
A Noval Load-Dependent Multimodal Vibration Signal Enhancement and Fusion Framework (LD-MVSEFF) for Load-Specific Condition Monitoring

[Shahd Ziad Hejazi](#) * and [Michael Packianather](#)

Posted Date: 13 January 2026

doi: 10.20944/preprints202601.0897.v1

Keywords: bearings fault classification; Customised Load Adaptive Framework (CLAF); Gramian Angular Difference Field (GADF); load variation; Machinery Fault Prevention Technology (MFPT) dataset; multimodal decision fusion; wavelet scalograms



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a [Creative Commons CC BY 4.0 license](#), which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

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.

Article

A Noval Load-Dependent Multimodal Vibration Signal Enhancement and Fusion Framework (LD-MVSEFF) for Load-Specific Condition Monitoring

Shahd Ziad Hejazi ^{1,*} and Michael Packianather ²

¹ Department of Industrial Engineering, Faculty of Engineering, King Abdulaziz University, P.O. Box 80204, Jeddah 21589, Saudi Arabia

² School of Engineering, Cardiff University, Queen's Buildings, 14–17 The Parade, Cardiff CF24 3AA, UK

* Correspondence: szhejazi@kau.edu.sa

Abstract

This paper presents a Load-Dependent Multimodal Vibration Signal Enhancement and Fusion Framework (LD-MVSEFF) for load-specific condition monitoring, building on the Customised Load Adaptive Framework (CLAF). The proposed approach enhances the classification of CLAF load-dependent fault subclasses namely Healthy, Mild, Moderate, and Severe by integrating complementary information from raw vibration signals and signal-encoded representations. Three input channels are employed, combining time–frequency domain features with Continuous Wavelet Transform (CWT) and Gramian Angular Difference Field (GADF) image encodings, with each channel independently trained and evaluated to identify its most effective classifiers. To address the reduced separability of the Mild and Moderate fault subclasses under varying load conditions, a weighted decision fusion strategy is introduced, assigning classifier contributions according to their class-specific strengths. Experimental evaluation over five runs demonstrates high and stable performance, with the best configuration achieving an overall accuracy of $99.04\% \pm 0.22\%$ and an average training time of 18 min and 30 s. The results confirm the effectiveness of LD-MVSEFF as a robust multimodal methodology for load-specific condition monitoring.

Keywords: bearings fault classification; Customised Load Adaptive Framework (CLAF); Gramian Angular Difference Field (GADF); load variation; Machinery Fault Prevention Technology (MFPT) dataset; multimodal decision fusion; wavelet scalograms

1. Introduction

Rotating machinery, including pumps, electric motors, ventilators, and wind turbines, plays a vital role across numerous industrial applications. As such, failures in these systems carry significant operational and economic implications. Among their components, bearings are particularly susceptible to degradation due to prolonged operational stress, often resulting in suboptimal conditions and reduced system efficiency [1]. In Induction Motors (IMs), bearing faults account for approximately 40%–50% of total failures [2]. Various diagnostic techniques have been developed to detect bearing faults, including temperature monitoring, acoustic emission analysis, vibration signal analysis, and more recently, data-driven methods such as neural networks [3]. Detecting these faults early is essential for maintaining reliability and avoiding costly downtime. Traditional approaches rely on time-domain, frequency-domain, or time–frequency analysis of vibration signals [4], but often overlook the complementary nature of these domains [5].

Recent advances in Machine Learning (ML) and Deep Learning (DL) have improved fault classification performance [3,6,7]. Convolutional Neural Networks (CNNs), in particular, are widely used for fault diagnosis using both 1D vibration signals [8] and 2D signal encodings such as Continuous Wavelet Transform (CWT) images [9–11] or thermal images [1]. Pre-trained CNNs and

Transfer Learning (TL) have further enhanced diagnostic performance across domains [12–16]. However, most studies assume access to multiple sensor modalities or use single-view representations, limiting their generalization under varying load conditions.

A key challenge remains: how to effectively capture rich, discriminative features from a single signal source (vibration) to improve fault subclass classification under different loads. While data fusion techniques have been explored [1,17–21], their application has largely focused on multi-sensor setups or generic fault types. Meanwhile, emerging 2D signal encoding techniques such as Gramian Angular Fields (GAF), including Gramian Angular Summation Field (GASF) and Gramian Angular Difference Field (GADF), and CWT have shown potential for improving feature extraction from vibration signals [1,11,22–25]. Although GADF has demonstrated strong potential, it is still less commonly utilized than CWT in the context of vibration signal encoding for in vibration-based condition monitoring frameworks.

This study builds on these findings by proposing a deep learning–based fusion framework aligned with the Customized Load Adaptive Framework (CLAF) [7], which enables more granular load-dependent subclassification (Healthy, Mild, Moderate, Severe). Unlike conventional binary classification (“Normal” vs “Faulty”), CLAF provides a structured way to model fault severity under varying load conditions.

To the best of the authors’ knowledge, no existing work systematically integrates multi-view representations derived from vibration signals within a unified fusion architecture explicitly tailored for CLAF-based subclass classification. This paper addresses that gap through a multichannel architecture that combines statistical, time–frequency, and 2D image-based features using performance-weighted decision fusion.

The proposed Load-Dependent Multimodal Vibration Signal Enhancement and Fusion Framework (LD-MVSEFF) employs a three-channel decision fusion technique, integrating GADF, CWT, and Time and Frequency Domain (TFD) features through three dedicated feature extraction channels. Each channel is paired with an optimized classifier, and their outputs are fused using a weighted strategy to improve robustness and classification performance under load variation. The contributions of this paper are summarized as follows:

(a) Multimodal fusion and decision fusion: The proposed framework LD-MVSEFF combines features from GADF, CWT, and TFD data to enhance the Load-Dependent Fault Classification builds on the CLAF. By integrating these complementary patterns and using a weighted decision fusion approach, the framework assigns classifier weights based on performance, helping to improve accuracy, particularly in the more challenging Mild and Moderate fault subclasses.

(b) Comprehensive data integration: Insights from both 1D vibration signals and 2D RGB images (CWT and GADF) were combined to capture complementary patterns, enhancing the classification.

The paper’s structure is as follows: Section 2 covers the theoretical background and state-of-the-art research. Section 3 details the proposed framework and the dataset. Section 4 discusses the experimental results and evaluation. Lastly, Section 5 concludes the paper with suggestions for future research directions.

2. Background and Related Work

2.1. Features Extraction Domains in Signal Processing

Feature extraction operates within three primary domains: temporal, spectral, and time-frequency. These distinct domains serve as tools to capture distinctive aspects of signal behavior. The section starts with Time-Frequency Domain (TFD) feature extraction and moves to the 2D TFD features.

The Feature extraction from vibration signals in the time domain is a crucial component of machinery fault diagnosis, enabling the early detection and continuous monitoring of machinery faults. This method entails computing diverse statistical parameters from the original vibration

signal, which can subsequently be employed to assess the machinery's condition and detect potential problems. Various key parameters are utilized in vibration signal analysis to extract vital information. These parameters include the Peak or Max value, which denotes the highest observed amplitude in the signal, and the Root Mean Square (RMS), which provides insights into signal magnitude. Skewness assesses distribution asymmetry, whereas Standard Deviation (std) quantifies average deviation from the mean. Kurtosis indicates distribution "tailedness," potentially identifying outliers or impulses. The Crest Factor, calculated as the peak amplitude-to-RMS ratio, reflects peak sharpness. Peak-to-peak measures the range between maximum and minimum values, whereas the Impulse Factor accentuates impulsive behaviors often linked to machinery faults. These parameters contribute to a comprehensive understanding of vibration signal characteristics, facilitating effective fault diagnosis and condition monitoring [26–29].

On the other hand, extracting features from the frequency domain can provide insights into the data's periodic components and harmonic structures. The frequency domain analysis of vibration signals involves examining the amplitude changes for different frequencies [30]. These features capture frequency-specific aspects of the signal and contribute to a better understanding of the vibration behavior [31]. Analyzing the frequency domain of vibration signals is crucial for understanding periodic components and harmonic structures. Key features include Root Mean Square Frequency (RMSF), Centre Frequency (CF), Mean Square Frequency (MSF), Frequency Variance (FV), and Root Frequency Variance (RVF), providing insights into signal characteristics and power distribution [31]. Standard harmonic features, such as Total Harmonic Distortion (THD), quantify frequency content [32,33]. Signal-to-Noise Ratio (S/N) and Signal-to-Noise and Distortion Ratio (SINAD) assess signal quality, particularly in gearbox fault analysis [34]. Spectral analysis transforms signals from the time domain to the frequency domain, with the AR model being a popular choice. Various methods, like Yule-Walker and Burg's, compute AR coefficients, whereas the forward-backwards approach enhances classification, especially in machinery fault diagnosis [35,36]. Spectral features like Peak Amplitude, Peak Frequency, and Band Power offer comprehensive insights into frequency characteristics [30–33,35,37].

2.2. Two-Dimensional (2D) Signal Encoding Techniques

(a) Gramian Angular Field (GAF) Signal Encoding

Wang and Oates introduced the concept of GAF encoding, a method that transforms time series data into images. GAF's distinctive matrix construction maintains the integrity of the original data while capturing relationships between neighboring elements. This methodology proves beneficial for CNN models, enabling automatic feature extraction and enhancing classification performance [38]. The core concept behind converting time-series data into images using GAF involves creating a matrix based on polar coordinates. This matrix preserves the temporal relationships within the one-dimensional (1D) time-series signal, maintaining accurate temporal correlations compared to Cartesian coordinates. The process yields two types of GAF images: Gramian Angular Summation Field (GASF) and Gramian Angular Difference Field (GADF) [22].

