Evaluation of the Use of the Angular Domain and Order Domain in a Bearing Fault Detection Framework using Deep Learning

Racquel Knust Domingues[1], Júlio A. Cordioli[2], Danilo Silva[3], Danilo de Souza Braga[4] and Guilherme Cartagena Miron[5]

[1,2,3] *Federal University of Santa Catarina, Florianópolis, Santa Catarina, 88040-900, Brazil*
racquel.knust@lva.ufsc.br
julio.cordioli@ufsc.br
danilo.silva@ufsc.br

[4,5] *Dynamox SA, Florianópolis, Santa Catarina, 88034-110, Brazil*
danilo@dynamox.net
guilherme.miron@dynamox.net

ABSTRACT

Bearing failures are very common in the industrial environment, requiring effective fault detection methods, which can be categorized into physics-based, knowledge-based and data-driven types. Data-driven methods are efficient in differentiating healthy conditions from faulty conditions by characterizing machine signals, involving stages of data acquisition, feature extraction, and condition determination. Traditionally, feature extraction and condition determination were manual, but advances in artificial intelligence and machine learning, especially deep learning, have automated this process. Although deep learning can automatically learn features from input data, the signal domain can affects model performance. Time and frequency domain representations are widely used in fault detection methodologies using vibration signals. In contrast, angular and order domains are more common in variable operating conditions, but their direct use with deep learning remains rare in the literature. Considering this, this study evaluates a bearing fault detection methodology using vibration signals in different domains (time, frequency, angular, and order) under various rotational conditions. Three distinct approaches were tested to assess the effectiveness of these representations. Results showed the frequency domain had the best performance, and the study concluded that angular and order domains offer no significant advantage over it. Nonetheless, it is recommended to conduct a more in-depth analysis with more diverse datasets, especially those containing early-stage bearing fault signals.

1. INTRODUCTION

Industrial plants are commonly composed of rotating machinery, resulting in the extensive use of bearings, as they are fundamental elements for reducing friction in rotary motion (Lei, 2016). In addition to this extensive use, bearings are often exposed to harsh operating conditions, such as high loads or high temperatures. Consequently, bearing failures account for a significant portion of mechanical failures in the industry, creating the need for methodologies to detect these failures, which can contribute to efficient and optimized maintenance.

Failure detection methodologies can be classified into three categories: model-based methodologies, knowledge-based methodologies, and data-based methodologies (X. Zhang, Zhao, & Lin, 2021). Unlike the first two, which require a deep understanding of the behavior of the element, the data-based methodology offers a more efficient approach by distinguishing between a healthy and a faulty condition through the characterization of a signal measured from the machine.

Data-based methodologies are basically composed of three main stages: data acquisition, extraction of useful features, and determination of the current condition of the element (Mushtaq, Islam, & Sohaib, 2021). In the first stage, an internal parameter of the element must be measured and stored. Machine vibration is a widely used parameter in this application because any change in the element will cause an immediate modification in its dynamic response (Domingues, 2023). In the feature extraction process, the goal is to extract representations and metrics that allow for the distinction between data from different conditions (Guyon & Elisseeff, 2006). From these extracted features, the current condition of the element can then be determined, which will serve as the basis for the proper planning of maintenance actions.

Traditionally, the feature extraction stage and the current condition determination stage were performed manually, utilizing the knowledge of field experts. However, with the advancement of research in artificial intelligence, specifically in machine learning algorithms, this process began to be automated (Hoang & Kang, 2019). Initially, traditional shallow learning algorithms were applied only in the stage of determining the current state of the element. In this context, various bearing fault detection methodologies based on algorithms such as Support Vector Machines (SVM), K-Nearest Neighbor (KNN), and Random Forests have been proposed in the literature, as observed in the works of X. Zhang et al. (2021) and Peng, Bi, Xue, Zhang, and Wan (2022).

However, this process still required the manual extraction of features using signal processing techniques and mathematical transformations specific to each situation. With the advanced development of deep learning methods, the feature extraction process also became automated. Deep learning architectures have the ability to automatically extract the most useful representations from the data for the task to which they are being applied, using the concept of layered learning, which can abstract from simple characteristics of the data to the most complex ones (Goodfellow, Bengio, & Courville, 2016).

Several deep learning architectures have been applied in bearing fault detection methodologies, as pointed out by Hoang and Kang (2019) and S. Zhang, Zhang, Wang, and Habetler (2020). However, convolutional neural networks (CNNs) have gained prominence in the field (Mushtaq et al., 2021). This is because their learning is based on applying filters to input data to recognize relevant features that can contribute to the model's final task. Feature extraction in these architectures is done locally, allowing for hierarchical learning, where complex features are built based on simpler ones. Thus, by using vibration signals along with CNN models, it is possible to automatically extract intrinsic characteristics of these signals that characterize the current condition of the analyzed element.

Although the principle of deep learning is to automatically learn the best features from the input data for each task, some studies have shown that the signal domain can interfere with a model's performance (Rosa, Borges, de S. Braga, & Silva, 2023). The concept of using different representations for the same signal is based on the fact that each domain presents different characteristics of the signal, and therefore, different information can be extracted. Evaluating which domain produces the most appropriate representation for each application can, therefore, be a key factor in achieving good performance in fault detection.

Representations in the time domain and frequency domain are already widely used in fault detection methodologies that utilize vibration signals, both with deep learning and shallow learning. In cases where the machine's operating condition involves rotational variation, angular domain and order domain representations are more common (Lu, Yan, Liu, & Wang, 2019). However, in most cases, such representations are employed contributing to the feature extraction stage. Wang et al. (2024) proposed a methodology to assist in the extraction of features from vibration signals, based on the use of the order domain. The approach demonstrated the advantages of this domain in identifying fault frequencies in bearings, even in signals with rotational variation. X. Huang, Zhang, Xu, and Xu (2024) also employed the order domain, presenting a technique for extracting multiple time-frequency curves, which identifies features associated with bearing faults.

