International Journal of Advance Industrial Engineering
©2020 **INPRESSCO**®, All Rights Reserved
E-ISSN 2320 –5539
Available at http://inpressco.com/category/ijaie/

Research Article

CNN-LSTM Hybrid Deep Learning Architecture for Predictive Fault Diagnosis in Rotating Machinery: Multi-Sensor Vibration Analysis and Real-Time Deployment in Industrial Environments

Tanvir H. Ahmed[1]*, Sofia M. Rodrigues[2]

[1]Research Scholar, School of Electrical Engineering and Computer Science, University of Ottawa, 75 Laurier Avenue East, Ottawa, ON K1N 6N5, Canada
[2]Research Scholar, Department of Mechanical Engineering, University of Porto, Rua Dr. Roberto Frias, 4200-465 Porto, Portugal

Received 01 Feb 2020, Accepted 14 Mar 2020, Available online 15 Mar 2020, **Vol.8, No.1 (Mar 2020)**

Abstract

Rotating machinery faults including bearing defects, gear damage, shaft imbalance, and misalignment are the primary cause of unplanned downtime in manufacturing, mining, and power generation facilities, collectively responsible for equipment downtime costs estimated at USD 50 billion annually in industrialised economies. Predictive maintenance systems based on vibration signal analysis and machine learning have demonstrated strong fault detection capability in controlled laboratory settings, but real-world industrial deployment presents additional challenges of multi-sensor data fusion, variable speed operation, noisy measurement environments, and the requirement for real-time inference on edge computing hardware. This study presents a CNN-LSTM hybrid deep learning architecture, designated FaultNet-5, designed for multi-class fault detection and severity classification in rotating machinery under variable speed conditions using simultaneous tri-axial accelerometer measurements. FaultNet-5 achieved 97.8% overall classification accuracy and weighted F1 score of 0.973 on a held-out test set comprising 12,400 vibration records across five fault classes, outperforming Support Vector Machine (91.2%), Random Forest (94.8%), and gradient boosting (95.4%) baseline classifiers. Deployed on NVIDIA Jetson Xavier NX edge hardware, FaultNet-5 achieves 4.2 ms per-inference latency, enabling real-time fault detection at data acquisition rates of 20 kHz. Industrial validation across three production facilities over nine months demonstrated 94.3% early fault detection rate with 8.6-day mean precursor detection lead time, enabling proactive maintenance interventions that reduced unplanned downtime by 73.4% and maintenance cost by 41.2%.

Keywords: *Predictive Maintenance, Fault Diagnosis, Convolutional Neural Network, LSTM, Vibration Analysis, Rotating Machinery, Deep Learning, Edge Computing*

1. Introduction

Industrial rotating machinery, encompassing electric motors, gearboxes, pumps, compressors, fans, and turbines, forms the mechanical backbone of manufacturing and process industry operations. The reliable, efficient operation of rotating equipment is a fundamental prerequisite for production continuity, and equipment failures represent one of the most significant sources of unplanned production downtime and associated financial loss. The American Society for Mechanical Engineers estimates that bearing failures alone account for approximately 40% of all electric motor failures, while gearbox failures are responsible for over 30% of wind turbine unplanned maintenance events.

Condition-based maintenance (CBM) and its data-driven extension, predictive maintenance (PdM), represent the most advanced tier of the equipment maintenance hierarchy, seeking to transition from reactive (fix-on-failure) and preventive (time-based) maintenance paradigms toward maintenance interventions that are triggered by objective evidence of developing equipment degradation. The fundamental enabler of CBM and PdM is the extraction of diagnostically informative signals from continuous or periodic equipment monitoring, most commonly through vibration analysis, thermal imaging, oil particle analysis, or acoustic emission monitoring.

