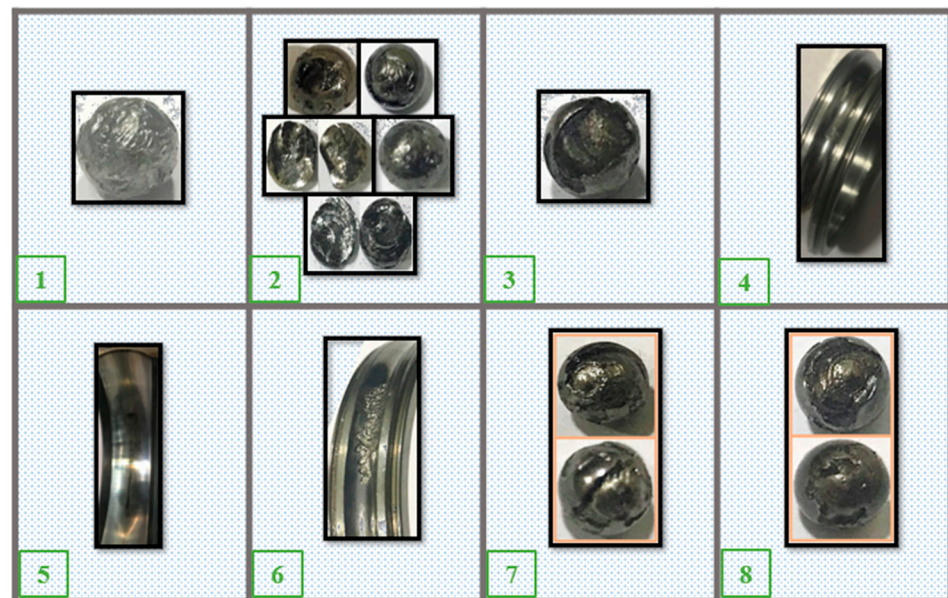


Figure 21. Visual inspection results of faults on the elements of accelerated life test REBs.

The trend of RMS for eight laboratory tests from the start to the end of life, are shown in Figure 22.

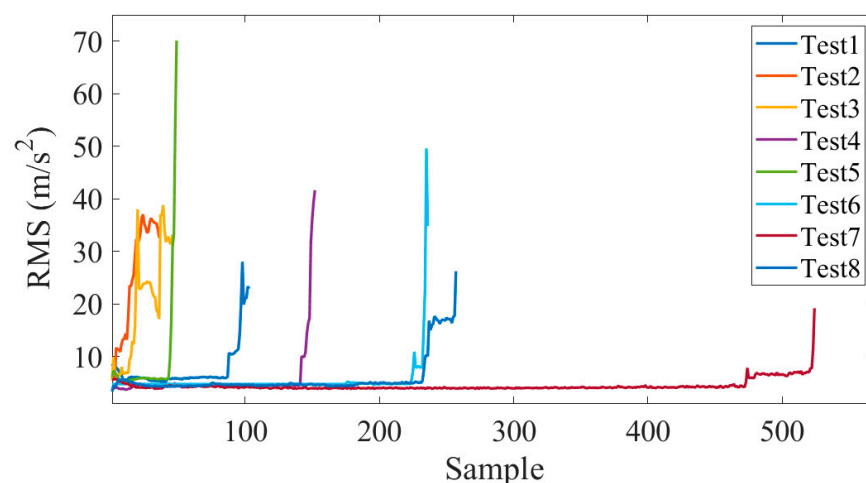


Figure 22. The trend of RMS for eight accelerated life tests.

In this study, the performance of two models, the conventional SCNN and the proposed TL-CNN, was evaluated for REB fault detection using three inputs, including time-domain signal, FFT, and signal envelope, with a second laboratory dataset of 75 samples. Five samples from each REB condition (H, ORF, IRE, and REF) were used for re-training, and the remaining samples were used for testing. The results, as presented in Table 10, indicate that the SCNN model performs optimally with the signal envelope input, achieving an accuracy of 77.8%. The FFT input ranks second with an accuracy of 73.6%, while the time-domain signal input demonstrates the lowest performance at 68.1% accuracy. In contrast, for the TL-CNN model, as illustrated in Table 11, the signal envelope input yields the highest performance, reaching an accuracy of 88.5%, followed by FFT at 86.8%, and time-domain signals at 74.7%. These findings clearly demonstrate that the TL-CNN model outperforms the SCNN model across all input types, particularly when utilizing the signal envelope. Furthermore, the TL-CNN model exhibits significant potential for industrial applications due to its ability to generalize to new, unseen data with minimal retraining.

