Data-Driven Fault Diagnosis of Rolling Bearings Using Neural Networks Optimized via Bayesian, Particle Swarm, and Genetic Algorithm Methods with Time–Frequency Features

Khoualdia Kaaïs [1], Khoualdia Tarek [1,2], and Nahal Mourad [1,2]

[1] *Department of Mechanical Engineering, University of Souk-Ahras, 41000, Algeria*
k.khoualdia@univ-soukahras.dz

[2] *Laboratory of Research on Electromechanical and Dependability (LRESF), University of Souk-Ahras, 41000, Algeria*
t.khoualdia@univ-soukahras.dz
m.nahal @univ-soukahras.dz

Abstract

This study presents an applied comparative evaluation of automated rolling bearing fault identification using Artificial Neural Networks (ANNs) optimized through Bayesian Optimization (BO), Particle Swarm Optimization (PSO), and Genetic Algorithm (GA). Vibration signals were collected from bearings operating under five health conditions (healthy, outer ring fault, inner ring fault, ball fault, and combined faults), at three rotational speeds, and along three measurement directions. The acquired signals were preprocessed using filtering, normalization, and segmentation. Time-domain and Fast Fourier Transform (FFT)-based frequency-domain features were extracted and used to train ANN models. The ANN architectures, including hidden layers, neurons, and activation functions, were optimized using BO, PSO, and GA, resulting in six configurations. Since the problem is a multi-class classification task, performance was assessed using F1-score, Accuracy, precision, and recall. The optimized ANN models were also benchmarked against Support Vector Machine (SVM), K-Nearest Neighbors (kNN), and Random Forest (RF) classifiers using the same feature sets. Results show that FFT -based features consistently outperformed time-domain features**,** and ANN-PSO with FFT-based features achieved the best performance, with F1-score = 0.982, Accuracy = 0.998, precision = 0.982, and recall = 0.982**.** This work contributes a systematic applied comparison of ANN optimization strategies rather than a fundamentally new machine-learning architecture, highlighting optimized ANNs as competitive and computationally efficient solutions for bearing fault diagnosis.

https://doi.org/10.36001/IJPHM.2026.v17i2.4768

1. Introduction

Rolling bearings play a crucial role in rotating machinery, as their condition significantly influences system reliability, operational continuity, and maintenance costs. Effective fault diagnosis requires robust methods capable of detecting faults at an early stage while remaining resilient to noise and variations in operating conditions. This need is central to the framework of predictive maintenance and health monitoring, which offers a comprehensive approach, from data acquisition to the estimation of the remaining useful life of components, to support proactive and efficient maintenance strategies in industrial environments (Lee et al., 2013).

Recent years have witnessed a surge in data-driven techniques, driven by advances in deep learning and signal processing. Yang et al. (2025) introduced a lightweight hybrid Transformer, Long Short-Term Memory (LSTM) model, achieving around 99 % accuracy while maintaining rapid convergence and practical scalability. Meanwhile, Gao et al. (2022) combined parameter optimized Maximum Correlated Kurtosis Deconvolution (MCKD) with Convolutional Neural Network (CNN) architectures to improve the detection of compound bearing faults, thereby enhancing sensitivity to subtle signal features under noisy conditions. Similarly, Liao et al. (2024) proposed a physics informed neural blind deconvolution module that guided denoising through kurtosis and classifier-based feedback, significantly improving signal clarity in harsh environments.

Time-frequency representation methods have also shown notable effectiveness. Tang et al. (2022) developed a Time-Frequency Region-based CNN that isolates discriminative time-frequency representation patches through region attention modules, enabling accurate multiclass diagnosis. Chen et al. (2025) built on wavelet-transformed inputs to create an Efficient Cross-Space Multiscale CNN Transformer

Parallelism architecture (ECMCTP), which extracts complementary spatiotemporal features from different network branches. Fu et al. (2024) introduced a parallel CNN–LSTM structure that jointly processes continuous wavelet transform (CWT) images, capturing both spatial representations and temporal dependencies. Siddique et al. (2025) further refined this concept through a hybrid continuous-wavelet and attention model, achieving strong diagnostic precision from small datasets.

Handling imbalanced data and multiple operating conditions has become increasingly important. Zhang et al. (2025) proposed a signal processing augmented deep learning pipeline designed to generalize across varied fault and load conditions in industrial applications. An et al. (2022) employed a Periodic Sparse Attention LSTM to exploit inherent signal periodicity, enhancing robustness while simplifying sequence modeling. Other studies have focused on richer feature learning. Xiao et al. (2025) introduced Variable Filtered feature extraction and Quantitative fusion using bidirectional LSTM (VFQB), which improved classification by combining advanced feature extraction with hierarchical learning. Similarly, Ding et al. (2021) proposed a time–frequency Transformer architecture based on self-attention to capture temporal and spectral dependencies more effectively. Furthermore, Tama et al. (2022) reviewed these developments and emphasized the growing role of vibration-based deep-learning methods in improving diagnostic accuracy and reliability.

Beyond end-to-end architectures, feature engineering and preprocessing remain critical. Pan et al. (2023) integrated kernel principal component analysis (KPCA) with a Spatial Group-wise Enhance CNN (SGECNN) to reduce redundancy and emphasize localized fault patterns. Jiang et al. (2023) adopted Capsule Networks to capture part–whole relationships in vibration images, supporting robust generalization. Moreover, Rajabioun et al. (2023) combined a multi-axis sensory board with hybrid deep models to aggregate raw and engineered features for distributed classification. These developments extend earlier hybrid diagnostic approaches, such as the wavelet-transform and fuzzy-inference method proposed by Lou and Loparo (2003). More recently, Huang et al. (2025) introduced an enhanced dual-channel CNN with bidirectional LSTM fusion and attention mechanisms for aero-engine bearing diagnosis, enabling more effective spatiotemporal feature integration. Qiu et al. (2025) developed a dynamic adaptive fault diagnosis framework based on multi-channel image fusion and deep learning that remained operational under partial sensor failure. Raj et al. (2024) provided a systematic review of bearing component failure modes and diagnostic strategies, covering both traditional and data-driven approaches for fault detection and prognosis in rotating machinery.

In this study, an applied comparative framework is presented for automated rolling bearing fault identification using ANNs optimized through PSO, BO, and GA. Rather than introducing a fundamentally new machine-learning architecture, this study systematically evaluates established optimization strategies for improving ANN structure and diagnostic performance. The optimization focuses on the ANN architecture, including hidden layers, number of neurons, and activation functions.

Vibration signals were collected from bearings under five health conditions and three rotational speeds, with measurements acquired in axial, vertical, and horizontal directions. After filtering, normalization, and segmentation, time-domain and frequency-domain features were extracted to train several ANN configurations. The objective is to identify the most effective combination of feature representation and optimization algorithm for reliable rolling bearing fault diagnosis.

Among the evaluated configurations, PSO combined with frequency-domain features showed the best overall diagnostic performance. This configuration is presented as a comparatively simple and lightweight alternative to more complex deep-learning approaches for PHM applications. However, comparisons with previously reported state-of-the-art studies should be interpreted cautiously because differences in datasets, fault classes, preprocessing procedures, data-partitioning strategies, experimental conditions, and evaluation protocols may limit direct performance equivalence. Therefore, these comparisons provide a contextual reference rather than evidence of the absolute superiority of the proposed approach.

The remainder of the paper is organized as follows: Section 2 presents the ANN modeling framework and the associated hyperparameter optimization procedures using BO, PSO, and GA. Section 3 describes the experimental setup and vibration signal acquisition process. Section 4 presents the performance evaluation of the optimized ANN architectures using convergence analysis, classification metrics, and confusion matrices. It also includes two benchmarking levels: a direct comparison with conventional classifiers, namely SVM, kNN, and RF, and a broader comparison with recent state-of-the-art studies reported in the literature. Finally, Section 5 summarizes the main conclusions and highlights potential directions for future research.

2. METHODOLOGY

2.1. Overview of the proposed approach

Figure 1 illustrates the overall methodology adopted for rolling bearing fault classification using an ANN. The approach begins with vibration signal acquisition under five bearing conditions at three rotational speeds and along three measurement directions.

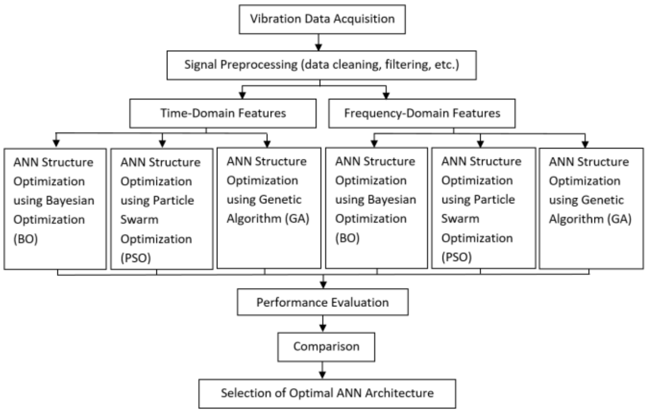


Figure 1. Workflow of ANN architecture optimization using BO, PSO, and GA with time and frequency domain vibration features.

