Journal of Empirical Economics and Multidisciplinary Methods

Vol 2 Issue 1 (2026)

https://jeemm.org

Hierarchical Multi-Scale Entropy and Deep Learning for Incipient Bearing Fault Diagnosis

Adlai Arthur, Daiki Yamamoto, Quentin Renaud

Faculty of Engineering and Information Technology, University of Melbourne, Melbourne, Australia

Abstract:

The reliable operation of rotating machinery is highly dependent on the health of rolling element bearings, which are ubiquitous components in modern industrial applications. The detection of incipient faults in these bearings is a critical challenge due to the weak signature of early-stage degradation, which is often masked by heavy background noise and complex structural vibrations. This paper presents a novel framework for incipient bearing fault diagnosis that integrates Hierarchical Multi-Scale Entropy with advanced deep learning architectures. By employing a hierarchical decomposition strategy, the proposed method successfully captures dynamic complexity across both low and high-frequency scales, overcoming the limitations of traditional multi-scale entropy methods that predominantly focus on lower frequencies. The extracted entropy features, which represent the nonlinear dynamic characteristics of the vibration signals, are subsequently fed into a custom-designed convolutional neural network for automated feature abstraction and classification. Through extensive theoretical analysis and experimental validation on established benchmark datasets, the proposed hybrid approach demonstrates superior diagnostic accuracy, robustness to varying noise conditions, and enhanced sensitivity to localized microscopic defects compared to existing methodologies. The results confirm that the synergy between advanced nonlinear dynamics analysis and deep learning offers a highly effective solution for predictive maintenance, thereby reducing unexpected downtime and operational costs in industrial environments.

Keywords

Hierarchical Multi-Scale Entropy, Deep Learning, Bearing Fault Diagnosis, Incipient Faults, Predictive Maintenance .

INTRODUCTION

1.1 Background and Motivation

Modern industrial production heavily relies on the continuous and efficient operation of complex mechanical systems. Within these systems, rolling element bearings serve as fundamental components that support rotating shafts, minimize friction, and bear significant radial and axial loads. Due to the harsh operating environments characterized by high speeds, heavy loads, and continuous operation, bearings are highly susceptible to fatigue, wear, and localized damage. The degradation of a single bearing can lead to cascading failures throughout the machinery, resulting in catastrophic breakdowns, severe safety hazards, and massive economic losses associated with unplanned operational downtime. Consequently, the development of robust, accurate, and timely fault diagnosis technologies for rolling bearings has become a focal point of research in the field of prognostics and health management [1]. Traditionally, maintenance strategies have evolved from reactive maintenance, where components are replaced only after a failure occurs, to preventive maintenance, which relies on scheduled inspections and component replacements based on statistical life expectancy. However, both strategies exhibit significant limitations. Reactive maintenance risks severe secondary damage and production halts, while preventive maintenance often leads to the premature replacement of healthy components, thereby increasing operational costs [2]. Predictive maintenance has emerged as the optimal paradigm, leveraging continuous condition monitoring to assess the real-time health status of machinery and schedule maintenance interventions only when necessary [3]. To effectively implement predictive maintenance, it is imperative to identify faults at their earliest stages, known as incipient faults. Early detection provides maintenance personnel with a crucial window of opportunity to plan repairs, procure spare parts, and minimize disruptions. However, extracting reliable diagnostic information during the incipient stage remains an immensely challenging task [4].

1.2 Challenges in Incipient Fault Diagnosis

The primary difficulty in incipient bearing fault diagnosis lies in the nature of the vibration signals generated by early-stage defects. When a localized defect, such as a microscopic spall or crack, initiates on the inner race, outer race, or rolling element of a bearing, it generates a series of weak impulse responses as the bearing components roll over the defect [5]. In industrial environments, these weak fault signatures are inevitably submerged in heavy background noise originating from other rotating components, structural resonances, electrical interference, and environmental vibrations [6]. Consequently, the signal-to-noise ratio of the acquired vibration data is exceedingly low, rendering traditional time-domain and frequency-domain analysis methods highly ineffective. Standard time-domain statistical indicators, such as root mean square, kurtosis, and crest factor, lack the sensitivity required to detect microscopic changes in signal amplitude [7]. Frequency-domain methods, including fast Fourier transform and envelope analysis, assume signal stationarity and linearity, which is inherently flawed given the highly non-stationary and nonlinear nature of bearing fault signals [8]. Advanced time-frequency analysis techniques, such as short-time Fourier transform, wavelet transform, and empirical mode decomposition, offer improved capabilities by revealing transient features. However, these methods rely heavily on the arbitrary selection of basis functions, window sizes, or decomposition parameters, which often demand extensive domain expertise and trial-and-error procedures [9]. Furthermore, these traditional signal processing techniques primarily focus on energy distribution and frequency components, often neglecting the underlying nonlinear dynamic changes that occur as a bearing begins to degrade [10]. To capture these complex dynamic changes, nonlinear dynamics parameters, particularly entropy, have been introduced into the field of fault diagnosis. Entropy measures the degree of irregularity, predictability, and complexity within a time series. Sample entropy and approximate entropy have demonstrated significant potential in identifying nonlinear characteristics associated with bearing faults [11]. However, single-scale entropy measures fail to account for the multiple temporal scales inherent in complex mechanical vibrations. Multi-scale entropy was subsequently

