

Article

A New Deep Learning Framework for Imbalance Detection of a Rotating Shaft

Muhammad Wisal and Ki-Yong Oh * 

Department of Mechanical Convergence Engineering, Hanyang University, 222, Wangsimni ro, Seongdong gu, Seoul 04763, Republic of Korea

* Correspondence: kiyongoh@hanyang.ac.kr

Abstract: Rotor unbalance is the most common cause of vibration in industrial machines. The unbalance can result in efficiency losses and decreased lifetime of bearings and other components, leading to system failure and significant safety risk. Many complex analytical techniques and specific classifiers algorithms have been developed to study rotor imbalance. The classifier algorithms, though simple to use, lack the flexibility to be used efficiently for both low and high numbers of classes. Therefore, a robust multiclass prediction algorithm is needed to efficiently classify the rotor imbalance problem during runtime and avoid the problem's escalation to failure. In this work, a new deep learning (DL) algorithm was developed for detecting the unbalance of a rotating shaft for both binary and multiclass identification. The model was developed by utilizing the depth and efficacy of ResNet and the feature extraction property of Convolutional Neural Network (CNN). The new algorithm outperforms both ResNet and CNN. Accelerometer data collected by a vibration sensor were used to train the algorithm. This time series data were preprocessed to extract important vibration signatures such as Fast Fourier Transform (FFT) and Short-Time Fourier Transform (STFT). STFT, being a feature-rich characteristic, performs better on our model. Two types of analyses were carried out: (i) balanced vs. unbalanced case detection (two output classes) and (ii) the level of unbalance detection (five output classes). The developed model gave a testing accuracy of 99.23% for the two-class classification and 95.15% for the multilevel unbalance classification. The results suggest that the proposed deep learning framework is robust for both binary and multiclass classification problems. This study provides a robust framework for detecting shaft unbalance of rotating machinery and can serve as a real-time fault detection mechanism in industrial applications.



Citation: Wisal, M.; Oh, K.-Y. A New Deep Learning Framework for Imbalance Detection of a Rotating Shaft. *Sensors* **2023**, *23*, 7141. <https://doi.org/10.3390/s23167141>

Academic Editor: Hossam A. Gabbar

Received: 11 July 2023

Revised: 6 August 2023

Accepted: 10 August 2023

Published: 12 August 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: unbalance detection; artificial neural network; deep learning; statistical property; Short-Time Fourier Transform; optimization

1. Introduction

The predominant factor contributing to mechanical failures in rotating machinery is the presence of vibrations [1], commonly caused by unbalance in the rotor [2,3]. “Rotor unbalance is a condition in which the Centre of mass of a rotating assembly, typically the shaft and its fixed components like disks, and blades etc. is not coincident with the Centre of rotation” [3]. Frequent causes of rotor unbalance are vibrations due to some externally applied load, bent shafts, or asymmetric mass distribution [4,5]. Other causes of rotor unbalance include bearing damage [6] and casting inaccuracies like porosity, non-uniform material density, asymmetric shaft fatigue, or faults in ball bearings that support the shaft [7,8]. Vibrations due to rotor unbalance can easily escalate to the failure of critical machine components such as bearings, gears, couplings, etc. [9]. Constant health monitoring in rotating machinery is needed to avoid structural damage.

Usually, three types of machinery maintenance approaches are adopted in industry: **Reactive maintenance**, often referred to as the “run till failure” approach;

Preventive maintenance, when machinery is periodically overhauled regardless of the condition of parts;

Predictive maintenance, which is the process of machinery health monitoring during runtime in order to predict when and which parts are likely to fail [10].

Reactive and preventive maintenance do not require any sensors to constantly monitor the condition of the machinery, whereas predictive maintenance does need health monitoring sensors like acoustic, vibration, shock pulse monitoring, wear debris, and thermal sensors [11]. Vibration sensor-based condition monitoring is common in industry as it allows for identifying 90% of faults or failures in machines [1].

In contrast to preventive and reactive maintenance, predictive maintenance is less costly and provides a constant analysis of machinery conditions [10]. Predictive-maintenance-based fault identification methods can be classified into two groups according to their dependence on the structural model: *model-based methods and signal-based methods* [12,13]. Model-based methods can estimate the damage location and severity by improving the mathematical model of the structure using experimental data or Finite Element Analysis data [12]. On the other hand, *signal-based methods* detect damage by comparing the structural responses before and after the damage. The method makes use of recorded signals rather than explicit input–output models for fault prediction. Damage is defined by damage indices, which are determined from the time–frequency domain analysis or by modal analysis [14,15]. Artificial Intelligence (AI), as a class of signal-based methodology, has emerged as a powerful tool for the identification of multivariable nonlinear systems.

Exploiting AI for motor fault detection has been of interest since 1993. Mo-Yuen Chow et al. [16] concluded that the size of training data directly affects the accuracy of the AI system. Bo Li et al. (1998) showed that FFT values of a vibration signal spectrum can be used as relevant features of the vibration dataset [17]. The authors of [4,16] proved that optimizing network parameters like learning rate, momentum, and neuron size is important for increasing accuracy. S. Rajakarunakaran et al. used an Adaptive Resonance Theory (ART) network as a vector classifier. It accepts time series vibration signals as the input vector and classifies them into one of the categories (clusters) depending on which of the stored patterns it most resembles [18]. The unsupervised method is good for clearly distinguished classes of balanced and highly unbalanced machine signals where the difference is high. However, the accuracy decreases fast when the unbalance factor is reduced. Chih-Hao Chen et al. (2008) proposed a new fault diagnosis procedure for rotating machinery fault detection using wavelet packet–fractal technology and a radial basis function neural network. The fault categories considered were unbalance, shaft misalignment, and looseness [19]; however, the level of unbalance was not considered. For diagnosing the type of unbalance, Sendhilkumar et al. proposed feeding vibration acceleration data features as input into Elman neural networks [20]. Weining Lu et al. [21] introduced a deep neural network for Domain Adaptation in Fault Diagnosis (DAFD), which increases classifier accuracy by adapting the neurons' weights to new values by deploying the weight regularization term, hence strengthening the representative features of the original data. Tang et al. [22] trained an SVM model using chaotic particle swarm optimization (PSO) to classify inputs into the multiple faults of rotating machines. Similar to the references [22,23], several other authors adopted SVM in fault classification models [24–27]. However, SVM runs into target class overlapping problems when faced with noisy and industrial-level large-scale data [28].

Rotor imbalance has been studied by various researchers recently. The authors of [29] presented an online unbalance rotor fault detection technique. The method carries out statistical and frequency domain analyses of the vibration data as well as current signatures; however, unlike machine learning methods, the proposed method requires expert-level knowledge of spectral analysis. Brando et al. used SVM for multiclass rotor misalignment prediction of an induction motor [30]; however, the methodology uses complex and extensive preprocessing steps such as the coefficient of statistical variation, the Boruta algorithm, and the Recursive Feature Elimination (RFE) algorithm. Wang et al. have shown that CNN performs significantly better and reduces preprocessing as compared

to SVM [31]. The authors used a three-dimensional STFT feature of the vibration data to classify the imbalanced, broken, and normal rotor. However, it is important to note that the use of 3D signals as input for CNN can be computationally expensive. In [32], the authors focus on a new CNN-based method for detecting various bearing faults of a planetary gearbox. The developed method specializes in cases where datasets are unbalanced or too noisy. The algorithm is promising for data with medium and high signal-to-noise ratios; however, its performance for large-scale industrial data is not known. Zhao et al. (2022) developed a rotor fault diagnostic framework based on the Normalized Conditional Variational Auto-encoder (NCVAE) [33]. The basic theme is to enhance and exploit the feature-learning ability of the NCVAE. Simulation-based bearing and rotor fault data are generated for classification purposes. The model focuses on increasing accuracy when data are imbalanced, as fewer data samples are available for certain classes. The model works well on simulated bearing and rotor fault data; however, as suggested by the authors, the actual industrial data, being complex and noisy, may need a more sophisticated model. All these models use variable depth models as the number of classes and data increases. A single model which can deal with fewer data and fewer classes, as well as multiclass problems and large datasets, has not been optimized for industrial rotor fault applications, which is the focus of this work.

A de facto research trend for prognostic health monitoring (PHM) problems is the use of shallower AI algorithms for binary class prediction with relatively small training data and deeper networks for multiclass high-level data [34]. These algorithms have shown promising results for the specific test conditions; however, shallower networks fail to adapt to complex data in multiclass problems [35], while deep networks degrade in performance on a small number of classes and run into overfitting [36]. This paper proposes a new deep learning algorithm that is deep enough to extract deep features of large datasets used in multiclass problems and also performs well on simple binary classification problems without running into overfitting. The model was developed for a special use case of an industrial rotor problem, using the same data and the same machine learning (ML) model for both binary class and multiclass prediction problems. For training the ML algorithms on vibration data, PHM engineers mostly prefer to use frequency components of signals obtained by Fast Fourier Transform (FFT) [34]. To address this problem, it is a good practice to analyze portions of the signal at various intervals with the help of STFT [37]. The feature-rich STFT has been used in fault diagnosis by several researchers to detect broken rotors [38], faulty bearings [39,40], and motor winding faults [31]. However, it has not been given much attention in the prediction of the unbalance level. FFT is effective for stationary signals but it is not quite powerful enough to analyze non-stationary signals [41]. To exploit the benefits of feature-rich STFT, we utilized STFT as the training input data for the classification of the rotor unbalance. This work focused on developing an online prediction model for multiclass rotor unbalance. When trained with different levels of unbalanced data, the algorithm can be applied to industrial rotors during runtime to predict the type and level of unbalance signaled by the runtime vibration data of the rotor. Hence, the unbalance problems identified in real time are dealt with before escalating into damage and complications. This is of significant help to the operator.