The acquired signals are then preprocessed through noise reduction, normalization, and segmentation. After preprocessing, two types of features are extracted: time-domain features and frequency-domain features based on the FFT.

These feature sets are used as inputs to optimized ANN models using three algorithms: BO, PSO, and GA. The resulting ANN configurations are evaluated using classification metrics, including F1-score, Accuracy, precision, recall, and confusion matrices, since rolling bearing fault diagnosis is formulated as a multi-class classification problem**.** Regression-based indicators such as mean squared error (MSE) and coefficient of determination (R^2) are retained only as complementary measures to analyze training behavior and convergence. A comparative analysis is finally conducted to identify the most effective model for rolling bearing fault diagnosis, including comparisons with conventional classifiers and recent diagnostic methods reported in the literature.

2.2. Data preparation

The dataset used in this study comprises time and frequency domain features extracted from vibration signals recorded under five bearing health conditions: healthy, outer ring fault, inner ring fault, ball fault, and combined faults. Measurements were conducted in three orthogonal directions: axial, vertical, and horizontal, using appropriate vibration sensors. To simulate diverse operational conditions, the data were collected at three different rotational speeds: 1800 rpm, 2400 rpm, and 3000 rpm and three directions: axial, vertical and horizontal. Each sample in the dataset represents a segment of vibration data processed into a set of discriminative features. These features were structured into input vectors, while the corresponding output labels denote the specific bearing fault class. This multi-condition, multi-speed, and multi-directional dataset was prepared to train and evaluate the model's robustness in real-world diagnostic scenarios.

2.3. ANN Architecture and hyperparameter definition

A multi-layer feedforward artificial neural network was used to capture the nonlinear relationships within the bearing datasets. The ANN architecture was characterized by several tunable hyperparameters: the number of hidden layers (from 1to 4), the number of neurons per layer (from 5 to 50), and the activation functions at each layer. The activation functions used in the hidden layers are not fixed a priori; instead, they are selected using optimization algorithms to identify the most suitable nonlinearity for each layer. The candidate set includes the hyperbolic tangent (tansig), logistic sigmoid (logsig), rectified linear unit (poslin), the linear "identity" function (purelin) and the softmax (normalized exponential) function. This optimization-driven selection enables the network to better match the underlying data characteristics by assigning the most effective activation function to each hidden layer, improving overall predictive performance. This hierarchical structure, tailored via optimization, ensures that the model can generalize across varying fault conditions and feature distributions.

2.4. Bayesian optimization strategy

BO was employed as a probabilistic global optimization framework for efficient ANN hyperparameter tuning. A surrogate model was used to approximate the objective function, while the Expected Improvement Plus acquisition function guided the search by balancing exploration and exploitation. The optimization was conducted over 30 iterations, with each candidate hyperparameter configuration evaluated using a custom objective function based on validation loss. The optimized variables included the number of hidden layers, the number of neurons in each layer, and the activation functions. The search space was defined from 1 to 4 hidden layers, 5 to 50 neurons per layer, and five candidate activation functions: hyperbolic tangent sigmoid, logistic sigmoid, linear activation, rectified linear unit, and softmax. For each candidate architecture, a feedforward ANN was trained using the Levenberg–Marquardt algorithm, with a maximum of 500 epochs, a learning rate of 0.01, a performance goal of 10^{-6}, and an early-stopping criterion of 10 validation failures. The data were divided into 70% for training, 15% for validation, and 15% for testing, and Nguyen–Widrow initialization was used to initialize the network weights and biases. The optimal ANN architecture identified by BO was subsequently retrained, and its performance was assessed primarily using classification metrics, including F1-score, Accuracy, precision, recall, and confusion matrices, in accordance with the multi-class nature of the bearing fault diagnosis task. Regression-based indicators, such as MSE and the coefficient of determination (R^2), were retained only as complementary measures for analyzing training behavior and convergence trends.

Probabilistic optimization approaches such as BO have been widely reported to provide structured and efficient hyperparameter exploration in ANN-based fault diagnosis applications (Wu et al., 2019).

2.5. Particle swarm optimization strategy

PSO was implemented as a population-based metaheuristic inspired by collective social behavior. A swarm of 20 particles was initialized, with each particle encoding a complete ANN hyperparameter vector. Particle positions and velocities were iteratively updated based on individual best and global best solutions, using the same ANN objective function applied in BO. The hyperparameter vector encoded the number of hidden layers, neuron counts, and activation functions. The search bounds were set to 1–4 hidden layers, 5–50 neurons per layer, and five activation functions: hyperbolic tangent sigmoid, logistic sigmoid, linear activation, rectified linear unit, and softmax. The PSO parameters were set to an inertia weight of 0.7, a cognitive coefficient (c_1 = 1.5), and a social coefficient (c_2 = 1.5). Particle positions were constrained within the predefined lower and upper bounds. The PSO process was executed for 30 iterations, during which convergence characteristics and iteration-wise best configurations were recorded. The best-performing hyperparameter set was then used to train the final ANN model and evaluated using classification metrics, including F1-score, Accuracy, precision, recall, and confusion matrices. MSE and R^2 were retained only as complementary indicators for training behavior and convergence analysis. Swarm-based optimization techniques have demonstrated strong effectiveness for ANN tuning in machinery condition monitoring and bearing fault diagnosis (Yu et al., 2023b).

2.6. Genetic Algorithm optimization strategy

GA was adopted as an evolutionary optimization approach for ANN hyperparameter selection. Each individual in the population represented a candidate ANN architecture defined by the number of hidden layers, neuron counts, and activation functions. An initial population of 20 individuals was evolved over 30 generations using selection, crossover, and mutation operators to minimize validation loss. The search space was bounded between 1 and 4 hidden layers, 5 and 50 neurons per layer, and five candidate activation functions: hyperbolic tangent sigmoid, logistic sigmoid, linear activation, rectified linear unit, and softmax. Parallel computation was disabled to ensure sequential and repeatable evaluation of candidate architectures. Since the selection, crossover, mutation, and elitism parameters were not manually specified, the standard settings of the optimization routine were used for these operators. Fitness evaluation was performed using the same ANN objective function employed across all optimization methods, and the best solution from each generation was recorded to analyze convergence behavior. The optimal hyperparameter configuration obtained by GA was subsequently used to train the final ANN model and evaluated using classification metrics, including F1-score, Accuracy, precision, recall, and confusion matrices. MSE and R^2 were used only as complementary indicators for training behavior and convergence analysis. Evolutionary optimization methods have been extensively applied to ANN-based fault diagnosis problems due to their robustness to local optima and suitability for nonlinear search spaces (El-Hassani et al., 2024).

2.7. Evaluation of optimization strategies

The optimization strategies, BO, PSO, and GA, were evaluated based on their effectiveness in identifying optimal ANN structures. Specifically, the evaluation focused on each method's ability to automatically determine the number of hidden layers, the number of neurons per layer, and the activation functions used in the hidden layers.

All optimized ANN configurations were assessed under identical training and validation conditions using classification-based performance metrics, including F1-score, Accuracy, precision, recall, and confusion matrices, while MSE and R^2 were retained as complementary indicators of training behavior. In addition to diagnostic performance, the convergence behavior of each optimization strategy was analyzed to evaluate search efficiency, stability, and consistency in exploring the ANN architectural space. Monitoring the evolution of the objective function across iterations provided insight into the balance between exploration and exploitation achieved by each method.

This evaluation framework is motivated by prior studies indicating that PSO-based optimization can improve convergence and neural-network-based bearing fault diagnosis (Yu et al., 2023b), while Bayesian optimization strategies provide robust and structured hyperparameter exploration in ANN training (Onorato, 2024). In addition, Genetic Algorithms have been widely adopted for ANN structural optimization, demonstrating strong global search capability and effectiveness in determining optimal neuron configurations and activation functions for fault diagnosis and pattern recognition tasks (Chen et al., 2024).

3. EXPERIMENTAL SETUP

3.1. Test rig and measurement conditions

In this study, vibration signals were collected using a machinery diagnostic system from five U205 ball bearings, each representing a different health condition and corresponding classification label: Class 0: healthy bearing, Class 1: outer ring fault, Class 2: inner ring fault, Class 3: ball fault, and Class 4: combined faults. The experiments were conducted on a specialized experimental setup designed to simulate realistic operating conditions. The setup consisted of a rotating shaft supported by rolling element bearings and driven by a variable-speed motor. To reflect practical fault scenarios, the five bearing conditions were tested under three different rotational speeds: 1800 rpm, 2400 rpm, and 3000 rpm (Qian et al., 2024; Chi et al., 2022), representing typical ranges found in industrial machinery.