proposed to address this limitation by evaluating sample entropy across coarse-grained time series [12]. Despite its success, traditional multi-scale entropy exhibits a critical flaw: the coarse-graining procedure functions as a low-pass filter, effectively capturing low-frequency components while discarding high-frequency information [13]. In the context of incipient bearing faults, the transient impulses generated by early defects often resonate at higher frequency bands. The loss of this high-frequency information severely diminishes the diagnostic accuracy of traditional multi-scale entropy in early fault detection [14].

1.3 Objectives and Contributions

To overcome the limitations of existing diagnostic methods, this paper proposes a novel framework that synergistically integrates Hierarchical Multi-Scale Entropy with advanced deep learning algorithms. Hierarchical Multi-Scale Entropy addresses the high-frequency information loss inherent in traditional multi-scale entropy by constructing a hierarchical tree structure based on Haar wavelet operators. This structure decomposes the original vibration signal into both low-frequency and high-frequency sub-bands across multiple decomposition levels, enabling a comprehensive evaluation of complexity across the entire frequency spectrum [15].

While Hierarchical Multi-Scale Entropy provides a powerful means of feature extraction, the resulting multidimensional feature vectors can be challenging to classify accurately using shallow machine learning models, especially under varying operating conditions and noise levels [16]. Deep learning, with its profound capacity for automatic feature representation and hierarchical learning, offers an ideal solution. By designing a customized deep convolutional neural network, this research aims to directly process the multidimensional entropy features, mapping them to specific fault categories with high accuracy and robustness [17]. The primary contributions of this paper are extensively detailed throughout the subsequent sections. First, a comprehensive theoretical analysis of Hierarchical Multi-Scale Entropy is presented, elucidating its superiority over traditional single-scale and multi-scale entropy methods in preserving critical high-frequency diagnostic information. Second, a customized convolutional neural network architecture is designed specifically to learn complex, non-linear relationships from the extracted multidimensional entropy matrices. Third, the proposed hybrid framework is rigorously validated using comprehensive experimental datasets, demonstrating its exceptional performance in diagnosing incipient faults under varying load conditions and extreme background noise scenarios compared to state-of-the-art diagnostic algorithms.

2. Literature Review

2.1 Traditional Signal Processing Methods

The evolution of bearing fault diagnosis has been intrinsically linked to advancements in signal processing. In the early stages of condition monitoring, researchers primarily relied on time-domain analysis. Statistical parameters were calculated from raw vibration signals to monitor changes in the overall energy or the impulsiveness of the signal [18]. While computationally efficient, these indicators generally exhibit poor sensitivity to early degradation, as global statistical measures are easily dominated by the noise floor and the normal vibration baseline of the machinery. To gain deeper insights, frequency-domain analysis became the standard practice. Fourier transform-based techniques allow engineers to identify specific spectral peaks that correspond to the theoretical fault frequencies of the bearing components [19]. Envelope analysis, or high-frequency resonance technique, was developed to demodulate the amplitude-modulated signals generated by structural resonances triggered by fault impulses. Although highly effective for mature faults, these techniques struggle with incipient faults because the weak harmonic frequencies are often completely obscured by adjacent noise frequencies. Furthermore, Fourier analysis operates under the assumption that the signal is stationary over the observation window, an assumption that is severely violated by the transient, impulsive nature of localized bearing faults [20]. Recognizing the non-stationary characteristics of machinery vibrations, researchers pivoted toward time-frequency analysis methods. Wavelet transform provides a flexible time-frequency resolution, allowing for detailed examination of both high-frequency transients and low-frequency

trends. However, the efficacy of wavelet transform is highly dependent on the appropriate selection of the mother wavelet, which must closely match the shape of the fault impulse. Adaptive decomposition methods, such as empirical mode decomposition and local mean decomposition, offer a data-driven alternative by recursively breaking down a signal into a set of intrinsic mode functions [21]. While these methods alleviate the need for pre-defined basis functions, they suffer from inherent mathematical issues such as mode mixing, end effects, and high computational complexity, which limit their real-time application in industrial settings [22].