The database used in this study was downloaded from the Fraunhofer Fordatis Research Institute, Germany [42], and was collected on a DC motor setup connected with three vibration sensors. In total, five test condition datasets were recorded; one dataset was measured when there was no unbalance in the rotor, and the remaining four datasets were collected when different levels of unbalance were introduced to the rotor. The main contributions of this work are as follows:

- A new deep learning approach is developed for improved accuracy of binary and multiclass classification. Our model's basic architecture is derived from ResNet and CNN, and the developed model outperformed both algorithms. In addition, the results are compared with state-of-the-art ML algorithms, which shows the superiority of our

algorithm. To the best of our knowledge, this is the first study for the development of this model.

- A core dataset selection strategy is presented to speed up the training process by selecting fewer datasets for training. Among the four datasets for the unbalanced cases, only two were selected for training based on the statistical analysis by observing the standard deviations of the datasets. Two types of classifications were carried out. First, two-class classification for predicting balanced and unbalanced signals was performed (Analysis-1). Afterwards, in Analysis-2, a multiclass classification was performed to categorize the severity of rotor unbalance (refer to Table 1 for details). Analysis-2 is useful to predict the severity of the imbalance in the rotor (divided into four different classes).
- It demonstrates the feasibility of using the STFT feature map for better training, in contrast to conventional FFT as the main feature of data.

Table 1. Parameters of the dataset used. Data classes and datasets vary according to the unbalance factor present in the rotor.

Training Dataset	Evaluation Dataset	Attached Mass (g)	Radius (mm)	Unbalance Factor (mmg)	Analysis-1	Analysis-2
					Class (Total 2)	Class (Total 5)
D0	E0	0	0	0	Balanced (B)	Balanced (B)
D1	E1	3.281 ± 0.003	14 ± 0.1	45.9 ± 1.4	Unbalanced (U)	Unbalanced-1 (U1)
D2	E2	3.281 ± 0.003	18.5 ± 0.1	60.7 ± 1.9		Unbalanced-2 (U2)
D3	E3	3.281 ± 0.003	23 ± 0.1	75.5 ± 2.3		Unbalanced-3 (U3)
D4	E4	6.661 ± 0.007	23 ± 0.1	152.1 ± 2.3		Unbalanced-4 (U4)

The rest of the paper is organized as follows. Section 2 details the system under discussion. Section 3 describes the methodology in detail. An improved deep learning model is proposed in Section 4. Section 5 deals with the results. Finally, the conclusion of this study is given in Section 6.

2. System Overview

Rotary machines frequently experience imbalances as a result of high speeds, misalignment, and inappropriate loading or mass deposition. This imbalance is often the starting point of machine failure. One of the frequent causes of mechanical failures is rotor imbalance [8].

Imbalance in Rotary Machines

Mathematically, rotor imbalance can be expressed as follows [8]:

$$\vec{U} = m \times \vec{r}_u \quad (1)$$

where m is the added mass in grams and \vec{r}_u is the distance in millimeters of added mass from the axis of rotation. The common types of imbalances in rotating machinery are (i) static, (ii) coupled, (iii) quasi-static, and (iv) dynamic [29]. In this article, only static imbalance is considered, which is the most common type of imbalance in industrial machines. Static imbalance is the condition when the center of gravity of a rotor shaft is not aligned with its axis of rotation, primarily due to asymmetric mass distribution [43], as shown in Figure 1.

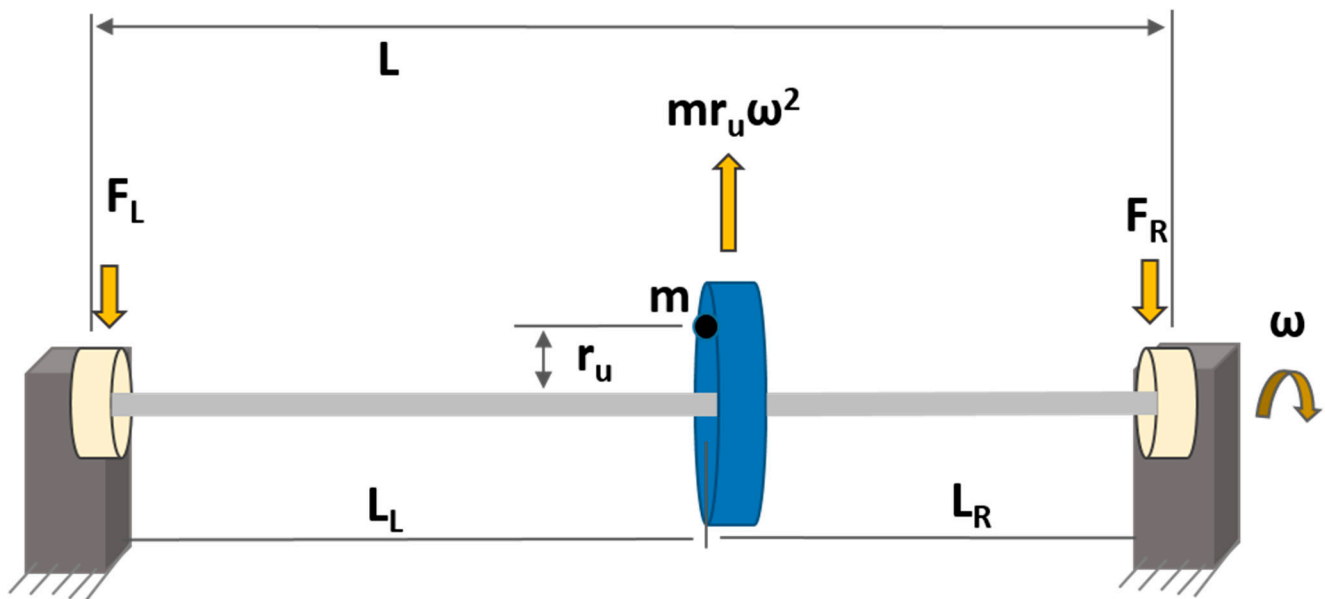


Figure 1. Graphical representation of static imbalance in rotors.

It is called static because it is always present in rotors, even if they are stationary [43]. Static imbalance produces the centrifugal force induced by mass imbalance when a rotor is in motion, which is stated as follows [44]:

$$F = m \times r_u \times \omega^2 \quad (2)$$

where F is the force in Newtons and ω is the speed in radians per second (rad/s). The centrifugal force is balanced by the reaction force on the bearing(s) on the endpoint of the rotor.

3. Methodology

3.1. System Parameters

The data used in this study were collected from a 130-watt DC motor (type UE 511TM, manufactured by WEG GmbH, Kerpen, Germany) [45] connected to a steel shaft 12 mm in diameter and 75 mm in length. The shaft was guided by a roller bearing in a galvanized steel bearing block. The experimental setup is depicted in Figure 2. A 3D-printed bracket for holding the load mass was inserted at the end of the shaft. Unbalance was created by inserting a mass at different radii in the 3D-printed bracket. Three vibration sensors (M001AC) attached to the motor mounting and bearing block were read out using a four-channel data acquisition device (DT9837). The mass (m) and the distance from the axis of rotation (r_u) are the two basic parameters considered for the imbalance factor. Table 1 shows the test conditions used in this study. For each test condition, the motor was swept through a rotation speed between 300 and 2300 revolutions per minute. Motor speed was controlled by a 24 V controller (WEG GmbH, type W2300) mounted on a plate of galvanized steel. A total of 10 datasets were recorded, 5 for training purposes (dataset D) and 5 for evaluation (dataset E). Dataset 0 D, 0 E was recorded when there was no unbalanced load acting on the shaft. Datasets 1–4 were recorded when loads acted with various configurations of size and location.

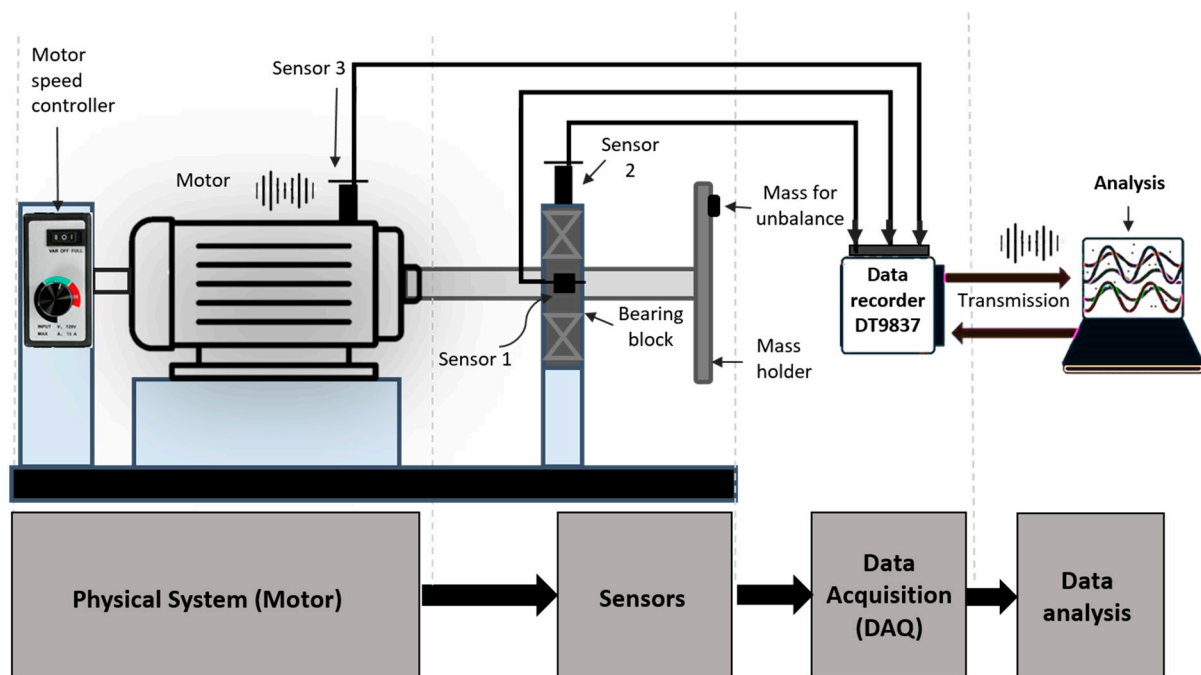


Figure 2. Conceptual design layout of the data acquisition setup.