The direct use of signals in these domains in conjunction with deep learning is relatively scarce in the literature, especially when considering bearing fault detection. A exception is the work by He et al. (2023), which addresses this gap by proposing a method that leverages order domain information in conjunction with deep learning for fault diagnosis in rotating machinery. Their approach uses vibration data in order domain, effectively reducing discrepancies in input data distribution across different working conditions.

Therefore, considering the aforementioned, this article aims to evaluate the performance of a deep learning-based fault detection methodology for bearings. It utilizes vibration signals across different domains and considers data from time-varying rotational conditions. The primary focus is to investigate whether representations in the angular domain or order domain provide benefits in detection compared to representations in the time domain or frequency domain, which are widely used in existing literature methodologies.

This paper is structured as follows. In Section 2, the proposed methodology for fault detection is detailed, indicating the classification model used as well as a brief description of the representations employed. Section 3 presents the dataset used in this study, describing the main characteristics of signal acquisition and the operating conditions employed. In Section 4, the methodology for evaluating the performance of fault detection is described, presenting the procedures used at each stage of implementation. Finally, Section 5 presents the obtained results, and Section 6 provides the final conclusions of the study.

2. FAULT DETECTION METHODOLOGY IN BEARINGS

The fault detection methodology used in this work consists of two main stages: data preparation and fault detection. In the first stage, the vibration signal is segmented, transformed into a specific domain, and normalized. In the second stage, these data are fed into a deep learning model based on CNN architecture. The architecture was implemented to perform a classification task, where each input data point is assigned to a class for fault detection purposes.

Following the main objective of this work, we propose the use of signal representations in four different domains: time, frequency, angle, and order. The representation in the time domain refers to the raw signal obtained directly from experimental measurements without applying any transformations. On the other hand, the representation in the frequency domain can be obtained through the Fourier transform (Shin & Hammond, 2008). Representing a signal in the frequency domain allows us to observe which frequencies contain higher energy concentration. This characteristic is very useful because localized faults in bearings often have characteristic frequencies associated with the faults (Randall, 2021).

The angular domain is analogous to the time domain, where the signal is represented "instantaneously" in terms of the angular position of a rotating element. This way, the signals maintain their periodic characteristics, even if there are fluctuations in rotational speed, because periodicity, in this case, is related to angle (Rémond, Antoni, & Randall, 2014). In this representation, the signal must be sampled to ensure an equal number of samples per revolution, meaning the sampling should occur at regular angular increments.

To obtain the representation in the angular domain, this study employed the technique of angular resampling. In this method, a mapping is constructed that relates time to angular position, using a reference signal containing rotation information, which can be generated by an external device such as a shaft encoder or a tachometer. Using this signal, the moments in time when a complete rotation occurs are identified, corresponding to an increment of 2π in angle. With this relationship, it is possible to determine the times when there are regular angular increments and subsequently resample the signal in the angular domain using numerical interpolation (Domingues, 2023).

Just as the frequency domain complements the time domain, the order domain complements the angular domain (Borghesani, Pennacchi, Chatterton, & Ricci, 2014). Therefore, one way to obtain the representation of a signal in the order domain is by applying the Fourier transform to this signal when it is in the angular domain. This representation reveals spectral information concerning the reference rotation and its harmonics, referred to as orders. With this definition, this representation allows visualization of the energy concentration associated with multiples of the reference rotation. It is worth noting that if there are no fluctuations in the rotation speed, the shape of the spectrum in frequency will be the same as the spectrum in order.

In Figure 1, the four representations described earlier for the same signal are presented. The representations in the time, frequency, angular, and order domains are expressed in seconds, Hertz, radians, and nX, respectively. It is important to emphasize that the unit of measure representing the order, nX, corresponds to the number of cycles per revolution, quantifying the proportionality relationship to multiples of the reference rotation. It is observed that the representations in the time and angular domains are quite similar, differing only in the periodicity of the waveform. However, the representations in the frequency and order domains exhibit notable differences. Considering that this signal was measured under a variable rotation condition, the frequency spectrum shows energy spread around the rotation frequency (17 Hz) and its harmonics, in contrast with the order spectrum, which only exhibits peaks at integer orders.

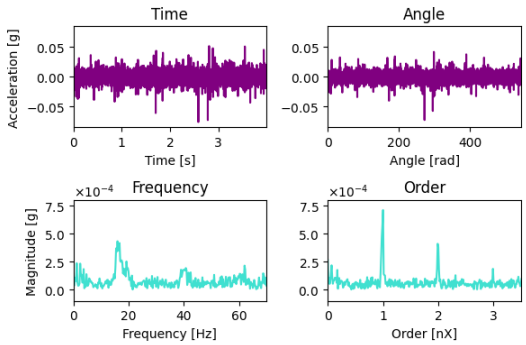

Figure 1. Representations of a Signal in Different Domains

Within the scope of preparing the input data, this work proposes the use of two techniques: signal segmentation and data normalization. Signal segmentation is proposed to increase the dataset and thereby contribute to reducing model overfitting (Géron, 2022). For this purpose, the sliding window segmentation algorithm was employed, where a window moves through the signal, typically with overlap, extracting segments. Details regarding the window size and overlap parameters used will be provided in Section 4.

Data normalization was proposed to reduce instabilities, as attributes with different scales can directly interfere with training convergence (Géron, 2022). In this work, normalization was achieved through standardization, where the data is processed to have zero mean and unit variance. It is important to note that the mean and standard deviation were calculated separately for each dataset used during model training and evaluation (training set, validation set, and test set).