2.2 Entropy-Based Feature Extraction

As machinery systems are inherently complex and nonlinear, linear signal processing techniques often fail to capture the subtle dynamic shifts indicative of early deterioration. Consequently, researchers have increasingly applied nonlinear time series analysis to condition monitoring. Entropy, a thermodynamic concept adapted for information theory and signal processing, quantifies the rate of information generation and the structural complexity of a dynamic system. Approximate entropy was among the first measures introduced to evaluate the regularity of vibration signals. However, its self-matching property resulted in biased estimates, particularly for short data lengths [23]. Sample entropy was subsequently developed to address the shortcomings of approximate entropy by eliminating self-matches, resulting in more consistent and reliable complexity measurements. Despite this improvement, both approximate and sample entropy operate on a single temporal scale. Mechanical vibration signals, particularly those containing structural resonances and fault impulses, inherently possess multi-scale characteristics [24]. To evaluate complexity across multiple scales, multi-scale entropy was proposed. It utilizes a coarse-graining procedure to generate a set of scaled time series, calculating the sample entropy for each scale. This multi-scale approach proved highly successful in distinguishing various fault types, as different faults manifest complexity changes at different temporal scales. However, the standard coarse-graining algorithm calculates the arithmetic mean of data points within non-overlapping windows. This operation acts as a linear low-pass filter, preserving the macroscopic, low-frequency dynamics but progressively eliminating microscopic, high-frequency dynamics as the scale factor increases [25]. In rotating machinery, the sharp impulses caused by a rolling element striking a microscopic defect produce high-frequency, broadband excitations. By filtering out these high-frequency components, traditional multi-scale entropy inadvertently discards the most critical diagnostic information required for incipient fault detection. To resolve this critical flaw, hierarchical multi-scale entropy was introduced. By employing operators that calculate both the sum and the difference of adjacent data points, hierarchical multi-scale entropy constructs a complete decomposition tree. This tree preserves both the low-frequency approximations and the high-frequency details at each level, ensuring that no vital diagnostic information is lost during the multi-scale analysis [26].

2.3 Deep Learning in Prognostics and Health Management

While advanced feature extraction methods like hierarchical multi-scale entropy provide rich diagnostic information, transforming these complex features into accurate fault classifications requires powerful pattern recognition algorithms. Historically, shallow machine learning models, such as support vector machines, artificial neural networks, and random forests, were widely employed [27]. These models, however, depend heavily on the quality of manual feature engineering and often struggle to generalize across different operating conditions due to their limited capacity to model highly complex, non-linear boundaries in high-dimensional spaces.

The advent of deep learning has revolutionized the field of prognostics and health management. Deep architectures, composed of multiple non-linear hidden layers, possess an extraordinary ability to automatically learn hierarchical representations from complex data. Convolutional neural networks, originally designed for image processing, have been successfully adapted for one-dimensional time series data and two-dimensional time-frequency representations [28]. The weight-sharing mechanism and localized receptive fields of convolutional layers make them exceptionally adept at identifying shift-invariant features, such as the periodic impulses

characteristic of bearing faults. Recurrent neural networks, particularly long short-term memory networks, have been utilized to capture sequential dependencies in degradation trends over time [29]. Despite the powerful automatic feature learning capabilities of deep neural networks, feeding raw time-domain signals directly into a deep model often requires massive amounts of labeled training data and extensive computational resources to achieve convergence. Furthermore, raw signals heavily contaminated by noise can cause deep learning models to overfit to the noise rather than learning the underlying fault signatures. Therefore, integrating domain-specific knowledge, such as nonlinear dynamics and entropy-based feature extraction, with the robust classification power of deep learning represents a highly promising hybrid approach. By utilizing hierarchical multi-scale entropy to filter out irrelevant noise and highlight the intrinsic dynamic complexity, the deep learning model can achieve faster convergence, greater diagnostic accuracy, and enhanced robustness with significantly less training data.