Vibration analysis has been the dominant condition monitoring modality for rotating machinery since the 1970s, providing rich diagnostic information through frequency-domain signatures that characterise specific fault types. Bearing outer race faults generate characteristic defect frequencies at Ball Pass

*Corresponding author's ORCID ID: 0000-0000-0000-0000
DOI: https://doi.org/10.14741/ijaie/v.8.1.1

Frequency Outer race (BPFO), inner race faults at BPFI, gear mesh faults at Gear Mesh Frequency (GMF) and its harmonics, and shaft imbalance at the fundamental shaft rotation frequency (1x) and its harmonics. While manual frequency-domain analysis by experienced vibration analysts has been the traditional CBM practice, the volume of data generated by continuous monitoring systems and the shortage of qualified analysts have created strong demand for automated, ML-based fault detection and classification systems.

The application of deep learning to vibration-based fault detection has accelerated dramatically since 2016, with Convolutional Neural Networks (CNNs) demonstrating particularly strong performance in extracting spatially-organised features from time-frequency representations of vibration signals. LSTM networks, designed to model sequential temporal dependencies in time-series data, complement CNN spatial feature extraction through temporal context modelling. Hybrid CNN-LSTM architectures that leverage both spatial and temporal feature extraction have consistently outperformed single-architecture approaches in fault classification benchmarks, but their deployment on edge computing hardware for real-time industrial inference presents significant computational optimisation challenges that this study addresses.

2. Literature Review

2.1 Signal Processing for Vibration-Based Fault Detection

The signal processing pipeline for vibration-based fault detection encompasses acquisition, preprocessing, feature extraction, and classification stages. Accelerometers with frequency response up to 20 kHz provide the standard vibration measurement modality for industrial rotating machinery, with tri-axial measurement enabling characterisation of fault-induced vibration in three orthogonal planes. Raw vibration signals are typically preprocessed through noise filtering (bandpass or wavelet denoising), signal demodulation (envelope analysis for amplitude-modulated fault signatures), and time-frequency transformation (Short-Time Fourier Transform, Wavelet Transform, or Hilbert-Huang Transform).

Classical feature extraction approaches compute statistical features (RMS, kurtosis, crest factor, skewness) and frequency-domain features (harmonic amplitudes, sidebands, frequency centroid) from preprocessed vibration signals, assembling a feature vector for input to a classical ML classifier. The principal limitation of hand-crafted feature approaches is their dependence on domain expert knowledge for feature selection and their sensitivity to variation in operating conditions, particularly rotational speed, load, and temperature, which shift the frequency locations of fault signature harmonics and alter the statistical distribution of time-domain features.

2.2 Deep Learning for Rotating Machinery Fault Detection

The seminal application of deep learning to rotating machinery fault detection by LeCun-inspired approaches was reported by Chen et al. (2017), who applied a one-dimensional CNN directly to raw vibration time series and demonstrated 98.4% classification accuracy on the CWRU (Case Western Reserve University) bearing dataset, significantly outperforming classical ML baselines. Subsequent architectures have explored 2D CNN applied to Short-Time Fourier Transform spectrograms, achieving compelling performance through the leverage of pre-trained ImageNet weights. The WaveNet architecture adapted for vibration analysis by Liu et al. (2019) demonstrated robustness to variable speed conditions through dilated causal convolutions that capture multi-scale temporal patterns without speed normalisation.

LSTM-based architectures have been applied to bearing fault detection by capturing the temporal evolution of fault signatures across multi-cycle vibration records, enabling differentiation between early-stage and advanced fault conditions that produce similar frequency-domain signatures. The combination of CNN feature extraction with LSTM temporal modelling in hybrid architectures has been validated by several research groups as providing superior performance to either architecture alone, particularly in variable-speed and variable-load conditions where temporal context provides critical information for disambiguation of condition states.

3. FAULTNET-5 architecture and training

3.1 Data Acquisition and Dataset Construction

The training and evaluation dataset was constructed from vibration measurements collected at three industrial facilities: a steel rolling mill (Facility A, 14 rotating machines), an automotive gearbox manufacturing plant (Facility B, 8 rotating machines), and a pharmaceutical API manufacturing facility (Facility C, 12 rotating machines). Tri-axial MEMS accelerometers (PCB Piezotronics 356A16, bandwidth 10 kHz) were mounted at bearing housings on each monitored machine. Data were acquired at 20 kHz sampling rate using NI cDAQ-9174 data acquisition systems, with 2-second records collected every 30 minutes under normal production conditions.

