A Monitoring Framework for Multi-Fault Detection in Rotating Shafts Integrating Deep Learning and Signal Reconstruction

EMANUELE PETRICONI, VASILIKI PANAGIOTOUPOULOU, MARCO GIGLIO and CLAUDIO SBARUFATTI

ABSTRACT

Shafts and rotors are essential components in many engineering applications, often playing a critical role in system functionality. Unexpected failures can result in costly downtime or pose risks to human safety, with both situations to be prevented. To mitigate these risks, there is a growing interest in monitoring systems that can proactively detect potential faults and provide real-time diagnostic insights to operators. However, rotor damage often leads to imbalances that can adversely affect other components, such as supports, or induce secondary faults, like cracks caused by high forces associated with rotation. Detecting both the imbalance and the presence of additional faults enhances the assessment of operational risks, enabling early alerts for catastrophic failures and safety-critical concerns. This study presents a monitoring framework for detecting multiple sources of damage in a transmission shaft by analyzing the accelerations at the support bearings during rotation. The rotor is experimentally tested on a test rig, with imbalance masses placed at various locations to simulate the effects of bullet impacts. In addition, a second damage source, specifically a support fault, is also artificially introduced and tested.. The proposed methodology combines a CNN for imbalance diagnosis with a signal reconstruction model for anomaly detection, both using the same acceleration data. This dual approach enables accurate localization of imbalances and detection of unknown faults while minimizing data requirements. The framework is designed to encourage future implementations in real-world applications potentially subjected to multiple faults.

INTRODUCTION

Condition monitoring (CM) and fault diagnosis of rotating machinery are key aspects of modern industrial maintenance strategies, especially within the broader context of Structural Health Monitoring (SHM). These techniques aim to detect and characterize mechanical faults at an early stage, enabling predictive maintenance, minimizing unexpected downtimes, and ultimately ensuring the safety and efficiency of critical mechanical systems. Rotating components such as shafts and bearings are integral to a wide variety of industrial equipment used in sectors including power generation, aerospace, automotive, and manufacturing. Due to their continuous rotation and the loads generated during operations, the driveline is prone to degradation mechanisms such as imbalances and support defects, which can turn into severe failures if not immediately addressed.

Traditional fault detection techniques are typically based on vibration analysis and signal processing methods, including Fourier transforms, envelope analysis, and wavelet decomposition [1, 2]. Although effective, these approaches often require expert interpretation and heavily rely on hand-crafted features, which may not generalize well under varying operational conditions. Moreover, the increasing complexity of modern machinery has created a demand for more robust and automated diagnostic tools that can operate reliably under non-stationary conditions and varying load profiles.

Condition monitoring can rely on either model-based or data-driven approaches.

Emanuele Petriconi, PhD Student, Email: emanuele.petriconi@polimi.it, Department of Mechanical Engineering, Politecnico di Milano, Milan, MI, Italy

Model-based methods require accurate physical modeling of the system and can provide deep insight into the nature of damage [3, 4]. However, they are often impractical for complex structures where detailed modeling is not feasible. Data-driven approaches, on the other hand, leverage numerical and experimental data, often through deep learning techniques, to detect and classify damage without necessarily requiring an explicit physical model [5–7]. Convolutional Neural Networks (CNNs), originally developed for image recognition, have been widely adopted in Structural Health Monitoring (SHM) due to their ability to automatically extract relevant features from input data. While CNNs are commonly used with 2D inputs such as greyscale images derived from Short-Time Fourier Transform (STFT) or Continuous Wavelet Transform (CWT), they can also be adapted to process 1D signals [8, 9]. In this case, the input is a time-series signal—such as acceleration or strain—from which the network learns to identify patterns associated with specific fault conditions. Another Artificial Neural Network (ANN) architecture that has proven effective for supervised monitoring is the Recurrent Neural Network (RNN). In particular, Long Short-Term Memory (LSTM) networks and Non-linear Autoregressive Exogenous (NARX) models have shown considerable potential in modeling the temporal dynamics of vibration signals [10, 11]. These architectures are well-suited for capturing long-term dependencies and the dynamic behavior associated with a health conditions. Unlike Convolutional Neural Networks (CNNs), these models can reconstruct the signals and evaluating the reconstruction error can provide an anomaly detection. The main limitation, which to the best of the authors' knowledge has not yet been adequately addressed, concerns the degradation of diagnostic performance when a different kind of damage is present. This scenario does not simply involve extrapolation beyond the training data, as a novel fault type can alter the structural mechanical behavior in a fundamentally different manner. As a consequence, the monitoring ANN may consistently produce inaccurate predictions without issuing any warning to the operators, or to the crew in the case of aircraft, potentially leading to incorrect decisions and significant safety risks.