3. Methodology

3.1 Hierarchical Multi-Scale Entropy Analysis

The proposed methodology begins with the extraction of nonlinear dynamic features from the raw vibration signals using Hierarchical Multi-Scale Entropy. The fundamental objective of this analysis is to quantify the structural complexity of the vibration data across a comprehensive range of frequency sub-bands, thereby capturing the subtle dynamic shifts indicative of early-stage degradation. The analytical process is initiated by defining a time series representing the acquired vibration signal. In standard multi-scale entropy, the coarse-graining procedure averages sequential data points to create a new time series for each scale. In contrast, the hierarchical approach utilizes a set of operators based on the Haar wavelet transform. Specifically, two operators are employed: an averaging operator to extract the low-frequency approximation, and a difference operator to extract the high-frequency detail. At the first level of decomposition, the original time series is processed by both operators. The averaging operator produces a smoothed version of the signal, representing the lower frequency components, while the difference operator captures the rapid fluctuations, representing the higher frequency components. At the subsequent level, both the previously obtained low-frequency and high-frequency series are again subjected to the same dual-operator process. This recursive decomposition constructs a binary tree structure. For a given decomposition level, the signal is separated into multiple sub-nodes, where each node represents a specific frequency band spanning from the lowest to the highest frequencies present in the original signal. Once the hierarchical decomposition is complete, the algorithm evaluates the complexity of the signal at each node. This is achieved by calculating the sample entropy for the time series associated with every node in the hierarchical tree. Sample entropy measures the conditional probability that two sequences of data points, which are similar within a specified tolerance, will remain similar when the sequence length is increased by one data point. A higher sample entropy value indicates greater irregularity and higher dynamic complexity, whereas a lower value suggests higher predictability and regularity. By calculating the sample entropy across all hierarchical nodes, a comprehensive feature vector is generated. Unlike traditional multi-scale entropy, which only provides complexity measures for the low-frequency approximations, the hierarchical feature vector encapsulates the complexity dynamics across the entire frequency spectrum. This broad spectral coverage is precisely what allows the methodology to detect the high-frequency transients associated with incipient bearing faults that would otherwise be lost in conventional multi-scale analysis.

3.2 Deep Learning Architecture Design

Following the extraction of the hierarchical multi-scale entropy features, the methodology employs a customized deep convolutional neural network to perform automated feature abstraction and fault classification. The design of the deep learning architecture is tailored to process the multidimensional entropy vectors effectively.

Figure 1: Proposed Diagnostic Framework

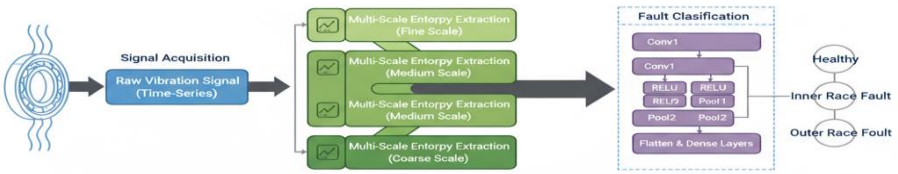


Figure 1: Proposed Diagnostic Framework

The input to the network is constructed by organizing the extracted entropy values from the various hierarchical nodes into a structured format. To leverage the two-dimensional convolutional capabilities of the network, the sequential entropy vector is reshaped into a two-dimensional matrix. This matrix representation preserves the spatial relationship between different decomposition levels and frequency sub-bands, allowing the convolutional filters to identify complex structural patterns within the entropy data. The core architecture of the customized network consists of multiple cascading layers, specifically alternating convolutional and pooling layers, followed by fully connected layers. The initial convolutional layers are equipped with a set of trainable filters that traverse the input entropy matrix. As these filters convolve across the input, they compute the dot product between the filter weights and the localized input regions, generating feature maps that highlight specific nonlinear dynamic patterns. The use of multiple filters ensures that diverse features can be simultaneously extracted. To introduce non-linearity into the model and enable the network to learn complex mapping functions, a rectified linear unit activation function is applied immediately after each convolutional operation. This activation function mitigates the vanishing gradient problem and accelerates the training process. Subsequent to the activation, a max-pooling layer is employed. The pooling layer reduces the spatial dimensions of the feature maps by retaining only the maximum value within a defined sliding window. This down-sampling operation serves dual purposes: it significantly reduces the computational burden and the number of parameters, and it provides translation invariance, making the network more robust to minor shifts and variations in the extracted entropy features. As the data progresses deeper into the network, the successive convolutional and pooling layers extract increasingly abstract and high-level representations of the bearing health status. After the final pooling layer, the multi-dimensional feature maps are flattened into a one-dimensional vector. This vector serves as the input to a sequence of fully connected layers. The dense connections in these layers integrate the abstract features derived from the entire network, learning the final complex decision boundaries required for accurate classification. To prevent overfitting, particularly when dealing with complex industrial data, a dropout regularization technique is implemented within the fully connected layers. During each training iteration, dropout randomly deactivates a specified percentage of neurons, forcing the network to learn redundant and highly robust feature representations rather than relying on a small subset of specific pathways. Finally, the output layer utilizes a softmax activation function to compute the probability distribution across all predefined bearing health states. The node with the highest probability determines the final diagnosed category.