Given a time series $X = \{x_1, x_2, \dots, x_n\}$, the signal is first normalised and rescaled to the interval $[-1,1]$ to ensure a bijective mapping during polar coordinate transformation, as in (1) [22,39]:

$$\bar{x}_i = \frac{((x_i - \max(X)) + (x_i - \min(X)))}{\max(X) - \min(X)} \quad (1)$$

The time series is then mapped into polar coordinates by computing the angular component ϕ , defined as the inverse cosine of the normalized signal \bar{x}_i , while the polar coordinate r encodes the temporal position of each sample. The transformation is given in (2) [22,39]:

$$\begin{cases} \phi = \arccos(\bar{x}_i), & -1 \leq \bar{x}_i \leq 1, \bar{x}_i \in \bar{X} \\ r = \frac{t_i}{N}, & t_i \in N \end{cases} \quad (2)$$

Where t_i denotes the timestamp (sample index), N is a scaling constant, and r represents the polar coordinate used to encode the temporal progression in the polar coordinate space. Restricting

ϕ to the interval $[0, \pi]$ ensures a bijective angular mapping, preserving unique temporal relationships.

Unlike Cartesian representations, GAF preserves temporal structure by encoding time progression along the main diagonal of the resulting matrix. Temporal correlations between samples are quantified through angular relationships, using either the summation term $\cos(\phi_i + \phi_j)$ for GASF or the difference $\cos(\phi_i - \phi_j)$ for the GADF [1].

(b) Continuous Wavelet Transform (CWT)

The Wavelet Transform (WT) provides an alternative to the Short-Time Fourier Transform (STFT) for analyzing non-stationary signals, as it can represent both temporal and spectral characteristics with variable time–frequency resolution [9,24]. The CWT maps a time-domain signal into a time–frequency representation via convolution with a scaled and translated mother wavelet, producing correlation coefficients between the wavelet function and the original signal [10]. By adjusting the scale and translation parameters, CWT enables precise correlation measurement and energy-distribution mapping of the waveform, typically visualized as a scalogram [1]. In machine fault diagnosis, the Morlet wavelet is commonly combined with CWT to analyze vibration signals and generate time–frequency images that can be used as inputs to CNN-based classifiers for fault identification [40].

2.3. Customised Load Adaptive Framework (CLAF)

The CLAF is a two-phase methodology designed to enhance fault classification in induction motors, particularly under varying radial loads. It is tailored specifically for the MFPT bearing dataset. Phase 1 involves load-dependent pattern analysis in the time and frequency domains, including data preprocessing, segmentation, feature extraction, and validation using one-way ANOVA. Phase 2 customizes the methodology for the dataset by using Wavelet Singular Entropy and the CWT to classify faults into load-dependent subclasses: Normal (fault-free) or Healthy, Mild, Moderate, and Severe. This approach provides a detailed understanding of how load variations affect induction motor defects and introduces a new dimension to traditional fault classification by focusing on load variation and dataset customization. The CLAF is validated through classifier training and classification accuracy analysis for the proposed subclasses [7].

2.4. State-of- The Art and Research Gaps

The field of fault detection in manufacturing systems has seen significant advances, particularly in the use of vibration signals for condition monitoring and bearing health assessment. A wide range of techniques has been developed to extract informative features from vibration signals, including time-domain, frequency-domain, and spectral analyses, as well as Autoregressive (AR) models. More recently, Machine Learning (ML) and Deep Learning (DL), especially Convolutional Neural Networks (CNNs), together with fusion strategies and advanced signal encoding methods such as Gramian Angular Fields (GAF) and Continuous Wavelet Transform (CWT), have further improved fault classification capabilities.

Conventional vibration-based approaches typically rely on time-domain features such as RMS, variance, and kurtosis, and frequency-domain spectral attributes [29]. Time-domain features have proven effective for early fault detection, but real-world signals often exhibit strong non-stationarity, making consistent feature extraction challenging [41]. Moreover, identifying informative features can be labor-intensive and sometimes infeasible, particularly in the case of complex machinery or rare fault types [42]. To improve fault characterization under noisy conditions, Lai et al. (2025) proposed a diagnosis method for rolling bearings that combines spectral kurtosis analysis and Hilbert envelope demodulation, integrated with Least-Squares Support Vector Machines (LS-SVM), demonstrating robust performance in challenging environments [3]. AR models have also been explored for spectral feature extraction [7,43,44]. In response to these challenges, statistical feature selection techniques have been employed to isolate the most discriminative indicators for classification. For example,

independent t-tests have identified kurtosis, skewness, and maximum value as key features [6], while one-way ANOVA has been used to rank feature significance [7].

ML advancements have significantly improved fault classification performance. A variety of algorithms has been applied, including Support Vector Machines (SVM), Multilayer Neural Networks (MNN), Random Forest (RF) [6], Least-Squares Support Vector Machine (LS-SVM) [3], as well as CubicSVM and WNN [7]. Building on these developments, Deep Learning (DL), particularly CNNs, has demonstrated strong potential in fault diagnosis. When the extracted features fail to capture fault-relevant patterns, DL models may misclassify subtle faults or background noise, compromising classification accuracy [39]. To address this, Transfer Learning (TL) using pre-trained CNNs has been widely adopted. Some studies fine-tune shallow layers while reusing deeper layers from large-scale datasets [13]. Notably, AlexNet [13,16] and ResNet [14,15] have shown strong performance in vibration-based diagnostics. AlexNet, due to its simplicity and lower computational cost, suits lightweight classification tasks [45], while ResNet, with its deeper architecture and residual connections, provides more powerful feature extraction but requires higher computational resources [46,47]. CNNs have been implemented in both 1D and 2D architectures, depending on the data representation. 1D CNNs are typically trained on raw vibration signals, whereas 2D CNNs are applied to signal encodings such as CWT images [9–11] or thermal images [1]. One notable 1D approach is the One-Dimensional Ternary Pattern (1D-TP) method, which extracts statistical features across time and frequency domains and achieves strong diagnostic performance when used with classifiers like RF, k-NN, SVM, BayesNet, and ANN [48].

Signal encoding techniques such as GAF, including GASF and GADF, and CWT have shown great potential for feature extraction. GAF has achieved high precision in fault classification [11] and has been found to be more discriminative than CWT in low signal-to-noise conditions [1]. Its effectiveness in bearing diagnostics has been validated [22], while CWT has been further improved through multiscale feature fusion and channel attention mechanisms [23]. Recent work has combined GASF with CWT for fault detection in wind turbine gearboxes [24], and in 2025, combining GASF and GADF as input images significantly improved diagnostic accuracy [25]. Nonetheless, GADF remains underutilized compared to more established methods like CWT—particularly in load-dependent subclassification—indicating potential for further exploration of 2D signal encoding for vibration signal analysis.

A key challenge in fault diagnosis is identifying which sensor signals or data modalities provide the most relevant information for accurate pattern recognition. In multi-sensor systems, suboptimal input selection can hinder classification performance. To address this, recent studies have focused on multi-sensor data fusion, which combines complementary information from different sources to improve diagnostic accuracy and reduce reliance on manual sensor selection [17]. Multi-domain fusion and advanced deep learning models have also been explored to further enhance accuracy [39]. Common fusion strategies include sensor-level [17,19,20], feature-level [1,8,21,25,49], and decision-level fusion [50–52]. These strategies reduce dependence on any single sensor or representation and provide a more comprehensive view of system health. Closely related to decision-level fusion is ensemble learning, which aggregates outputs from multiple classifiers to improve robustness, typically using the same input representation [53].

In 2025, a study further highlighted the importance of operating load in fault analysis by introducing a model bank-based approach in which independent classifiers are trained for each load condition, improving accuracy and efficiency compared with a single global model [6]. Despite such advances, most existing work still focuses on binary or generic multi-class classification and does not address load-dependent fault subclass separation, which is essential for practical fault classification. To the best of the authors' knowledge, no load-dependent fault classification framework that extracts features independently from multiple data representations within a single source has yet combined statistical, time–frequency, and 2D signal encodings within a unified decision-fusion architecture explicitly aligned with the CLAF approach for fault severity classification under varying loads. Although GADF and GASF have shown strong potential, they remain underutilized in vibration-

based diagnostic systems, and decision-fusion strategies that combine classical ML models with pre-trained CNNs, conditioned on load-dependent severity levels, are still underexplored in the current literature.

3. Proposed Framework

This section outlines the systematic approach of the proposed Load-Dependent Multimodal Vibration Signal Enhancement and Fusion Framework (LD-MVSEFF) for Load-Specific Condition Monitoring, building upon the CLAF load-dependent fault subclasses introduced in [7]. The proposed framework was applied to the MFPT-bearing dataset. It involves the independent extraction of from different data representations within this single data source, implemented across three separate feature extraction channels. Features extracted from each channel are then directed to their respective classification modules, where individual classification decisions are made. Subsequently, a fusion module consolidates these individual decisions into a unified classification result. The processing for the proposed methodology was conducted using MATLAB R2023a software. This section provides an overview of the methodology framework and details the data used.

3.1. Methodology

The proposed LD-MVSEFF incorporates multiple data channels and decision fusion approaches, complemented by the CLAF for creating load-dependent fault subclasses. Various data sources are integrated, including Gradient Angular Difference Field (GADF) images, Continuous Wavelet Transform (CWT) images, and features from the time and frequency domains. These inputs enhance the efficacy of condition monitoring by leveraging complementary patterns across different modalities for improved fault classification. The outputs from multiple classifiers are consolidated using decision fusion techniques, ensuring robust and accurate classification. The methodology includes six detailed steps, as presented in Figure 1.

1. Data Preprocessing with CLAF:

In this stage, the MFPT bearing vibration signals are segmented and prepared using the Customized Load Adaptive Framework (CLAF). The raw vibration data are first divided according to fault conditions—Normal (fault-free) or Healthy, Inner Race Fault (IRF), and Outer Race Fault (ORF)—and further organized based on load factors. This process enables the formation of load-dependent fault subclasses corresponding to different operating conditions. The structured segmentation ensures that the data are consistently prepared for subsequent feature extraction, graph construction, and classification.