Table 10. Accuracy of SCNN with various inputs, including time signal, FFT and envelope on second laboratory data.

Input	Accuracy of Test Data for 10 Scenarios										Average
	Scenario 1	Scenario 2	Scenario 3	Scenario 4	Scenario 5	Scenario 6	Scenario 7	Scenario 8	Scenario 9	Scenario 10	
Time wave	72.2	71.1	70	66.4	65.1	64.2	65.3	66.1	63.2	78.1	68.1
FFT	76.5	75.1	77.3	74.2	73.1	72.1	69.4	65.2	71.4	81.8	73.6
Envelope	80.2	79.1	81.2	79.4	76.6	80.1	72.4	68.2	74.1	87.2	77.8

Table 11. Accuracy of TL-CNN with various inputs, including time signal, FFT and envelope on second laboratory data.

Input	Accuracy of Test Data for 10 Scenarios										Average
	Scenario 1	Scenario 2	Scenario 3	Scenario 4	Scenario 5	Scenario 6	Scenario 7	Scenario 8	Scenario 9	Scenario 10	
Time wave	82.7	80.5	79.2	72.9	69.4	74.3	66.5	67.2	65.6	89	74.7
FFT	89.1	90.2	89.2	85.1	84.1	84.2	85.6	83.7	83.1	93.7	86.8
Envelope	91.2	92.8	90.1	86.7	85.4	87.3	86.1	85.1	84.7	96.3	88.5

Figure 23 shows the performance of the SCNN and TL-CNN models in fault detection of REBs using the second laboratory dataset. For time-domain signal input, CNN showed low accuracy, while TL-CNN performed better, with more predictions on the main diagonal. The FFT input improved accuracy for both models, with TL-CNN concentrating more predictions along the diagonal and reducing errors. The best performance was achieved with the signal envelope input, where TL-CNN demonstrated superior accuracy, correctly classifying most cases and significantly minimizing errors.

Table 12 presents the average performance results of the TL-CNN and SCNN models across 10 laboratory scenarios using the second experimental dataset. The TL-CNN model consistently outperforms the SCNN model across all evaluation metrics and input types. Specifically, for the signal envelope input, the TL-CNN model demonstrates higher sensitivity and specificity while maintaining a lower false negative rate compared to the SCNN model. A similar trend is observed for the FFT and time-domain signal inputs, where TL-CNN achieves superior performance. These findings emphasize the effectiveness of the TL-CNN model in fault detection, particularly in adapting to diverse laboratory conditions.

Table 12. Performance evaluation of TL-CNN and SCNN models using sensitivity, specificity, and FNR metrics based on average results from ten scenarios on the second laboratory dataset.

Metric	Sensitivity of Faulty Data ($\frac{TP}{TP+FN}$)		Specificity of Healthy Data ($\frac{TN}{TN+FP}$)		False Negative Rate of Faulty Data ($\frac{FN}{TP+FN}$)	
	TL_CNN	SCNN	TL_CNN	SCNN	TL_CNN	SCNN
Time wave	86.1	73.2	84.2	76.1	13.9	26.8
FFT	93.2	87.1	90.1	85.1	6.8	12.9
Envelope	98.4	92.2	92.8	89.1	1.6	7.8