For the fault detection stage, this work employed the WDCNN deep learning model proposed by W. Zhang, Peng, Li, Chen, and Zhang (2017). The model is a Convolutional Neural Network (CNN) designed specifically to process 1D inputs. The first convolutional layer of the model is sufficiently wide to extract patterns of larger dimensions, which may contain information from medium and low frequencies of the signal. Additionally, after this initial layer, the architecture comprises several narrow convolutional layers, which play the typical role of extracting local features from the signal, thereby enhancing the model's ability to extract useful representations.

The classification stage of the model consists of two fully connected layers, where the number of units in the last layer is defined according to the specific task. For binary classification tasks, this layer comprises only one unit followed by a sigmoid function, which is used to obtain the probability of the input belonging to either class. If the classification is multi-class, the number of units in this layer is determined by the number of classes in the problem, and the softmax function is used to obtain probabilities corresponding to each class. A schematic of the proposed architecture is shown in Figure 2. More detailed information regarding its implementation, such as kernel sizes and layer configurations, is provided in Section 4.

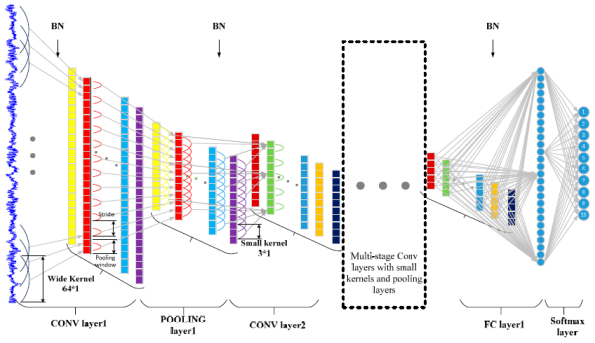

Figure 2. WDCNN Architecture (W. Zhang et al., 2017)

3. EXPERIMENTAL DATA

The dataset used in this work was collected and published by the University of Ottawa, where vibration signals from bearings operating under varying rotational conditions over time are made available (H. Huang & Baddour, 2019). These data were generated using a fault simulation test rig from SpectraQuest, model MFS-PK5M. The test rig consists of a motor driven by an AC Drive, rotating a shaft supported by two bearings fitted with ER16K ball bearings. Figure 3 depicts this test rig. For the construction of this dataset, the bearing closest to the motor was kept consistently healthy, while different conditions of bearings were tested in the bearing furthest from the motor, labeled as the experimental bearing.

The dataset comprises vibration signals and rotational speed signals. The vibration signals were acquired using a uniaxial ICP accelerometer, model 623C01, mounted in the vertical direction of the experimental bearing. Rotational speed signals were obtained through an incremental encoder from EPC, model 775, installed on the main shaft of the test rig. Both signals were captured using NI acquisition boards, model USB-6212, configured to capture signals with a duration of 10 seconds and a sampling frequency of 200 kHz. It's worth noting that the encoder produces a signal containing 1024 pulses per revolution.

In this dataset, four operational conditions related to shaft

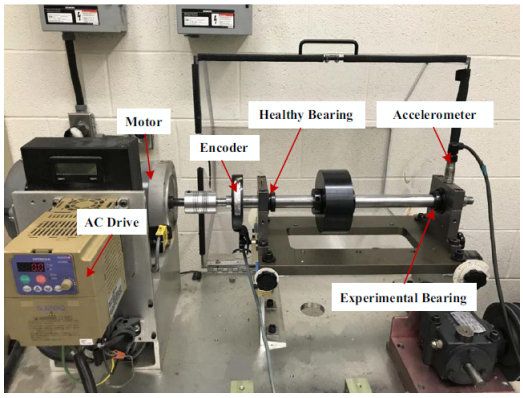

Figure 3. Experimental Set-up (H. Huang & Baddour, 2019)

rotation were evaluated, combined with five different bearing conditions. The rotation conditions are: a) increasing ramp, b) decreasing ramp, c) increasing ramp followed by decreasing ramp, and d) decreasing ramp followed by increasing ramp. It's important to note that "ramp" here refers to a linear variation in rotation speed. The bearing conditions include: i) healthy, ii) inner race fault, iii) outer race fault, iv) rolling element fault, and v) compound fault. For each combination of rotation condition and bearing condition, three trials were conducted to assess repeatability. Therefore, the dataset comprises 60 pairs of signals (vibration + rotational speed). In Figure 4, a schematic of the dataset composition is presented, and in Figure 5, the curves of instantaneous rotations are shown for the all trials of each rotation condition. It is noted that there is a certain randomness in the rotation operating condition for each test.

	Inner Race	Outer Race	Ball Elements	Compound	Healthy
Increasing	Trial 1 Trial 2 Trial 3	Trial 4 Trial 5 Trial 6	Trial 7 Trial 8 Trial 9	Trial 10 Trial 11 Trial 12	Trial 13 Trial 14 Trial 15
Decreasing	Trial 16 Trial 17 Trial 18	Trial 19 Trial 20 Trial 21	Trial 22 Trial 23 Trial 24	Trial 25 Trial 26 Trial 27	Trial 28 Trial 29 Trial 30
Increasing + Decreasing	Trial 31 Trial 32 Trial 33	Trial 34 Trial 35 Trial 36	Trial 37 Trial 38 Trial 39	Trial 40 Trial 41 Trial 42	Trial 43 Trial 44 Trial 45
Decreasing + Increasing	Trial 46 Trial 47 Trial 48	Trial 49 Trial 50 Trial 51	Trial 52 Trial 53 Trial 54	Trial 55 Trial 56 Trial 57	Trial 58 Trial 59 Trial 60