2. Multichannel Input Preparations:

In this stage, the first phase establishes three distinct data channels for comprehensive analysis. In Channel 1, raw vibration signals are processed. Channel 2 generates two-dimensional (2D) CWT images from these signals, and Channel 3 produces 2D encoded GADF images. After splitting, the raw vibration signals are encoded into equivalent image formats:

- Channel 1: Raw vibration signal.
- Channel 2: The class-specific raw vibration signals are encoded into CWT images using the Amor technique.
- Channel 3: The class-specific raw vibration signals are encoded into 2D GADF images.
- Applying the CLAF to create load-dependent fault subclasses— Normal (fault-free) or Healthy, Mild, Moderate, and Severe—tailored to specific datasets, forming the foundation for subsequent analysis.

3. Feature Extraction and Classifier Selection for Channel 1 (Raw Vibration Signal):

For Channel 1, time- and frequency-domain (TFD) features, including spectral features derived from AR modelling, are extracted from the raw vibration signals. Feature relevance is assessed using one-way Analysis of Variance (ANOVA), and an optimal feature subset is selected. To address class imbalance, oversampling is applied to ensure that all fault subclasses contain an equal number of

samples. The selected TFD features are then aligned with their corresponding CWT and GADF representations to maintain consistency across channels.

4. Channels Classification Approaches and Training Methods:

Training and Selection of Classifiers for TFD features, including spectral features using Autoregression (Channel 1) and CNN Architectures for Channels 2 and 3 (CWT and GADF images):

- For Channel 1 (TFD features, including spectral features using Autoregression), classifiers such as Cubic Support Vector Machine (CubicSVM) and Wide Neural Network (WNN) are trained on the extracted features. The best-performing model is selected for further analysis.
- For Channels 2 and 3 (CWT and GADF images), pre-trained Convolutional Neural Networks (CNNs), such as AlexNet and ResNet-18, originally trained on the ImageNet dataset, are fine-tuned on the 2D encoded images. The final fully connected layer of each network is replaced with a new layer containing four output neurons, where each neuron corresponds to one of the four CLAF load-dependent subclasses: Normal (fault-free) or Healthy, Mild, Moderate, and Severe. The images, including CWT spectrograms and 2D GADF-encoded images, are resized to match the input dimensions of the CNN architectures: $227 \times 227 \times 3$ for AlexNet and $224 \times 224 \times 3$ for ResNet-18.

5. Single Channel Performance Analysis:

The classification performance of each channel is evaluated independently in terms of its ability to distinguish between CLAF load-dependent subclasses. For each channel, the classifier achieving the highest overall accuracy is selected for use in the subsequent fusion stage. This step provides insight into the strengths and limitations of individual signal representations.

6. Weighted Decision Fusion:

In the final stage, decision-level fusion is applied to combine the outputs of the selected classifiers. Two weighting strategies are investigated: (a) adaptive weighting, where classifier weights are assigned based on class-specific performance, and (b) equal weighting, where all channels contribute equally. Both two-channel and three-channel fusion configurations are evaluated to identify the most effective combination for improving classification accuracy and robustness.

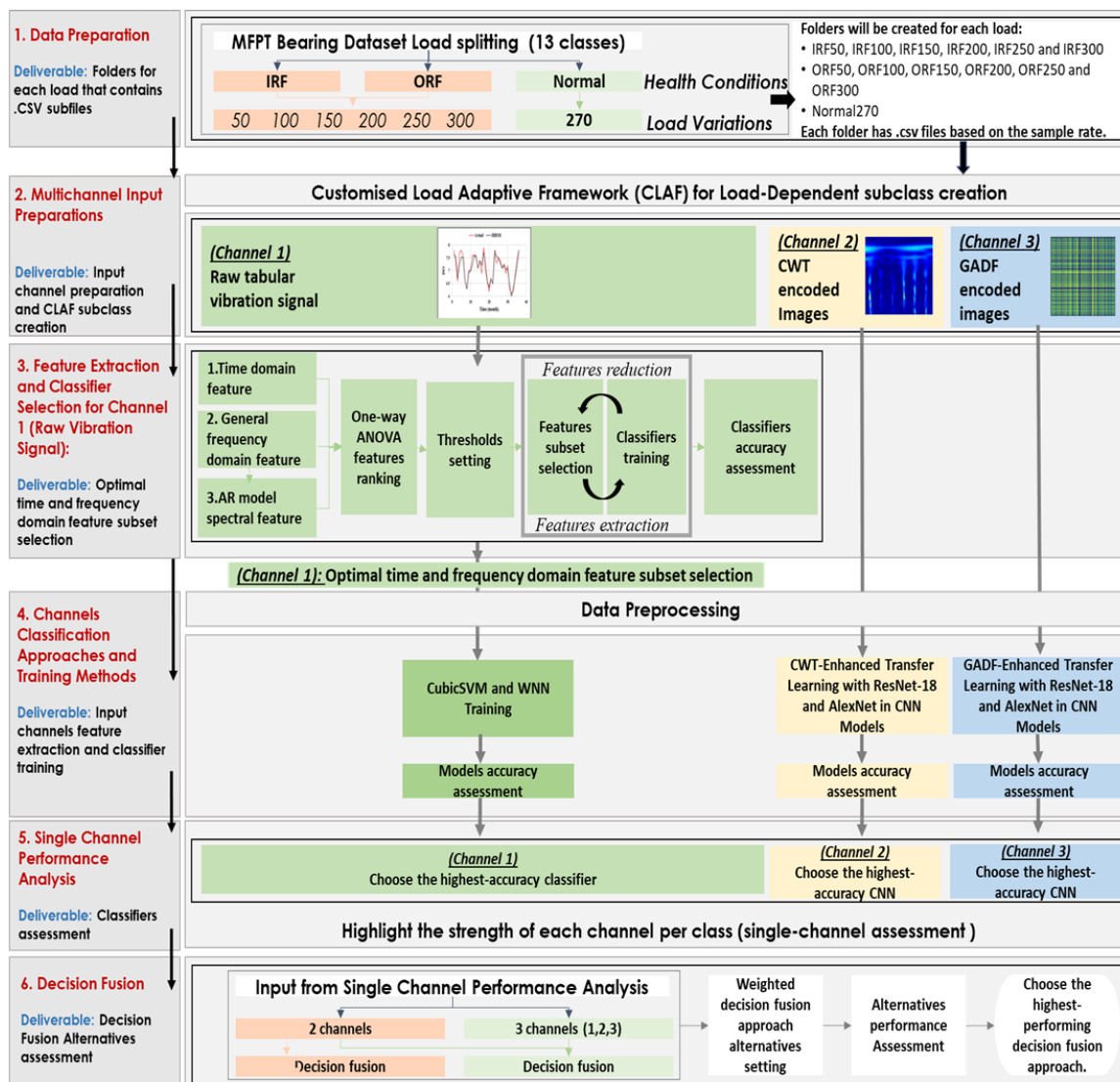


Figure 1. The proposed Load-Dependent Multimodal Vibration Signal Enhancement and Fusion Framework (LD-MVSEFF).

3.2. Dataset

This study is divided into two phases, focusing on the radial impacts of loads under various operational conditions using the Machinery Fault Prevention Technology (MFPT) Bearings Dataset. The setup included a test rig with a NICE bearing featuring a roller diameter of 5.969 mm (0.235 inches), a pitch diameter of 31.623 mm (1.245 inches), and eight rolling elements at a contact angle of zero degrees. Vibration data were collected under varying loads to simulate bearings with and without faults for detailed analysis. Normal (fault-free) or Healthy data were gathered under a load of 27.5 kg (approximately 60.6 lbs), with a sampling frequency of 97,656 Hz over 6 seconds. Additionally, fault signals from Inner Race Defect (IRD) or Inner Race Fault (IRF) and Outer Race Defect (ORD) or Outer Race Fault (ORF), shown in Figure 2 were recorded under six load conditions: 22.7 kg (50 lbs), 45.4 kg (100 lbs), 68.1 kg (150 lbs), 90.8 kg (200 lbs), 113.5 kg (250 lbs), and 136.2 kg (300 lbs), while maintaining a constant speed of 25 Hz. This dataset serves as a standardized benchmark, providing essential information such as radial load, shaft speed, and signal characteristics while maintaining a consistent shaft speed of 1500 rpm (25 Hz) [54].

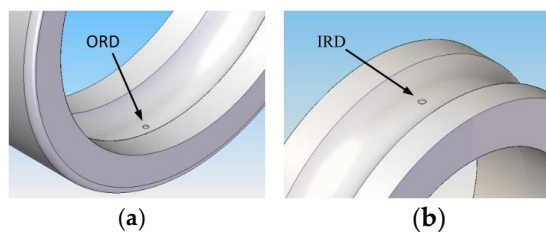


Figure 2. Computer-aided drawings of defects made on (a) ORF; (b) IRF [55].

4. Results and Discussion

This section presents a comprehensive analysis and interpretation of the outcomes obtained from the experimental study. The focal point of the analysis revolves around the performance evaluation of various fusion techniques utilized for the classification of CLAF load-dependent fault subclasses. These techniques encompass diverse feature representations and models. The overarching goal of the current study is to discern effective strategies for enhancing the accuracy of CLAF load-dependent fault subclasses prediction, thereby improving the reliability and robustness of machinery fault classification. Steps 3 and 4 show the three approaches used in the single channel, starting with TFD extraction features on the original vibration signal, then with pre-trained CNNs (AlexNet and Residual Network-18 (ResNet-18)) on encoded vibration signals into two forms: CWT vibration-encoded images and GADF vibration-encoded images.

4.1. Data Preparation

This section presents the preprocessing conducted on the MFPT-bearings dataset, detailing dividing the data according to load variations. This division is critical for applying the CLAF, designed to identify patterns dependent on load, diverging from the conventional fault classification methods used in IM bearings. The dataset has been systematically split following the CLAF approach, which marks a significant departure from traditional fault classification by factoring in load fluctuations and adapting the dataset for specialized analysis, as indicated in [7]. The research compares six load values—22.7 kg (50 lbs), 45.4 kg (100 lbs), 68.1 kg (150 lbs), 90.8 kg (200 lbs), 113.5 kg (250 lbs), and 136.2 kg (300 lbs)—against a Normal (fault-free) or Healthy condition set at 122.5 kg (270 lbs). This results in 13 categories: six for each Inner Race Fault (IRF) and Outer Race Fault (ORF) under varying loads and one for the Normal (fault-free) or Healthy condition of the data for further analysis, such as fault detection or machine learning applications.