The main objective of this work is to develop a framework that addresses these limitations, enabling reliable diagnosis even in the presence of different sources of damage. The proposed approach combines CNNs and RNNs to diagnose faults in a rotating shaft affected by various issues, such as imbalances and support defects, while minimizing the amount of required data without sacrificing diagnostic performance. In the case of imbalances, the framework provides a more detailed diagnosis also evaluating the fault location.

METHODOLOGY

This work aims to develop a monitoring framework capable of detecting and localizing shaft imbalances, while also providing alerts when different faults are present. Following experimental testing to collect data for training the proposed data-driven Deep Learning (DL) models, the methodology is structured in two phases: (i) the development of a CNN for imbalance detection and localization, and (ii) the implementation of LSTM and NARX models for anomaly detection related to support faults, providing a comparative study to evaluate the obtained performance. The approach is entirely data-driven, developed using experimental data acquired from the shaft test rig described in the following section. The primary damage condition targeted by the framework is the shaft imbalance, with its location treated as discrete, posing a classification challenge for the network. The imbalance alters the mechanical response of the shaft during rotation, and the acceleration signal captured at the support is used to train the data-driven models which solve the inverse problem to identify the responsible damage condition. The CNN architecture is selected for this task due to its proven ability to autonomously identify relevant features from input data and associate them with different classes. The CNN is structured to receive a 1D acceleration signal as input, as the network can operate effectively with vectors rather than images, reducing the number of parameters to be optimized. CNN is trained solely on experimental data from healthy and unbalanced conditions. It is important to note that including data from the support fault condition would necessitate a more complex classification function, requiring a larger dataset. Using a minimal amount of data to train an ANN while ensuring reliable predictions is crucial, particularly in rotating machinery, where experimental data can be costly. Rather than increasing model complexity, this approach decouples the shaft diagnosis problem into two stages: (i) imbalance detection and localization, addressed by the CNN, and (ii) anomaly detection for other damages exploiting signal reconstruction. For this purpose, a second ANN is developed. It must predict whether the shaft is healthy, imbalanced, or affected by another type of fault—specifically, a support defect, which represents the anomaly in this case. Both LSTM and NARX architectures are tested for this purpose, since they allow a classification through an unsupervised learning. The training for the reconstruction takes in input two accelerations captured at both shaft's ends in both cases. Evaluating the reconstruction Mean Square Error (MSE) is possible to determine whether the accelerations belongs to an anomaly condition, or to an imbalanced or healthy shaft. The imbalance is not treated as an anomaly since the CNN already monitors it, meanwhile the second network must only verify that the shaft's healthy status falls within the ones seen in training by the first network. Finally, online monitoring is carried out according to the following logic: the acquired acceleration signals are preprocessed and then used as input for both CNN-based classification and signal reconstruction. CNN is responsible for diagnosing whether an imbalance is present and identifying its location. At the same time, the reconstruction error serves as an indicator of other potential faults and provides a warning regarding the reliability of the CNN's prediction.

CASE STUDY

The proposed methodology is developed and applied to the following case study. The subject of the investigation is a rotating shaft subjected to various fault conditions, installed in the center of the experimental setup shown in Figure 1.

The tested shaft is made of C40 steel, with a length of $1024mm$ and a diameter of $35mm$. It is manufactured with five threaded M8 holes equally spaced along its longitudinal axis, which serve as mounting points for imbalance masses. The test rig is remotely controlled via a LabVIEW program, allowing for both speed regulation and sensor data acquisition. The sensing system includes four vertically oriented monoaxial accelerometers (PCB Piezotronics 333B30) mounted on top of each bearing housing, an encoder (SICK DGS80) to monitor rotational speed, and a thermocouple (Tersid FT905) placed near the motor to measure the external temperature of the closest bearing. Tests are performed keeping a constant rotational speed of $2500 rpm$, with the system operating within a measured temperature range of $35^{\circ}C$ to $40^{\circ}C$. Artificial damage conditions are introduced to simulate both imbalance and support defects. The imbalancing mass, approximately $60g$, generates a centrifugal force of approximately $72N$ at $2500rpm$. This weight is selected to replicate the imbalance caused by a projectile impact on a drive shaft [10]. Support defects are simulated by loosening the bolts of a bearing housing, with different bearings tested separately. Examples of both damage conditions are illustrated in Figure 2.