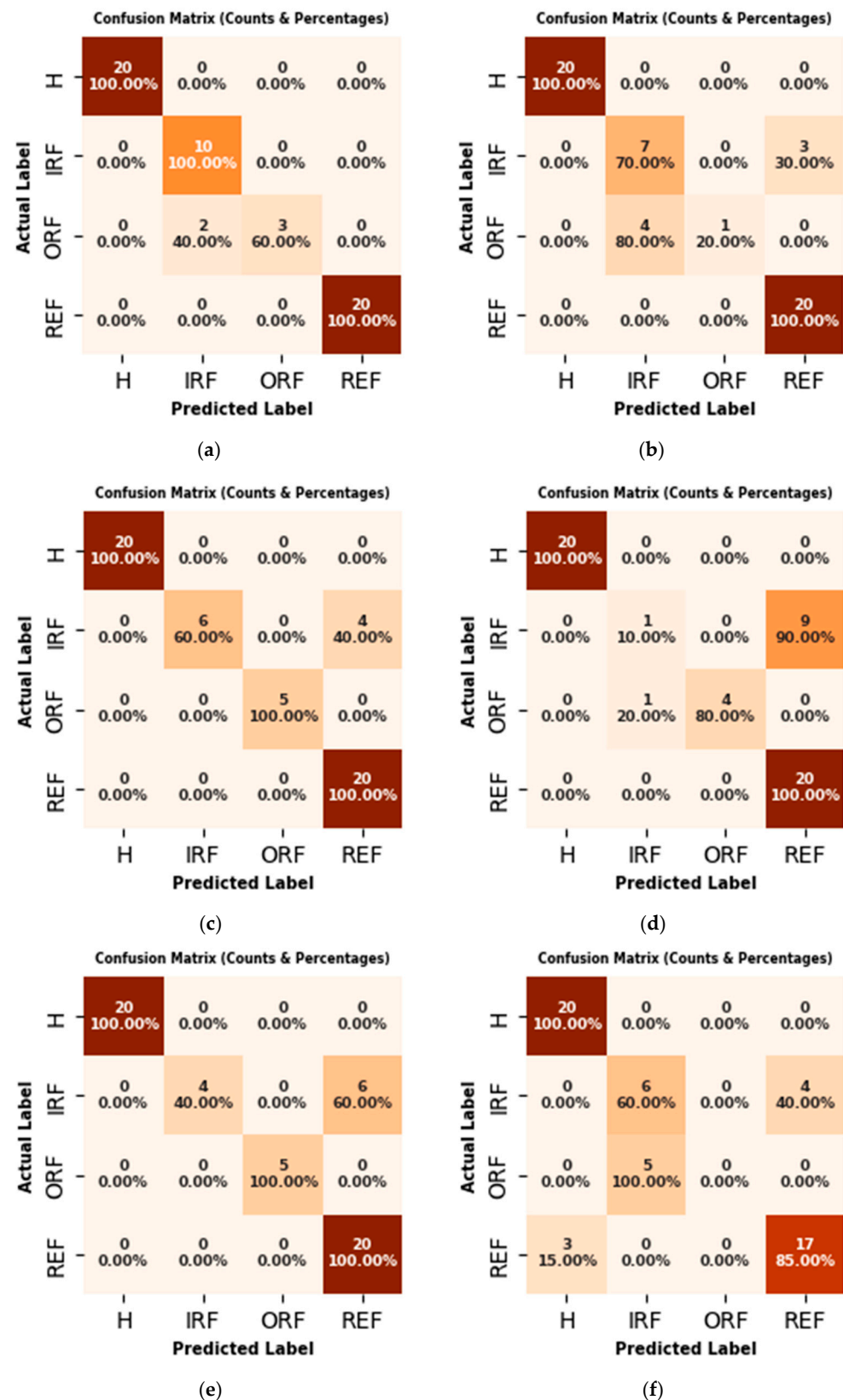


Figure 23. The results correspond to Scenario 10 and include confusion matrices for the test data obtained from both SCNN and TL-CNN models using experimental data, with the following input types: (a) TL-CNN model with envelope input, (b) SCNN model with envelope input, (c) TL-CNN model with FFT input, (d) SCNN model with FFT input, (e) TL-CNN model with time-domain wave input, and (f) SCNN model with time-domain wave input.

5. Discussion and Future Research Directions

Section 5 presents the analysis and conclusions of the study, with Section 5.1 focusing on the discussion of the results and Section 5.2 addressing the limitations and future research directions.

5.1. Discussion

In the evaluation of the performance of the SCNN and TL-CNN models using three different input types, namely time-domain signals, FFT, and signal envelope, as illustrated in Figures 24–26 with box plot diagrams, it was consistently observed that the TL-CNN model outperformed the SCNN model across all scenarios. This performance advantage was particularly evident when applied to industrial datasets, which exhibit distinct characteristics compared to laboratory data. The TL-CNN model effectively leveraged TL techniques, enabling the extraction of critical features related to REB faults with higher accuracy.