4.2. Multichannel Input Preparations

This section outlines the creation of three distinct data channels from the raw vibration signal for analysis. Channel 1 contains the raw segmented vibration signals, Channel 2 encodes the signals into CWT images, and Channel 3 encodes them into GADF images. The dataset was analyzed using the CLAF, focusing on load-dependent subclasses: Normal (fault-free) or Healthy, Mild, Moderate, and Severe.


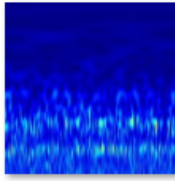
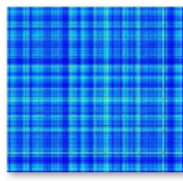

To ensure consistency and fairness in classifier performance evaluation, the datasets for Channels 1, 2, and 3—derived from the same pool of 813 subfiles—were divided in a uniform manner. Figure 3 shows the structure of the MATLAB datastore. In this datastore, the Index column, as shown in the attached image, represents the unique subfolder name used to encode both the CWT images (stored in ImagePath_cwt) and the GADF images (stored in ImagePath_GADF). This ensures that the files corresponding to each load condition (e.g., IRF_50) are consistently linked across all three channels

	1	2	3	4	5	6
	Subfiles	LoadFactor	CLAF	Index	ImagePath_cwt	ImagePath_GADF
154	2500x1 time...	'IRF_50'	"Mild"	154	'C:\Users\Shahd\Docu...	'C:\Users\Shahd\Docu...
155	2500x1 time...	'IRF_50'	"Mild"	155	'C:\Users\Shahd\Docu...	'C:\Users\Shahd\Docu...
156	2500x1 time...	'IRF_50'	"Mild"	156	'C:\Users\Shahd\Docu...	'C:\Users\Shahd\Docu...
157	2500x1 time...	'IRF_50'	"Mild"	157	'C:\Users\Shahd\Docu...	'C:\Users\Shahd\Docu...

Figure 3. Datastore structure linking raw vibration signals with CWT and GADF images.

As shown in Table 1, this uniform approach is critical for a thorough and unbiased evaluation across all three channels. It ensures that the performance of the CNN models, which are trained on various types of encoded image data such as CWT and GADF in Channels 2 and 3, and tabular features extracted for each segment in Channel 1, is evaluated under similar conditions.

Table 1. Multichannel input preparations.

Subfiles	Channel 1	Channel 2	Channel 3	CLAF
	Tabular features extracted from the raw vibration signal	CWT 2D encoded image	GADF 2D encoded image	Load-dependent fault subclasses
 Normal-1	The time and frequency domain features.			Normal (fault-free) or Healthy condition.
		 Normal-1	Normal-1	

4.2.1. Channel 1: Raw Tabular Vibration Signal

For Channels 2 and 3, the dimensions of each encoded vibration image are set to $227 \times 227 \times 3$ and $224 \times 224 \times 3$, respectively. These size specifications align with the input requirements of the AlexNet and ResNet-18 architectures, respectively. Figure 4 visually displays the connection between each channel and outlines the process of creating each channel, starting with the raw vibration signal. For Channel 1, Time and Frequency Domain (TFD) features, including spectral features using Autoregression (AR), are extracted from the segmented vibration signal and used as the input in the proposed methodology. The extracted features are detailed in Section 4.3, as shown in Table 2.

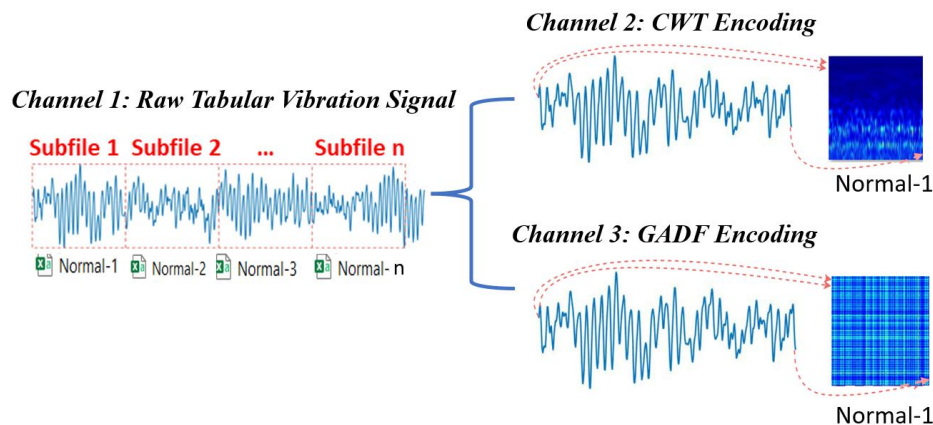


Figure 4. Input channels general overview.

Table 2. One-way ANOVA ranking including spectral features extracted by Autoregressive (AR) Model.

Feature Rank	One-way ANOVA Score	Feature Rank	One-way ANOVA Score
1. Mean	316.44	13. PeakAmplitude 5	84.33
2. ShapeFactor	288.42	14. Skewness	73.13
3. PeakValue	245.43	15. PeakAmplitude 2	70.50
4. RMS	240.93	16. PeakFreq1	69.14
5. Std	240.27	17. SINAD	58.72
6. ClearanceFactor	235.23	18. SNR	58.61
7. ImpulseFactor	225.26	19. PeakAmplitude 4	51.39
8. Kurtosis	211.94	20. PeakAmplitude 3	38.77
9. CrestFactor	198.26	21. PeakFreq4	25.18
10. PeakAmplitude	161.22	22. PeakFreq2	17.64
11. BandPower	126.85	23. PeakFreq5	13.9307
12. PeakFrequency 3	116.80	24. THD	0

4.2.2. Channel 2: Continuous Wavelet Transform

Converting vibration signals to scalogram images in MATLAB involves several systematic steps. The dataset comprises Normal (fault-free) or Healthy condition and IRF and ORF types and is first partitioned based on distinct LF conditions. Each signal subset is then processed through a Wavelet Transform (WT) using the CWT method with the 'Amor' wavelet. The transformed signals are converted into scalogram images by taking the absolute values of the CWT coefficients, flipping and scaling them. These images are then colour-mapped using the 'jet' colour map and resized to a uniform size of 224x224 pixels for consistency. The images of some of the generated samples are presented in Figure 5.

Each processed signal subset and its corresponding scalogram image are saved as an image file and a CSV file, categorically organised in folders named after the ensemble types and indices. This meticulous process is repeated for each subset of the signal, ensuring that every part of the signal is represented as a distinct image. This approach visualises the time-frequency information of vibration signals and prepares the data for further analysis, such as fault detection or ML applications. Top of Form

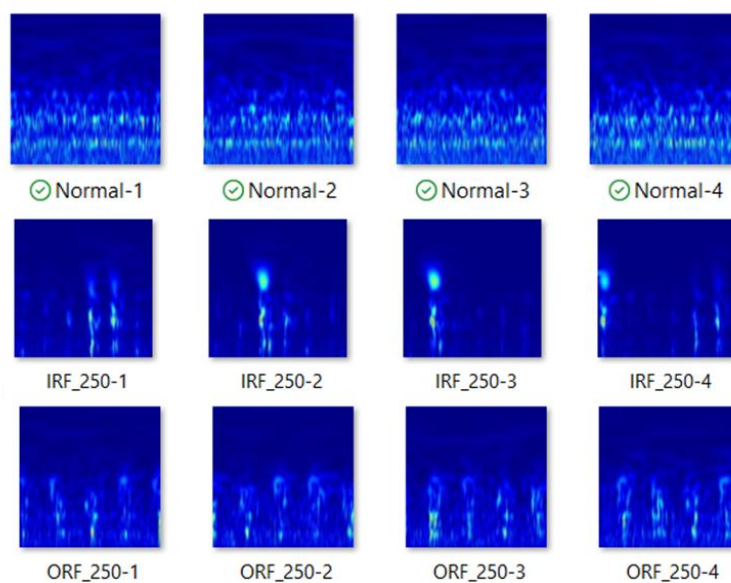


Figure 5. CWT 2D encoded image examples showing the Normal (fault-free) or Healthy condition, and the faults IRF and ORF.

4.2.3. Channel 3: Gramian Angular Difference Field (GADF)

Creating GADF images from vibration signals involves several key steps. Initially, the time series signal is segmented into smaller subsets. For each subset, a Gramian matrix is computed using the GADF algorithm, which involves calculating the pairwise dot product of the signal and then manipulating the resulting sine and cosine matrices. The Gramian matrix is then transformed into a GADF image. This transformation includes scaling the matrix values to a range between 0 and 1, inverting this scaled matrix and resizing the image to a specified size. This process is iteratively applied to the entire signal, converting each segment into a GADF image representing the underlying time series data. This method offers an alternative way to analyse and interpret vibration signals, facilitating more profound insights into their characteristics. GADF encoding produces distinct patterns for various health conditions, which need further analysis to validate their ability to differentiate between health conditions, as illustrated in Figure 6.