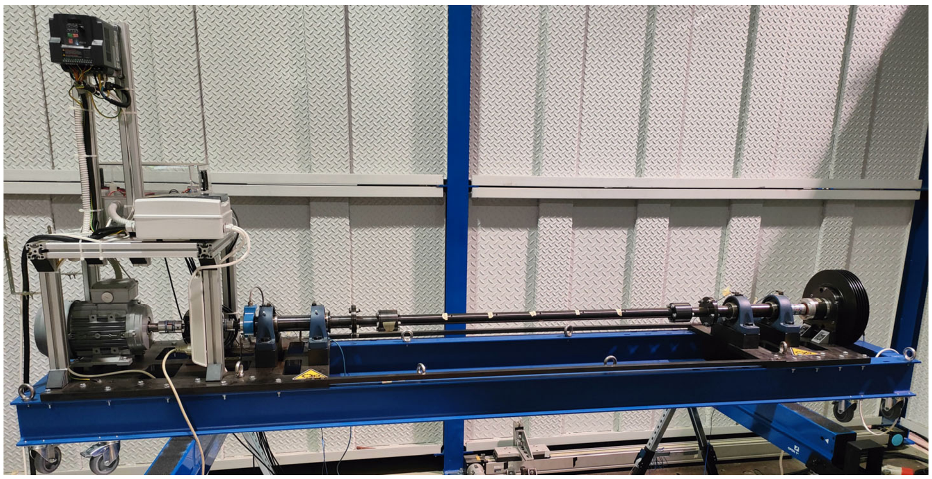

Figure 1. Experimental setup.

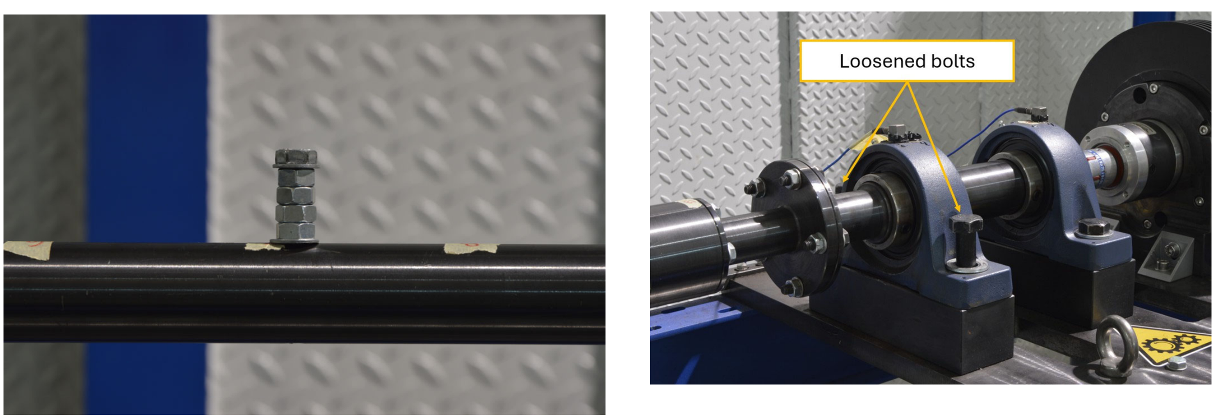

Figure 2. Damaged conditions: Imbalance (*left*) and Support fault (*right*).

The experimental tests are grouped into four main categories: (case 1) baseline, (case 2) imbalance only, (case 3) support fault only, and (case 4) combined imbalance and support fault. The baseline condition includes tests performed without any artificial damage. In case 2, the added mass placement varies across the five threaded holes along the shaft. The support fault condition is simulated by loosening the bolts of one bearing, with different bearings tested in separate runs to evaluate the influence of defect location. Finally, the experimental campaign includes tests with both faults combined. Acceleration data are acquired at a sampling frequency of $51.2kHz$ and preprocessed using a 5^{th}-order

Butterworth low-pass filter with a cut-off frequency of $1500Hz$. The filtered signals are then resampled to $40kHz$ and segmented into windows of 1000 samples, which approximately corresponds to one shaft revolution. The proposed framework begins with the training of the CNN for the diagnosis of shaft imbalance. The network is trained using data collected under baseline and imbalance-only conditions. The imbalance detection task is formulated as a classification problem, given that the imbalance mass can only be located at discrete positions along the shaft. This formulation aligns with realistic scenarios, such as those found in gas or steam turbines, where imbalances typically occur at specific stages. The dataset is divided into training and validation subsets following an $80/20$ split.