The presented box plot diagrams provide a detailed representation of the distribution of accuracies for each input type. Notably, the signal envelope input demonstrated the best performance in both models, with the proposed TL-CNN method achieving the highest accuracy in fault detection. In contrast, the time-domain signal input exhibited the poorest performance, likely due to its limited ability to extract discriminative features required for accurate fault classification. Additionally, the mean accuracy values for each input type are clearly depicted in the diagrams, effectively illustrating the overall performance trends of the models.

Regarding the evaluation metrics, including sensitivity, specificity, and FNR, the TL-CNN model consistently outperformed the SCNN model across all input types and datasets. The superior sensitivity of the TL-CNN model indicates its enhanced ability to correctly identify faulty data instances. Additionally, the TL-CNN model exhibited improved specificity, highlighting its effectiveness in distinguishing healthy data more accurately. Furthermore, the TL-CNN model recorded a significantly lower FNR, indicating a reduced likelihood of misclassifying faulty data as healthy. This performance discrepancy was particularly pronounced when the signal envelope was used as input, where TL-CNN demonstrated superior results across all three metrics. These findings emphasize the power of transfer learning in generalizing across diverse data distributions and conditions, highlighting its significant potential in improving intelligent fault detection systems, particularly in applications involving unseen and varying data.

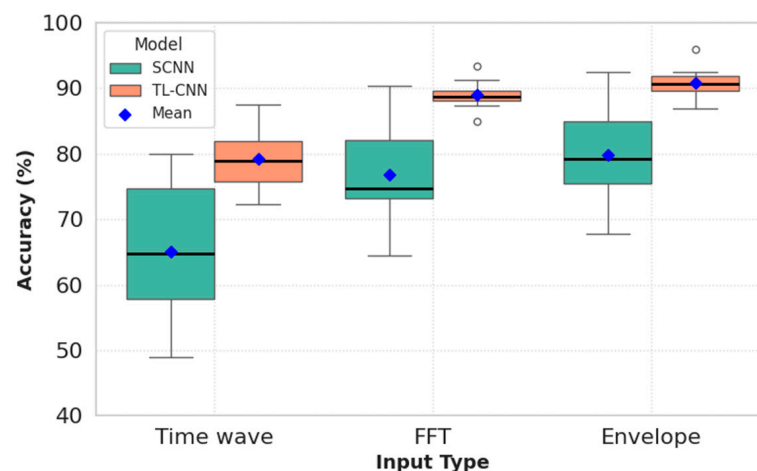


Figure 24. Box plot of the average accuracy of the two models for ten scenarios with three different inputs for the first laboratory dataset.

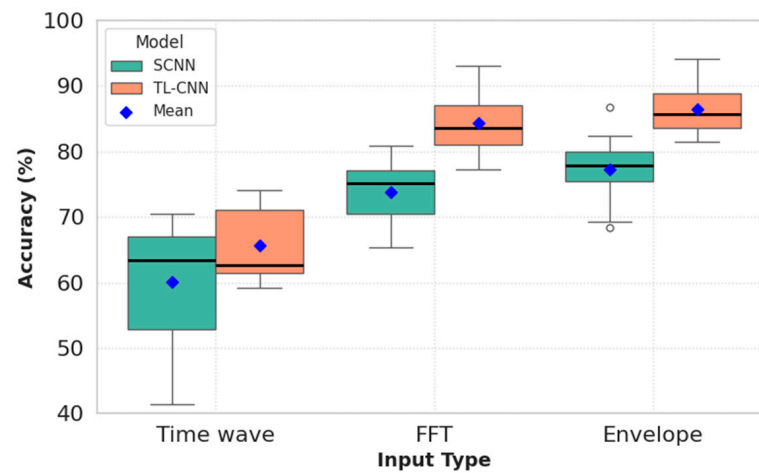


Figure 25. Box plot of the average accuracy of the two models for ten scenarios with three different inputs for the industrial dataset.