Figure 4. Composition of Dataset

4. EVALUATION METHODOLOGY

To evaluate the performance of the proposed bearing fault detection methodology, two distinct approaches were adopted: a multiclass classification and a binary classification. In the first approach, the methodology was configured to distinguish, from vibration signals, between healthy bearings and

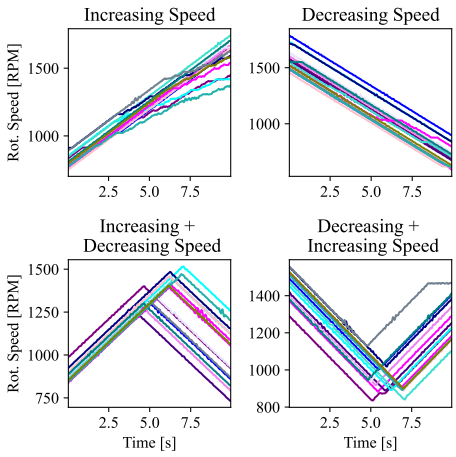

Figure 5. Instantaneous Rotations - Dataset

faulty bearings, as well as to identify the specific types of faults present in the defective bearings: inner race, outer race, rolling elements, or compound faults. In the second approach, the methodology was used to differentiate only between healthy and faulty bearings without specifying the type of fault. The motivation for these approaches lies in evaluating whether the analyzed representations offer advantages both in the generic characterization of bearing defects (binary classification) and in the identification of specific localized bearing defects (multiclass classification). For the implementation of the methodology and the evaluation experiments, a Python framework[1] with GPU processing[2] was used. The preparation of the input data as well as the training of the model were implemented with Pytorch Lightning library[3].

The input data preparation stage was divided into three steps, as illustrated in Figure 6. Within this stage, two distinct approaches were proposed, varying the sampling frequency of the raw signals, the size of the segments used, and the rotation signal. In the first approach, the initial data acquisition configuration was adopted, with a window size set to 400,000 points for segmentation, equivalent to a 2-second segment. This size was chosen due to significant rotation variation within the segment period, contributing to the analysis of differences between the proposed representations. A 95% overlap was established to maximize the number of resulting segments. Additionally, the rotation signal used was the encoder signal obtained without modifications. This first approach was exclusively applied to the multiclass problem.

Subsequently, the methodology was evaluated using a lower acquisition rate. This second approach was proposed to align the data acquisition configuration with common practices in the industry. To achieve this, vibration signals were down-

[1]Version 3.11.9
[2]RTX 3090 GPU
[3]Version 2.2.5

Figure 6. Sample Pre-Processing Flowchart

sampled, reducing the sampling rate to 2 kHz. Prior to this process, a low-pass filter of 1 kHz was applied to the signals to prevent possible aliasing effects. With this new configuration, a window size of 8000 points was defined for segmentation, equivalent to longer segments than those used in the first approach (4 seconds). The overlap was maintained at 95%. Regarding the rotation signal, the signals were modified to resemble a tachometer signal, i.e., a signal with a pulse for each revolution. This second approach was evaluated for both the multiclass and binary classification problems.

Within the segment transformation stage, certain points regarding the procedures used should be highlighted. Firstly, it is noted that the time-domain representation required no processing, and thus the raw signal segment was used directly. For the angular-domain representation, an angular resampling algorithm was implemented. In the scope of this work, the algorithm was developed to ensure that the mapping between time and angle was constructed using the entire extent of the vibration and rotation signals, while resampling was applied to each segment. This approach was chosen to ensure that the same information contained in the time segment was present in the angular segment, albeit with a different waveform due to the angular periodicity introduced by resampling. It is noted that the segments for both representations remained with the same number of points. For representations in the frequency and order domains, the FFT module from the numpy library[4] was used. In these domains, respective unilateral spectra were employed, resulting in segments with half the number of points compared to their counterparts in the time and angular domains (200,000/4,000 points).

The deep learning model proposed for fault detection in this work was implemented using the PyTorch library[5]. The model can be characterized by two stages: the convolutional layers, responsible for extracting useful features, and the fully connected layers, responsible for classification. The convolutional layers were built from cells (ConvBnPool1D) composed of a 1D convolutional layer, a batch normalization layer, a pooling layer, and an activation layer. Five cells were used, and their parameters can be found in Table 1. In the second stage of the model, there is a fully connected layer followed by a normalization layer, followed by a layer composed of five units for multiclass classification, or a single unit for binary classification. The classes for the multiclass task cor-

[4]Version 1.26.4
[5]Version 2.3.1

respond to the five conditions of the bearing: 1) healthy, 2) inner race fault, 3) outer race fault, 4) roller fault, and 5) combined fault. The parameters of these layers can also be found in Table 1. It is noted that the ReLU function was used for activation.

Table 1. Model Parameters

Layer	Type	Parameters
ConvBnPool1D_1	Convolutional + BN + Pooling	Conv1D: in=1, out=16, kernel_size=64, stride=16, padding=24 BN: out=16 MaxPool1D: kernel_size=2, stride=2
ConvBnPool1D_2	Convolutional + BN + Pooling	Conv1D: in=16, out=32, kernel_size=3, stride=1, padding=1 BN: out=32 MaxPool1D: kernel_size=2, stride=2
ConvBnPool1D_3	Convolutional + BN + Pooling	Conv1D: in=32, out=64, kernel_size=3, stride=1, padding=1 BN: out=64 MaxPool1D: kernel_size=2, stride=2
ConvBnPool1D_4	Convolutional + BN + Pooling	Conv1D: in=64, out=64, kernel_size=3, stride=1, padding=1 BN: out=64 MaxPool1D: kernel_size=2, stride=2
ConvBnPool1D_5	Convolutional + BN + Pooling	Conv1D: in=64, out=64, kernel_size=3, stride=1, padding=0 BN: out=64 MaxPool1D: kernel_size=2, stride=2
BatchNorm1D	Batch Normalization	out=100
LazyLinear	Linear	out=100
Linear	Linear	in=100, out=5 or 1

For evaluating the methodology's performance and training the model, this work proposes two data splits to obtain a training set and a test set. The first proposed split involves dividing the data based on their rotation condition, which was used for the multiclass approach. For the binary approach, a split based on bearing conditions was proposed. For both splits, a k-fold cross-validation strategy was employed. In this type of strategy, the data is divided into k subsets, or folds, and the model is trained on all but one subset, which is used for performance evaluation. This process is repeated until all folds have been used for evaluation, and the model's performance is estimated from the average of all evaluations (Géron, 2022).