The CNN architecture comprises a sequence of convolutional layers, each followed by a ReLU activation function to introduce non-linearity into the model. Following each convolutional block, a max pooling layer is applied to progressively reduce the temporal resolution while preserving the most salient features. This design enhances computational efficiency and improves the network's robustness to temporal translations. The input layer of the network has dimensions 1×1000, corresponding to one-dimensional acceleration signals consisting of 1000 samples, acquired from the accelerometer located to the left of the analyzed shaft. The final part of the network consists of fully connected layers, culminating in a softmax classifier that outputs the predicted class corresponding to the detected imbalance location. To prevent overfitting and enhance generalization, two dropout layers are incorporated at key points in the architecture, before the first fully connected layer and before the output layer. These layers randomly deactivate a subset of neurons during training, encouraging the network to learn more robust and distributed representations. The respective dropout rates are 0.1 and 0.4. The model is trained using the Adam optimizer over 100 epochs, which adaptively adjusts the learning rate during training to facilitate efficient convergence. A categorical cross-entropy loss function is employed, as it is well suited for multi-class classification tasks. The CNN is trained to classify the input signal into six distinct classes: class '0' corresponds to a healthy, undamaged shaft, while classes '1' through '5' represent various imbalanced conditions, with each class indicating the specific hole where the additional mass is mounted. The model's performance is evaluated by predicting the shaft condition on five new data acquisitions for each true class. These additional tests are designed to assess the network's extrapolation capability by gradually reducing the imbalance mass by up to 20%, either by removing a nut or replacing it with lighter washers. This experimental setup allows for evaluating the model's robustness when exposed to damage levels not included in the training dataset. As shown in Table I, the prediction accuracy remains above 80% for all classes, despite the reduced imbalance levels. To further evaluate the network's extrapolation limits, additional acquisitions are performed under support fault conditions, both with and without the presence of an imbalance. As expected, the classification performance significantly deteriorates in the presence of this additional fault, as also visible in Table I. This result highlights the need for an auxiliary monitoring mechanism to distinguish between predictable and unforeseen conditions.

It is important to note that extending the CNN to directly account for additional damage conditions would require a substantially larger dataset. CNNs used for classification tasks typically demand balanced datasets, and including further fault classes would significantly increase data acquisition costs,a critical limitation in many real-world scenar-

	Prediction accuracy [%]					
	Class 0	**Class 1**	**Class 2**	**Class 3**	**Class 4**	**Class 5**
No support fault	100	82	84	91	90	97
With support fault	30	14	12	19	9	4

TABLE I. Prediction accuracy for different classes and conditions.