As shown in Equation (1), the product of the mass m and the radius r is a direct measure of the unbalance strength. A large mass and radius will therefore produce the highest unbalance factor (unbalance U_4 in our study). Two types of classifications are performed. First, in Analysis-1, a binary classification of balanced (B) and unbalanced (U) is performed, as shown in Table 1. Second, in Analysis-2, a 5-label classification of balance (B) and the level of unbalance (U_1, U_2, U_3, U_4) is performed.

3.2. Data Selection and Preprocessing

It is important to emphasize that two types of classification analyses (shown in Table 1) are performed in this study. Analysis-2, a multiclass prediction problem, is relatively simple in terms of input data, as for each class label, exactly one dataset is available. Meanwhile, Analysis-1, a binary classification problem—“Balanced” and “Unbalanced”—has one training dataset D_0 for the balanced case and four training datasets (D_1 – D_4) for the unbalanced case. Using all these datasets (D_1 – D_4) in training for the class label “Unbalanced” in Analysis-1 will result in overfitting as well as high computational costs. The large datasets in Analysis-1 will push the model to overfit towards the minute behavior of the data, and hence, a better generalization cannot be achieved. This results in a high computational cost and decreased accuracy. Therefore, only two of the four unbalanced case datasets were used so that the computational advantage could be achieved without compromising the accuracy.

By choosing the core set from huge datasets, it is possible to considerably increase the computational effectiveness of ML models. Even if we omit them from training, the model’s evaluation of D_3 and D_4 as unbalanced is obvious due to their high data distribution (as shown in Table 2 and Figure 3). On the contrary, D_1 and D_2 have low standard deviations, similar to the balanced dataset D_0 . In order for the model to distinguish the unique feature qualities of D_1 and D_2 from those of D_0 , it is crucial to choose these datasets for training. This will help the model attain the traits of small local features of data and increase the sensitivity and accuracy of the model. This will also reduce the training time for Analysis-1 and better suit industrial needs when training data are small.

Table 2. Mean and standard deviation of the training datasets. Distributions of D0, D1, and D2 do not vary much as compared to D3 and D4.

Training Dataset	Minimum	Maximum	Mean	Std. Deviation
D0	−0.10675	0.101037	0.000664	0.00838
D1	−0.09483	0.084269	0.000715	0.007499
D2	−0.11995	0.118957	0.000571	0.010673
D3	−0.1188	0.127205	0.000689	0.010989
D4	−0.12719	0.125183	0.000684	0.013755

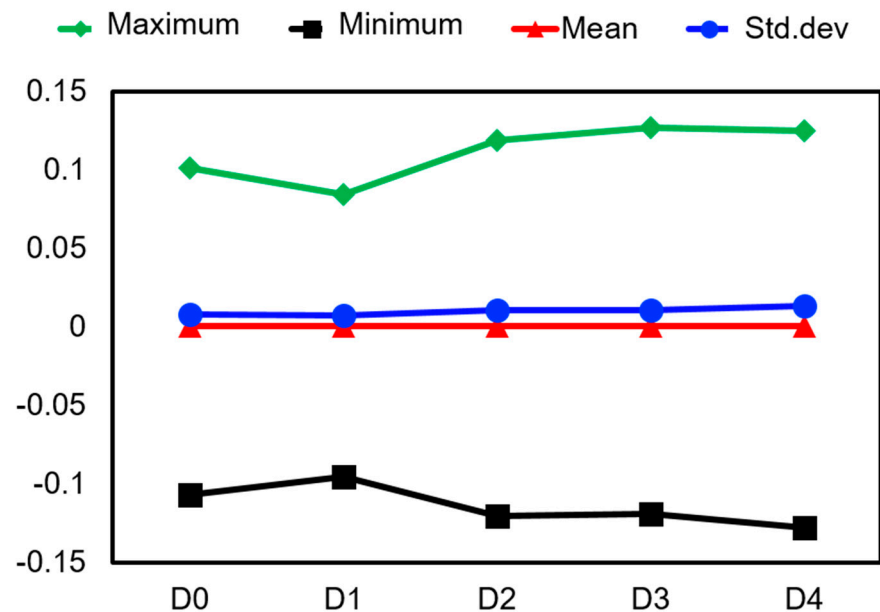


Figure 3. Statistical distribution of the training datasets.

There is always some degree of noise in the signals measured from a real-life mechanical system. Figure 4c shows how the random noise effect is reduced by passing the vibration data through a Gaussian moving average filter. A two-sided moving-average Gaussian filter can be written as follows:

$$y(i) = \frac{1}{N} \sum_{j=-(N-1)/2}^{+(N-1)/2} w(i-j) \times x(i-j) \quad (3)$$

where $y(i)$ is the smoothed value for the i th data point, N is the window size, and $w(i-j)$ are the weights associated to each data point. The filter reduces the noise effect and unwanted data edges in the time domain. A sampling frequency (F_s) of 4096 samples per second is used in the data acquisition setup. For simplicity, the filtered raw data are broken into samples of size F_s . Each sample is called a window. Each window has 4096 data points. Data from vibration sensor 1, as well as samples of raw and filtered signals, are presented in Figure 4.

3.3. Feature Extraction

The filtered data were examined for important features. The most important characteristics of vibration signals are the frequency components. By frequency analysis, it can be seen that there are certain frequencies present in both balanced and unbalanced vibration data that correspond to the operational parameters such as rotation frequency and the number of rotating components such as bearings. Furthermore, in the “Unbalanced case”, additional frequencies are present as well, which are the result of unbalance present in

the rotation system. Fourier Transform is an important tool to analyze the characteristic features of a signal.

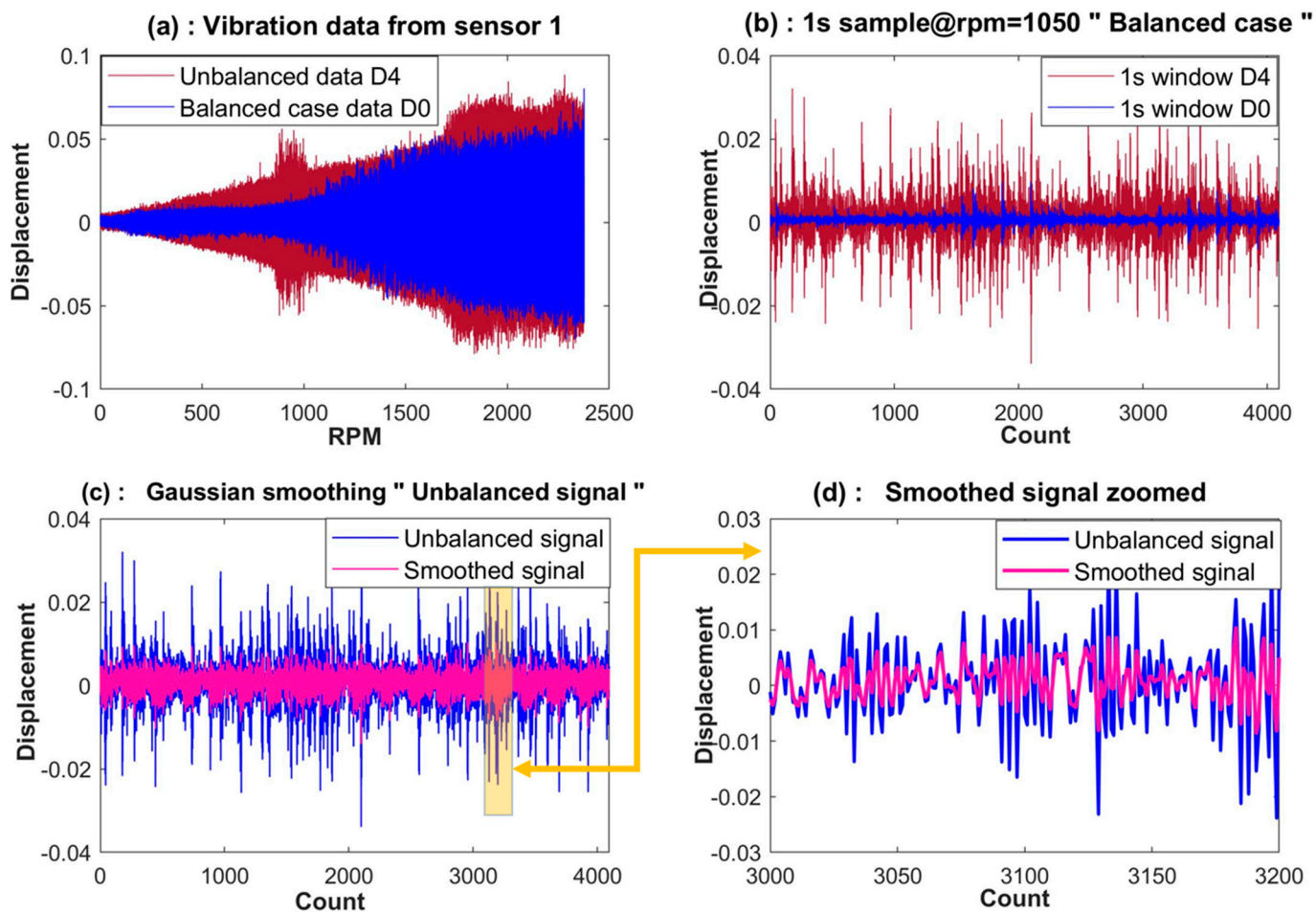


Figure 4. Signal preconditioning: (a) balanced dataset D0 and unbalanced dataset D4; (b) window sample of D0 and D4; (c) smoothing for filtering out noise; (d) a portion of the smoothing process zoomed in.