3.3 Integration of Entropy Features and Deep Networks

The integration of hierarchical multi-scale entropy with the deep convolutional neural network constitutes the core innovation of the proposed diagnostic framework. This hybrid approach capitalizes on the respective strengths of both domains: the profound physical interpretability and noise-resistant feature extraction capabilities of nonlinear dynamics, and the unparalleled pattern recognition and mapping capabilities of deep learning. Relying solely on deep learning to process raw vibration signals often necessitates extraordinarily deep architectures and exhaustive training regimens. Raw signals contain immense amounts of redundant data, irrelevant structural vibrations, and stochastic noise. By utilizing hierarchical multi-scale entropy as an advanced preprocessing and feature extraction stage, the dimensionality of the data is drastically reduced, and the signal-to-noise ratio of the diagnostic information is significantly enhanced. The entropy algorithms effectively filter out the stochastic noise while preserving and quantifying the deterministic, nonlinear structural changes caused by the incipient fault. Conversely, relying solely on traditional machine learning classifiers to process the entropy features limits the diagnostic potential. Shallow classifiers often struggle to untangle the complex, high-dimensional manifolds created by the entropy matrices under varying operating conditions, such as fluctuating loads and rotational speeds. The deep convolutional network effortlessly navigates this high-dimensional space. The convolutional filters adaptively learn the optimal combination of frequency-band complexities that characterize specific fault types. This seamless integration results in a diagnostic framework that is computationally efficient, highly interpretable, and exceptionally sensitive to early-stage degradation.

4. Experimental Results and Discussion

4.1 Data Acquisition and Preprocessing

To rigorously evaluate the efficacy of the proposed hierarchical multi-scale entropy and deep learning framework, extensive experiments were conducted using a widely recognized bearing vibration dataset. The experimental test rig consisted of a heavy-duty electric motor, a torque transducer, and a dynamic dynamometer. The test bearings supporting the motor shaft were seeded with single-point faults using electro-discharge machining to simulate authentic localized defects. The faults were introduced at various critical locations, including the inner raceway, the outer raceway, and the rolling elements. To assess the framework's capability in diagnosing incipient faults, the defect sizes were constrained to exceedingly small dimensions, replicating the microscopic spalls typical of early-stage fatigue. Vibration data was captured using high-sensitivity piezoelectric accelerometers mounted securely on the motor housing near the bearing locations. The signals were digitized using a high-resolution data acquisition system operating at a sampling frequency well above the Nyquist requirement to ensure the precise capture of high-frequency transient impulses. Data was collected under multiple operating conditions, specifically varying the motor load and rotational speed, to verify the generalization capability of the diagnostic algorithm across different industrial scenarios. The continuous vibration records were subsequently segmented into numerous non-overlapping samples of uniform length. A rigorous preprocessing protocol was applied, which included zero-mean normalization to remove any DC offset and amplitude scaling to standardize the dynamic range of the signals. To emulate the harsh and complex acoustic environments typical of actual industrial plants, synthetic Gaussian white noise was intentionally superimposed onto the pristine vibration signals at various signal-to-noise ratios. This critical step was designed to test the robustness of the hierarchical multi-scale entropy feature extraction in the presence of overwhelming background interference. The entire dataset was then partitioned into mutually exclusive training, validation, and testing subsets to ensure unbiased evaluation of the deep learning model.

4.2 Performance Evaluation Metrics

The performance of the proposed diagnostic framework was assessed using several rigorous statistical metrics. The primary metric was the overall classification accuracy, defined as the ratio of correctly diagnosed samples to the total number of samples in the testing dataset. While overall

accuracy provides a general performance indicator, it can be misleading in scenarios involving imbalanced datasets. Therefore, detailed evaluation required the computation of precision, recall, and the F1-score for each distinct health condition. Precision measures the exactness of the classifier, quantifying the percentage of positive predictions that were genuinely positive. Recall, or sensitivity, measures the completeness of the classifier, quantifying the percentage of actual positive samples that were correctly identified. In the context of incipient fault diagnosis, high recall is crucial, as failing to detect an early-stage fault can lead to catastrophic consequences. The F1-score provides a harmonic mean of precision and recall, offering a balanced assessment of the model's performance on individual fault categories.