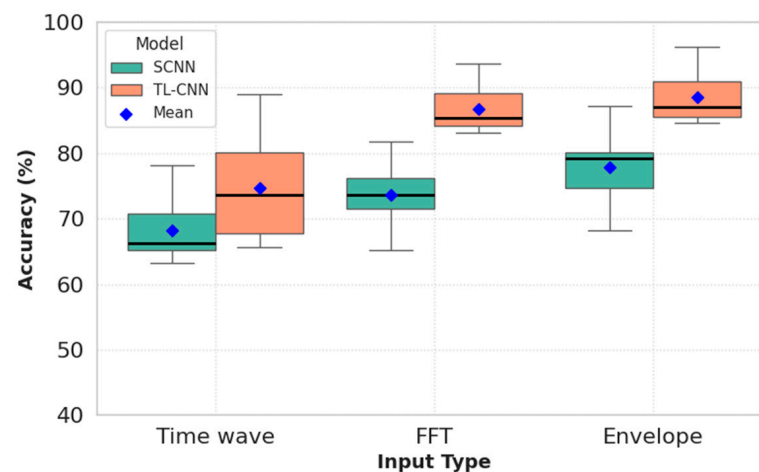


Figure 26. Box plot of the average accuracy of the two models for ten scenarios with three different inputs for the second laboratory dataset.

Furthermore, as explained in the manuscript, transfer learning was employed in the TL-CNN model to address domain shifts in the data. In this model, the weights of the convolutional layers from the pre-trained SCNN model (trained on the initial laboratory data) were used as the starting weights. Then, only the fully connected layers were fine-tuned using the new data. These new data included samples from both the industrial dataset and the second laboratory dataset, each containing various samples from faulted and healthy conditions. This approach allowed the model to retain the general features extracted from the initial laboratory data while only updating the final layers of the network to adapt to the specific features of the new datasets. This process was applied to both the industrial and second laboratory datasets across all ten predefined experimental scenarios.

5.2. Limitation and Future Research Directions

The models presented in this study are primarily designed for the detection of independent faults. However, in some industrial equipment, REBs may simultaneously experience multiple types of faults, which can potentially impact the accuracy of the models. Additionally, collecting real-world data that involves combinations of multiple faults in a single bearing is challenging, as such faults rarely occur under actual operating conditions. To overcome these limitations, it is suggested that future research focus on the collection of combined fault data. These data could either be gathered from specific in-

dustrial equipment prone to such conditions or generated through accurate simulations in controlled environments. The development of models capable of detecting combined faults could significantly enhance fault detection accuracy in industrial applications, leading to improved monitoring systems and reduced operational costs.

6. Conclusions

This study evaluated and compared the performance of two deep learning models, SCNN and TL-CNN, for fault detection of REB. Three different input types, including time-domain signals, FFT, and signal envelope, were utilized to assess the models. The results demonstrated that the TL-CNN with fine-tuning outperformed the SCNN model, significantly improving fault detection accuracy. Transfer learning enables the model to leverage features extracted from diverse datasets and effectively adapt to new and more complex data, which is particularly beneficial in enhancing fault detection performance under variable operating conditions.

Another key innovation of this study was the introduction of an algorithm for identifying resonance zones to determine the optimal frequency range for applying the envelope technique. By focusing on extracting fault-specific frequency features and reducing noise effects, this algorithm effectively identified resonance zones, thereby improving fault detection accuracy. Additionally, a preprocessing step was introduced to eliminate the effects of system speed variations from the data, which played a crucial role in enhancing the precision and efficiency of the models. This step facilitated the accurate extraction of fault-related features and mitigated the influence of speed fluctuations on fault detection, ultimately improving model performance.

The findings revealed that employing TL, particularly in combination with signal envelope and FFT inputs, significantly enhanced fault detection accuracy compared to the SCNN model. The accuracy improvement in the TL-CNN model ranged between 5% and 15%, depending on the input type. Furthermore, an evaluation of model performance using statistical parameters such as sensitivity, specificity, and FNR indicated that the TL-CNN model consistently outperformed the SCNN model across all input types and scenarios. The TL-CNN model exhibited higher sensitivity and specificity, accurately identifying faulty and healthy data while maintaining a lower FNR, demonstrating its superior capability in distinguishing between defective and non-defective samples. This improvement is particularly crucial in industrial applications, where precise and timely fault detection is essential.

The findings of this study highlight that selecting appropriate input representations and employing advanced machine learning techniques, such as transfer learning and fine-tuning, can substantially enhance fault detection systems, especially when new data distributions differ significantly from the training data. These results can serve as a foundation for future research in the development of condition monitoring and fault detection systems for industrial and mechanical equipment, particularly REB.