In the first split, the dataset was divided into four folds based on the rotation condition: A) increasing ramp, B) decreasing ramp, C) increasing followed by decreasing ramp, and D) decreasing followed by increasing ramp. For the second split, the proposed folds are : I) Inner Race Faults, O) Outer Race Faults, B) Roller Faults, and C) Combined Faults. It is worth noting that in this second split, the signals from healthy bearings were divided such that half of the signals, corresponding to rotation conditions A and B, were included in the three training folds, while the other half, corresponding to rotation conditions C and D, were included in the test fold.

There are hyperparameters that need to be optimized to achieve the best model performance. For this work, we chose to optimize two main hyperparameters: batch size and learning rate. To prevent information leakage from the test set to the training set during the optimization process, an additional split is performed within the training set to extract a validation set, which will be used to assess the model's performance. In this work, two fixed splits were proposed for this process, following the earlier proposals. For the rotation condition split, it is proposed to use conditions A and B for training and condition C for validation. For the fault condition split, it is proposed to use conditions I and O, along with a portion of the healthy condition (rotation A), for training, and condition B, along with the other portion of the healthy condition (rotation B), for validation.

The optimization strategy used was a systematic search within a predefined set of values. The values tested for the learning rate were 10^{-1}, 10^{-2}, 10^{-3} and 10^{-4}. For the batch size, the values tested were 16, 32, 64 and 128. In order to reduce processing time, the batch size was optimized only once, using the time domain representation and a fixed learning rate of 10^{-4}. The value obtained in this process was used for all other representations. For each value tested, both for the learning rate and the batch size, five repetitions were performed, and the average performance results were considered as the validation outcome. This process was used to account for the influence of randomness on the results.

In this work, no specific optimization was performed to determine the number of epochs. However, within the optimization of the learning rate, a strategy combining the concept of early stopping with model checkpoints was used to obtain this value. In this strategy, the model's performance is evaluated at each epoch, and its current state, as well as its performance on the validation set, are saved. When performance on this set does not show significant improvement or begins to decline, training is stopped, and the best model is saved. From this model, the final number of epochs is extracted, and this value is associated with the performance of the tested learning rate. Thus, at the end of the learning rate optimization, along with the best value for this hyperparameter, an optimized value for the number of epochs is obtained.

After the hyperparameter optimization process, a training and final model evaluation pipeline was proposed. The signal preparation step follows the same guidelines described earlier and shown in Figure 6. For training and evaluating the model, the optimized hyperparameters obtained from the optimization were used, and 10 repetitions were performed for each of the four folds defined in the proposed divisions. The average performance of the repetitions was associated with each fold, and the average performance across folds was as-

sociated with the model. This process was repeated for each signal representation domain, and the results obtained, which will be shown in Section 5, were used to assess the influence of the domain on the methodology's performance.

Some inherent characteristics of the proposed methodology are highlighted here. The model error, i.e., the cost function used to approximate the model predictions to the true labels, was calculated using the cross-entropy loss function for the multiclass task and binary cross-entropy for the binary task. The Adam optimizer was used for weight updates during the model training process. For evaluating the model performance, both in the hyperparameter optimization process and the final evaluation, the area under the ROC curve (AUC ROC) was used as the metric. It is emphasized that for the multiclass approach, AUC ROC is calculated using a one-vs-rest approach and averaging across all classes. Furthermore, in the final evaluation of the model, the false positive rate obtained by setting a true positive rate at 90% is also calculated.

Considering all the issues addressed in this section, the evaluation of this work can be summarized in three distinct experiments. The first involves the use of the multiclass approach along with the original acquisition setup provided in the dataset. In the second experiment, the multiclass approach was adopted with undersampling applied to the signal. Finally, in the last experiment, the same acquisition setup as the second experiment was used, but with a binary approach. Table 2 provides a summary of the three experiments, including the main characteristics of each.

Table 2. Summary of Evaluation Experiments

Experiment	1	2	3
Classification Task	Multiclass	Multiclass	Binary
Sampling Rate	F_s = 200 kHz	F_s = 2 kHz	F_s = 2 kHz
Segment Size	T_{seg} = 2 s	T_{seg} = 4 s	T_{seg} = 4 s
Data Division	By rotation condition (A/B/C/D)	By rotation condition (A/B/C/D)	By bearing condition (I/O/B/C)

5. RESULTS AND DISCUSSION

In this section, we present the results obtained using the methodology proposed in this work. This section is structured in two parts. Firstly, a brief initial exploratory analysis of the signals that compose the dataset is presented. In this analysis, the behavior of the signals under different rotation conditions and bearing conditions is shown. Subsequently, the results obtained in the fault detection process are presented, along with a brief analysis of them.

5.1. Initial Exploratory Analysis

Firstly, the signals in the time domain for the four faults present in the dataset are presented in Figure 7, along with the signals from the healthy bearing. The rotation condition A (increasing ramp) and the first trial for each condition were used. It is noticeable that in the signals from bearings with faults in the inner race and the compound fault, the characteristic impulsive response typical of bearing faults is clearly evident, with a significant difference in amplitude compared to the healthy signal. For failures in the outer race and rolling elements, although the difference in amplitude between the healthy and defective signals is much smaller, it is still possible to identify sections in the signals that indicate typical defect responses.