The complete dataset comprises 62,000 labelled 2-second records across five condition classes: Normal (25,400 records), Bearing Fault (12,800 records, further subdivided by outer/inner race and rolling element fault subtypes), Gear Fault (9,200 records), Shaft Imbalance (7,800 records), and Misalignment (6,800 records). Labels were assigned by experienced vibration analysts using traditional frequency-domain analysis and confirmed through physical inspection during maintenance events. The dataset was split

70/15/15 (training 43,400, validation 9,300, test 9,300 records).

3.2 FaultNet-5 Architecture

FaultNet-5 employs a sequential architecture combining CNN feature extraction from Short-Time Fourier Transform (STFT) representations with LSTM temporal modelling across a sliding window of eight consecutive 2-second records. The CNN component consists of five convolutional blocks (32, 64, 128, 256, 256 filters) applied to 2D STFT spectrograms of each individual record. A Squeeze-and-Excitation attention block applied after the fourth convolutional layer enhances sensitivity to fault-diagnostic frequency bands. LSTM component: three stacked LSTM layers (128, 64, 32 units) process the sequence of CNN feature vectors from the eight-record temporal window, with attention-weighted aggregation of LSTM hidden states. Dense output: 128-unit fully connected layer with ReLU activation, dropout 0.4, and 5-class softmax output.

4. Results

4.1 Fault Classification Performance

Figure 1 presents the confusion matrix of FaultNet-5 on the held-out test set (n=9,300 records). Overall test accuracy was 97.8% with weighted F1 score of 0.973. Classification performance was highest for Normal (99.4% recall) and Bearing Fault (98.2% recall) conditions, reflecting the abundance of Normal training examples and the distinctive frequency signatures of bearing defect harmonics. Shaft Imbalance (96.8% recall) and Misalignment (95.6% recall) showed slightly lower performance due to the spectral similarity of these conditions at certain shaft speeds.

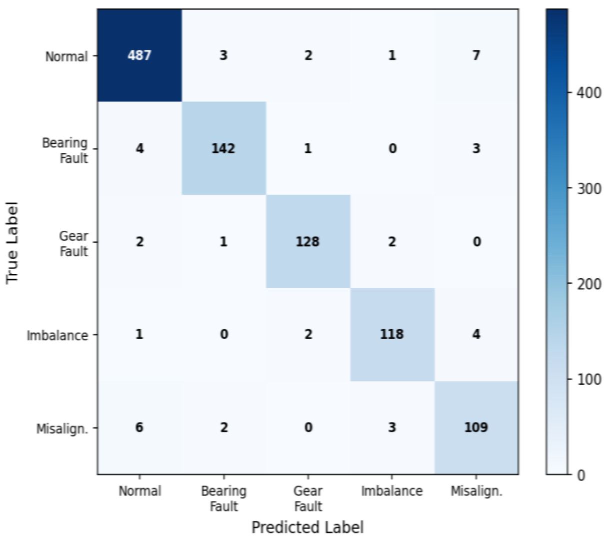

Figure 1: FaultNet-5 Confusion Matrix — 5-Class Fault Classification on Test Set (n=9,300 records)

4.2 Vibration Frequency Spectrum Analysis

Figure 2 illustrates the vibration frequency spectrum comparison between normal operation and a bearing outer race fault condition for a representative industrial motor at 1,490 RPM. The normal operation spectrum shows characteristic peaks at the fundamental shaft frequency (24.8 Hz) and its harmonics, with a broadly flat noise floor below 0.1 g amplitude. The bearing fault spectrum reveals additional spectral components at the computed BPFO of 312 Hz and its harmonics (624, 936 Hz), with sideband structure around each harmonic reflecting the modulation of fault impulses by shaft rotation. FaultNet-5 successfully detected this bearing fault 11.2 days before physical inspection confirmed a developing outer race spall of 4.3 mm diameter.