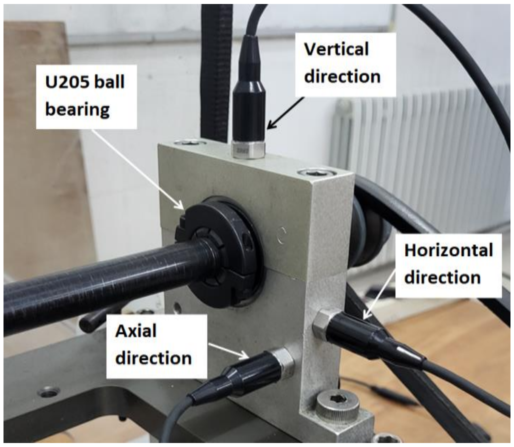


Figure 2. Experimental Setup for Vibration Measurement Using Accelerometers in Three Directions

The experimental setup, illustrated in Figure 2, includes three accelerometers mounted on the bearing housing to measure vibrations in the vertical, axial, and horizontal directions. Data were collected at a sampling frequency of 10400 Hz, enabling high-resolution capture of the vibration signals. Each test was carried out under controlled load and speed conditions to ensure consistency and repeatability. The collected signals form the primary dataset for subsequent frequency domain analysis and fault diagnosis.

3.2. Sensor configuration and data acquisition

The vibration signals were acquired using three piezoelectric accelerometers mounted on the bearing in the vertical, axial, and horizontal directions, along with a tachometer to monitor rotational speed and ensure synchronization. As shown in Figure 2, the multi-directional sensor arrangement enabled simultaneous acquisition of vibration data along the three axes. All signals were sampled at 10400 Hz, ensuring accurate representation of both time and frequency domain characteristics.

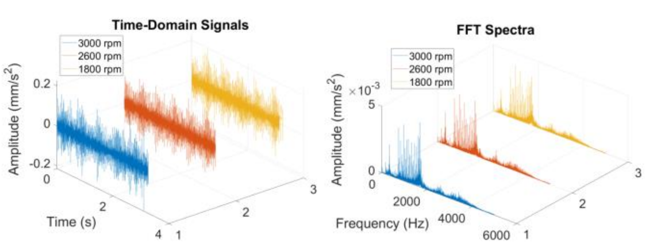


Figure 3. Time and frequency domain representations of vertical vibration signals at three rotational speeds (1800, 2400, and 3000 rpm)

The effect of rotational speed on vibration behavior is illustrated in Figure 3. As the speed increased from 1800 to 3000 rpm, the vibration amplitudes became more pronounced in the time domain, while the FFT spectra exhibited stronger and higher-frequency peaks. This indicates that higher rotational speeds intensify the dynamic response of the bearing and make fault-related signatures more visible.

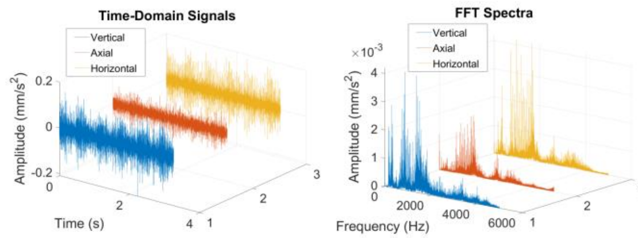


Figure 4. Time and frequency domain representations of vibration signals in three spatial directions (axial, horizontal, and vertical) at 3000 rpm

A comparison of the vibration responses in the three spatial directions at 3000 rpm is presented in Figure 4. The vertical direction showed the highest amplitudes and the most significant spectral peaks, followed by the horizontal direction, whereas the axial direction displayed the weakest response. This confirms that radial directions are more sensitive to typical bearing defects such as spalling, which mainly generate impact forces in the radial plane.

Finally, Figure 5 compares the healthy bearing condition with different fault states, including ball faults, outer race faults, inner race faults, and combined faults. The healthy bearing exhibited a relatively smooth time-domain signal and a sparse frequency spectrum, while faulty bearings generated periodic impacts and clear characteristic frequency peaks. Combined faults produced multiple spectral components corresponding to the coexistence of several defect types.

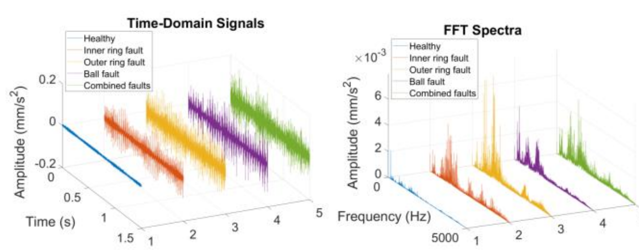


Figure 5. Time and frequency domain comparison of vibration signals across bearing conditions (healthy, inner ring fault, outer ring fault, ball fault, combined faults) at 3000 rpm

Overall, the results confirm that combined time-domain and frequency-domain vibration analysis, supported by multi-directional sensing, is an effective tool for bearing fault diagnosis and condition monitoring.

3.3. Signal preprocessing and feature extraction

The raw vibration signals acquired from the rolling bearings were subjected to a preprocessing stage prior to feature extraction in order to improve data quality and ensure consistency across all operating conditions. First, band-pass filtering was applied to eliminate unwanted high-frequency noise and low-frequency drift while preserving the frequency components relevant to bearing fault diagnosis.

Table 1. Time-domain features

Feature	Mathematical Expression
RMS (Root Mean Square)	$RMS = \sqrt{\frac{1}{N}\sum_{i=1}^{N} x_i^2}$
Peak Value	$Peak = \max(\|x_i\|)$
Peak-to-Peak	$PP = max(x_i) - min(x_i)$
Mean	$\mu = \frac{1}{N}\sum_{i=1}^{N} x_i$
Mean Absolute Value	$MAV = \frac{1}{N}\sum_{i=1}^{N} \|x_i\|$
Standard Deviation	$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N} (x_i - \mu)^2}$
Skewness	$Skewness = \frac{\frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^3}{\sigma^3}$
Kurtosis	$K = \frac{\frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^4}{\sigma^4}$
Shape Factor	$SF = RMS/MAV$
Impulse Factor	$IF = Peak/MAV$
Clearance Factor	$CF = \frac{Peak}{\left(\frac{1}{N}\sum_{i=1}^{N}\sqrt{\|x_i\|}\right)^2}$
Crest Factor	$Crest = Peak/RMS$

The maximum analyzed frequency was limited to 5000 Hz, while the sampling frequency was set to 10,400 Hz, satisfying the Nyquist–Shannon sampling criterion. The filtered signals were then normalized to standardize their amplitude ranges and reduce the influence of measurement variability. Subsequently, each vibration signal was acquired over a total duration of 6 s and divided into eight consecutive blocks, resulting in 0.75 s per block. At the selected sampling frequency, each block contained 7800 samples, while the complete acquisition contained 62,400 samples. No overlap was applied between consecutive blocks, and each block was used as a fixed-length segment for extracting time-domain and frequency-domain features. The segment length was selected to ensure uniformity and to include complete vibration cycles at different rotational speeds, thereby preserving the characteristic patterns associated with bearing faults. Overall, these preprocessing steps provided clean, consistent, and reliable input data for the subsequent diagnostic analysis. To prevent data leakage, all eight segments from each six-second recording were assigned to the same training, validation, or test subset. The dataset was divided into 70%, 15%, and 15% for training, validation, and testing, respectively.

Table 2. Frequency-domain features

Feature	Mathematical Expression
Dominant Frequency	$f_{dom} = arg\max(\|X(f)\|)$
Spectral Centroid	$C = \frac{\sum_i f_i \|X(f_i)\|}{\sum_i \|X(f_i)\|}$
Spectral Spread	$SS = \sqrt{\frac{\sum_i (f_i - C)^2 \|X(f_i)\|}{\sum_i \|X(f_i)\|}}$
Spectral Entropy	$H = -\sum_i \frac{\|X(f_i)\|^2}{\sum_j \|X(f_j)\|^2} log_2\left(\frac{\|X(f_i)\|^2}{\sum_j \|X(f_j)\|^2}\right)$
Spectral Flatness	$SF = \frac{exp\left(\frac{1}{N}\sum_i log\,\|X(f_i)\|\right)}{\frac{1}{N}\sum_i \|X(f_i)\|}$
Peak Amplitude	$A_{peak} = \max(\|X(f)\|)$
FFT Energy	$E_{FFT} = \sum_i \|X(f_i)\|^2$
FFT Mean	$\mu_{FFT} = \frac{1}{N}\sum_i \|X(f_i)\|$
FFT Std	$\sigma_{FFT} = \sqrt{\frac{1}{N}\sum_i (\|X(f_i)\| - \mu_{FFT})^2}$
FFT Skewness	$\gamma_{skew} = \frac{\frac{1}{N}\sum_i (\|X(f_i)\| - \mu_{FFT})^3}{\sigma_{FFT}^3}$
FFT Kurtosis	$\gamma_{kutr} = \frac{\frac{1}{N}\sum_i (\|X(f_i)\| - \mu_{FFT})^4}{\sigma_{FFT}^4}$
Frequency Variance	$V_{arf} = \frac{1}{N}\sum_i (f_i - \mu_f)^2 \|X(f_i)\|$
Band Power 0–100 Hz	$P_{0-100} = \sum_{f=0}^{100} \|X(f)\|^2$
Band Power 100–300 Hz	$P_{100-300} = \sum_{f=100}^{300} \|X(f)\|^2$
Number of Peaks	$N_{peaks} = Count\ of\ local\ maxima\ in\ \|X(f)\|$
Ratio Low/High Band Power	$R_{low/high} = P_{0-100}/P_{100-300}$

Considering five bearing conditions, three rotational speeds, three measurement directions, and eight blocks per acquisition, the resulting dataset included 360 feature vectors for each feature domain.