3.3.1. Fast Fourier Transform

Fast Fourier Transform (FFT) resolves a signal into its frequency components. FFT gives us acting frequencies hidden in the signal. The unbalance in the machine, due to mass asymmetry, adds extra frequency components and amplitude spikes to the vibration signal. The set of active frequencies and their amplitudes deduced from the FFT signal can be selected as a promising feature for ML algorithms. Figure 5a (blue) and Figure 5b (pink) illustrate sample FFT for balanced and unbalanced case signals, respectively. It is evident from Figure 5b that in the unbalanced case, there are more frequency components and high amplitudes as compared to the balanced case in Figure 5a. Regression-based ML algorithms work well with FFT; however, the efficiency degrades when data complexity and the number of classes increase, as is the case for Analysis-2 (5 classes). For such cases, along with deeper networks, a comprehensive set of features or signal characteristics like STFT can perform better.

3.3.2. Short-Time Fourier Transform

Short-Time Fourier Transform (STFT), which is an extension of FFT, allows us to view the frequency characteristics of a signal as a function of time. STFT captures frequency characteristics as a function of time by using sliding windows in time. As a result, a spectrum is formed at discrete time intervals. STFT has better temporal and frequency localization

properties compared with Fourier Transform. It is used to generate representations that capture both the local time and frequency content in the signal. When the rotor along with the added mass spins at high speeds, sharp energy gusts inevitably happen in the vibration data. Bright notches, as shown in Figure 6, that change in the time domain are an indication of this effect in the STFT histogram.

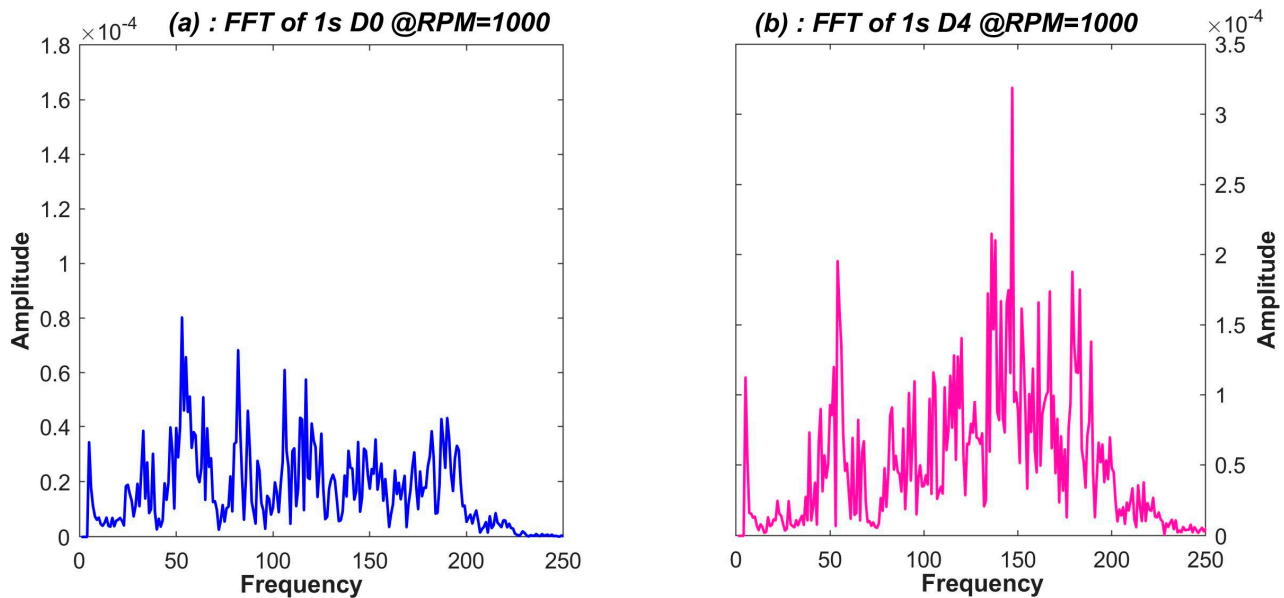


Figure 5. Fast Fourier Transform showing various acting frequencies in (a) balanced case, shown in blue, (b) unbalanced case, shown in pink.

Figure 6a is an example of how STFT spectral images were created from signals with a sampling rate of 4096 data points per second. The STFT images are transformed into grayscale 2D images with half the resolution (438×328) of the original signal because the STFT images' high resolution (3D RGB channel; 875×656 for one sample) consumes a significant amount of memory and is burdensome for academic-scale machines with memory constraints. This makes the computation process faster. Figure 6a,b display the 3D STFT images for the balanced and unbalanced examples, respectively, whereas Figure 6c,d displays the corresponding grayscale images.

3.4. Classification Models

In this work, we focus on the modified ResNet model for our data. For comparison purposes, 5 state-of-the-art classification algorithms are also employed: Artificial Neural Network, Convolution Neural Network, Random Forest, Support Vector Machine, and Xtreme Gradient Boost. These classification techniques are often called algorithm adaptation techniques.

3.4.1. Artificial Neural Network

Artificial Neural Network (ANN) is a machine learning algorithm that mimics the structure and function of the human brain [46]. ANNs are composed of interconnected node layers, containing an input layer, one or more hidden layers, and an output layer. When a node value reaches a certain threshold, the node is activated, and its value is passed on to the next layer of the network. ANNs have been extensively used by many researchers, including [24,47] and others, for fault detection of rotors and bearings.

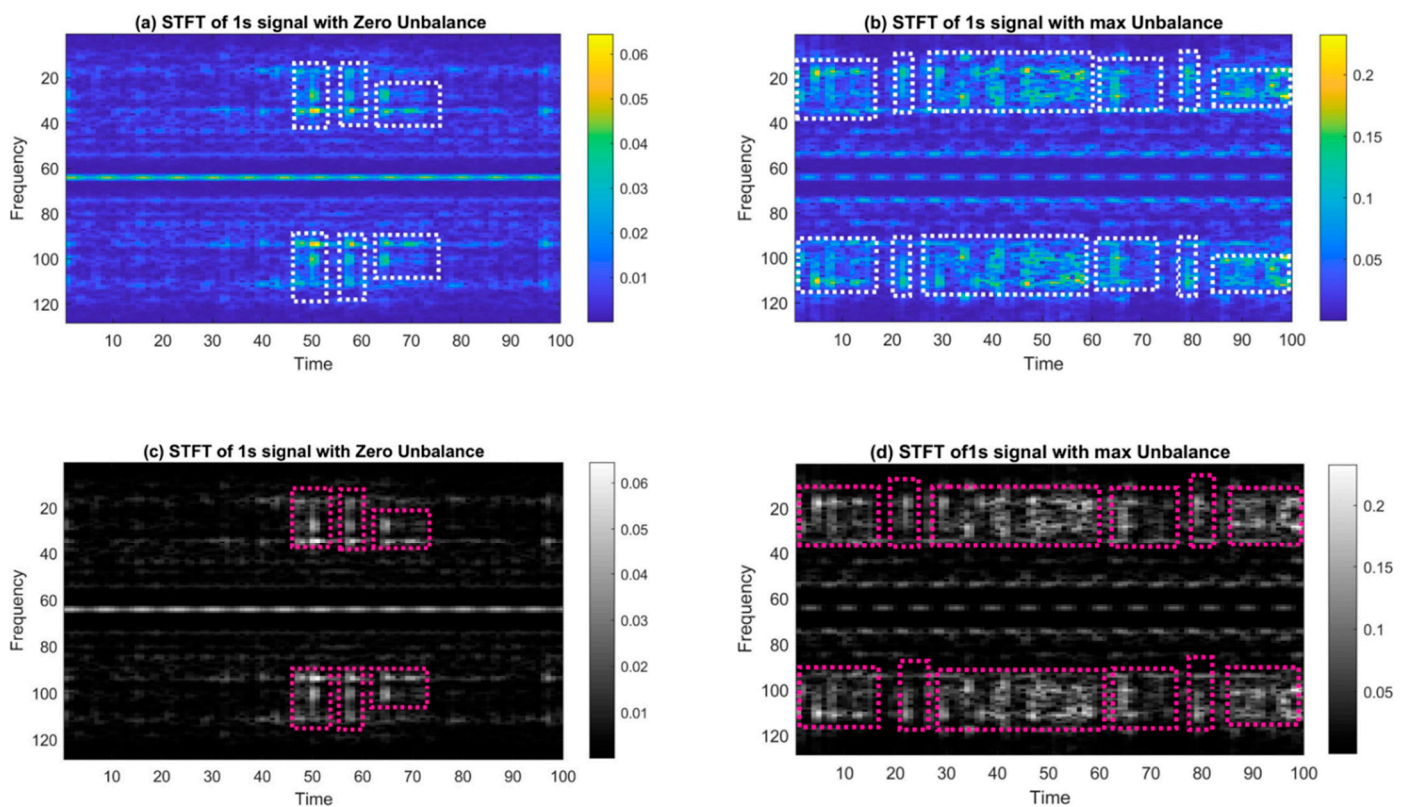


Figure 6. Short-Time Fourier Transform: (a) 3D sample for balanced case, (b) 3D sample for unbalanced case, active frequency components encircled in white triangles (c) grayscale STFT for balanced case, and (d) Grayscale STFT for unbalanced case., active frequency components encircled in pink triangles.

3.4.2. Random Forest

Random Forest (RF) is a supervised learning algorithm that utilizes ensemble learning techniques to create a robust classifier by combining weaker classifiers [48]. This approach involves training models using a bagging method, which is responsible for the improved performance of the algorithm. As the name suggests, the algorithm builds a “forest” of decision trees, each of which serves as a weak classifier. The output of each decision tree is then merged in parallel to form a strong classifier. Random Forest has been used by many authors [29,42,48] for vibration-based damage classification of rotary machines.

3.4.3. Xtreme Gradient Boost

Xtreme Gradient Boosting (XGBoost) is a decision tree technique based on an ensemble learning algorithm that uses a gradient boosting framework. Its most appealing features are its execution speed and model performance. Many writers [49,50] proved that XGBoost outperforms other ensemble-based approaches such as FFT in terms of vibration characteristics.

3.4.4. Convolutional Neural Network

Convolutional Neural Network (CNN or ConvNets) is a deep learning network architecture that can recognize patterns in input images by assigning importance (learnable weights and biases) to various aspects of images to differentiate one object from another. As CNN learns directly from the data, the data preprocessing for ConvNets is substantially reduced as compared to other classification algorithms. CNNs are frequently employed in the fault detection of bearings [47] and rotors [39].