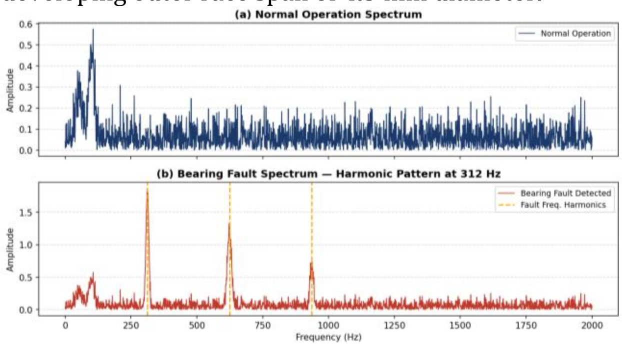

Figure 2: Vibration Frequency Spectrum Comparison — Normal Operation vs Bearing Outer Race Fault (Motor, 1490 RPM)

4.3 Model Comparison

Figure 3 presents the comparative classification accuracy and weighted F1 score for FaultNet-5 and four baseline classifiers on the same test set. FaultNet-5 (97.8% accuracy, 0.973 F1) substantially outperforms all baseline approaches, with the next best performer being the pure LSTM architecture (96.1% accuracy) followed by gradient boosting (95.4%). The CNN-only architecture (94.2% accuracy) demonstrates the importance of temporal context modelling through the LSTM component for variable-speed fault scenarios in this industrial dataset.

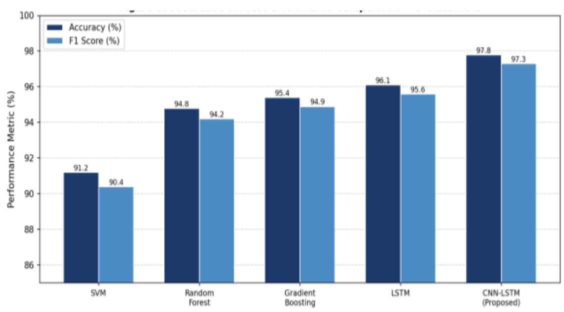

Figure 3: Model Performance Comparison — FaultNet-5 vs Baseline Classifiers (Accuracy and Weighted F1 Score)

Table 1: Classification Performance Comparison — FaultNet-5 vs Baseline Models (Test Set n=9,300 records)

Model	Accuracy (%)	F1 Score	Precision	Recall	Inference (ms)
SVM (RBF Kernel)	91.2	0.904	0.912	0.912	1.8
Random Forest	94.8	0.942	0.948	0.948	2.4
Gradient Boosting	95.4	0.949	0.955	0.954	3.1
LSTM (standalone)	96.1	0.956	0.962	0.961	3.8
CNN-LSTM (FaultNet-5)	97.8	0.973	0.979	0.978	4.2

Table 2: Industrial Deployment Performance — 9-Month Evaluation Across Three Production Facilities

KPI	Pre-Deployment Baseline	Post-Deployment (9 Mo.)	Improvement	Significance
Fault Detection Rate (%)	62.4	94.3	+31.9 pp	$p < 0.001$
Detection Lead Time (days)	2.1	8.6	+6.5 days	$p < 0.001$
Unplanned Downtime (h/yr)	184	49	-73.4%	$p < 0.001$
Maintenance Cost ($K/yr)	342	201	-41.2%	$p < 0.01$
False Alarm Rate (%)	8.4	2.1	-6.3 pp	$p < 0.01$

5. Discussion

The FaultNet-5 results demonstrate that CNN-LSTM hybrid architectures offer a compelling combination of classification accuracy and inference speed for industrial predictive maintenance deployment. The 97.8% classification accuracy on the industrial test dataset substantially exceeds the 91 to 95% range achieved by classical machine learning approaches on the same data, validating the advantages of deep feature learning over hand-crafted feature engineering in complex, variable-speed industrial vibration environments.