As summarized in Table 1, the selected twelve indicators provide reliable discrimination capability and contribute to early fault detection and improved system reliability (Jain & Bhosle, 2022; Khoualdia et al., 2025b).

Frequency-domain features derived from vibration signals play a fundamental role in rolling bearing fault analysis by providing spectral information associated with bearing health and defect-related vibration behavior. These indicators, extracted through Fourier transform analysis and summarized in Table 2, include parameters such as dominant frequency, spectral centroid, spectral spread, and spectral entropy. When used as inputs to ANNs, they allow accurate differentiation between normal and faulty bearing states and contribute significantly to fault diagnosis, condition assessment, and predictive maintenance (Aiordăchioaie, 2024; Sharma et al., 2018).

For training and testing the ANN models, both time domain and frequency domain features were organized into input vectors. These vectors, encapsulating the key aspects of vibration signals, allowed the ANN models to learn the correlation between the signal patterns and fault types. During training, the models associated these features with specific faults, and their ability to generalize and accurately diagnose faults was tested.

4. PERFORMANCE EVALUATION OF OPTIMIZED ANN CONFIGURATIONS

4.1. Architectural configurations of optimized ANNs

Table 3 summarizes the optimized ANN architectures obtained for two feature domains, namely time-domain and frequency-domain features, using three optimization algorithms: BO, PSO, and GA. For time-domain features, BO yielded a two-hidden-layer network with 30 and 15 neurons, using softmax activation and hyperbolic tangent sigmoid functions. This configuration indicates that BO selected a mixed activation structure for representing time-domain vibration patterns. In contrast, PSO selected a three-hidden-layer architecture with 47, 47, and 13 neurons, where all hidden layers employed logistic sigmoid activation functions, suggesting a consistent nonlinear mapping strategy for temporal features. Using GA with time-domain features produced a single-hidden-layer structure with 34 neurons and a logistic sigmoid activation function, indicating a more compact architecture for this feature set.

For frequency-domain features, BO identified a single-hidden-layer network with 23 neurons and a linear activation function, suggesting that the extracted spectral features provide a relatively informative representation for class separation. PSO optimized a three-hidden-layer model with 19, 13, and 43 neurons, using two softmax activation functions followed by a logistic sigmoid function, reflecting a more layered transformation of FFT-based patterns. Finally, GA selected a three-hidden-layer structure with 12, 17, and 6 neurons, where all hidden layers used hyperbolic tangent sigmoid activation functions, indicating a stable nonlinear configuration for frequency-domain inputs. The output layer consisted of five neurons, each representing one bearing-condition class. The target classes were one-hot encoded. The Softmax activation function was applied to the output layer, and the network was trained using the cross-entropy loss function. Each observation was assigned to the class corresponding to the neuron with the highest output value.

Table 3. Optimized Artificial Neural Network Structures for Bearing Fault Detection

Method and features	Hidden layers	Hidden layers Activation functions
BO with Time Domain	{30, 15}	{softmax, hyperbolic tangent sigmoid}
BO with Frequency Domain	{23}	{linear activation function}
PSO with Time Domain	{47, 47, 13}	{logistic sigmoid, logistic sigmoid, logistic sigmoid}
PSO with Frequency Domain	{19, 13, 43}	{softmax, softmax, logistic sigmoid}
GA with Time Domain	{34}	{logistic sigmoid }
GA with Frequency Domain	{12, 17, 6}	{hyperbolic tangent sigmoid, hyperbolic tangent sigmoid, hyperbolic tangent sigmoid}

Overall, the results in Table 3 show that both the feature domain and the optimization algorithm substantially influence the ANN depth, neuron distribution, and activation-function selection, leading to distinct architectural preferences for time versus frequency-based bearing fault representations.

4.2. Optimization dynamics for time and frequency domain features

The convergence behavior of the optimization algorithms provides insight into their effectiveness in identifying ANN architectures that support accurate fault identification. Figure 6 illustrates the convergence characteristics of BO, PSO, and GA when applied to time-domain and frequency-domain vibration features.

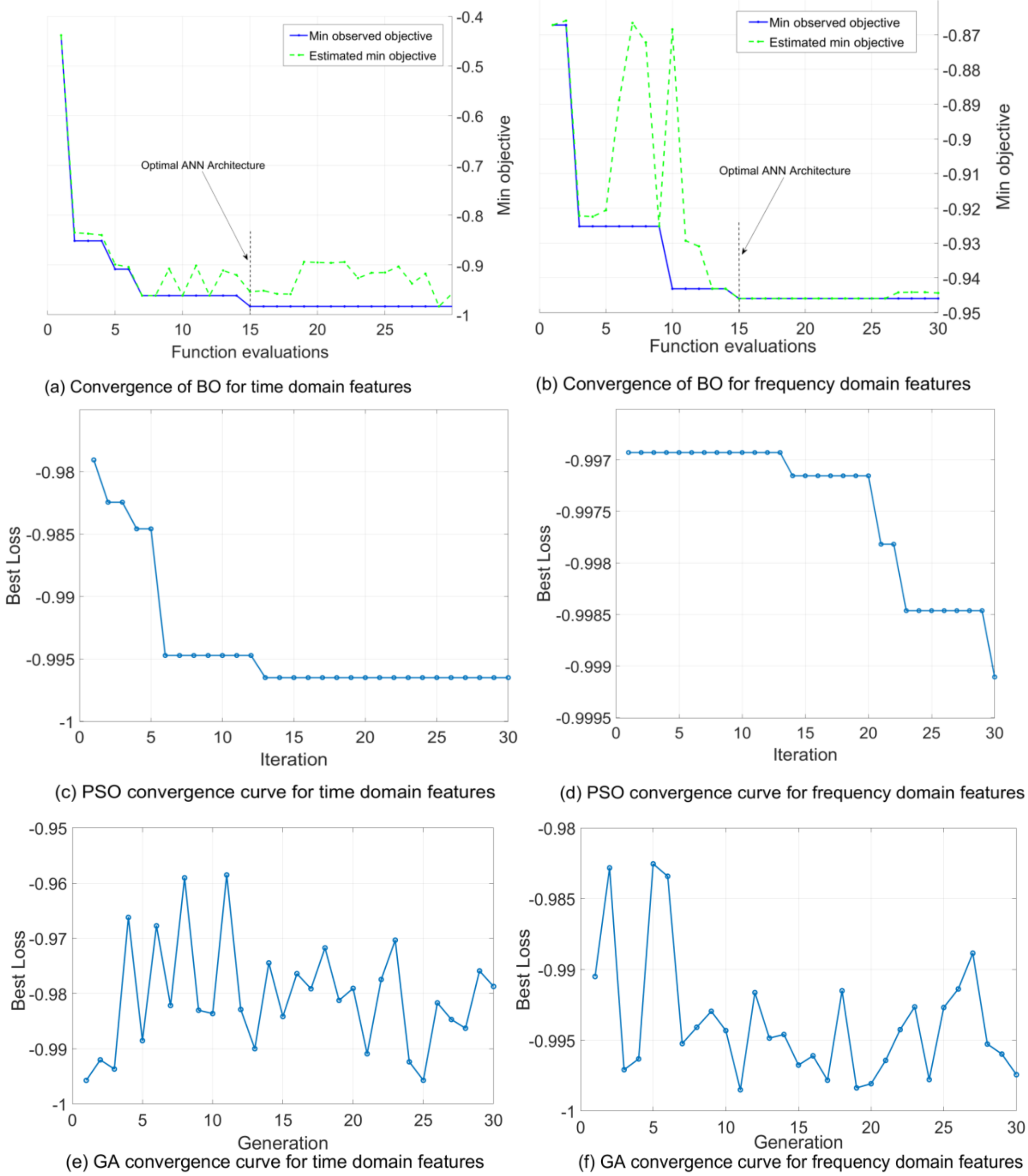
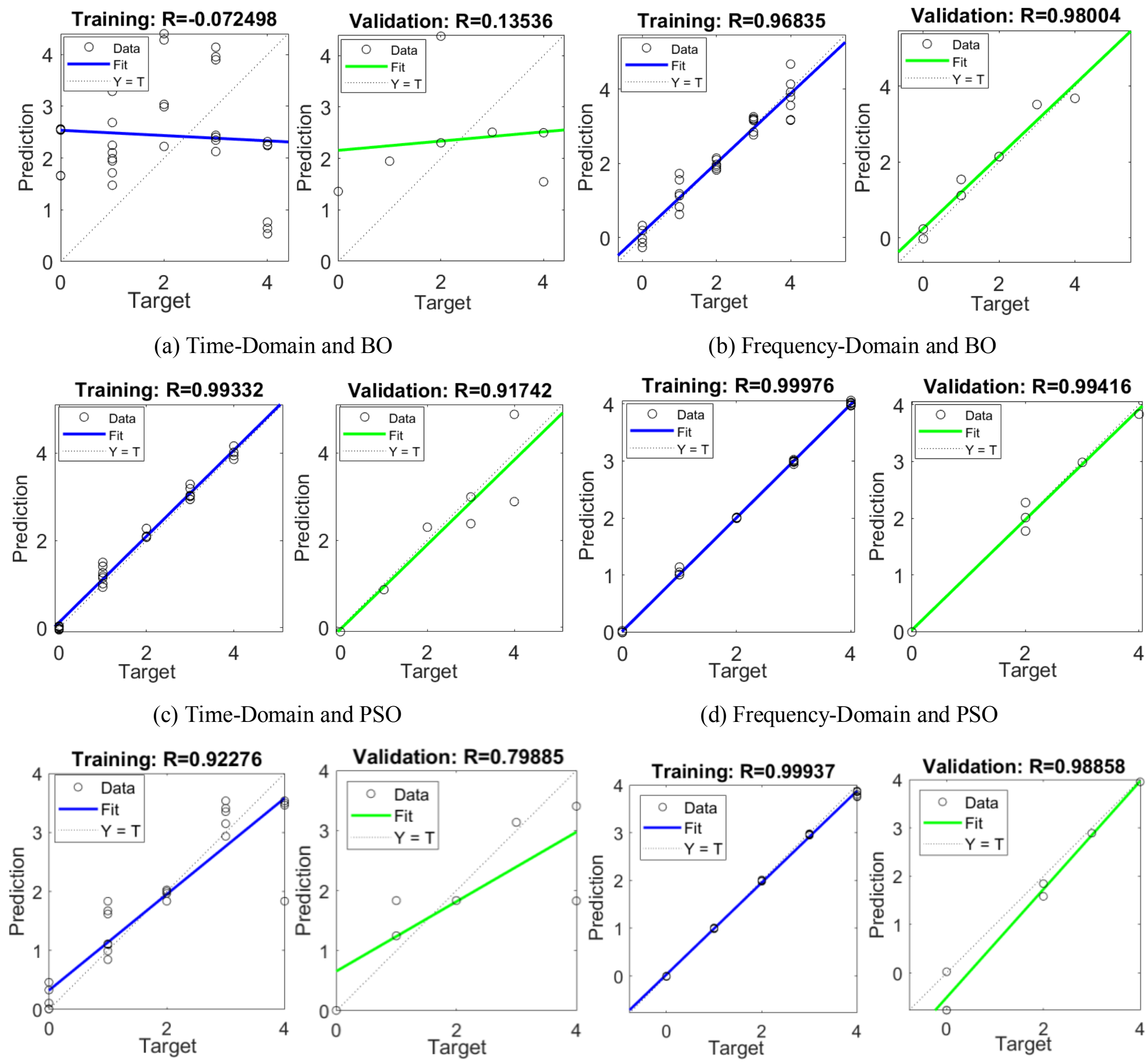