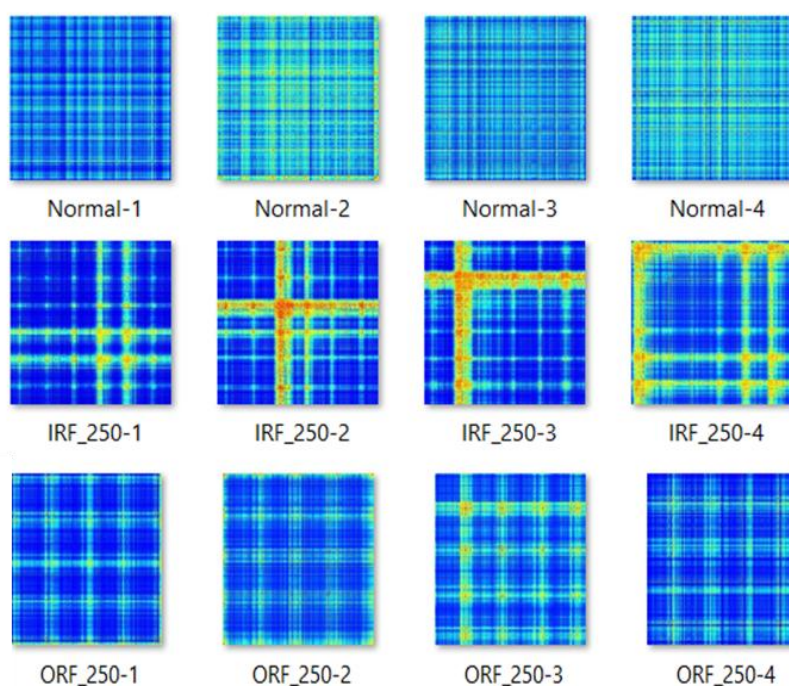


Figure 6. GADF 2D encoded image examples showing the Normal (fault-free) or Healthy condition, and the faults IRF and ORF.

4.3. Feature Extraction and Classifier Selection for Channel 1 (Raw Vibration Signal)

This section conducts a one-way ANOVA test to rank the extracted general TFD features. Additionally, spectral features are extracted using an AR model of order 15 and a maximum of 5 peaks, creating 24 features. The selection of features for data representation plays a crucial role in the model's performance. Choosing the most relevant features is essential to ensure the model effectively captures the critical information related to the fault. On the other hand, including irrelevant features can sometimes result in overfitting or decreased performance [56]. One-way ANOVA feature selection involves comparing the means of each feature across different target classes to determine if there is a statistically significant difference. Features are ranked based on their p-values from the ANOVA test; the lower the p-value, the more likely the feature is to be influential in distinguishing between classes. These p-values are often transformed into scores by taking the negative logarithm, with higher scores indicating more significant features for classification. In Table 2, the first column represents the extracted features, and the second column represents the one-way ANOVA scores, ranked from highest to lowest significance. Features that scored less than 26 (Peak frequency 4, Peak

frequency 2, Peak frequency 5, and Total Harmonic Distortion (THD) were not included in the current study due to their low scores, which could lead to confused training.

A critical analysis of classifier performance, based on the top 20 feature sets ranked by the one-way ANOVA score, provides diverse insights, as presented in Table 2. Various classifiers, including Support Vector Machines (SVMs), Neural Networks (NN), and Ensembles, were employed using MATLAB 2023a [57]. The efficacy of SVM hinges on how effectively the input data are represented in this new space, a determination often made through the utilisation of diverse kernels like Linear, Polynomial (including quadratic and cubic), Gaussian, and others [58]. The CubicSVM is a classifier that falls under the umbrella of supervised learning. SVMs are effective for high-dimensional data and are versatile in handling various structured datasets. Hence, they are widely used for classification and regression tasks [57].

The objective was to identify the classifier with the highest accuracy, making it a strong candidate for the proposed LD-MVSEFF load-dependent fault classification framework. The training dataset, comprising 813 subfolders, was divided as follows: 60.00% for training, 20.00% for validation, and 20.00% for testing. Five-fold cross-validation was implemented to ensure a robust performance assessment (see Table 3), divided by the load-dependent fault subclasses.

The Ensemble: The Boosted Trees classifier recorded a notable 94.40% accuracy, demonstrating its ability to effectively harness a larger feature set. Reducing the feature set to the top 17 had a minimal effect on accuracy, which consistently remained above 90.00%, demonstrating the classifiers' robustness and efficiency with a minor feature set. A further reduction in the feature set to the top 10, 7, and 5 revealed a nuanced interplay between feature count and accuracy. The WNN, using the top 10 features, outperformed its counterparts with a peak accuracy of 92.02%, suggesting its superior capability in working with a more compact yet pertinent feature set.

Classifier performance exhibited considerable variation in the Mild class, with the following Ensemble: Boosted Trees classifier's accuracy ranging from 89.20% with the top 17 features to 95.40% with the top 20 features. The CubicSVM classifier recorded the lowest accuracy in the Moderate class, scoring 85.70% and 91.40% with the top 7 and 5 feature subsets, respectively. In contrast, WNN achieved 91.40% with the top 10 feature subset. Remarkably, the Normal (fault-free) or Healthy condition class maintained a stable 100% accuracy across all classifiers and feature subsets, underscoring the classifiers' consistent ability to identify Normal (fault-free) or Healthy condition accurately. This consistency indicates a shared strength among the classifiers. At the same time, the variability in the load-dependent fault subclasses (Mild and Moderate classes) underscores the critical importance of appropriate feature subset selection for optimal classifier performance.

In a direct comparison, the accuracy of the CubicSVM and the WNN was closely matched. However, the selection of the top 10 features by one-way ANOVA demonstrated a well-calibrated compromise between training feature quantity and test dataset accuracy. The CubicSVM and WNN achieved 94.60% and 93.40% overall testing accuracies, respectively. Breaking this down further, the CubicSVM recorded 92.50% in the Mild class, 85.70% in the Moderate class, and 100% in the Severe class. Meanwhile, the WNN scored 90.3% in the Mild, 91.40% in the Moderate, and 91.70% in the Severe class. Notably, the CubicSVM outperformed the WNN in the Severe class by 8.30% and the Mild class by 2.20%. Conversely, the WNN outperformed the CubicSVM in the Moderate class by 5.70%. As a result, these two classifiers were selected as the top performers for Channel 1. CubicSVM is designated as Channel 1a, while WNN is designated as Channel 1b.

Table 3. Classifier performance on Channel 1 across distinct feature sets ranked by One-Way ANOVA feature significance.

Classifier	ANOVA ranking	TTime ¹		Testing Dataset				
		(s)	VA ²	NA ³	MA ⁴	MoA ⁵	SA ⁶	Overall Accuracy
Ensemble:Boosted Trees	Top 20 >26	114.4	94.50%	100%	95.40%	88.50%	93.50%	94.40%
Ensemble: Boosted Trees	Top 17 >58.6	16.8	95.10%	100%	89.20%	85.70%	91.70%	91.70%
Cubic SVM	Top 10 (a) >161	5.9	94.30%	100%	92.50%	85.70%	100%	94.60%
WNN	Top 10 (b) >161	15.7	92.20%	100%	90.30%	91.40%	91.70%	93.40%
Cubic SVM	Top 7 >215	7.1	93.20%	100%	90.30%	85.70%	100%	94.00%
Cubic SVM	Top 5 >240	9.4	93.70%	100%	90.30%	85.70%	100%	94.00%

¹ TTime is the Training Time, ² VA is the Validation Accuracy, ³ NA is the Normal (fault-free) or Healthy condition Accuracy, ⁴ MA is the Mild state Accuracy, ⁵ MoA is the Moderate state Accuracy, ⁶ SA is the Severe state Accuracy.

4.4. Channels Classification Approaches and Training Methods

To evaluate classification performance across the CLAF load-dependent fault subclasses: 'Normal (fault-free) or Healthy condition,' 'Mild,' 'Moderate,' and 'Severe.' The dataset was balanced and uniformly split into training (60%), validation (20%), and testing (20%) sets using a fixed random seed to ensure reproducibility. All channels were derived from the same pool of vibration signal segments to enable fair and consistent performance comparison.

4.4.1. Channel 1: CubicSVM and WNN

As shown in Table 4, CubicSVM achieved a higher overall accuracy (96.28%) than WNN (94.95%), while also requiring substantially less training time (26.71 s compared to 48.63 s). In addition, CubicSVM demonstrated improved performance in the Mild and Moderate fault subclasses, and was therefore selected for Channel 1.

Table 4. Channel 1 classifiers training.

Classifier	TTime ¹		Test Dataset				Overall Accuracy
	(s)	VA ²	NA ³	MA ⁴	MoA ⁵	SA ⁶	
a. CubicSVM	26.71	96.80%	100%	89.36%	95.74%	100%	96.28%
b. WNN	48.63	96.70%	100%	95.74%	84.04%	100%	94.95%

¹ TTime is the Training Time, ² VA is Validation Accuracy, ³ NA is the Normal (fault-free) or Healthy condition Accuracy, ⁴ MA is the Mild state Accuracy, ⁵ MoA is the Moderate state Accuracy, ⁶ SA is the Severe state Accuracy.

4.4.2. Channels 2 and 3: Pre-trained CNN Selection

For Channels 2 and 3, transfer learning was applied using pre-trained AlexNet and ResNet-18 architectures to classify vibration signals encoded as Continuous Wavelet Transform (CWT) and Gramian Angular Difference Field (GADF) images, respectively. Both networks were fine-tuned to classify the four CLAF load-dependent fault subclasses: 'Normal (fault-free) or Healthy condition,' 'Mild,' 'Moderate,' and 'Severe.' The comparative performance of the CNN models for both channels is summarised in Table 5.

For Channel 2 (CWT images), ResNet-18 achieved a slightly higher overall accuracy (98.94%) compared to AlexNet (98.40%); however, AlexNet required substantially less training time (7.20 min versus 17.35 min). Given the marginal accuracy difference and the significantly lower computational cost, AlexNet was selected for Channel 2.

For Channel 3 (GADF images), AlexNet outperformed ResNet-18 in overall test accuracy (98.67% versus 95.21%) and demonstrated improved class-wise performance, particularly in the Mild (96.81% versus 89.36%) and Moderate (97.87% versus 92.55%) subclasses. In addition, AlexNet reduced training time from 18.50 min to 7.53 min. Based on both accuracy and efficiency considerations, AlexNet was selected as the preferred model for Channel 3.

Table 5. Pretrained CNNs performance on Channel 2 (CWT signal encoded images) and Channel 3 (GADF signal encoded images).