Author Contributions: Conceptualization, A.D., M.B., H.A.A. and S.M.; Methodology, A.D., H.A.A. and S.M.; Software, A.D.; Validation, A.D.; Formal analysis, A.D., H.A.A. and S.M.; Investigation, A.D., H.A.A. and S.M.; Resources, A.D. and H.A.A.; Data curation, A.D.; Writing—original draft, A.D.; Writing—review & editing, M.B., H.A.A., S.M. and L.G.; Visualization, A.D. and S.M.; Supervision, M.B., H.A.A. and S.M.; Project administration, M.B. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Data are contained within the article.

Acknowledgments: The authors sincerely thank Behravesh Vibration Engineering Company for their valuable assistance in data collection and their technical support throughout this research. The authors also thank the AI tool for helping them check the manuscript grammatically.

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

References

1. Mobley, R.K. *An Introduction to Predictive Maintenance*; Elsevier: Amsterdam, The Netherlands, 2002.
2. Cui, W.; Meng, G.; Gou, T.; Wang, A.; Xiao, R.; Zhang, X. Intelligent rolling bearing fault diagnosis method using symmetrized dot pattern images and CBAM-DRN. *Sensors* **2022**, *22*, 9954. [[CrossRef](#)] [[PubMed](#)]
3. Bhagwan, S. A review of various techniques used for shaft failure analysis. *Asian J. Multidimens. Res.* **2021**, *10*, 387–392. [[CrossRef](#)]
4. Guo, H.; Ping, D.; Wang, L.; Zhang, W.; Wu, J.; Ma, X.; Xu, Q.; Lu, Z. Fault Diagnosis Method of Rolling Bearing Based on 1D Multi-Channel Improved Convolutional Neural Network in Noisy Environment. *Sensors* **2025**, *25*, 2286. [[CrossRef](#)]
5. Song, X.; Cong, Y.; Song, Y.; Chen, Y.; Liang, P. A bearing fault diagnosis model based on CNN with wide convolution kernels. *J. Ambient. Intell. Humaniz. Comput.* **2022**, *13*, 4041–4056. [[CrossRef](#)]
6. Chen, X.; Zhang, B.; Gao, D. Bearing fault diagnosis based on multi-scale CNN and LSTM model. *J. Intell. Manuf.* **2021**, *32*, 971–987. [[CrossRef](#)]
7. Jia, F.; Lei, Y.; Lin, J.; Zhou, X.; Lu, N. Deep neural networks: A promising tool for fault characteristic mining and intelligent diagnosis of rotating machinery with massive data. *Mech. Syst. Signal Process.* **2016**, *72*, 303–315. [[CrossRef](#)]
8. Xu, G.; Liu, M.; Jiang, Z.; Söffker, D.; Shen, W. Bearing fault diagnosis method based on deep convolutional neural network and random forest ensemble learning. *Sensors* **2019**, *19*, 1088. [[CrossRef](#)]
9. Eren, L. Bearing fault detection by one-dimensional convolutional neural networks. *Math. Probl. Eng.* **2017**, *2017*, 8617315. [[CrossRef](#)]
10. Chen, J.; Jiang, J.; Guo, X.; Tan, L. An efficient CNN with tunable input-size for bearing fault diagnosis. *Int. J. Comput. Intell. Syst.* **2021**, *14*, 625–634. [[CrossRef](#)]