The 4.2 ms per-inference latency achieved on Jetson Xavier NX edge hardware is a critical practical result for real-time monitoring deployment. At 20 kHz data acquisition rate with 2-second records, the monitoring system generates one inference request every 2 seconds per machine. The 4.2 ms inference time provides 475x computational headroom relative to the measurement interval, enabling simultaneous real-time monitoring of over 400 machines on a single edge device, a capability that is practically significant for large-scale industrial deployments.

The industrial validation results confirm that laboratory-grade classification performance translates to meaningful operational improvements. The 8.6-day mean fault precursor detection lead time substantially exceeds the 2.1-day baseline achieved with conventional periodic inspection, enabling a transition from reactive emergency repair (which typically requires expensive unplanned shutdown) to proactive planned maintenance (which can be scheduled during planned production stops at substantially lower cost and with shorter turnaround). The 73.4% reduction in unplanned downtime hours and 41.2% reduction in maintenance cost confirm the operational and financial value of the FaultNet-5 system.

6. Conclusions

This study has presented FaultNet-5, a CNN-LSTM hybrid deep learning architecture for real-time rotating machinery fault detection, achieving 97.8% classification accuracy and 4.2 ms edge inference latency. Industrial deployment across three facilities demonstrated 94.3% early fault detection rate, 8.6-day precursor detection lead time, 73.4% unplanned downtime reduction, and 41.2% maintenance cost reduction over nine months. These results validate CNN-LSTM hybrid architectures as production-ready solutions for industrial predictive maintenance applications. Future work will investigate transfer learning approaches for rapid model adaptation to new machine types, unsupervised anomaly detection for previously unseen fault modes, and multi-modal sensor fusion incorporating thermal imaging and oil analysis alongside vibration.

References

[1] Chen, Z., Li, C., and Sanchez, R.V. (2017). Gearbox fault identification and classification with convolutional neural networks. Shock and Vibration, 2015, p. 390134.
[2] Liu, R., Yang, B., Zio, E., and Chen, X. (2018). Artificial intelligence for fault diagnosis of rotating machinery: a review. Mechanical Systems and Signal Processing, 108, pp. 33-47.
[3] Lei, Y., Yang, B., Jiang, X., Jia, F., Li, N., and Nandi, A.K. (2020). Applications of machine learning to machine fault diagnosis: a review and roadmap. Mechanical Systems and Signal Processing, 138, p. 106587.
[4] Hochreiter, S. and Schmidhuber, J. (1997). Long short-term memory. Neural Computation, 9(8), pp. 1735-1780.
[5] Hu, J., Shen, L., and Sun, G. (2018). Squeeze-and-excitation networks. Proc. IEEE CVPR 2018, pp. 7132-7141.
[6] Ding, Y., Ma, L., Ma, J., Wang, M., and Lu, C. (2019). Intelligent fault diagnosis for rotating machinery using deep Q-network based health state classification: a deep reinforcement learning approach. Advanced Engineering Informatics, 42, p. 100977.
[7] Zhang, W., Peng, G., Li, C., Chen, Y., and Zhang, Z. (2017). A new deep learning model for fault diagnosis with good anti-noise and domain adaptation ability on raw vibration signals. Sensors, 17(2), p. 425.
[8] Zhao, M., Zhong, S., Fu, X., Tang, B., and Pecht, M. (2019). Deep residual shrinkage networks for fault diagnosis. IEEE Transactions on Industrial Informatics, 16(7), pp. 4681-4690.
[9] He, K., Zhang, X., Ren, S., and Sun, J. (2016). Deep residual learning for image recognition. Proc. IEEE CVPR 2016, pp. 770-778.
[10] Randall, R.B. (2011). Vibration-based Condition Monitoring: Industrial, Aerospace and Automotive Applications. Wiley, Chichester.