Channel	Pre-trained CNN	TTime ¹ (min)	Testing Dataset					
			VA ²	NA ³	MA ⁴	MoA ⁵	SA ⁶	Overall Accuracy
Channel 2	1. ResNet-18	17.35	97.55%	100%	95.74%	100%	100%	98.94%
	2. AlexNet	7.20	97.28%	100%	96.81%	96.81%	100%	98.40%
Channel 3	1. ResNet-18	18.5	96.47%	100%	89.36%	92.55%	98.94%	95.21%
	2. AlexNet	7.53	96.47%	100%	96.81%	97.87%	100%	98.67%

¹TTime is the Training Time, ²VA is the Validation Accuracy, ³NA is the Normal (fault-free) or Healthy condition Accuracy, ⁴MA is the Mild state Accuracy, ⁵MoA is the Moderate state Accuracy, ⁶SA is the Severe state Accuracy.

4.5. Single-Channel Performance Analysis

Figure 7 presents the load-dependent subclass accuracy assessment for each channel. All classifiers performed well for the Normal (fault-free) or Healthy and Severe condition classes, with each achieving 100% accuracy. This high performance, while expected for these extreme conditions where the patterns are more distinct and more straightforward to differentiate, could be attributed to the more apparent fault or non-fault signals in the data. The clear distinction between the Healthy and Severe conditions allowed the classifiers to identify them without error consistently.

In the Mild condition class, Channel 2, using CWT (AlexNet), showed the best performance with an accuracy of 96.81%, followed closely by Channel 3 (GADF with AlexNet) at 95.74%. However, Channels 1a and 1b, which utilise CubicSVM and WNN classifiers, struggled more in detecting Mild conditions, with accuracies of 89.36% and 84.04%, respectively. This suggests that the Mild class presents more challenges for accurate classification, likely due to the less distinct signal patterns associated with early or mild faults.

While performance remained strong for the Moderate condition class, there were noticeable differences between the channels. Channel 3 (GADF with AlexNet) showed the highest accuracy at 97.87%, followed by Channel 2 (CWT with AlexNet) at 96.81% and Channel 1b (WNN) at 95.74%. Channel 1a (CubicSVM) exhibited the lowest performance in this category, with an accuracy of

84.04%. This indicates that Moderate conditions are more difficult to classify than extremes as the signal patterns become less clear.

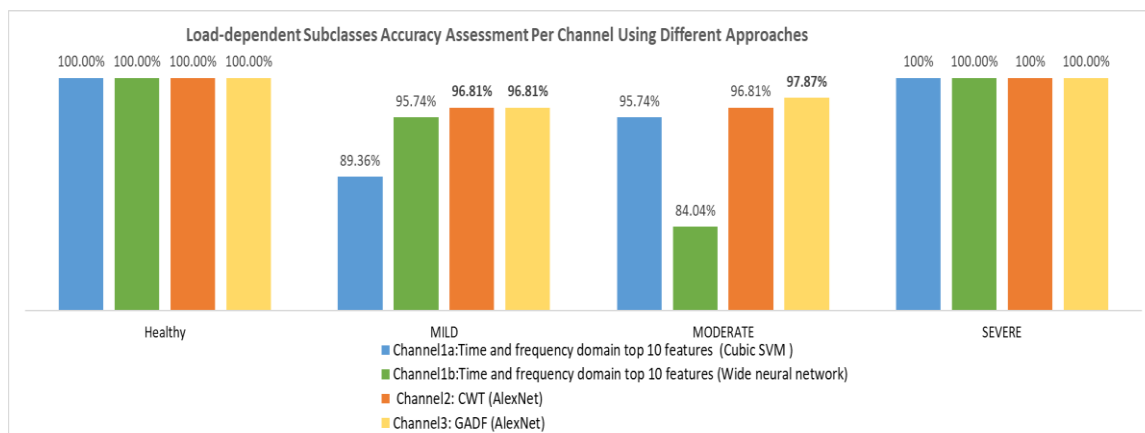


Figure 7. CLAF load-dependent fault subclass accuracy assessment per channel using different approaches.

4.6. Decision Fusion

4.6.1. Weighted Decision Fusion Approach (Alternatives Setting)

This section investigates two weighted decision-fusion schemes across three alternatives, each using a different combination of channels within the LD-MVSEFF. In all cases, the weights assigned to the channels for a given CLAF load-dependent fault subclass (Healthy, Mild, Moderate, Severe) sum to 1, ensuring a balanced contribution in the final decision.

Weighting System 1 (adaptive weighting) assigns channel weights according to the classification accuracy achieved for each fault subclass and load condition. Channels that perform better for a given subclass receive higher weights, allowing the fusion mechanism to emphasize the most reliable sources of information in a load- and severity-dependent manner. Weighting System 2 (equal weighting) assigns the same weight to all channels for all subclasses, providing a baseline in which no classifier dominates the decision process. The decision-fusion weights are guided by per-channel accuracy for each CLAF subclass, as shown in Figure 7. Here, TFDa denotes the CubicSVM classifier trained on the top 10 TFD features, and TFDb denotes the WNN trained on the same feature set; the corresponding alternatives and weighting schemes are summarized in Table 6.

- For Alternative 1, the fusion combines Channel 1b (TFD features with WNN) and Channel 2 (CWT–AlexNet). Under Weighting System 1, Channel 1b receives higher weights for the Healthy and Severe subclasses, where it achieves perfect accuracy, while Channel 2 is emphasized for Mild and Moderate subclasses, where it outperforms Channel 1b. Under Weighting System 2, both channels receive equal weights across all subclasses.
- Alternative 2 fuses Channel 1a (TFD features with CubicSVM) with Channel 2 (CWT–AlexNet). In Weighting System 1, Channel 1a is favored for Healthy and Severe subclasses due to its perfect accuracy, whereas Channel 2 receives higher weights in Mild and Moderate subclasses, where it attains superior performance. Under Weighting System 2, both channels contribute equally for all subclasses.
- Alternative 3 extends the fusion to three channels: Channel 1a (TFD–CubicSVM), Channel 2 (CWT–AlexNet), and Channel 3 (GADF–AlexNet). With Weighting System 1, Channel 1a is down-weighted for Mild and Moderate subclasses, where its accuracy is lower, while Channels 2 and 3 receive higher weights reflecting their stronger performance. For Healthy and Severe subclasses, all three channels achieve perfect accuracy and are therefore assigned equal weights. Under Weighting System 2, all three channels receive equal weights for every subclass.

Table 6. Alternatives setting and decision fusion weighting system.

Channel No.	Input	Classifier	Weighting System 1				Weighting System 2				
			Health y	MA ²	MoA ³	SA ⁴	Health y	M A	Mo A	SA	
Alternative No.			1.1 (TFDb -CWT)				1.2 (TFDb-CWT)				
Alternative 1	1	TFD ¹	WNN	0.5	0.4	0.3	0.5	0.5	0.5	0.5	
	2	CWT	AlexNet	0.5	0.6	0.7	0.5	0.5	0.5	0.5	
Alternative No.			2.1 (TFDa -CWT)				2.2 (TFDa -CWT)				
Alternative 2	1a	TFD	CubicSV M	0.5	0.3	0.4	0.5	0.5	0.5	0.5	
	2	CWT	AlexNet	0.5	0.7	0.6	0.5	0.5	0.5	0.5	
Alternative No.			3.1 (TFDa -CWT-GADF)				3.2 (TFDa -CWT-GADF)				
Alternative 3	1a	TFD	CubicSV M	0.33	0.2	0.2	0.3	0.33	0.3	0.33	0.3
	2	CWT	AlexNet	0.33	0.4	0.4	0.3	0.33	0.3	0.33	0.3
	3	GAD F	AlexNet	0.33	0.4	0.4	0.3	0.33	0.3	0.33	0.3

¹ TFD is the Time and Frequency Domain extracted features, ² MA is the Mild state Accuracy, ³ MoA is the Moderate state Accuracy, ⁴ SA is the Severe state Accuracy.

4.6.2. Choosing The Highest-Performing Weighted Decision Fusion Approach

To ensure robustness and reproducibility, each fusion alternative (Alternatives 1–3) was evaluated over five independent runs using different random seeds, with regenerated training, validation, and testing splits for each run. All reported results therefore correspond to mean test accuracy with the associated standard deviation.

Across all alternatives, classification performance was consistently high for the Normal (fault-free) or Healthy and Severe subclasses. However, differences were more pronounced for the Mild and Moderate subclasses, where fault signatures are less distinct. Among the evaluated configurations, Alternative 3.1, which corresponds to the three-channel fusion (TFDa–CWT–GADF) under Adaptive Weighting, delivered the most balanced and reliable performance. This configuration achieved an overall accuracy of $99.04\% \pm 0.22\%$, with strong subclass performance in the Mild ($97.20\% \pm 1.75\%$) and Moderate ($99.15\% \pm 0.89\%$) classes, while maintaining an average training time of 18 min 30 s. Table 7 summarises the overall test accuracy across the five runs for all decision-fusion alternatives.

Based on these results, Alternative 3.1 (TFDa–CWT–GADF) was selected as the final decision fusion configuration and forms the basis of the proposed LD-MVSEFF.

Table 7. Decision fusion overall test accuracy over the 5 runs.

Alternatives	1	2	3	4	5	Avg.
1.1 (TFDb -CWT)	98.67%	98.94%	98.94%	98.67%	99.07%	98.86%
1.2 (TFDb-CWT)	98.94%	97.08%	98.67%	98.67%	98.67%	98.40%
2.1 (TFDa -CWT)	98.40%	98.67%	96.54%	98.67%	98.54%	98.16%

2.2 (TFDa -CWT)	98.67%	98.94%	97.07%	98.67%	98.94%	98.46%
3.1 (TFDa -CWT-GADF)	98.67%	98.94%	99.47%	98.94%	99.20%	99.04%
3.2 (TFDa - CWT - GADF)	98.40%	97.51%	98.94%	98.94%	99.20%	98.60%

5. Conclusion

This paper introduced the novel Load-Dependent Multimodal Vibration Signal Enhancement and Fusion Framework (LD-MVSEFF), designed for dealing with load-dependent fault classification using the MFPT bearing dataset. The LD-MVSEFF incorporates load-dependent fault subclasses derived from the CLAF, shifting the focus from traditional fault classification to a load-specific approach. It integrates three distinct channels for analysis: Channel 1 extracts TFD features, including spectral features using Autoregression; Channel 2 converts vibration signals into CWT images; and Channel 3 encodes the signals into GADF 2D images.