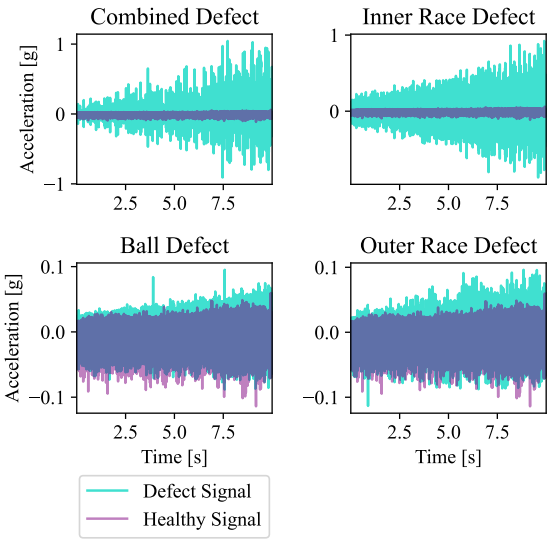

Figure 7. Signals in Time - Bearing Conditions

Subsequently, the signals in the time domain for the four rotation conditions used are presented in Figure 8. In the plots shown, signals from the bearing with faults in the inner race and signals from the healthy bearings were used, all from the first trial. It is noticeable that the amplitude of the faulty signals varies with rotation, but this variation is not as evident in the healthy signals.

Finally, the frequency spectra and order spectra for the four faults are presented in Figure 9. For this analysis, rotation condition A (increasing ramp) and the first trial were also used. As discussed earlier, it is noticeable that all frequency spectra exhibit a spread characteristic, in contrast to the order spectra, which are mainly composed of discrete peaks. Furthermore, it can be observed that the distinction between defective spectra and healthy spectra is much more evident in cases of inner race defect and combined defect, especially in a higher frequency range.

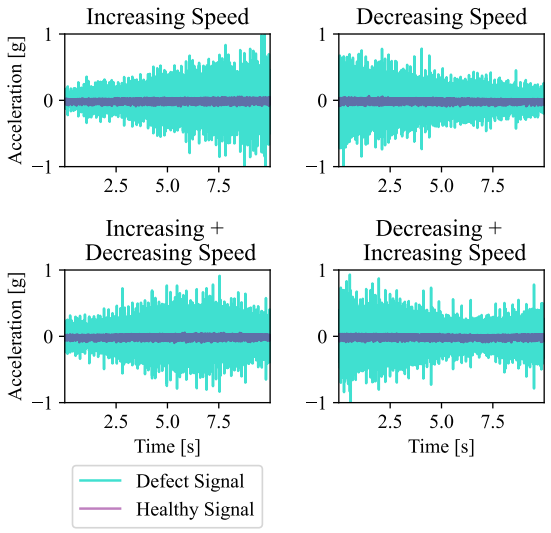

Figure 8. Signals in Time - Rotation Conditions

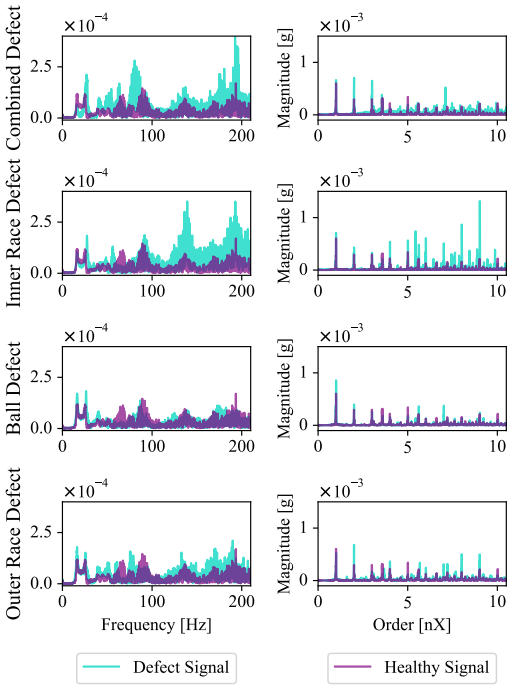

Figure 9. Frequency and Order Spectra - Bearing Conditions

5.2. Detection Performance

The results obtained with the methodology proposed are presented. Three evaluations were performed as outlined earlier: 1) multiclass approach using the original acquisition configuration, 2) multiclass approach using downsampling and 3) binary approach using downsampling. First, the hyperparameters obtained in the optimization processes are presented. As previously reported, batch size optimization was performed at the beginning of each evaluation, using the time domain representation and a learning rate of 10^{-4}, and the values obtained were: 1) 16, 2) 128 and 3) 32.

The learning rates (LR) and the number of epochs (E) associated, obtained in the optimization process for each of the experiments, are shown in Table 3. As previously mentioned, the result shown is the average of five repetitions, using the proposed splits for each approach. It is emphasized that one value was optimized for each proposed representation.

Table 3. Optimized Learning Rate and Number of Epochs

Experiment	Domain	LR	E
1	Time	10^{-2}	4
	Frequency	10^{-4}	2
	Angle	10^{-3}	2
	Order	10^{-3}	3
2	Time	10^{-3}	3
	Frequency	10^{-3}	3
	Angle	10^{-3}	2
	Order	10^{-3}	3
3	Time	10^{-2}	2
	Frequency	10^{-2}	7
	Angle	10^{-4}	2
	Order	10^{-2}	2

Next, the results of the methodology evaluation for the first experiment are presented. It is recalled that, in this experiment, we have the multiclass approach using the original acquisition configuration and that the proposed data split was based on the rotation condition. The evaluation results for each fold in the cross-validation and the average of the folds for each representation used are presented in Table 4. It is emphasized that 10 repetitions were performed for each fold, and the mean AUC ROC values obtained (multiplied by 100) and the standard deviation are presented.