11. Peng, D.; Wang, H.; Liu, Z.; Zhang, W.; Zuo, M.J.; Chen, J. Multibranch and multiscale CNN for fault diagnosis of wheelset bearings under strong noise and variable load condition. *IEEE Trans. Ind. Inf.* **2020**, *16*, 4949–4960. [[CrossRef](#)]
12. Zhu, X.; Luo, X.; Zhao, J.; Hou, D.; Han, Z.; Wang, Y. Research on deep feature learning and condition recognition method for bearing vibration. *Appl. Acoust.* **2020**, *168*, 107435. [[CrossRef](#)]
13. Peng, D.; Wang, H.; Desmet, W.; Gryllias, K. RMA-CNN: A residual mixed-domain attention CNN for bearings fault diagnosis and its time-frequency domain interpretability. *J. Dyn. Monit. Diagn.* **2023**, *2*, 115–132. [[CrossRef](#)]
14. Wang, B.; Feng, G.; Huo, D.; Kang, Y. A bearing fault diagnosis method based on spectrum map information fusion and convolutional neural network. *Processes* **2022**, *10*, 1426. [[CrossRef](#)]
15. Pham, M.T.; Kim, J.M.; Kim, C.H. Accurate bearing fault diagnosis under variable shaft speed using convolutional neural networks and vibration spectrogram. *Appl. Sci.* **2020**, *10*, 6385. [[CrossRef](#)]
16. Zhang, W.; Peng, G.; Li, C. Rolling element bearings fault intelligent diagnosis based on convolutional neural networks using raw sensing signal. In *Advances in Intelligent Information Hiding and Multimedia Signal Processing, Proceedings of the Twelfth International Conference on Intelligent Information Hiding and Multimedia Signal Processing, Kaohsiung, Taiwan, 21–23 November 2016*; Springer International Publishing: Berlin/Heidelberg, Germany, 2016; Volume 2, pp. 77–84.
17. Hasan, M.J.; Islam, M.M.; Kim, J.M. Bearing fault diagnosis using multidomain fusion-based vibration imaging and multitask learning. *Sensors* **2021**, *22*, 56. [[CrossRef](#)] [[PubMed](#)]
18. Qin, Y.; Shi, X. Fault diagnosis method for rolling bearings based on two-channel CNN under unbalanced datasets. *Appl. Sci.* **2022**, *12*, 8474. [[CrossRef](#)]
19. Xin, Y.; Li, S.; Wang, J.; An, Z.; Zhang, W. Intelligent fault diagnosis method for rotating machinery based on vibration signal analysis and hybrid multi-object deep CNN. *IET Sci. Meas. Technol.* **2020**, *14*, 407–415. [[CrossRef](#)]
20. Kannan, V.; Zhang, T.; Li, H. A review of the intelligent condition monitoring of rolling element bearings. *Machines* **2024**, *12*, 484. [[CrossRef](#)]
21. Wang, Z.; Xu, X.; Song, D.; Zheng, Z.; Li, W. A Novel Bearing Fault Diagnosis Method Based on Improved Convolutional Neural Network and Multi-Sensor Fusion. *Machines* **2025**, *13*, 216. [[CrossRef](#)]
22. Kiakojouri, A.; Wang, L. A Generalized Convolutional Neural Network Model Trained on Simulated Data for Fault Diagnosis in a Wide Range of Bearing Designs. *Sensors* **2025**, *25*, 2378. [[CrossRef](#)]
23. Tayyab, S.M.; Chatterton, S.; Pennacchi, P. Intelligent defect diagnosis of rolling element bearings under variable operating conditions using convolutional neural network and order maps. *Sensors* **2022**, *22*, 2026. [[CrossRef](#)] [[PubMed](#)]