ios where faults are difficult or expensive to reproduce. Moreover, incorporating additional classes would inherently raise the complexity of the classification task, requiring a greater number of trainable parameters and potentially compromising the model's generalization if not supported by sufficient training data. These considerations motivate a problem decomposition strategy, where the same dataset used for CNN training is also employed to train a second network. For this reason, two signal reconstruction models, NARX and LSTM, are trained and tested to evaluate their effectiveness in detecting anomalous conditions not seen by the CNN. Both models are trained using the same dataset employed for CNN training, segmented into batches of 50 samples. By learning to accurately reconstruct signals under known healthy or imbalanced conditions, the network can flag deviations, such as those caused by unseen support faults—based on reconstruction errors, thereby enabling detection without explicit supervision for every possible fault type. For both NARX and LSTM models a total of $48 \cdot 10^5$ samples are randomly partitioned into training and testing sets using a $80/20$ split, and the Levenberg-Marquardt (LM) algorithm is used as optimization method. Training is carried out in four batches over 10^3 epochs for the NARX model and 100 epochs for the LSTM model, respectively. These hyperparameters are optimized through a dedicated tuning process to ensure effective model convergence. The maximum recorded mean squared error (MSE) loss for the NARX model is $4.4 \cdot 10^{-6}$, while the LSTM model achieves a higher loss of $1.3 \cdot 10^{-5}$. Upon completion of the training phase, the centroid of the estimation error distribution is calculated for each approximated time series and stored in a vector ϵ. A threshold is then established based on the mean value of this vector. To evaluate the performance of both models in damage detection, their capabilities are assessed using a data set in which support faults are introduced into the rotary shaft system, thereby altering its dynamic behavior and vibration characteristics. This damage data set comprises 10 acceleration signal recordings from each sensor, obtained by testing the support fault with and without the shaft's imbalance. Following the same signal processing procedure described above, the resulting time series are segmented into data batches of 50 samples and subsequently input into the trained models. For each approximated acceleration signal, the centroid of the estimation error, indicated as e, was continuously monitored. When this centroid exceeded the predefined threshold, an alarm was triggered, indicating the presence of structural damage. In Figure 3, the ability of the two RNN models to detect the faulty state of the system is evaluated through two distinct dynamic test scenarios, known as Case 1 and Case 2. Case 1 corresponds to the dataset previously described, whereas Case 2 comprises one acquisition at the baseline state and one at the bearing fault state for each of the two accelerometers, resulting in a total of four acquisitions. The notation *damaged* is referred to a condition where the support fault is present, independently if an imbalance is present or not. Conversely, the notation *healthy* is referred to a configuration without artificial support damages, and the imbalance can be present or not. It should be noted that Case 2 is excluded from the training dataset; therefore, it serves to evaluate and validate the generalization performance of both RNN models under varying operational and environmental conditions. Overall, both models offer robust and early-stage alarms for the detection of bearing faults. The NARX model, in particular, demonstrates a clearer distinction between baseline and faulty system states, thereby ensuring reliable fault identification. However, in more complex scenarios, this can lead to missed fault detections—referred to as false negatives—which can result in the continued operation of the rotary system despite the presence of actual faults. Conversely, the LSTM model may produce false positive alarms, potentially triggering unnecessary interruptions in system operation. Such unwarranted halts, in the absence of real faults or emergencies, can lead to significant downtime and associated financial losses. The selection of the best model depends on the case study and the specific industrial application.

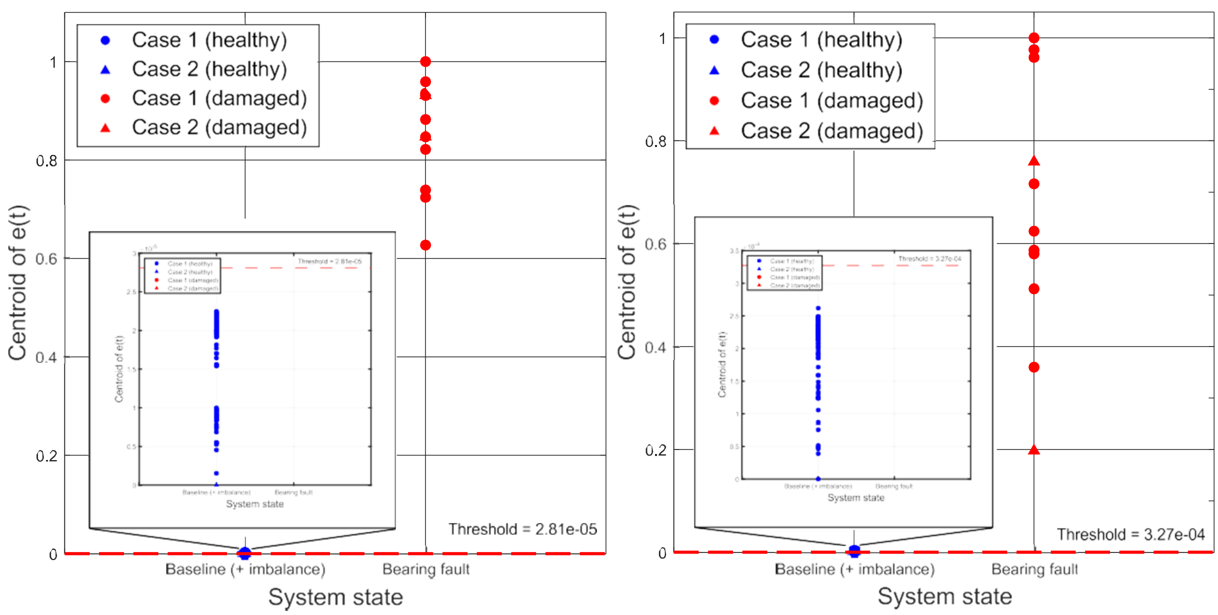

Figure 3. Normalized centroids of the error distribution $e(t)$ during time series estimation under baseline and faulty system states, using the NARX model (*left*) and the LSTM model (*right*).