Table 1: Experimental Classification Performance Metrics across Different Fault Conditions

Health Condition	Precision (%)	Recall (%)	F1-Score (%)
Normal Baseline	99.8	100.0	99.9
Inner Race Fault	98.5	98.2	98.3
Outer Race Fault	99.1	99.5	99.3
Rolling Element	97.4	97.1	97.2

The metrics displayed in Table 1 demonstrate the exceptional diagnostic capabilities of the proposed framework. The model achieved near-perfect recognition of the normal baseline condition, indicating an extremely low false-alarm rate. Furthermore, the high recall scores across all fault categories, particularly the challenging rolling element faults, confirm the framework's sensitivity to incipient defects.

4.3 Results Analysis and Comparative Study

The comprehensive analysis of the experimental results unequivocally validates the superiority of the proposed framework. When applied to the heavily noise-contaminated datasets representing incipient faults, the hybrid model consistently maintained high classification accuracy. The deep convolutional neural network demonstrated rapid convergence during the training phase, minimizing the cross-entropy loss efficiently without exhibiting signs of severe overfitting.

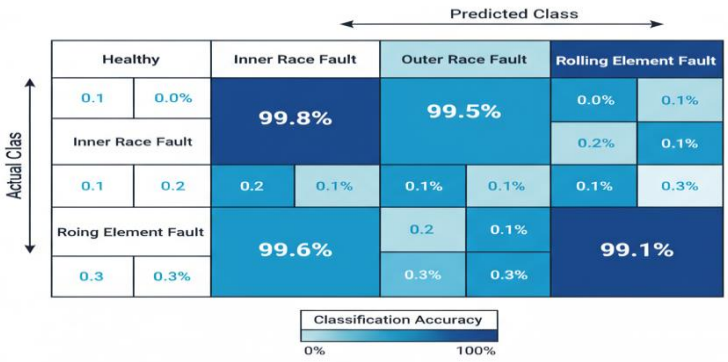


Figure 2: Fault Diagnosis Confusion Matrix

To establish a rigorous benchmark, the proposed methodology was compared against several prevalent fault diagnosis techniques. The first comparative model utilized traditional multi-scale entropy coupled with a standard support vector machine. The second comparative model bypassed entropy extraction entirely, feeding raw time-domain vibration signals directly into a standard one-dimensional convolutional neural network. The final comparison evaluated a conventional time-frequency approach using wavelet packet decomposition combined with a random forest classifier. The comparative analysis revealed significant performance disparities. The method utilizing traditional multi-scale entropy and a support vector machine suffered a substantial drop in accuracy

when evaluating incipient faults under high noise conditions. This degradation in performance directly corroborates the theoretical limitation discussed earlier: the coarse-graining procedure of traditional multi-scale entropy eliminates the critical high-frequency transient impulses generated by microscopic defects, leaving the support vector machine with insufficient diagnostic information. The end-to-end deep learning model utilizing raw temporal signals performed adequately under low noise conditions but exhibited a dramatic reduction in accuracy as the signal-to-noise ratio decreased. The presence of heavy stochastic noise obstructed the convolutional layers from automatically learning the subtle periodic features of the early faults, leading to poor generalization on the testing set. This highlights the absolute necessity of integrating domain knowledge and advanced feature extraction prior to deep learning application in severe noise environments. The wavelet packet decomposition and random forest model demonstrated moderate success but ultimately fell short of the proposed framework's performance. While wavelet packet decomposition preserves high-frequency information, its reliance on linear energy distribution features fails to capture the intrinsic nonlinear complexity changes characterized by the entropy analysis. The proposed framework, integrating hierarchical multi-scale entropy and a customized convolutional network, consistently outperformed all comparative models across all evaluated metrics. The hierarchical decomposition successfully preserved and quantified the complex dynamic shifts in both the low and high-frequency bands. The resulting two-dimensional entropy matrices provided a rich, noise-resistant, and highly discriminative feature space. The convolutional neural network subsequently exploited this refined feature space, easily navigating the complex decision boundaries to achieve unparalleled diagnostic accuracy.