24. Shen, F.; Chen, C.; Yan, R.; Gao, R.X. Bearing fault diagnosis based on SVD feature extraction and transfer learning classification. In Proceedings of the 2015 Prognostics and System Health Management Conference (PHM), Beijing, China, 21–23 October 2015; IEEE: Piscataway, NJ, USA, 2015; pp. 1–6.
25. Zhang, W.; Peng, G.; Li, C.; Chen, Y.; Zhang, Z. A new deep learning model for fault diagnosis with good anti-noise and domain adaptation ability on raw vibration signals. *Sensors* **2017**, *17*, 425. [\[CrossRef\]](#)
26. Zhang, B.; Li, W.; Tong, Z.; Zhang, M. Bearing fault diagnosis under varying working condition based on domain adaptation. *arXiv* **2017**, arXiv:1707.09890.
27. Zhang, R.; Tao, H.; Wu, L.; Guan, Y. Transfer learning with neural networks for bearing fault diagnosis in changing working conditions. *IEEE Access* **2017**, *5*, 14347–14357. [\[CrossRef\]](#)
28. Qian, W.; Li, S.; Wang, J.; Xin, Y.; Ma, H. A new deep transfer learning network for fault diagnosis of rotating machine under variable working conditions. In Proceedings of the 2018 Prognostics and System Health Management Conference (PHM-Chongqing), Chongqing, China, 26–28 October 2018; IEEE: Piscataway, NJ, USA, 2018; pp. 1010–1016.
29. Tong, Z.; Li, W.; Zhang, B.; Zhang, M. Bearing fault diagnosis based on domain adaptation using transferable features under different working conditions. *Shock Vib.* **2018**, *2018*, 6714520. [\[CrossRef\]](#)
30. Tong, Z.; Li, W.; Zhang, B.; Jiang, F.; Zhou, G. Bearing fault diagnosis under variable working conditions based on domain adaptation using feature transfer learning. *IEEE Access* **2018**, *6*, 76187–76197. [\[CrossRef\]](#)
31. Yang, B.; Lei, Y.; Jia, F.; Xing, S. An intelligent fault diagnosis approach based on transfer learning from laboratory bearings to locomotive bearings. *Mech. Syst. Signal Process.* **2019**, *122*, 692–706. [\[CrossRef\]](#)
32. Li, X.; Jiang, H.; Zhao, K.; Wang, R. A deep transfer nonnegativity-constraint sparse autoencoder for rolling bearing fault diagnosis with few labeled data. *IEEE Access* **2019**, *7*, 91216–91224. [\[CrossRef\]](#)
33. Shao, S.; McAleer, S.; Yan, R.; Baldi, P. Highly accurate machine fault diagnosis using deep transfer learning. *IEEE Trans. Ind. Informatics.* **2018**, *15*, 2446–2455. [\[CrossRef\]](#)
34. Zhiyi, H.; Haidong, S.; Lin, J.; Junsheng, C.; Yu, Y. Transfer fault diagnosis of bearing installed in different machines using enhanced deep auto-encoder. *Measurement* **2020**, *152*, 107393. [\[CrossRef\]](#)
35. Zhao, B.; Zhang, X.; Zhan, Z.; Pang, S. Deep multi-scale convolutional transfer learning network: A novel method for intelligent fault diagnosis of rolling bearings under variable working conditions and domains. *Neurocomputing* **2020**, *407*, 24–38. [\[CrossRef\]](#)
36. Li, X.; Jia, X.D.; Zhang, W.; Ma, H.; Luo, Z.; Li, X. Intelligent cross-machine fault diagnosis approach with deep auto-encoder and domain adaptation. *Neurocomputing* **2020**, *383*, 235–247. [\[CrossRef\]](#)
37. Wu, Z.; Jiang, H.; Lu, T.; Zhao, K. A deep transfer maximum classifier discrepancy method for rolling bearing fault diagnosis under few labeled data. *Knowl.-Based Syst.* **2020**, *196*, 105814. [\[CrossRef\]](#)
38. Dong, Y.; Li, Y.; Zheng, H.; Wang, R.; Xu, M. A new dynamic model and transfer learning based intelligent fault diagnosis framework for rolling element bearings race faults: Solving the small sample problem. *ISA Trans.* **2022**, *121*, 327–348. [\[CrossRef\]](#) [\[PubMed\]](#)
39. Chen, X.; Yang, R.; Xue, Y.; Huang, M.; Ferrero, R.; Wang, Z. Deep transfer learning for bearing fault diagnosis: A systematic review since 2016. *IEEE Trans. Instrum. Meas.* **2023**, *72*, 3508221. [\[CrossRef\]](#)
40. Zhong, Z.; Xie, H.; Wang, Z.; Zhang, Z. Domain Adversarial Transfer Learning Bearing Fault Diagnosis Model Incorporating Structural Adjustment Modules. *Sensors* **2025**, *25*, 1851. [\[CrossRef\]](#)
41. Bearing Data Center. Available online: <https://engineering.case.edu/bearingdatacenter> (accessed on 14 April 2025).
42. McFadden, P.D.; Smith, J.D. Vibration monitoring of rolling element bearings by the high-frequency resonance technique—A review. *Tribol. Int.* **1984**, *17*, 3–10. [\[CrossRef\]](#)
43. SDT, Vigilant. Available online: <https://sdtultrasound.com/products/permanent-monitoring/vigilant/> (accessed on 14 April 2025).
44. SKF Bearing Select. Available online: <https://skfbearingselect.com/#/size-lubrication/single-bearing> (accessed on 14 April 2025).

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