Each channel was trained over five separate runs, and the best-performing classifiers were selected based on their accuracy in classifying four load-dependent fault subclasses: Healthy, Mild, Moderate, and Severe. In Channel 1, CubicSVM and WNN classifiers achieved average accuracies of $96.43\% \pm 0.76\%$ and $97.50\% \pm 1.60\%$, respectively. For Channels 2 and 3, pre-trained AlexNet and ResNet-18 models were used, with AlexNet performing the best, achieving accuracies of $97.76\% \pm 1.33\%$ on Channel 2 (CWT images) and $95.95\% \pm 2.05\%$ on Channel 3 (GADF images).

One of the main challenges observed was the classification of the Mild and Moderate fault subclasses, which presented subtler signal variations compared to the Healthy and Severe conditions. The proposed LD-MVSEFF addressed these challenges by employing a weighted decision fusion approach, where decisions were tailored according to the strengths of each channel for specific fault subclass. For instance, Channel 2 (CWT with AlexNet) performed well in classifying the Moderate class, while Channel 3 (GADF with AlexNet) showed high accuracy in both the Mild and Moderate conditions. By assigning dynamic weights to each classifier based on their strengths, the LD-MVSEFF improved the classification of these more challenging subclasses.

The proposed weighted decision fusion approach demonstrated excellent performance across all fault conditions. Alternative approach indicated 3.1 in Table 7 (TFDa - CWT - GADF) achieved the highest overall accuracy of $99.04\% \pm 0.22\%$ across five runs, with an average training time of 18 min and 30 s. This approach minimized the limitations of individual classifiers and effectively handled load-specific fault classification. Future work will explore advanced graph-based and deep learning feature representations to further enhance fault discrimination under varying load conditions.

Author Contributions: Conceptualisation, S.Z.H. and M.P.; methodology, S.Z.H. and M.P.; software, S.Z.H.; validation, S.Z.H. and M.P.; formal analysis, S.Z.H. and M.P.; investigation, M.P.; resources, S.Z.H.; data curation, S.Z.H.; writing—original draft preparation, S.Z.H.; writing—review and editing, S.Z.H. and M.P.; visualisation, S.Z.H.; supervision, M.P.; project administration, S.Z.H. and M.P.; funding acquisition, S.Z.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Condition Based Maintenance Fault Database for Testing of Diagnostic and Prognostics Algorithms Available online: <https://www.mfpt.org/fault-data-sets/> (accessed on 25 January 2024) [54].

Acknowledgements: Special thanks to the Society for Machinery Failure Prevention Technology for providing the publicly available dataset used in this study.

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

References

1. Hejazi, S.; Packianather, M.; Liu, Y. Novel Preprocessing of Multimodal Condition Monitoring Data for Classifying Induction Motor Faults Using Deep Learning Methods. In Proceedings of the 2022 IEEE 2nd International Symposium on Sustainable Energy, Signal Processing and Cyber Security (iSSSC); IEEE: Gunupur Odisha, India, 15–17 December 2022, December 15 2022; pp. 1–6.
2. Hejazi, S.; Packianather, M.; Liu, Y. A Novel Approach Using WGAN-GP and Conditional WGAN-GP for Generating Artificial Thermal Images of Induction Motor Faults. In Proceedings of the 27th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems (KES 2023); Elsevier B.V., 2023; Vol. 225, pp. 3681–3691.
3. Lai, L.; Xu, W.; Song, Z. A Novel Fault Diagnosis Method for Rolling Bearings Based on Spectral Kurtosis and LS-SVM. *Electronics (Switzerland)* 2025, 14, doi:10.3390/electronics14142790.
4. Jang, J.G.; Noh, C.M.; Kim, S.S.; Shin, S.C.; Lee, S.S.; Lee, J.C. Vibration Data Feature Extraction and Deep Learning-Based Preprocessing Method for Highly Accurate Motor Fault Diagnosis. *J Comput Des Eng* 2023, 10, 204–220, doi:10.1093/jcde/qwac128.
5. Xie, F.; Li, G.; Fan, Q.; Xiao, Q.; Zhou, S. Optimizing and Analyzing Performance of Motor Fault Diagnosis Algorithms for Autonomous Vehicles via Cross-Domain Data Fusion. *Processes* 2023, 11, 2862, doi:10.3390/pr11102862.
6. Lee, H.G.; Yoo, S.M.; Hao, W.K.; Lee, I.S. Time-Domain and Neural Network-Based Diagnosis of Bearing Faults in Induction Motors Under Variable Loads. *Machines* 2025, 13, 1055, doi:10.3390/machines13111055.
7. Hejazi, S.Z.; Packianather, M.; Liu, Y. A Novel Customised Load Adaptive Framework for Induction Motor Fault Classification Utilising MFPT Bearing Dataset. *Machines* 2024, 12, 44, doi:10.3390/machines12010044.
8. Alam, T.E.; Ahsan, M.M.; Raman, S. Multimodal Bearing Fault Classification under Variable Conditions: A 1D CNN with Transfer Learning. *Machine Learning with Applications* 2025, 21, 100682, doi:10.1016/j.mlwa.2025.100682.
9. Nishat Toma, R.; Kim, C.-H.; Kim, J.-M. Bearing Fault Classification Using Ensemble Empirical Mode Decomposition and Convolutional Neural Network. *Electronics (Basel)* 2021, 10, 1248, doi:10.3390/electronics10111248.
10. Kaji, M.; Parvizian, J.; van de Venn, H.W. Constructing a Reliable Health Indicator for Bearings Using Convolutional Autoencoder and Continuous Wavelet Transform. *Applied Sciences* 2020, 10, 8948, doi:10.3390/app10248948.
11. Zhang, J.; Kong, X.; Cheng, L.; Qi, H.; Yu, M. Eksploatacja i Niezawodność – Maintenance and Reliability Transform-Multiscale Feature Fusion and Improved Channel Attention Mechanism. 2023, 25, 0–2.
12. Zhong, H.; Lv, Y.; Yuan, R.; Yang, D. Bearing Fault Diagnosis Using Transfer Learning and Self-Attention Ensemble Lightweight Convolutional Neural Network. *Neurocomputing* 2022, 501, 765–777, doi:10.1016/j.neucom.2022.06.066.
13. Asutkar, S.; Tallur, S. Deep Transfer Learning Strategy for Efficient Domain Generalisation in Machine Fault Diagnosis. *Sci Rep* 2023, 13, 1–9, doi:10.1038/s41598-023-33887-5.
14. Wu, G.; Ji, X.; Yang, G.; Jia, Y.; Cao, C. Signal-to-Image: Rolling Bearing Fault Diagnosis Using ResNet Family Deep-Learning Models. *Processes* 2023, 11, doi:10.3390/pr11051527.
15. Chang, M.; Yao, D.; Yang, J. Intelligent Fault Diagnosis of Rolling Bearings Using Efficient and Lightweight ResNet Networks Based on an Attention Mechanism. *IEEE Sens J* 2023, 23, 9136–9145, doi:10.1109/JSEN.2023.3251654.

16. Lu, T.; Yu, F.; Han, B.; Wang, J. A Generic Intelligent Bearing Fault Diagnosis System Using Convolutional Neural Networks with Transfer Learning. *IEEE Access* 2020, 8, 164807–164814, doi:10.1109/ACCESS.2020.3022840.
17. Cinar, E. A Sensor Fusion Method Using Deep Transfer Learning for Fault Detection in Equipment Condition Monitoring. In Proceedings of the 2022 International Conference on INnovations in Intelligent SysTems and Applications (INISTA); IEEE: Biarritz, France, 8–12 August, August 8 2022; pp. 1–6.
18. Cui, D.; Zhang, T.; Zhang, M.; Liu, X. Feature Extraction and Severity Identification for Autonomous Underwater Vehicle with Weak Thruster Fault. *Journal of Marine Science and Technology (Japan)* 2022, 27, 1105–1115, doi:10.1007/s00773-022-00891-9.
19. Kullu, O.; Cinar, E. A Deep-Learning-Based Multi-Modal Sensor Fusion Approach for Detection of Equipment Faults. *Machines* 2022, 10, 1105, doi:10.3390/machines10111105.
20. Pan, Z.; Zhang, Z.; Meng, Z.; Wang, Y. A Novel Fault Classification Feature Extraction Method for Rolling Bearing Based on Multi-Sensor Fusion Technology and EB-1D-TP Encoding Algorithm. *ISA Trans* 2023, 142, 427–444, doi:10.1016/j.isatra.2023.07.015.
21. Ye, Z.; Yu, J. Multi-Level Features Fusion Network-Based Feature Learning for Machinery Fault Diagnosis. *Appl Soft Comput* 2022, 122, 108900, doi:10.1016/j.asoc.2022.108900.
22. Toma, R.N.; Piltan, F.; Im, K.; Shon, D.; Yoon, T.H.; Yoo, D.; Kim, J. A Bearing Fault Classification Framework Based on Image Encoding Techniques and a Convolutional Neural Network under Different Operating Conditions. *Sensors* 2022, 22, 4881, doi:10.3390/s22134881.
23. Xiao, R.; Zhang, Z.; Wu, Y.; Jiang, P.; Deng, J. Multi-Scale Information Fusion Model for Feature Extraction of Converter Transformer Vibration Signal. *Measurement (Lond)* 2021, 180, 109555, doi:10.1016/j.measurement.2021.109555.
24. Yang, Q.; Tang, B.; Shen, Y.; Li, Q. Self-Attention Parallel Fusion Network for Wind Turbine Gearboxes Fault Diagnosis. *IEEE Sens J* 2023, 23, 23210–23220, doi:10.1109/JSEN.2023.3308971.
25. Guo, Q.; Yao, H.; Xu, Y.; Lu, B.; Ma, Z.; Huang, Y.; Shi, M. Transformer Fault Diagnosis Method Based on Gramian Angular Field and Optimized Parallel ShuffleNetV2. *Sci Rep* 2025, 15, doi:10.1038/s41598-025-08590-2.
26. Narayan, Y. Hb VsEMG Signal Classification with Time Domain and Frequency Domain Features Using LDA and ANN Classifier Materials Today : Proceedings Hb VsEMG Signal Classification with Time Domain and Frequency Domain Features Using LDA and ANN Classifier. *Mater Today Proc* 2021, 37, 3226–3230, doi:10.1016/j.matpr.2020.09.091.
27. Liu, M.K.; Weng, P.Y. Fault Diagnosis of Ball Bearing Elements: A Generic Procedure Based on Time-Frequency Analysis. *Measurement Science Review* 2019, 19, 185–194, doi:10.2478/msr-2019-0024.
28. Jain, P.H.; Bhosle, S.P. Study of Effects of Radial Load on Vibration of Bearing Using Time-Domain Statistical Parameters. *IOP Conf Ser Mater Sci Eng* 2021, 1070, 012130, doi:10.1088/1757-899x/1070/1/012130.
29. Pinedo-Sánchez, L.A.; Mercado-Ravell, D.A.; Carballo-Monsivais, C.A. Vibration Analysis in Bearings for Failure Prevention Using CNN. *Journal of the Brazilian Society of Mechanical Sciences and Engineering* 2020, 42, 628, 1–16, doi:10.1007/s40430-020-02711-w.
30. Ahmed, H.; Nandi, A.K. Compressive Sampling and Feature Ranking Framework for Bearing Fault Classification With Vibration Signals. *IEEE Access* 2018, 6, 44731–44746, doi:10.1109/ACCESS.2018.2865116.
31. Shi, Z.; Li, Y.; Liu, S. A Review of Fault Diagnosis Methods for Rotating Machinery. In Proceedings of the 2020 IEEE 16th International Conference on Control & Automation (ICCA); IEEE: Singapore, 9–11 October 2020, October 9 2020; pp. 1618–1623.