Convolutional Neural Networks are very effective in extracting important image features in computer vision applications. A typical CNN model is composed of convolu-

tional layers (hidden layers), activation layers, pooling layers, fully connected layers, and a classification layer. When an input feature vector x of size $m \times m \times 1$ is passed through a convolutional layer, the convolutional block passes the input through various convolutional filters of size $n \times n \times 1$, which then transforms the input vector into $n \times n \times 1$ convolved feature vectors. The CNN kernels (filters) aim to reduce the number of features present in a dataset by creating a new subset of features that summarizes the overall original set of features. When an input vector x passes through a convolution block, the convolution operations $F(x)$ are applied to it. The output $H(x)$, also shown in Figure 7 (left), can be represented as follows:

$$y = H(x)_{CNN} = F(x) \quad (4)$$

where $H(x)$ is the desired mapping of the stacked convolutional blocks. It can be deduced from the above equation that CNN output is dependent upon the selection of kernels used for the convolution process in the layers. The output mapping $F(x)$ is passed through a ReLU activation function to classify inputs appropriately, producing a probability from 0 to 1.

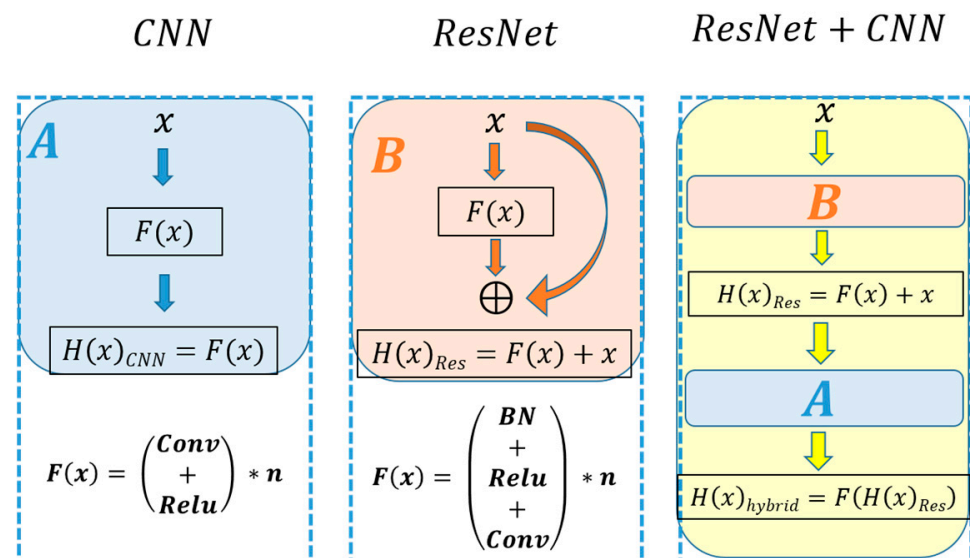


Figure 7. Network architecture comparison: (left) conventional CNN, the block function (light blue rectangle) is shown as block (A), (middle) ResNet, the block function (light orange rectangle) is shown as block (B), and (right) proposed hybrid framework, a combination of ResNet and CNN blocks (B and A).

Usually, the presence of more convolutional layers in the CNN model means a better feature vector of the input data (image) can be extracted. However, because of the multiplication of very small gradients, a deep plain neural network also suffers from the vanishing gradient problem or sometimes gradient explosion. The gradient explosion or vanishing problem limits the depth of the network. The ResNet architecture uses a shortcut pass, called a residual connection, to solve the vanishing gradient problem.

3.4.5. ResNet-152

Residual Neural Network (ResNet) is a CNN architecture that makes use of residual connections to overcome the vanishing gradient problem that is typical in deep neural networks [51]. ResNet-152 is an extension of the original ResNet architecture, with a deeper network that has 152 layers. It has been pre-trained on the large-scale ImageNet dataset, which contains over a million images and thousands of categories. This pre-training allows ResNet-152 to be fine-tuned for other image recognition tasks with relatively few training data.

In a residual connection, shown in the middle diagram of Figure 7, the output $H(x)_{Res}$ is the sum of the CNN's output $F(x)$ and the identity mapping of input x . The output y of a building block of the residual network is as follows:

$$y = H(x)_{Res} = F(x) + x \quad (5)$$

The formulation in Equation (5) can be realized by feedforward neural networks with "shortcut connections", as shown in Figure 7. Shortcut connections in ResNet serve the purpose of preserving input by skipping one or more layers in the residual block. For normalizing the shifted mean and covariance of the feature maps, batch normalization (BN) is applied before the activation function. Research has shown that among the various deeper versions of ResNet-50/101/152/1202 layers, the 152-layer ResNet offers both depth and high accuracy for better classification. Hence, this research adopted ResNet-152 architecture as the base model of choice.

In the first stage of our investigation for the model selection, the above-mentioned networks were trained with numerical FFT feature data and image-based STFT data. The two best-performing networks were picked to continue with the development of a robust framework. Figure 8 shows that CNN and ResNet-152 (trained with STFT) outperform the other supervised learning models in terms of accuracy. This highlights the fact that STFT is a desirable feature for our classification problems of Analysis-1 and Analysis-2.

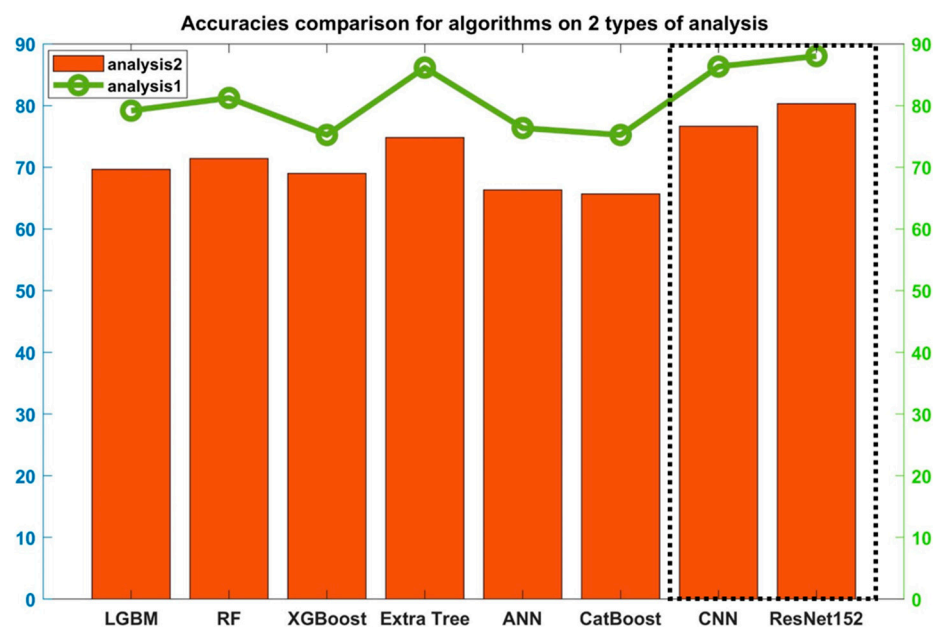


Figure 8. Test accuracies of ResNet CNN and other algorithms on the test data.

Almost similar to Extra Tree, CatBoost, and RF in terms of accuracy, it can be seen that CNN and ResNet-152 outperform all the algorithms in Figure 8. Both CNN and ResNet have their unique advantages, and it can be hypothesized that a hybrid model with the residual properties of ResNet and the classification accuracy of CNN will perform better than ResNet or CNN alone. A hybrid algorithm with base model ResNet (represented as B in Figure 7) and classification model CNN (represented as A in Figure 7) will have the advantage of depth (of ResNet) for better feature extraction and the advantage of promising classifiers (of CNN). In the following section, a hybrid version of ResNet-152 and CNN for this purpose is described.

4. Proposed Framework

We aimed to develop a neural network that is deep enough to successfully extract key features from complex vibration data in both binary and multiclass scenarios. The

problem with deep networks is that they often fall victim to overfitting due to the vanishing gradient problem. ResNet is one of the promising deep neural networks that, by using skip connections, can successfully overcome the vanishing gradient problem. ResNet-152 was chosen as the base model of our architecture because of its depth and ability to perform better on complicated multiclass data than other versions of ResNet. This enables the framework to use the feature extraction capability of very deep neural networks without encountering the vanishing gradient problem. ResNet's skip connections allow information to bypass one or more layers in the network, which can help reduce the vanishing gradient problem and improve gradient flow during training. However, if too many skip connections are employed, the network may struggle to learn increasingly complicated data, and the model may suffer from difficulties such as exploding gradients or overfitting, as evident in Figure 11a. The addition of plain neural network layers will eliminate the problem of exploding gradients while also assisting in the extraction of complex data features. The KERAS library was employed to determine the optimum number of plain neural network layers. Our optimized model has 3 neural layers on top of the 152 residual layers of ResNet. Output values from neural network layers are passed onto the flattened layer, which serves as a horizontal simplified representation of the neuron values. The flattened layer is followed by the fully connected or dense layer, which processes all the information and returns only a few values to determine feature-related values in the image. This output is then condensed in the next step according to the number of output classes and passed to the sigmoid layer for the binary class problem and to the softmax layer for the five-class problem.

Figure 9 depicts the proposed deep learning framework. Using an image generator, the training data comprising 3D STFT images are resized to $3 \times 224 \times 224$, which is the ResNet input image size. The final layer of ResNet, that is, the fully connected (FC) layer, is removed, and plain convolutional network layers are added.