Figure 6. Convergence behavior of ANN hyperparameter optimization using BO, PSO and GA for time domain and frequency domain vibration features.

For BO, distinct convergence patterns are observed across feature domains. When time-domain features are used, the objective value decreases rapidly during the initial evaluations but exhibits noticeable fluctuations after approximately 14 function evaluations, indicating instability in the surrogate model. This behavior is consistent with weaker classification performance, as reflected by lower F1-score, precision, recall, and Accuracy values. In contrast, BO applied to frequency-domain features converges smoothly after approximately 16 evaluations, with close agreement between observed and estimated minima. This stable convergence reflects a more reliable search process and corresponds to the improved diagnostic performance achieved with spectral features.

The PSO convergence curves further emphasize the impact of feature representation. For time-domain features, PSO shows rapid early convergence followed by a plateau, suggesting premature convergence to a suboptimal solution. Conversely, when frequency-domain features are employed, PSO demonstrates continuous improvement throughout the optimization process, achieving progressively lower objective values. This sustained convergence behavior indicates an effective balance between exploration and exploitation, enabling PSO to identify ANN architectures that yield superior fault identification performance.



(a) Time-Domain and BO

(b) Frequency-Domain and BO

(c) Time-Domain and PSO

(d) Frequency-Domain and PSO

(e) Time-Domain and GA

(f) Frequency-Domain and GA

Figure 7. Regression Performance of Optimized ANN Models Across Training, Validation, Test, and Overall Sets for Different Feature Domains and Optimization Methods (BO, PSO, and GA)

GA exhibits higher variability across generations due to its stochastic evolutionary operators. For time-domain features, pronounced oscillations in the best loss indicate difficulty in consistently refining candidate solutions, which aligns with reduced fault identification accuracy. When GA is applied to frequency-domain features, convergence becomes more stable and lower objective values are achieved, although the refinement process remains less smooth than that of PSO. These observations indicate that GA provides robust global exploration, but its convergence stability and fine-tuning efficiency depend strongly on feature representation.

Overall, the convergence analysis demonstrates that frequency-domain features facilitate faster and more stable optimization across all methods, with PSO exhibiting the most consistent refinement behavior and BO achieving rapid convergence when supported by informative spectral features.

4.3. Influence of feature domain and optimization method on classification accuracy

As shown in Figure 7, the influence of feature domain and optimization method on ANN-based fault identification performance was quantitatively evaluated using regression correlation coefficients (R-values) and best MSE, as summarized in Table 4 and illustrated in Figures 8 and 9. The results indicate that both factors play a decisive role in determining identification accuracy and generalization capability.

Across all optimization methods, frequency-domain features consistently outperform time-domain features. For BO, the use of time-domain features results in poor fault identification capability, with an overall R value of 0.044 and a high validation MSE of 2.4285, indicating ineffective learning and weak feature discrimination. When frequency-domain features are employed, BO performance improves substantially, achieving an overall R of 0.939 and a reduced validation MSE of 0.1074, confirming the advantage of spectral features for Bayesian-guided ANN optimization.

PSO demonstrates strong fault identification performance across both feature domains, with a clear advantage for frequency-domain features. The time-domain PSO configuration achieves an overall R of 0.956 and a validation MSE of 0.3609, indicating effective learning but limited refinement. In contrast, PSO combined with frequency-domain features yields the best performance among all evaluated models, achieving an overall R of 0.998 and the lowest validation MSE of 0.0223. These results reflect excellent consistency across training, validation, and test sets, as also observed in the regression plots and comparative performance charts.

For GA optimization, performance again depends strongly on the feature domain. The time-domain GA model achieves an overall R of 0.808 with a validation MSE of 0.5271, reflecting unstable learning and reduced generalization. When frequency-domain features are used, GA performance improves markedly, achieving an overall R of 0.987 and a validation MSE of 0.1998. Although this represents strong fault identification capability, GA remains slightly inferior to PSO in terms of both accuracy and convergence efficiency.

Table 4. Performance Comparison of Optimized ANN Architectures Based on Feature Domain and Optimization Method

Method and features	**Training R**	**Validation R**	**Test R**	**Overall R**	**Best Validation MSE**	**Epoch**	**Performance Rating**
BO with Time Domain	-0.072	0.135	0.698	0.044	2.4285	1	Poor
BO with Frequency Domain	0.968	0.980	0.884	0.939	0.1074	5	Very Good
PSO with Time Domain	0.993	0.917	0.864	0.956	0.3609	6	Excellent
PSO with Frequency Domain	0.999	0.994	0.991	0.998	0.0223	12	Outstanding (Optimal)
GA with Time Domain	0.923	0.799	0.148	0.808	0.5271	4	Good
GA with Frequency Domain	0.999	0.989	0.960	0.987	0.1998	10	Excellent

Bar charts in Figure 8 reinforce these findings, showing the Freq-PSO model consistently outperformed other configurations across all metrics. Conversely, Time-BO showed the weakest performance, with low R-values and the highest error.

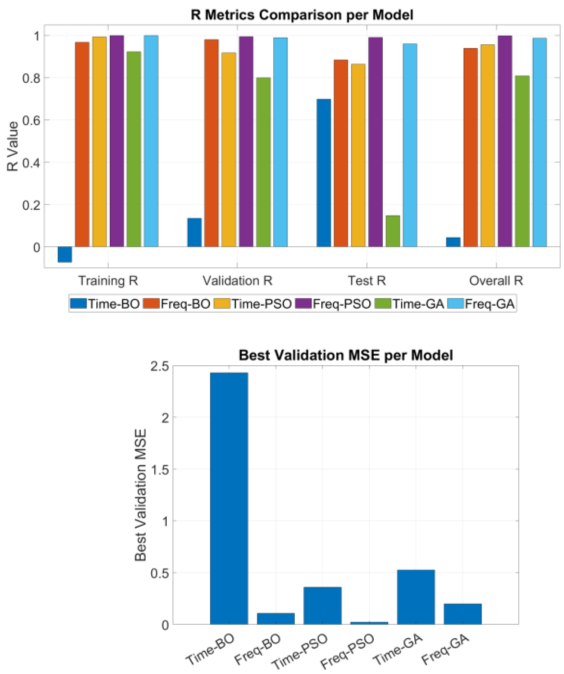


Figure 8. Comparative R-Metric Performance of Optimized ANN Models

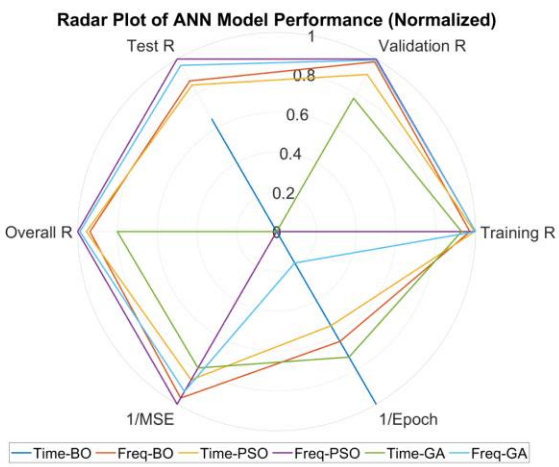


Figure 9. Radar Plot Comparing Normalized R Metrics, MSE, and Training Epochs for Optimized ANN Models

The radar plots further consolidate these findings by showing that the frequency-domain PSO configuration consistently achieves the highest normalized scores across multiple performance indicators, while the time-domain BO configuration exhibits the weakest performance. Overall, the results demonstrate that frequency-domain features provide superior discriminative information for rolling bearing fault identification, and that PSO is the most effective optimization method for ANN architecture selection in this application, followed by GA and BO.