Both models are able to detect the presence of an anomaly and thus provide a reliability feedback to the monitoring system, indicating that the CNN predictions may not be accurate. This mechanism is particularly valuable given the significant drop in CNN performance under damaged conditions, as shown in Table I.

CONCLUDING REMARKS

The study demonstrates that dividing the diagnostic task into two separate problems, imbalance identification and support fault detection, allows effective monitoring of multiple damages without the need to expand the training dataset. The approach relies on the initial assumption that imbalance requires a more detailed assessment, while support faults only need to be detected. This assumption is valid in scenarios where certain fault conditions are either particularly safety-critical or difficult to reproduce, and a balanced dataset is not available.

Future developments may aim to extend the current imbalance diagnosis by incorporating quantitative estimation of the imbalance level, in addition to its localization. On the side of unsupervised anomaly detection, while NARX and LSTM models have shown promising results, other approaches, such as Autoencoders, also represent viable and potential alternatives for this task. In conclusion, the key contribution of this work lies in the effectiveness of decomposing the monitoring task into two sub-problems, addressed with two complementary algorithms: supervised learning for detailed classification, and unsupervised learning for generalized fault detection in complex structures subjected to multiple faults, without requiring additional amount of data.

REFERENCES

1. D., G. and P. B. S. 2016. “The Vibration Monitoring Methods and Signal Processing Techniques for Structural Health Monitoring: A Review,” *Archives of Computational Methods in Engineering*, 23:585–594.
2. Chen, B., W. Zhang, J. Xi Gu, D. Song, Y. Cheng, Z. Zhou, F. Gu, and A. D. Ball. 2023. “Product envelope spectrum optimization-gram: An enhanced envelope analysis for rolling bearing fault diagnosis,” *Mechanical Systems and Signal Processing*, 193:110270.
3. Prabhat, K. and T. Rajiv. 2023. “A review: multiplicative faults and model-based condition monitoring strategies for fault diagnosis in rotary machines,” *Journal of the Brazilian Society of Mechanical Sciences and Engineering*, 45:282.
4. Souflas, I., A. Pezouvanis, and K. Ebrahimi. 2018. “Health monitoring system for transmission shafts based on adaptive parameter identification,” *Mechanical Systems and Signal Processing*, 104:673–687.
5. Gangsar, P., A. R. Bajpei, and R. Porwal. 2022. “A review on deep learning based condition monitoring and fault diagnosis of rotating machinery,” *Noise & Vibration Worldwide*, 53(11):550–578.
6. Calderon Hurtado, A., K. Kaur, M. Makki Alamdari, E. Atroshchenko, K. Chang, and C. Kim. 2023. “Unsupervised learning-based framework for indirect structural health monitoring using adversarial autoencoder,” *Journal of Sound and Vibration*, 550:117598.
7. Zhou, Q. and J. Tang. 2024. “An Interpretable Parallel Spatial CNN-LSTM Architecture for Fault Diagnosis in Rotating Machinery,” *IEEE Internet of Things Journal*, 11(19):31730–31744.
8. Parziale, M., Y. F. Yeung, K. Youcef-Toumi, M. Giglio, and F. Cadini. 2025. “Anomaly characterization for the condition monitoring of rotating shafts exploiting data fusion and explainable convolutional neural networks,” *Structural Health Monitoring*, 1:21.
9. Shahid, S. M., S. Ko, and S. Kwon. 2022. “Performance Comparison of 1D and 2D Convolutional Neural Networks for Real-Time Classification of Time Series Sensor Data,” in *2022 International Conference on Information Networking (ICOIN)*, pp. 507–511.
10. Vasiliki, P., B. Lorenzo, P. Emanuele, B. Andrea, M. Ugo, G. Marco, and S. Claudio. 2025. “Comparative study on ballistic impact detection in helicopter transmission shafts using NARX and LSTM models,” *Applied Intelligence*, 55:264.
11. Vasiliki, P., P. Emanuele, G. Marco, and S. Claudio. 2025. “Deep Learning-Based Identification of Shaft Imbalance Faults in Rotating Machinery Using the NARX Models,” *Journal of Vibration Engineering & Technologie*, 13:261.