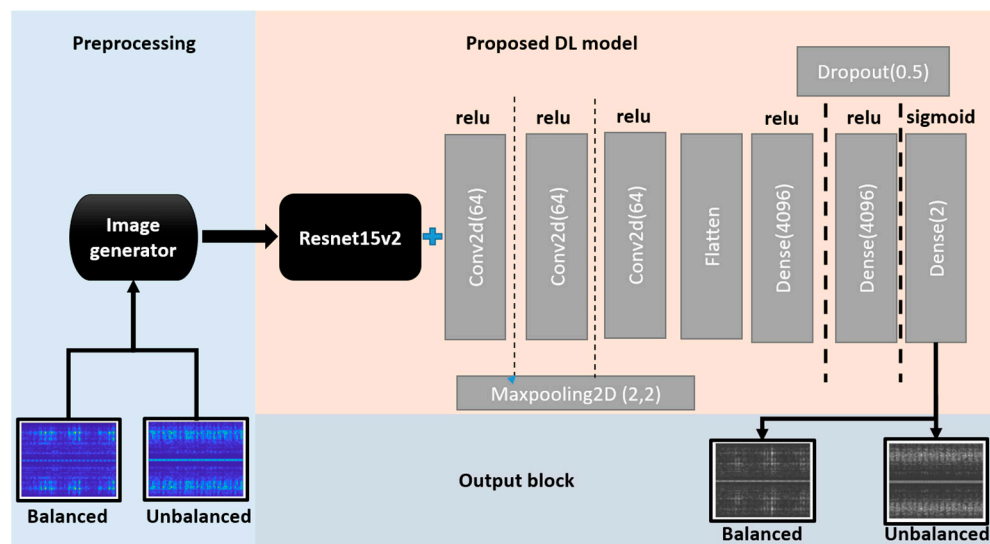


Figure 9. Proposed deep learning framework.

The categorical cross-entropy loss function *equation* employed for the multiclass problem is denoted as follows:

$$L = -\frac{1}{N} \sum_{i=1}^N \sum_{j=1}^M y_{ij} \times \log(p_{ij}) \quad (6)$$

where N is the number of training samples, M is the number of classes, y_{ij} represents the true label of the i th sample for the j th class (1 if the sample belongs to the class, else 0), and

p_{ij} is the predicted probability of the i th sample belonging to the j th class. For the binary classification problem, the number of categories is $M = 2$.

To ensure that forward-propagated signals have non-zero variances and hence avoid becoming trivial, batch normalization is performed after each convolution and before the activation function. The Adam optimizer was used for the optimization of the loss function. A 50% dropout rate was selected to avoid overfitting. The k-fold ($k = 5$) training strategy was used to enable the model to learn and test the characteristics of each dataset by dividing it into five subfolders during training. The key training parameters are given in Table 3.

Table 3. Hyperparameters of the proposed network.

Epoch	Resize	Folds	Patience	Mode	Optimizer	Rate	Beta_1	Beta_2	Activation Function	
									For 2 Class	For 5 Class
20	224,224,3	5	10	max	Adam	0.0001	0.9	0.999	Sigmoid	Softmax

Before the vibration data are sent to the network, they are preprocessed for noise filtering and sampling, followed by extraction of the STFT features. The complete process is shown in Figure 10.

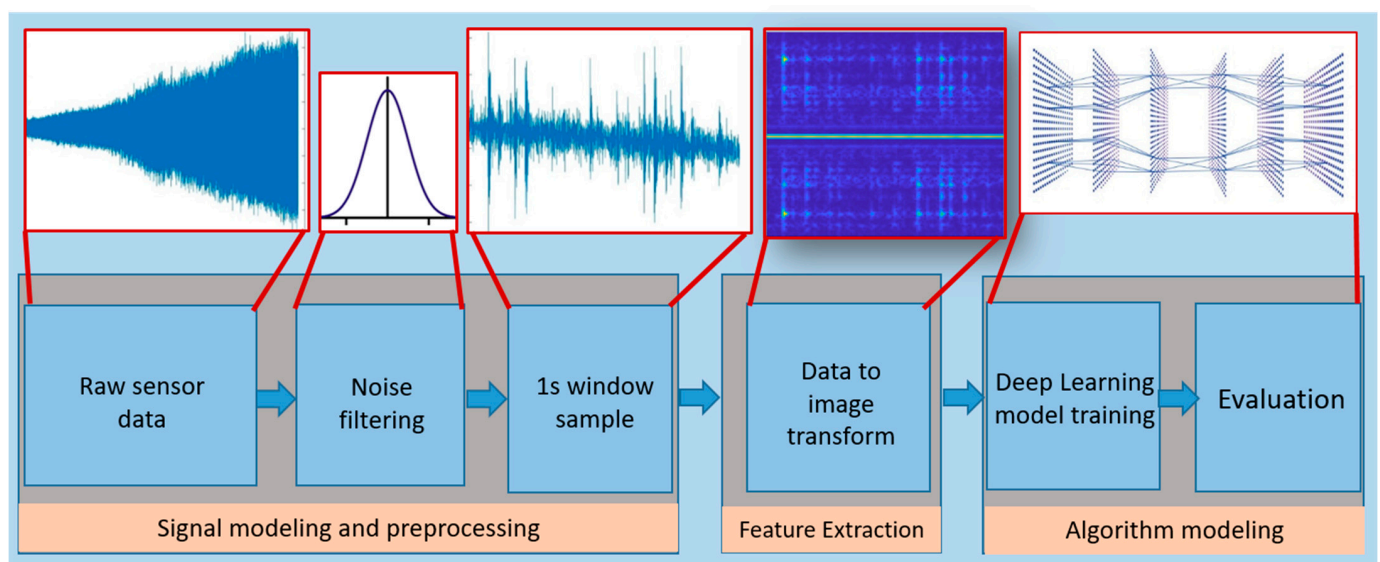


Figure 10. Proposed methodology.

5. Results

The proposed framework, trained with the STFT feature data, shows that our network performed best with three neural layers (ResNet-3N). Other neural network combinations with no added neural layers like ResNet-152 or more neural layers like ResNet-5N (with added layers) or ResNet-7N (with seven neural layers) suffer from the overfitting problem either for Analysis-1 or Analysis-2. Accuracy graphs for both Analysis-1 and Analysis-2 are plotted in Figure 11.

The Analysis-1 results (dotted lines) show that network accuracy decreases as the number of neural network (NN) layers increases beyond three. The difference between test and training accuracy increases with five NN layers, and going further, the network suffers from overfitting in the case of ResNet-7N (Figure 11d). The F1-score in Table 4 also follows a similar trend. This suggests that increasing the network depth beyond three NN layers pushes the model to generalize more toward the training data. This behavior is inevitable when the network is deep enough to memorize all data patterns, including noise

and random fluctuations. The three NN layers were found to be optimal and could indicate a better way to extract high-, medium-, and low-level features.

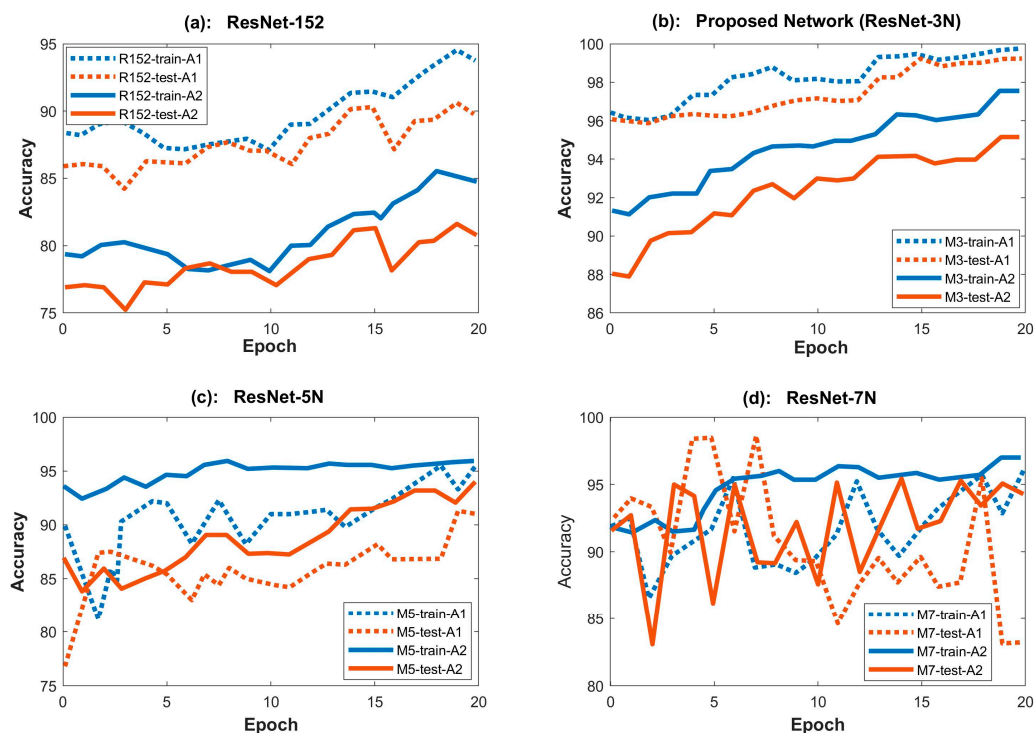


Figure 11. Training accuracies (blue) and test accuracies (orange) for Analysis-1 (solid lines) and Analysis-2 (dotted lines) of (a) ResNet-152, (b) ResNet-3N, (c) ResNet-5N, and (d) ResNet-7N.

Table 4. Test accuracies vs. F1-scores of the tested models. Note that the accuracy and F1-score of the proposed model are the highest.

Model	Accuracy		F1-Score	
	Analysis-1	Analysis-2	Analysis-1	Analysis-2
ResNet-152	88.01	80.31	0.8809	0.8405
ResNet-152-3N	98.71	94.15	0.9759	0.958
ResNet-152-5N	91.59	94.73	0.742	0.8847
ResNet-152-7N	83.2	94.540	0.845	0.910

In the case of Analysis-2 (solid lines), it is observed that as more NN layers are added to the ResNet-152 architecture, the model behavior is considerably stable and has increased accuracy and a higher F1-score. This is because the problem complexity increases with the increase in data classes (five classes). Consequently, the deeper network is better at identifying key features of each class without suffering from overfitting. However, going for a deeper network is useless as the network’s accuracy does not increase further than that of ResNet-3N for Analysis-2 and will overfit for Analysis-1.