4.4. Benchmarking Optimized ANN Models against Conventional Classifiers

For benchmarking purposes, three conventional classifiers were implemented using the same training and testing data as the optimized ANN models: SVM, kNN, and RF. The SVM classifier was configured with a radial basis function (RBF) kernel, a regularization parameter (C = 1.00), a numerical tolerance of 0.001, and an iteration limit of 100. The RBF kernel was selected because of its ability to model nonlinear decision boundaries, which are commonly encountered in vibration-based bearing fault classification. The kNN classifier was implemented with (k = 20) neighbors, providing a balanced compromise between local sensitivity and robustness to noise in the extracted feature space. The RF classifier was configured with 20 decision trees, replicable training enabled, a maximum tree depth limited to 3, and a minimum subset size of 5 samples before further splitting. These constraints were used to reduce overfitting and maintain a lightweight baseline model. Class balancing was not activated, so all classifiers were evaluated under the original class distribution. It should be emphasized that the hyperparameters of SVM, kNN, and RF were not optimized; these classifiers were included as fixed, lightweight, and reproducible conventional baselines. Therefore, the comparison provides a common reference under the same dataset and feature conditions and should not be interpreted as a comparison against fully optimized conventional classifiers. The confusion matrices obtained using time-domain vibration features are presented in Figure 10, while the corresponding performance metrics are included in the comparative results of Table 5, thus, Precision, recall, and F1-score were computed using macro-averaging across the five bearing classes, while Accuracy was calculated using a one-vs-rest strategy for the multi-class classification task. In addition, standard multiclass accuracy was calculated as the total number of correctly classified test samples divided by the total number of test samples (Accuracy = $\Sigma_i C_{ii} / \Sigma_i \Sigma_j C_{ij}$). The results show that the optimized ANN models provide better class discrimination than the conventional classifiers, namely SVM, kNN, and Random Forest. Among the optimized models, ANN-GA achieved the best performance for time-domain features, with a standard multiclass accuracy of 0.807, an Accuracy of 0.928, an F1-score of 0.801, a precision of 0.808, and a recall of 0.807. ANN-BO and ANN-PSO also outperformed the conventional classifiers, although their confusion matrices still reveal some misclassifications across several classes. For the traditional methods, SVM obtained the highest F1-score, precision, and recall, whereas kNN achieved the highest Accuracy. These findings indicate that ANN architecture optimization improves the classification capability when using time-domain features, but the overall performance remains moderate compared with the FFT-based representation reported in Table 5.

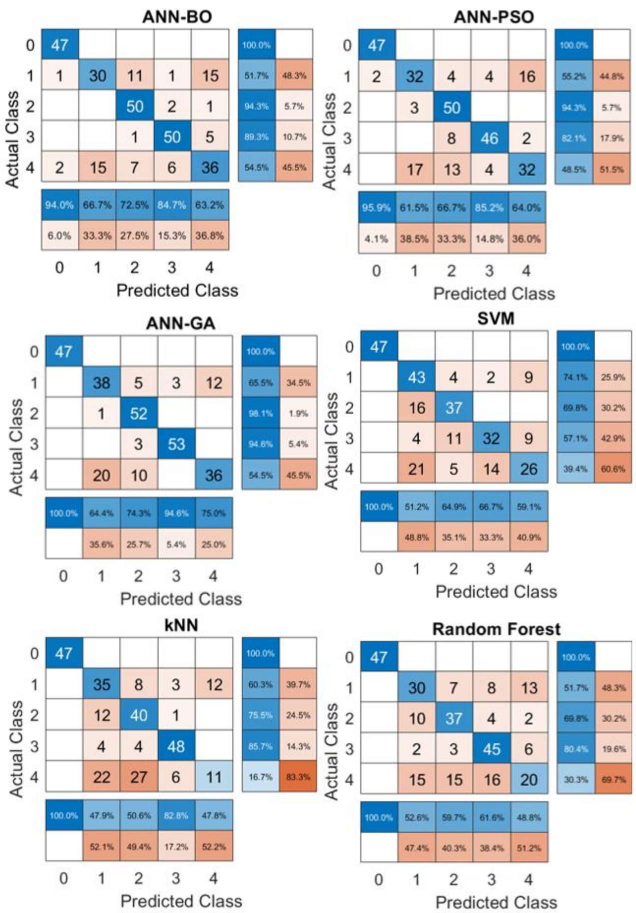


Figure 10. Confusion matrices of the optimized ANN models and conventional classifiers using time-domain vibration features

The confusion matrices based on FFT features, shown in Figure 11, indicate a clearer concentration of samples along the main diagonal compared with the time-domain case, reflecting improved class separation and reduced misclassification rates. This improvement is also confirmed by the metrics reported in Table 5, where all models generally achieved higher performance with FFT-based features. Among the optimized ANN models, ANN-PSO provided the best overall results, reaching a standard multiclass accuracy of 0.997, an Accuracy of 0.998 and F1-score, precision, and recall values of 0.982. ANN-BO and ANN-GA also demonstrated strong performance, with standard multiclass accuracies of 0.936 and 0.921, respectively, and F1-scores of 0.936 and 0.922, respectively. Although the conventional classifiers also benefited from the FFT representation, their performance remained below that of the optimized ANN models, particularly in terms of F1-score and recall. Overall, the joint analysis of Figures 10 and 11 and Table 5 confirms that frequency-domain features contain more discriminative information and that optimized ANN architectures are more effective for the considered classification task.

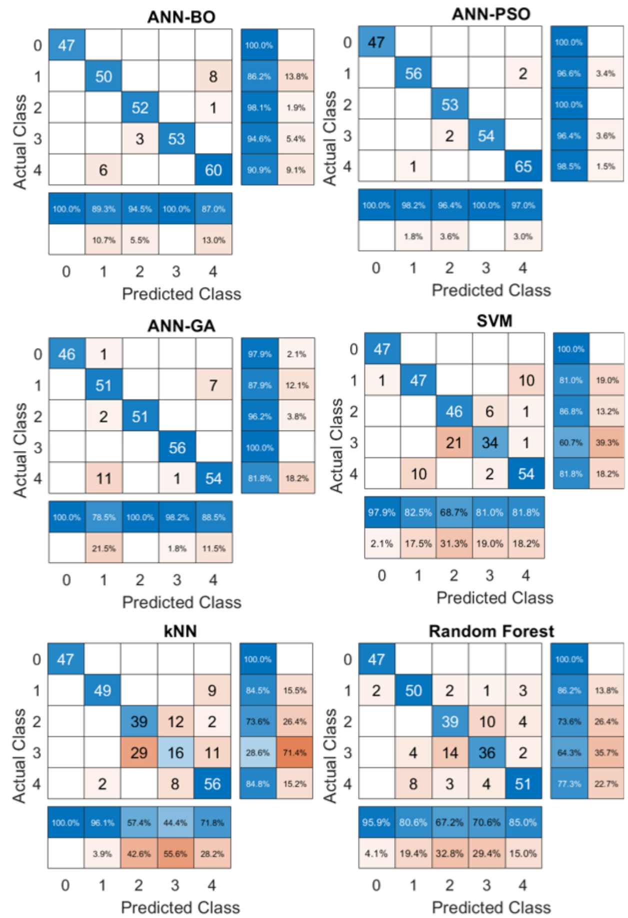


Figure 11. Confusion matrices of the optimized ANN models and conventional classifiers using FFT-based vibration features

Although the optimized ANN-PSO configuration contains three hidden layers, it remains a compact feedforward neural network compared with multi-component deep-learning architectures involving convolutional, recurrent, or attention-based modules.