32. Granados-Lieberman, D.; Huerta-Rosales, J.R.; Gonzalez-Cordoba, J.L.; Amezcua-Sanchez, J.P.; Valtierra-Rodriguez, M.; Camarena-Martinez, D. Time-Frequency Analysis and Neural Networks for Detecting Short-Circuited Turns in Transformers in Both Transient and Steady-State Regimes Using Vibration Signals. *Applied Sciences* 2023, *13*, 12218, doi:10.3390/app132212218.
33. Tian, B.; Fan, X.; Xu, Z.; Wang, Z.; Du, H. Finite Element Simulation on Transformer Vibration Characteristics under Typical Mechanical Faults. In Proceedings of the Proceedings of the 9th International Conference on Power Electronics Systems and Applications, (PESA 2022); IEEE: Hong Kong, China, 20–22 September 2022, 2022; pp. 1–4.
34. Kumar, V.; Mukherjee, S.; Verma, A.K.; Sarangi, S. An AI-Based Nonparametric Filter Approach for Gearbox Fault Diagnosis. *IEEE Trans Instrum Meas* 2022, *71*, 351661, 1–11, doi:10.1109/TIM.2022.3186700.
35. Hu, L.; Zhang, Z. *EEG Signal Processing and Feature Extraction*; Hu, L., Zhang, Z., Eds.; Springer Singapore: Singapore, 2019; ISBN 978-981-13-9112-5.
36. Metwally, M.; Hassan, M.M.; Hassaan, G. Diagnosis of Rotating Machines Faults Using Artificial Intelligence Based on Preprocessing for Input Data. In Proceedings of the In Proceedings of the 26th IEEE Conference of Open Innovations Association FRUCT (FRUCT26), Yaroslavl, Russia, 23–25 April 2020.; 2020.
37. Djemili, I.; Medoued, A.; Soufi, Y. A Wind Turbine Bearing Fault Detection Method Based on Improved CEEMDAN and AR-MEDA. *Journal of Vibration Engineering & Technologies* 2023, 1–22, doi:10.1007/s42417-023-01117-x.
38. Wang, Z.; Oates, T. Imaging Time-Series to Improve Classification and Imputation. *IJCAI International Joint Conference on Artificial Intelligence* 2015, *2015-Janua*, 3939–3945.
39. Cui, J.; Zhong, Q.; Zheng, S.; Peng, L.; Wen, J. A Lightweight Model for Bearing Fault Diagnosis Based on Gramian Angular Field and Coordinate Attention. *Machines* 2022, *10*, 282, doi:10.3390/machines10040282.
40. Łuczak, D. Machine Fault Diagnosis through Vibration Analysis: Continuous Wavelet Transform with Complex Morlet Wavelet and Time–Frequency RGB Image Recognition via Convolutional Neural Network. *Electronics (Switzerland)* 2024, *13*, doi:10.3390/electronics13020452.
41. Sayyad, S.; Kumar, S.; Bongale, A.; Kamat, P.; Patil, S.; Kotecha, K. Data-Driven Remaining Useful Life Estimation for Milling Process: Sensors, Algorithms, Datasets, and Future Directions. *IEEE Access* 2021, *9*, 110255–110286, doi:10.1109/ACCESS.2021.3101284.
42. Resendiz-Ochoa, E.; Osornio-Rios, R.A.; Benitez-Rangel, J.P.; Romero-Troncoso, R.D.J.; Morales-Hernandez, L.A. Induction Motor Failure Analysis: An Automatic Methodology Based on Infrared Imaging. *IEEE Access* 2018, *6*, 76993–77003, doi:10.1109/ACCESS.2018.2883988.
43. Ganapathy, S.; Mallidi, S.H.; Hermansky, H. Robust Feature Extraction Using Modulation Filtering of Autoregressive Models. *IEEE Trans Audio Speech Lang Process* 2014, *22*, 1285–1295, doi:10.1109/taslp.2014.2329190.
44. Vaibhaw; Sarraf, J.; Pattnaik, P.K. Brain–Computer Interfaces and Their Applications. In *An Industrial IoT Approach for Pharmaceutical Industry Growth*; Elsevier, 2020; pp. 31–54.
45. Ramzan, F.; Khan, M.U.G.; Rehmat, A.; Iqbal, S.; Saba, T.; Rehman, A.; Mehmood, Z. A Deep Learning Approach for Automated Diagnosis and Multi-Class Classification of Alzheimer’s Disease Stages Using Resting-State fMRI and Residual Neural Networks. *J Med Syst* 2020, *44*, doi:10.1007/s10916-019-1475-2.
46. Thalagala, S.; Walgampaya, C. Application of AlexNet Convolutional Neural Network Architecture-Based Transfer Learning for Automated Recognition of Casting Surface Defects. In Proceedings of the 2021 International Research Conference on Smart Computing and Systems Engineering (SCSE); IEEE, September 16 2021; Vol. 4, pp. 129–136.

47. Kadam, V.; Kumar, S.; Bongale, A.; Wazarkar, S.; Kamat, P.; Patil, S. Enhancing Surface Fault Detection Using Machine Learning for 3D Printed Products. *Applied System Innovation* 2021, 4, 34, doi:10.3390/asi4020034.
48. Kuncan, M.; Kaplan, K.; Minaz, M.R.; Kaya, Y.; Ertunç, H.M. A Novel Feature Extraction Method for Bearing Fault Classification with One Dimensional Ternary Patterns. *ISA Trans* 2020, 100, 346–357, doi:10.1016/j.isatra.2019.11.006.
49. Ye, L.; Ma, X.; Wen, C. Rotating Machinery Fault Diagnosis Method by Combining Time-Frequency Domain Features and Cnn Knowledge Transfer. *Sensors* 2021, 21, 8168, doi:10.3390/s21248168.
50. Li, J.; Ying, Y.; Ren, Y.; Xu, S.; Bi, D.; Chen, X.; Xu, Y. Research on Rolling Bearing Fault Diagnosis Based on Multi-Dimensional Feature Extraction and Evidence Fusion Theory. *R Soc Open Sci* 2019, 6, 181488, doi:10.1098/rsos.181488.
51. Yang, D.; Karimi, H.R.; Gelman, L. A Fuzzy Fusion Rotating Machinery Fault Diagnosis Framework Based on the Enhancement Deep Convolutional Neural Networks. 2022, 1–15.
52. Wang, X.; Li, A.; Han, G. Applied Sciences A Deep-Learning-Based Fault Diagnosis Method of Industrial Bearings Using Multi-Source Information. 2023, 1–26.
53. Jose, J.P.; Ananthan, T.; Prakash, N.K. Ensemble Learning Methods for Machine Fault Diagnosis. In Proceedings of the 2022 Third International Conference on Intelligent Computing Instrumentation and Control Technologies (ICICT); IEEE, August 11 2022; pp. 1127–1134.
54. Bechhoefer, E. Condition Based Maintenance Fault Database for Testing of Diagnostic and Prognostics Algorithms Available online: <https://www.mfpt.org/fault-data-sets/> (accessed on 30 October 2023).
55. Jain, P.H.; Bhosle, S.P. Analysis of Vibration Signals Caused by Ball Bearing Defects Using Time-Domain Statistical Indicators. *International Journal of Advanced Technology and Engineering Exploration* 2022, 9, 700–715, doi:10.19101/IJATEE.2021.875416.
56. Kareem, A.B.; Hur, J.-W. Towards Data-Driven Fault Diagnostics Framework for SMPS-AEC Using Supervised Learning Algorithms. *Electronics (Basel)* 2022, 11, 2492, doi:10.3390/electronics11162492.
57. MathWorks-3 Choose Classifier Options Available online: <https://www.mathworks.com/help/stats/choose-a-classifier.html> (accessed on 01 January 2026).
58. Khanjani, M.; Ezoji, M. Electrical Fault Detection in Three-Phase Induction Motor Using Deep Network-Based Features of Thermograms. *Measurement* 2021, 173, 108622, doi:10.1016/j.measurement.2020.108622.

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.