The proposed model gives an increased accuracy of 11.22% for Analysis-1 and 14.84% for Analysis-2 when compared with the original ResNet-152. This is a significant increase.

When the accuracy is compared with CNN, there is a significant increase of 12.92% for Analysis-1 and 18.5% for Analysis-2. These results are given in Table 5.

Table 5. The proposed hybrid model has higher accuracy than CNN or ResNet alone. Percent increase is in the last row.

	CNN		ResNet-152		Proposed Model (ResNet-3N)	
	Analysis-1	Analysis-2	Analysis-1	Analysis-2	Analysis-1	Analysis-2
Accuracy	86.31	76.65	88.01	80.31	99.23	95.15
Percent increase (%)	12.92	18.5	11.22	14.84	n/a	n/a

Since our proposed hybrid model is better than individual CNN and ResNet-152 for industrial-scale rotor problems, the model is usable for multiclass prediction where a single model can be used for a varying number of classes.

6. Conclusions

This paper proposes a robust framework for the classification of balanced and unbalanced vibration signals of a rotor. For the given vibration data, the proposed model displays significantly improved performance in terms of accuracy and F1-score in comparison to other state-of-the-art algorithms used in this field. Other close versions of the model, namely, ResNet-5N and ResNet-7N, performed satisfactorily in Analysis-2 but poorly in Analysis-1, whereas the proposed ResNet-3N performed excellently in both Analysis-1 and Analysis-2. It is concluded that for a multiclass rotor unbalance problem where the number of classes may vary from binary to multiple, ResNet-3N is a robust algorithm that will give high accuracy for industrial rotary machine data. Figure 12 implies that ResNet-3N outperforms similar experimental designs as well as parent architectures of CNN and ResNet-152.

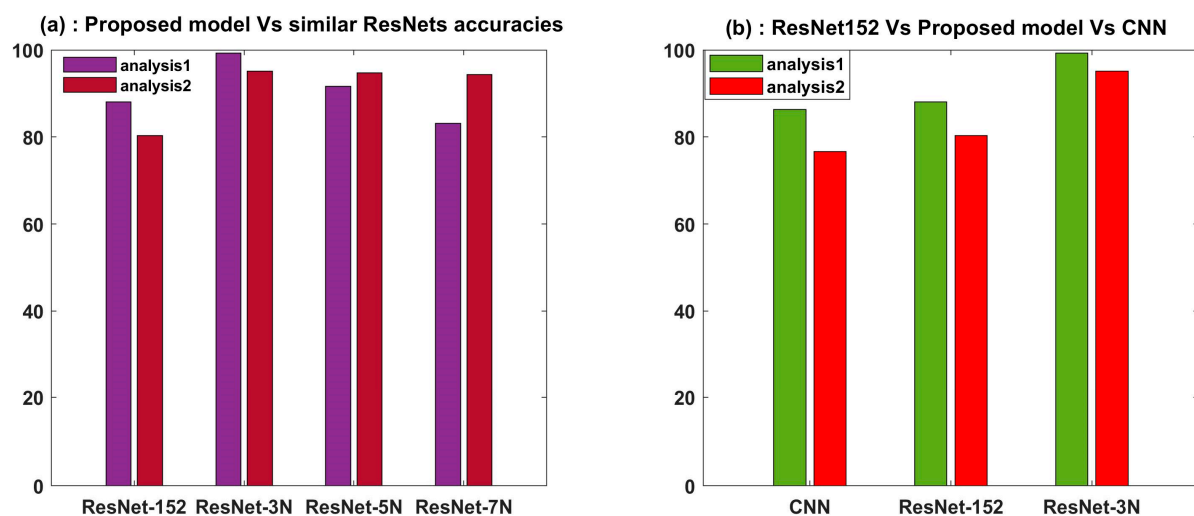


Figure 12. Accuracy comparison of the proposed model architecture (a) with added similar experimental designs and (b) with the base architectures. It is evident that framework ResNet-3N outperforms other similar designs in (a) as well as the parent architectures in (b).

The presented work also demonstrates a strategy to effectively select the core set of training data (from a pool of databases) by evaluating the statistical features of the datasets, which consequently speeds up the training process and increases accuracy. Also, it is deduced that similar to other applications, for rotor unbalance detection and quantification, a rich data representation scheme (such as STFT) performs well in comparison to single-feature FFT data. The STFT histogram (as shown in Figure 6) shows the effect of added mass in the unbalanced rotor in the form of varying frequencies in the time domain. The effect is prominent in transient state data, however is also visible if the rotor is in the steady-state operation. This effect is represented by yellow blotches indicative of energy

jumps of specific frequency components. This feature gives intrinsic properties to the STFT and helps deep neural networks characterize the signal with distinguishable properties, making STFT a more promising feature for our deep network.

Computational costs incurred during preprocessing can be a challenge to the real-time application of the method in industry. As image-processing vibration data for STFT feature maps need computational power (processor) and memory (RAM), a computer needs to be attached to the rotor machine for the real-time processing of the data (as shown in Figure 2). The grayscale STFT images can then be sent to a pre-programmed and pre-trained microchip for fault prediction.

For future work, it is possible to test other types of rotor faults on the proposed algorithm. Also, a combination of other data features, like power spectral density, voltage or current values at different faults, rotational speeds, etc., can be tested to see the effect on accuracy. Furthermore, a user interface strategy to transition between binary and multiclass models can be developed for industrial-scale applications in which, with the touch of a button, a user would be able to select the algorithm type for the number of classes to train.

Author Contributions: Conceptualization, M.W.; methodology, M.W.; software, M.W.; data curation, M.W.; validation, M.W.; formal analysis, M.W.; investigation, M.W.; resources, M.W.; writing—original draft preparation, M.W.; writing—review and editing, K.-Y.O.; visualization, M.W.; supervision, K.-Y.O.; project administration, K.-Y.O. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the Research and Development on Fire Safety Technology for ESS Hydrogen Facilities (Grant No. 20011568, Development of Automatic Extinguishing System for ESS Fire), funded by the National Fire Agency (NFA, Korea), and the Korea Institute of Energy Technology Evaluation and Planning (KETEP) (Grant No. 20213030020260, Development of Fire Detection and Protection System for Wind Turbine), funded by the Ministry of Trade, Industry, and Energy of Korea.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The datasets in this study were downloaded from Fraunhofer research institute and can be found at <https://fordatis.fraunhofer.de/handle/fordatis/151.2> (accessed on 9 August 2023).

Acknowledgments: The authors would like to acknowledge Fraunhofer Institute for open access to the “vibration data of the unbalanced rotor” database. Furthermore, the authors gratefully acknowledge the support of the Higher Education Commission (HEC) of Pakistan.

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

References

1. Malla, C.; Panigrahi, I. Review of Condition Monitoring of Rolling Element Bearing Using Vibration Analysis and Other Techniques. *J. Vib. Eng. Technol.* **2019**, *7*, 407–414. [[CrossRef](#)]
2. Alzghoul, A.; Jarndal, A.; Alsayouf, I.; Bingamil, A.A.; Ali, M.A.; AlBaiti, S. On the Usefulness of Pre-processing Methods in Rotating Machines Faults Classification using Artificial Neural Network. *J. Appl. Comput. Mech.* **2021**, *7*, 254–261. [[CrossRef](#)]
3. Saleem, M.A. Detection of Unbalance in Rotating Machines Using Shaft Deflection Measurement during Its Operation. *IOSR J. Mech. Civ. Eng.* **2012**, *3*, 8–20. [[CrossRef](#)]
4. Kalkat, M.; Yıldırım, Ş.; Uzman, I. Rotor Dynamics Analysis of Rotating Machine Systems Using Artificial Neural Networks. *Int. J. Rotating Mach.* **2003**, *9*, 255–262. [[CrossRef](#)]
5. Tiboni, M.; Remino, C.; Bussola, R.; Amici, C. A Review on Vibration-Based Condition Monitoring of Rotating Machinery. *Appl. Sci.* **2022**, *12*, 972. [[CrossRef](#)]
6. Vasić, M.; Stojanović, B.; Blagojević, M. Failure analysis of idler roller bearings in belt conveyors. *Eng. Fail. Anal.* **2020**, *117*, 104898. [[CrossRef](#)]
7. Reddy, M.C.S.; Sekhar, A.S. Identification of unbalance and looseness in rotor bearing systems using neural networks. In Proceedings of the 15th National Conference on Machines and Mechanisms NaCoMM, Chennai, India, 1–2 December 2011; pp. 69–84.
8. Pareesh Girdhar, S.M. *Practical Machinery Vibration Analysis and Predictive Maintenance*; Elsevier: Amsterdam, The Netherlands, 2004; Volume 4.