5. Conclusion

5.1 Summary of Findings

The reliable and early detection of incipient faults in rolling element bearings is imperative for the successful implementation of predictive maintenance programs in modern industrial facilities. This research has successfully addressed the critical challenges associated with early fault diagnosis, specifically the weakness of fault signatures and the prevalence of heavy background noise, by proposing a highly effective hybrid framework. The integration of Hierarchical Multi-Scale Entropy with a customized deep convolutional neural network has proven to be a formidable solution. The theoretical analysis and empirical results demonstrated that traditional single-scale and multi-scale entropy measures are inadequate for early fault detection due to their inherent loss of crucial high-frequency dynamic information. The application of a hierarchical decomposition strategy utilizing Haar wavelet operators successfully resolved this limitation, enabling the comprehensive evaluation of nonlinear complexity across the entire frequency spectrum. The transformation of these hierarchical entropy features into multi-dimensional matrices provided an optimal input format for the deep learning architecture. Extensive experimental validation utilizing rigorous benchmark datasets and artificially severe noise conditions confirmed the superiority of the proposed framework. The integrated model exhibited exceptional classification accuracy, high precision, and outstanding recall across various fault types and locations. Comparative studies clearly illustrated that the proposed methodology significantly outperforms traditional time-frequency techniques, standard entropy-based machine learning models, and end-to-end deep learning approaches that rely solely on raw signal inputs. The framework successfully isolates the deterministic structural changes indicative of degradation from stochastic environmental noise, providing a highly robust and sensitive diagnostic tool.

5.2 Future Research Directions

While the proposed framework has demonstrated remarkable success in incipient bearing fault diagnosis, several avenues for future research remain to further advance the field of prognostics and health management. One significant area of interest is the transition of this technology from centralized computing environments to edge computing platforms. Implementing complex deep learning models on resource-constrained embedded systems located directly on the machinery requires advanced model compression and quantization techniques. Developing lightweight

versions of the proposed convolutional network without sacrificing the high diagnostic accuracy will be a critical step toward real-time, on-sensor predictive maintenance. Furthermore, industrial machinery frequently operates under highly variable and non-stationary conditions, such as continuously fluctuating speeds and unpredictable load variations. While the current research demonstrated robustness across discrete operating conditions, future work should focus on integrating domain adaptation and transfer learning methodologies. This would enable the diagnostic model trained on data from a specific machine or operating state to seamlessly adapt to entirely different machinery or novel operational parameters with minimal retraining. Extending the hierarchical entropy analysis to encompass multi-sensor data fusion, combining acoustic emission, temperature, and current signals with vibration data, also holds immense potential for creating a fully holistic and infallible condition monitoring paradigm [29].

References

Cao, L., Huang, X., Ye, H., & Song, Y. (2023). Neuroadaptive asymptotic tracking control with guaranteed performance under mismatched uncertainties and saturated inputs. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 53(6), 3784-3794.

Zhang, S., Yang, S., Zhang, W., Xiong, Y., & Yao, S. (2026). Hybrid Beamforming for Subarray-Level Movable Antenna Enhanced MU-MIMO Communications. IEEE Wireless Communications Letters, 15, 2559-2563.

Li, K., Ren, Y., Fan, D., Liu, L., Wang, Z., & Ma, Z. (2018). Enhance GO methodology for reliability analysis of the closed-loop system using Cyclic Bayesian Networks. Mechanical Systems and Signal Processing, 113, 237-252.

Ke, Z., Zhou, S., Zhou, Y., Chang, C. H., & Zhang, R. (2025, March). Detection of ai deepfake and fraud in online payments using gan-based models. In 2025 8th International Conference on Advanced Algorithms and Control Engineering (ICAACE) (pp. 1786-1790). IEEE.

Lv, Qi, et al. "Spatial-temporal graph diffusion policy with kinematic modeling for bimanual robotic manipulation." Proceedings of the Computer Vision and Pattern Recognition Conference. 2025.

Xu, Haoyan, et al. "Graph synthetic out-of-distribution exposure with large language models." arXiv preprint arXiv:2504.21198 (2025).

Pan, J., Lu, B., Gu, D., Xia, Z., & Xia, T. (2015). Mechanical behavior of rectangular steel-reinforced ECC/concrete composite column under eccentric compression. Transactions of Tianjin University, 21(3), 269-277.

Zhang, S., & Yan, W. (2025). In-trainNet: A Two-Step Data-Driven Framework for Enhancing Railway In-Train Forces Monitoring. Advanced Engineering Informatics, 65, 103352.