4.5. Benchmarking against state-of-the-art diagnostic methods

As shown in Table 6, the comparative evaluation highlights the practical relevance of the ANN-PSO approach, where frequency-domain features are used for bearing fault diagnosis and the ANN structure is optimized in terms of hidden layers, neurons, and activation functions. Since this is a multi-class classification task, performance is assessed using classification metrics rather than regression indicators. The ANN-PSO model with frequency-domain features achieved an F1-score of 0.982, an Accuracy of 0.998, a precision of 0.982, and a recall of 0.982, demonstrating competitive diagnostic performance under the considered experimental conditions**.**

Table 5. Comparative performance metrics of optimized ANN models and conventional classifiers using time-domain and FFT-based vibration features

Model / Feature type	Accuracy	F1-Score	AUC	Precision	Recall
ANN-BO / Time-domain	0.761	0.730	0.921	0.736	0.739
ANN-BO / FFT-domain	0.936	0.936	0.988	0.937	0.936
ANN-PSO / Time-domain	0.739	0.750	0.913	0.751	0.761
ANN-PSO / FFT-domain	**0.997**	**0.982**	**0.998**	**0.982**	**0.982**
ANN-GA / Time-domain	0.807	0.801	0.928	0.808	0.807
ANN-GA / FFT- domain	0.921	0.922	0.983	0.925	0.921
SVM / Time-domain	0.661	0.655	0.798	0.669	0.661
SVM / FFT-domain	0.814	0.812	0.923	0.820	0.814
kNN / Time-domain	0.646	0.620	0.911	0.641	0.646
kNN / FFT-domain	0.739	0.729	0.912	0.734	0.739
RF / Time-domain	0.639	0.625	0.871	0.628	0.639
RF / FFT-domain	0.796	0.795	0.956	0.797	0.796

Compared with multi-component deep learning architectures, the proposed ANN-PSO model can be considered a lightweight alternative, since it focuses on optimizing the network structure, including the number of hidden layers, neurons, and activation functions, rather than increasing architectural depth or relying on multiple integrated learning modules. This structural optimization may contribute to faster convergence and reduced computational requirements under the considered experimental conditions**.** For instance, Wu et al. (2025) employed a hybrid framework combining, Fast Fourier Transformer, Squeeze-and-Excitation Temporal Convolutional Networks, and Support Vector Machine, which can improve diagnostic performance but involves convolutional feature extraction and classifier integration, thereby increasing the number of processing stages. Similarly, Liu et al. (2024) introduced a multi-scale quaternion convolutional neural network combined with Bidirectional Gated Recurrent Unit (BiGRU) and cross self-attention fusion, providing strong feature representation capabilities through several dedicated representation-learning components. Therefore, the ANN-PSO configuration is positioned not as a universally superior method, but as a computationally efficient and practically attractive solution for bearing fault diagnosis, especially when implementation simplicity and reduced model size are important.

More interpretability-driven methods, such as Kalay's (2025) uncertainty-aware 1D CNN-LSTM architecture and Singh's (2025) graph-based fault diagnosis model with structural feature extraction, emphasize confidence estimation and data structuring but introduce added layers of model design and post-processing, which limit real-time applicability.

Additionally, although Soares et al. (2025) proposed a hybrid ANN-PSO method with frequency-domain features for fault diagnosis, their focus was primarily on optimizing ANN weights, whereas the proposed work innovates by optimizing the network structure itself, enabling systematic and automated optimization of the ANN architecture and hyperparameters. Hatipoğlu et al. (2024) adopted a multi-domain LSTM-attention network with Least Absolute Shrinkage and Selection Operator (LASSO) regularization, which improves feature selection and model sparsity but still depends on intricate attention mechanisms and manually selected features.

Compared with recent multi-component deep-learning-based methods, the ANN-PSO configuration provides a comparatively simpler and lightweight alternative under the considered experimental conditions. However, this comparison should be interpreted cautiously because the reported studies use different datasets, fault classes, preprocessing procedures, data-partitioning strategies, experimental conditions, and evaluation protocols. Therefore, these results provide only a contextual reference and do not establish the absolute superiority of the proposed approach.

5. CONCLUSION

Rolling bearings are critical components in rotating machinery, and their degradation can cause downtime, operational disruptions, and increased maintenance costs. This study addressed rolling bearing fault diagnosis under five health conditions: healthy, outer ring fault, inner ring fault, ball fault, and combined faults.

To address this diagnostic task, an applied comparative study was conducted using ANN models optimized through PSO, BO, and GA. Rather than proposing a fundamentally new machine-learning architecture, this work focused on optimizing the ANN structure, including hidden layers, number of neurons, and activation functions. Six ANN configurations were evaluated by combining time-domain and frequency-domain features with the three optimization strategies.

The results showed that PSO combined with frequency-domain features achieved the best overall diagnostic performance. Because this is a multi-class classification problem, the evaluation was supported by classification metrics, including F1-score, Accuracy, precision, recall, and confusion matrices. The ANN-PSO model using frequency-domain features achieved F1-score = 0.982, Accuracy = 0.998, precision = 0.982, and recall = 0.982.

The comparative analysis further showed that the optimized ANN-PSO configuration achieved competitive performance compared with SVM, kNN, and RF using the same feature sets**.** Comparisons with recent literature were interpreted cautiously due to differences in datasets and evaluation protocols**.** Overall, this study contributes a systematic applied comparison of ANN optimization strategies for lightweight rolling bearing fault diagnosis**.** Future work should include validation on public benchmark datasets and broader operating conditions**.**

Table 6. Comparative Overview of Recent Methods for Rolling Bearing Fault Diagnosis

Study (Year)	**Method**	**Datasets**	**Fault Classes**	**Reported Performance**	**Key Highlights**
Hatipoğlu et al. (2024)	Multi-domain feature extraction + LSTM-Attention + LASSO feature selection	HUST & CWRU	5 classes	F1 ≈ 99.90% (HUST), 98.74% (CWRU)	Compact feature set; effective attention-based classification
Liu et al. (2024)	Multi-scale Quaternion CNN + BiGRU + Cross Self-Attention fusion	CWRU, MFPT, Ottawa	Multi-class faults	CWRU: ~99.90%, MFPT: 100%, Ottawa: ~99.21%	High accuracy with quaternion convolution and attention fusion
Guan & Wang (2025)	Improved high-dimensional Bayesian Optimization–based CNN	CWRU (wind turbine bearings)	Healthy and faulty states	≈ 99% diagnostic accuracy	Automated Bayesian optimization of CNN hyperparameters; improved robustness and generalization
Soares et al. (2025)	ANN optimized using PSO with Bayesian hyperparameter tuning	Custom experimental dataset (vibration & current)	Healthy and faulty states	Up to 97.5% accuracy	Hybrid PSO–Bayesian optimization; effective under varying torque and short signal durations
Wu et al. (2025)	FFT → SE-TCN (attention-based CNN) + PSO-optimized SVM	CWRU & lab data	Multi-class fault types	Accuracy > 99%	Strong performance in small-sample conditions using attention + optimized SVM
Kalay (2025)	1D CNN-LSTM (uncertainty-aware)	CWRU-like benchmark	Up to 6 classes	~99.75% accuracy	Incorporates uncertainty modeling for robustness
Singh et al. (2025)	Graph model + logistic regression	CWRU	5 classes	99.8% accuracy	Interpretable and generalizable; robust across data conditions
This Study	ANN + PSO using frequency-domain features selected from two feature types and three optimization methods	Custom dataset (5 states: healthy, inner ring, outer ring, ball, combined)	5 classes	99.8% accuracy	Optimized compact ANN architecture using FFT-based features.

NOMENCLATURE

f_i	the frequency at index i
$X(f_i)$	the Fourier Transform of the signal at frequency f_i
$\lvert X(f_i)\rvert$	the magnitude of $X(f_i)$
N	the total number of frequency bins
μ_{FFT}	the mean of the frequency spectrum
σ_{FFT}	the standard deviation of the frequency spectrum
μ_f	the mean frequency
P_{0-100}	band power in [0,100] Hz
$P_{100-300}$	band power in [100,300] Hz

REFERENCES

Aiordăchioaie, D. (2024). *A comparative analysis of fault detection and process diagnosis methods based on a signal processing paradigm*. Discover Applied Sciences, 7(1). https://doi.org/10.1007/s42452-024-06390-3

An, Y., Zhang, K., Liu, Q., Chai, Y., & Huang, X. (2022b). *Rolling bearing fault diagnosis method base on periodic sparse attention and LSTM*. IEEE Sensors Journal, 22(12), 12044–12053. https://doi.org/10.1109/jsen.2022.3173446

Chen, Q., Zhang, F., Wang, Y., Yu, Q., Lang, G., & Zeng, L. (2025). *Bearing fault diagnosis based on efficient cross space multiscale CNN transformer parallelism*. Scientific Reports, 15(1), 12344. https://doi.org/10.1038/s41598-025-95895-x

Chen, Z., Zhu, L., Lu, H., Chen, S., Zhu, F., Liu, S., Han, Y., & Xiong, G. (2024). *Research on bearing fault diagnosis based on improved genetic algorithm and BP neural network.* Scientific Reports, 14(1), 15527. https://doi.org/10.1038/s41598-024-66318-0

Chi, F., Yang, X., Shao, S., & Zhang, Q. (2022). *Bearing fault diagnosis for Time-Varying system using Vibration–Speed fusion network based on Self-Attention and sparse feature extraction.* Machines, 10(10), 948. https://doi.org/10.3390/machines10100948

Ding, Y., Jia, M., Miao, Q., & Cao, Y. (2021). *A novel time–frequency Transformer based on self–attention mechanism and its application in fault diagnosis of rolling bearings*. Mechanical Systems and Signal Processing, 168, 108616. https://doi.org/10.1016/j.ymssp.2021.108616