9. The Knowledge Engineering Company. *Bearing Damage and Failure Analysis Contents*; SKF: Gothenburg, Sweden, 2017.
10. Lacey, S.J. The role of vibration monitoring in predictive maintenance Part 2: Some illustrative examples. *Maint. Asset Manag.* **2010**, *25*, 38–44.
11. Bogue, R. Sensors for condition monitoring: A review of technologies and applications. *Sens. Rev.* **2013**, *33*, 295–299. [[CrossRef](#)]
12. Zou, Y.; Tong, L.; Steven, G.P. Vibration-based model-dependent damage (delamination) identification and health monitoring for composite structures—A review. *J. Sound Vib.* **2000**, *230*, 357–378. [[CrossRef](#)]
13. Gao, Z.; Cecati, C.; Ding, S.X. A Survey of Fault Diagnosis and Fault-Tolerant Techniques Part I: Fault Diagnosis with Model-Based and Signal-Based Approaches. *IEEE Trans. Ind. Electron.* **2015**, *62*, 3757–3767. [[CrossRef](#)]
14. Sampaio, R.P.C.; Maia, N.M.M.; Silva, J.M.M. Damage detection using the frequency-response-function curvature method. *J. Sound Vib.* **1999**, *226*, 1029–1042. [[CrossRef](#)]
15. Gawronski, W.; Sawicki, J.T. Structural Damage Detection Using Modal Norms. *J. Sound Vib.* **2000**, *229*, 194–198. [[CrossRef](#)]
16. Chow, M.Y.; Sharpe, R.N.; Hung, J.C. On the Application and Design of Artificial Neural Networks for Motor Fault Detection—Part II. *IEEE Trans. Ind. Electron.* **1993**, *40*, 189–196. [[CrossRef](#)]
17. Li, B.; Goddu, G.; Chow, M.Y. Detection of common motor bearing faults using frequency-domain vibration signals and a neural network based approach. In Proceedings of the 1998 American Control Conference. ACC (IEEE Cat. No.98CH36207), Philadelphia, PA, USA, 26 June 1998; Volume 4, pp. 2032–2036. [[CrossRef](#)]
18. Rajakarunakaran, S.; Venkumar, P.; Devaraj, D.; Rao, K.S.P. Artificial neural network approach for fault detection in rotary system. *Appl. Soft Comput. J.* **2008**, *8*, 740–748. [[CrossRef](#)]
19. Chen, C.H.; Shyu, R.J.; Ma, C.K. A new fault diagnosis method of rotating machinery. *Shock Vib.* **2008**, *15*, 585–598. [[CrossRef](#)]
20. Sendhilkumar, S.; Mohanasundaram, N.; Senthilkumar, M.; Sivanandam, S.N. Elman Neural Network for Diagnosis of Unbalance in a Rotor-Bearing System. *Int. J. Mech. Mechatron. Eng.* **2016**, *10*, 613–617.
21. Lu, W.; Liang, B.; Cheng, Y.; Meng, D.; Member, S.; Yang, J. DNN Based Domain Adaptation for Fault Diagnosis. *IEEE Trans. Ind. Electron.* **2017**, *64*, 2296–2305. [[CrossRef](#)]
22. Tang, X.; Zhuang, L.; Cai, J.; Li, C. Multi-fault classification based on support vector machine trained by chaos particle swarm optimization. *Knowl.-Based Syst.* **2010**, *23*, 486–490. [[CrossRef](#)]
23. Bordoloi, D.J.; Tiwari, R. Support vector machine based optimization of multi-fault classification of gears with evolutionary algorithms from time-frequency vibration data. *Measurement* **2014**, *55*, 1–14. [[CrossRef](#)]
24. Attaran, B.; Ghanbarzadeh, A. Bearing fault detection based on maximum likelihood estimation and optimized ANN using the bees algorithm. *J. Appl. Comput. Mech.* **2015**, *1*, 35–43. [[CrossRef](#)]
25. Fatima, S.; Guduri, B.; Mohanty, A.R.; Naikan, V.N.A. Transducer invariant multi-class fault classification in a rotor-bearing system using support vector machines. *Measurement* **2014**, *58*, 363–374. [[CrossRef](#)]
26. Wu, S.D.; Wu, P.H.; Wu, C.W.; Ding, J.J.; Wang, C.C. Bearing fault diagnosis based on multiscale permutation entropy and support vector machine. *Entropy* **2012**, *14*, 1343–1356. [[CrossRef](#)]
27. Shen, C.; Wang, D.; Kong, F.; Tse, P.W. Fault diagnosis of rotating machinery based on the statistical parameters of wavelet packet paving and a generic support vector regressive classifier. *Measurement* **2013**, *46*, 1551–1564. [[CrossRef](#)]
28. Gaye, B.; Zhang, D.; Wulamu, A. Improvement of Support Vector Machine Algorithm in Big Data Background. *Math. Probl. Eng.* **2021**, *2021*, 5594899. [[CrossRef](#)]
29. Rahman, M.M.; Uddin, M.N. Online Unbalanced Rotor Fault Detection of an IM Drive Based on Both Time and Frequency Domain Analyses. *IEEE Trans. Ind. Appl.* **2017**, *53*, 4087–4096. [[CrossRef](#)]
30. Brandao, I.M.; da Costa, C. Fault Diagnosis of Rotary Machines Using Machine Learning. *Eletrôn. Potência* **2022**, *27*, 236–243. [[CrossRef](#)]
31. Wang, L.H.; Zhao, X.P.; Wu, J.X.; Xie, Y.Y.; Zhang, Y.H. Motor Fault Diagnosis Based on Short-time Fourier Transform and Convolutional Neural Network. *Chin. J. Mech. Eng. Engl. Ed.* **2017**, *30*, 1357–1368. [[CrossRef](#)]
32. Xu, Y.; Yan, X.; Feng, K.; Zhang, Y.; Zhao, X.; Sun, B.; Liu, Z. Global contextual multiscale fusion networks for machine health state identification under noisy and imbalanced conditions. *Reliab. Eng. Syst. Saf.* **2023**, *231*, 108972. [[CrossRef](#)]
33. Zhao, X.; Yao, J.; Deng, W.; Jia, M.; Liu, Z. Normalized Conditional Variational Auto-Encoder with adaptive Focal loss for imbalanced fault diagnosis of Bearing-Rotor system. *Mech. Syst. Signal Process.* **2022**, *170*, 108826. [[CrossRef](#)]
34. Sunal, C.E.; Dyo, V.; Velisavljevic, V. Review of Machine Learning Based Fault Detection for Centrifugal Pump Induction Motors. *IEEE Access* **2022**, *10*, 71344–71355. [[CrossRef](#)]
35. Bejani, M.M.; Ghatee, M. Regularized Deep Networks in Intelligent Transportation Systems: A Taxonomy and a Case Study. *arXiv* **2019**, arXiv:1911.03010.
36. Pasupa, K.; Sunhem, W. A comparison between shallow and deep architecture classifiers on small dataset. In Proceedings of the 2016 8th International Conference on Information Technology and Electrical Engineering (ICITEE), Yogyakarta, Indonesia, 5–6 October 2016. [[CrossRef](#)]
37. Hassan, O.E.; Amer, M.; Abdelsalam, A.K.; Williams, B.W. Induction motor broken rotor bar fault detection techniques based on fault signature analysis—A review. *IET Electr. Power Appl.* **2018**, *12*, 895–907. [[CrossRef](#)]
38. Valtierra-Rodriguez, M.; Rivera-Guillen, J.R.; Basurto-Hurtado, J.A.; Jesus De-Santiago-Perez, J.; Granados-Lieberman, D.; Amezcua-Sanchez, J.P. Convolutional neural network and motor current signature analysis during the transient state for detection of broken rotor bars in induction motors. *Sensors* **2020**, *20*, 3721. [[CrossRef](#)]

39. Benkedjough, T.; Zerhouni, N.; Rechak, S. Deep Learning for Fault Diagnosis based on short-time Fourier transform. In Proceedings of the 2018 International Conference on Smart Communications in Network Technologies (SaCoNeT), El Oued, Algeria, 27–31 October 2018; pp. 288–293. [\[CrossRef\]](#)
40. He, M.; He, D.; Bechhoefer, E. Using deep learning based approaches for bearing fault diagnosis with AE sensors. In Proceedings of the Annual Conference of the PHM Society, Colorado, CO, USA, 3–6 October 2016; pp. 282–291.
41. Sapena-Baño, A.; Pineda-Sanchez, M.; Puche-Panadero, R.; Martinez-Roman, J.; Kanović, Ž. Low-Cost Diagnosis of Rotor Asymmetries in Induction Machines Working at a Very Low Slip Using the Reduced Envelope of the Stator Current. *IEEE Trans. Energy Convers.* **2015**, *30*, 1409–1419. [\[CrossRef\]](#)
42. Mey, O.; Neudeck, W.; Schneider, A.; Enge-Rosenblatt, O. Machine Learning-Based Unbalance Detection of a Rotating Shaft Using Vibration Data. *IEEE Int. Conf. Emerg. Technol. Fact. Autom. ETFA* **2020**, *1*, 1610–1617. [\[CrossRef\]](#)
43. Shortle, J.F.; Mendel, M.B. Predicting dynamic imbalance in rotors. *Probabilistic Eng. Mech.* **1996**, *11*, 31–35. [\[CrossRef\]](#)
44. MacCamhaoil, M. Static and Dynamic Balancing of Rigid Rotors. p. 2. Naerum, Denmark, 1989. Available online: <https://www.bksv.com/media/doc/bo0276.pdf> (accessed on 9 August 2023).
45. ANTRIEBE/DRIVES K AT A L O G/C AT A L O G. Edition 25, WEG, Germany. Available online: <https://pdf.directindustry.com/pdf/weg-antriebe/antriebe-drives/12376-993229.html> (accessed on 9 August 2023).
46. Patil, A.A.; Desai, S.S.; Patil, L.N.; Patil, S.A. Adopting Artificial Neural Network for Wear Investigation of Ball Bearing Materials Under Pure Sliding Condition. *Appl. Eng. Lett.* **2022**, *7*, 81–88. [\[CrossRef\]](#)
47. Samanta, B.; Al-Balushi, K.R.; Al-Araimi, S.A. Bearing Fault Detection Using ANN and Genetic Algorithm. *EURASIP J. Appl. Signal Process.* **2004**, *2004*, 785672. [\[CrossRef\]](#)
48. Waleed, M.; Um, T.W.; Kamal, T.; Usman, S.M. Classification of agriculture farm machinery using machine learning and internet of things. *Symmetry* **2021**, *13*, 403. [\[CrossRef\]](#)
49. Liang, Z.; Zhang, L.; Wang, X. A Novel Intelligent Method for Fault Diagnosis of Steam Turbines Based on T-SNE and XGBoost. *Algorithms* **2023**, *16*, 98. [\[CrossRef\]](#)
50. Gu, K.; Wang, J.; Qian, H.; Su, X. Study on Intelligent Diagnosis of Rotor Fault Causes with the PSO-XGBoost Algorithm. *Math. Probl. Eng.* **2021**, *2021*, 9963146. [\[CrossRef\]](#)
51. Lecun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436–444. [\[CrossRef\]](#) [\[PubMed\]](#)

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.