The evaluation results for the second and third experiments are presented in Tables 5 and 6, respectively. In these experiments, we have the multiclass approach (2) and the binary approach (3) using downsampling, meaning the sampling frequency was reduced to 2 kHz and the segment length was increased to cover 4 seconds. It should be noted that, for each of these experiments, the evaluation is carried out in folds, according to each previously proposed division. In other words, the evaluation is done by rotation for the multiclass experiment (2) and by rolling for the binary experiment (3).

It is noted that, in the first experiment, the average performance obtained for the four proposed domains was quite similar, with a slight superiority in the frequency domain. It is also observed that the values obtained for the average AUC ROC were very high, with this value being greater than 94% in all cases. To better understand the nature of these results, the average AUC ROC (multiplied by 100) for each of the

Table 4. Performance Results - Experiment 1

Fold	A/B/C → D		B/C/D → A		C/D/A → B		D/A/B → C		Average	
Domain	AUC ROC	FPR (TPR90%)	AUC ROC	FPR (TPR90%)	AUC ROC	FPR (TPR90%)	AUC ROC	FPR (TPR90%)	AUC ROC	FPR (TPR90%)
Time	97.27 ± 2.22	7.58 ± 13.81	85.23 ± 6.94	19.78 ± 20.17	98.43 ± 1.81	3.62 ± 7.03	98.54 ± 3.16	1.75 ± 7.05	94.86 ± 6.91	8.18 ± 14.95
Frequency	99.94 ± 0.18	0.27 ± 1.82	95.87 ± 3.48	7.2 ± 13.29	100.0 ± 0.00	0.00 ± 0.00	100.0 ± 0.00	0.00 ± 0.00	**98.95 ± 2.49**	**1.87 ± 7.38**
Angle	98.61 ± 2.82	2.71 ± 8.03	84.7 ± 7.64	19.55 ± 21.09	99.86 ± 0.42	0.39 ± 2.32	100.0 ± 0.00	0.00 ± 0.00	95.79 ± 7.61	5.67 ± 13.93
Order	94.51 ± 6.46	11.51 ±18.41	86.11 ± 4.72	21.25 ± 21.75	99.41 ± 1.23	1.08 ± 4.37	97.75 ± 2.94	3.22 ± 11.12	94.45 ± 6.69	9.26 ± 17.37

Table 5. Performance Results - Experiment 2

Fold	A/B/C → D		B/C/D → A		C/D/A → B		D/A/B → C		Average	
Domain	AUC ROC	FPR (TPR90%)	AUC ROC	FPR (TPR90%)	AUC ROC	FPR (TPR90%)	AUC ROC	FPR (TPR90%)	AUC ROC	FPR (TPR90%)
Time	81.08 ± 2.39	29.65 ± 20.64	60.43 ± 12.96	52.07 ± 29.48	80.16 ± 4.25	28.20 ± 20.46	82.85 ± 4.02	24.81 ± 21.86	76.13 ± 11.62	33.68 ± 25.76
Frequency	95.83 ± 2.24	9.74 ± 11.61	75.67 ± 4.41	37.13 ± 25.70	96.90 ± 2.36	6.30 ± 10.90	99.08 ± 2.10	1.80 ± 6.65	**91.87 ± 9.87**	**13.74 ± 20.73**
Angle	76.81 ± 9.47	37.29 ± 35.11	59.99 ± 9.48	54.45 ± 29.89	90.39 ± 9.26	30.17 ± 25.91	87.87 ± 5.70	20.84 ± 22.83	76.26 ± 13.37	35.69 ± 31.32
Order	88.13 ± 5.38	21.67 ±23.04	73.68 ± 7.1	38.47 ± 30.59	96.64 ± 4.06	5.98 ± 10.14	93.22 ± 5.11	9.50 ± 13.49	87.92 ± 10.36	18.90 ± 24.48

Table 6. Performance Results - Experiment 3

Fold	I/O/B/H → C/H		O/B/C/H → I/H		C/I/O/H → B/H		I/C/B/H → O/H		Average	
Domain	AUC ROC	FPR (TPR90%)	AUC ROC	FPR (TPR90%)	AUC ROC	FPR (TPR90%)	AUC ROC	FPR (TPR90%)	AUC ROC	FPR (TPR90%)
Time	71.31 ± 20.83	57.38 ± 41.67	36.69 ± 25.57	96.28 ± 10.78	48.54 ± 3.45	99.65 ± 1.05	68.24 ± 19.33	62.21 ± 40.24	56.20 ± 23.91	78.88 ± 35.17
Frequency	91.08 ± 17.61	19.88 ± 39.34	31.98 ± 42.35	72.85 ± 40.96	77.57 ± 19.47	43.9 ± 39.39	92.93 ± 10.74	23.72 ± 38.3	**73.39 ± 35.44**	**40.09 ± 44.74**
Angle	64.00 ± 28.36	71.22 ± 37.20	53.84 ± 33.30	85.12 ± 27.36	53.33 ± 5.39	85.76 ± 5.60	67.77 ± 16.21	67.21 ± 28.36	59.74 ± 24.30	77.33 ± 28.46
Order	95.58 ± 7.45	22.50 ±39.26	34.07 ± 40.59	80.29 ± 37.47	68.15 ± 14.27	60.64 ± 30.64	93.22 ± 1.30	9.94 ± 0.17	72.76 ± 33.04	43.34 ± 42.13

defined classes, both for the representations in the frequency domain and in the order domain, are presented in Table 7. It is noted that, when comparing the two domains, the classifications for the outer race defect and ball element defect classes, along with the healthy condition, showed a greater decline in performance. To understand this behavior, Figure 10 shows the spectra, in frequency and order, for these three conditions, in a segment using the acquisition condition of this experiment under the condition of increasing ramp rotation. It is observed that, although the order analysis is fulfilling its main objective of standardizing the spectrum based on a reference rotation, this behavior may be causing some difficulty for the model to distinguish between classes, as these signals exhibit similar behaviors in this domain.