El-Hassani, F. Z., Amri, M., Joudar, N., & Haddouch, K. (2024). *A new optimization model for MLP Hyperparameter tuning: Modeling and resolution by Real-Coded Genetic Algorithm.* Neural Processing Letters, 56(2). https://doi.org/10.1007/s11063-024-11578-0

Fu, G., Wei, Q., & Yang, Y. (2024b). *Bearing fault diagnosis with parallel CNN and LSTM*. Mathematical Biosciences & Engineering, 21(2), 2385–2406. https://doi.org/10.3934/mbe.2024105

Gao, S., Shi, S., & Zhang, Y. (2022b). *Rolling bearing compound fault diagnosis based on Parameter Optimization MCKD and convolutional neural network*. IEEE Transactions on Instrumentation and Measurement, 71, 1–8. https://doi.org/10.1109/TIM.2022.3158379

Guan, J., & Wang, Y. (2025). *Fault diagnosis strategy for wind turbine bearings based on improved high-dimensional Bayesian optimized CNN.* Journal of Mechanical Science and Technology, 39(11), 6655–6667. https://doi.org/10.1007/s12206-025-1012-9

Hatipoğlu, A., Süpürtülü, M., & Yılmaz, E. (2024). *Enhanced Fault Classification in Bearings: A Multi-Domain Feature Extraction Approach with LSTM-Attention and LASSO.* Arabian Journal for Science and Engineering. https://doi.org/10.1007/s13369-024-09842-5

Huang, D., Su, X., Yang, J., Du, S., Wang, D., & Ran, Q. (2025). *An improved dual-channel CNN-BILSTM fusion attention model for fault diagnosis of aero-engine bearings.* Measurement, 117761. https://doi.org/10.1016/j.measurement.2025.117761

Jain, P., & Bhosle, S. (2022). *Analysis of vibration signals caused by ball bearing defects using time-domain statistical indicators.* International Journal of Advanced Technology and Engineering Exploration, 9(90). https://doi.org/10.19101/ijatee.2021.875416

Jiang, G., Li, D., Feng, K., Li, Y., Zheng, J., Ni, Q., & Li, H. (2023b). *Rolling bearing fault diagnosis based on convolutional capsule network.* Journal of Dynamics Monitoring and Diagnostics. https://doi.org/10.37965/jdmd.2023.260

Kalay, O. C. (2025). *An optimized 1-D CNN-LSTM approach for fault diagnosis of rolling bearings considering epistemic uncertainty.* Machines, 13(7), 612. https://doi.org/10.3390/machines13070612

Khoualdia, T., Chelli, Z., & Boukhari, S. (2025b). *Multi-response optimization of diagnosis vibration parameters of broken rotor bars in induction motors using grey relational analysis in the Taguchi method.* Journal of Vibration Engineering & Technologies, 13(6). https://doi.org/10.1007/s42417-025-01949-9

Lee, J., Wu, F., Zhao, W., Ghaffari, M., Liao, L., & Siegel, D. (2013). *Prognostics and health management design for rotary machinery systems—Reviews, methodology and applications.* Mechanical Systems and Signal Processing, 42(1–2), 314–334. https://doi.org/10.1016/j.ymssp.2013.06.004

Liao, J., He, C., Li, J., Sun, J., Zhang, S., & Zhang, X. (2024a, April 11). *Classifier-guided neural blind deconvolution: a physics-informed denoising module for bearing fault diagnosis under heavy noise*. arXiv.org. https://arxiv.org/abs/2404.15341

Liu, H., Zhang, F., Tan, Y., Huang, L., Li, Y., Huang, G., Luo, S., & Zeng, A. (2024). *Multi-scale quaternion CNN and BiGRU with cross self-attention feature fusion for fault diagnosis of bearing.* Measurement Science and Technology, 35(8), 086138. https://doi.org/10.1088/1361-6501/ad4c8e

Lou, X., & Loparo, K. A. (2003). *Bearing fault diagnosis based on wavelet transform and fuzzy inference.* Mechanical Systems and Signal Processing, 18(5), 1077–1095. https://doi.org/10.1016/s0888-3270(03)00077-3

Onorato, G. (2024). *Bayesian optimization for hyperparameters tuning in neural networks.* arXiv (Cornell University). https://doi.org/10.48550/arxiv.2410.21886

Pan, H., Jiao, W., Yan, T., Rehman, A. U., Wan, A., & Yang, S. (2023b). *Combining kernel principal component* *analysis and spatial group-wise enhance convolutional neural network for fault recognition of rolling element bearings.* Measurement Science and Technology, 34(12), 125003. https://doi.org/10.1088/1361-6501/ace790

Qian, B., Cai, Y., Ran, Y., & Sun, W. (2024). *Nonlinear mechanical response analysis and convolutional neural network enabled diagnosis of single-span rotor bearing system.* Scientific Reports, 14(1). https://doi.org/10.1038/s41598-024-61180-6

Qiu, B., Liu, S., Li, W., Wang, X. V., & Wang, L. (2025). *Dynamic adaptive fault diagnosis using multi-channel image fusion and deep learning in channel failure occasions on rolling bearings.* Advanced Engineering Informatics, 68, 103631. https://doi.org/10.1016/j.aei.2025.103631

Raj, K. K., Kumar, S., & Kumar, R. R. (2024). *Systematic Review of Bearing Component Failure: Strategies for Diagnosis and Prognosis in Rotating Machinery.* Arabian Journal for Science and Engineering. https://doi.org/10.1007/s13369-024-09866-x

Rajabioun, R., Afshar, M., Atan, Ö., Mete, M., & Akin, B. (2023b). *Classification of Distributed Bearing Faults using a Novel Sensory Board and Deep Learning Networks with Hybrid Inputs.* IEEE Transactions on Energy Conversion, 39(2), 963–973. https://doi.org/10.1109/tec.2023.3338447

Sharma, A., Jigyasu, R., Mathew, L., & Chatterji, S. (2018). *Bearing fault diagnosis using frequency domain features and artificial neural networks.* In Smart innovation, systems and technologies (pp. 539–547). https://doi.org/10.1007/978-981-13-1747-7_52

Siddique, M. F., Saleem, F., Umar, M., Kim, C. H., & Kim, J. (2025c*). A hybrid deep learning approach for bearing fault diagnosis using continuous Wavelet transform and Attention-Enhanced spatiotemporal feature extraction.* Sensors, 25(9), 2712. https://doi.org/10.3390/s25092712

Singh, M. T. (2025, April 29). *Graph-Based fault diagnosis for rotating machinery: adaptive segmentation and structural feature integration.* arXiv.org. https://arxiv.org/abs/2504.20756

Soares, R. C., Silva, J. C., De Lucena, J. A., Filho, A. C. L., De Souza Ramos, J. G. G., & Brito, A. V. (2025). *Integration of Bayesian optimization into hyperparameter tuning of the particle swarm optimization algorithm to enhance neural networks in bearing failure classification.* Measurement, 242, 115829. https://doi.org/10.1016/j.measurement.2024.115829

Tama, B. A., Vania, M., Lee, S., & Lim, S. (2022). *Recent advances in the application of deep learning for fault diagnosis of rotating machinery using vibration signals.* Artificial Intelligence Review, 56(5), 4667–4709. https://doi.org/10.1007/s10462-022-10293-3

Tang, J., Wu, J., Hu, B., & Qing, J. (2022b). *Towards a Fault Diagnosis Method for Rolling Bearings with Time-Frequency Region-Based Convolutional Neural Network.* Machines, 10(12), 1145. https://doi.org/10.3390/machines10121145

Wu, J., Chen, X. Y., Zhang, H., Xiong, L. D., Lei, H., & Deng, S. H. (2019). *Hyperparameter optimization for machine learning models based on Bayesian optimization.* Journal of Electronic Science and Technology, 17(1), 26–40. https://doi.org/10.11989/jest.1674-862x.80904120

Wu, Y., Dai, J., Yang, X., Shao, F., Gong, J., Zhang, P., & Liu, S. (2025). *The fault diagnosis of rolling bearings based on FFT-SE-TCN-SVM.* Actuators, 14(3), 152. https://doi.org/10.3390/act14030152

Xiao, Z., Xu, Y., & Cui, J. (2025b). VFQB: *A novel deep learning model for rolling bearing fault diagnosis.* Sensors, 25(9), 2678. https://doi.org/10.3390/s25092678

Yang, P., Zhang, B., & Zhao, J. (2025b). *Efficient fault diagnosis in rolling bearings lightweight hybrid model.* Scientific Reports, 15(1). https://doi.org/10.1038/s41598-025-96285-z

Yu, Z., Zhang, L., & Kim, J. (2023b). *The performance analysis of PSO-RESNET for the fault diagnosis of vibration signals based on the pipeline robot.* Sensors, 23(9), 4289. https://doi.org/10.3390/s23094289

Zhang, B., Wang, W., & He, Y. (2025b). *A hybrid approach combining deep learning and signal processing for bearing fault diagnosis under imbalanced samples and multiple operating conditions.* Scientific Reports, 15(1). https://doi.org/10.1038/s41598-025-98138-1