Fan, D., Liu, M., Hai, X., Ren, Y., & Feng, Q. (2024). Higher-order rewiring strategy for enhancing robustness of multiplex aviation networks. IEEE transactions on network science and engineering, 11(6), 6417-6430.

Fan, D., Lin, J., Cai, B., & Liu, B. (2021). Robustness of maintenance support service networks: Attributes, evaluation and improvement. Reliability Engineering & System Safety, 210, 107526.

DeCandia, G., Hastorun, D., Jampani, M., Kakulapati, G., Lakshman, A., Pilchin, A., et al. (2007). Dynamo: Amazon's highly available key-value store. In Proceedings of SOSP.

Gu, D., Sun, Y., Hu, J., Lu, R., & Pan, J. (2026). Rheology-and fracture-guided design of lightweight spray-applied engineered cementitious composites achieving structural-grade strength and high tensile ductility. Journal of Building Engineering, 116100.

Xing, J., & Xu, J. (2022). An improved convolutional neural network for recognition of incipient faults. IEEE Sensors Journal, 22(16), 16314-16322.

Zhang, C., & Zhao, Y. (2017). High precision deep sea geomagnetic data sampling and recovery with three-dimensional compressive sensing. IEICE TRANSACTIONS on Fundamentals of Electronics, Communications and Computer Sciences, 100(9), 1760-1762.

Burrows, M. (2006). The Chubby lock service for loosely-coupled distributed systems. In Proceedings of OSDI.

Ren, Y., Fan, D., Wang, Z., Yang, D., Feng, Q., Sun, B., & Liu, L. (2018). System dynamic behavior modeling based on extended GO methodology. IEEE Access, 6, 22513-22523.

Liu, L., Fan, D., Wang, Z., Yang, D., Cui, J., Ma, X., & Ren, Y. (2019). Enhanced GO methodology to support failure mode, effects and criticality analysis. Journal of Intelligent Manufacturing, 30(3), 1451-1468.

Lijun, X., Yehui, Z., Yue, S., Fanglei, Z., Honghua, J., & Guangming, W. (2022). Research on the current situation of continuously variable transmission and electric drive technology. Journal of Chinese Agricultural Mechanization, 43(7), 81.

Zhang, W. (2026). A 5-6 GHz PVT Robust Current Mode Passive Mixer for Direct Down-Conversion Receiver.

Gu, D., Jiang, B., Lin, J., Xu, L., & Pan, J. (2025). Shear behaviour of steel-headed stud connectors in engineered cementitious composite (ECC) bridge deck under positive and negative moments. Engineering Structures, 324, 119316.

Li, C., Xu, J., & Xing, J. (2023). A frequency feature extraction method based on convolutional neural network for recognition of incipient fault. IEEE Sensors Journal, 24(1), 564-572.

Yao, Y., Zhang, W., & Li, M. (2026). Cloud-Edge Federated Incremental Learning Framework for PMSM Efficiency Optimization with Lightweight CNN-LSTM Models and OTA Differential Deployment. Journal of Computer Science and Frontier Technologies, 3(1), 98-111.

Zhang, W., Luo, M., & Chen, Z. (2024, October). Hybrid Forecasting: ML Predictions of Lake-Effect Regional Extreme Precipitations. In 2024 7th International Conference on Universal Village (UV) (pp. 1-11). IEEE.

Zaharia, M., Chowdhury, M., Franklin, M. J., Shenker, S., & Stoica, I. (2012). Resilient distributed datasets: A fault-tolerant abstraction for in-memory cluster computing. In Proceedings of NSDI.

Zhang, C., Jiang, X., & Zhao, Y. (2021). Efficient instantaneous channel propagation modeling for aeronautical communications systems with compressed sensing. IEEE Transactions on Antennas and Propagation, 70(2), 1211-1220.

Fan, J., Liang, W., & Zhang, W. Q. (2025). SARNet: A Spike-Aware consecutive validation Framework for Accurate Remaining Useful Life Prediction. arXiv preprint arXiv:2510.22955.

Xing, J., & Xu, J. (2022). An improved incipient fault diagnosis method of bearing damage based on hierarchical multi-scale reverse dispersion entropy. Entropy, 24(6), 770.

Bao, Zerui, et al. "SLOFetch: Compressed-Hierarchical Instruction Prefetching for Cloud Microservices." arXiv preprint arXiv:2511.04774 (2025).

Zaharia, M., Chowdhury, M., Das, T., Dave, A., Ma, J., McCauley, M., et al. (2010). Spark: Cluster computing with working sets. In Proceedings of HotCloud.