Table 7. Average AUC ROC for Experiment 1 - Each Class

	Outer Race	Inner Race	Ball	Combined	Healthy
Frequency	98	99	96	99	99
Order	92	99	89	97	93

Regarding the second experiment, it is noted that the average performance obtained for the frequency domain showed the best performance, although with a value very close to that of the order domain. Again, an attempt was made to better understand this prominence of the frequency domain by evaluating the average AUC ROC (multiplied by 100) for each class, which are shown in Table 8, for the frequency and order domains. In this case, it can be noted that, even using the frequency, the downsampling applied to the signal caused a decline in the model's performance in classifying all classes. Compared to the order domain, it is noted that, once again, the outer race defect and ball element defect classes were the most affected; however, in this experiment, the healthy condition remained similar. Seeking a better understanding of the results, the spectra in frequency and order, using the proposed acquisition configuration for this experiment, of the healthy bearing condition, with an outer race defect and ball element defect, are presented in Figure 11. It is noted that there is a behavior very similar to the previous experiment, where the order analysis standardizes the spectrum, which may be causing the model's difficulty in making distinctions.

The performances for the time and angle domains had lower and quite similar values. An attempt was made to understand the reason for the more pronounced drop in performance of

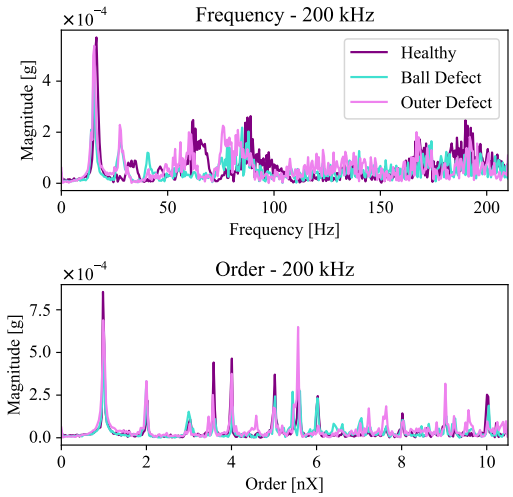

Figure 10. Spectras in Frequency and Order - Experiment 1

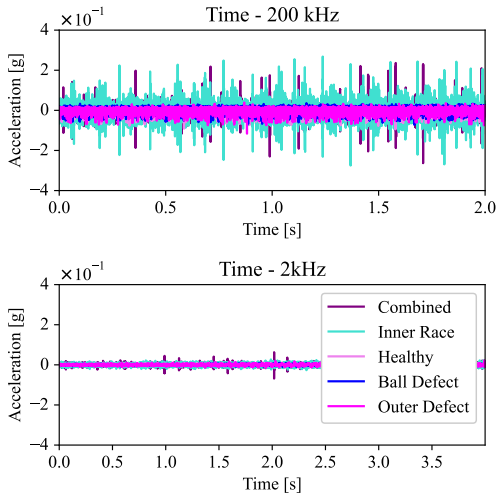

Figure 12. Times Signals - Experiment 1 and 2

the time and angle domains when changing the acquisition configuration. For this, Figure 12 shows the time signals for the five bearing conditions present in the dataset, for the condition of increasing ramp rotation, using the two proposed acquisition configurations. It is noted that, with the downsampling and, consequently, with the filtering process, the time signals, and consequently the angle signals, for the second experiment experienced a loss of energy, which may have resulted in the lower performance.

Table 8. Average AUC ROC for Experiment 2 - Each Class

	Outer Race	Inner Race	Ball	Combined	Healthy
Frequency	89	95	90	95	89
Order	80	95	78	97	88

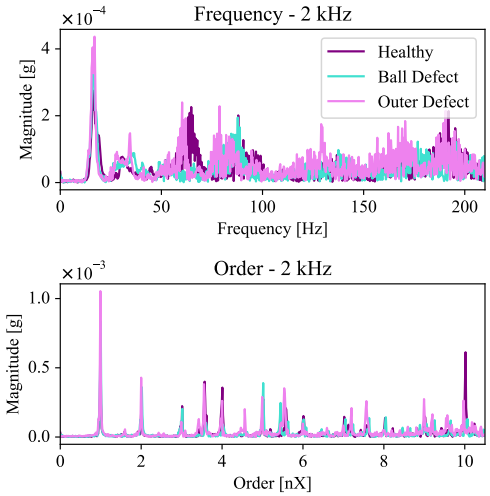

Figure 11. Spectras in Frequency and Order - Experiment 2

Finally, regarding the third experiment, it is noted that the performance trend followed the same pattern as the second experiment, with the frequency and order domains showing similar and superior performances compared to the others. However, in this experiment, there is a high variability associated with the results. It is noted that, in the proposed division of this experiment, the model is trained on certain defect conditions and tested on others, aiming to extract a useful and generalized representation for bearings and avoid possible information leakage between signals from the same bearing across the sets. However, this approach made the model's task more difficult, which may have resulted in these inferior and highly variable results.

5.3. Discussion of Hypotheses

Based on the results obtained from the three experiments conducted in this work, it was not possible to demonstrate a clear advantage of using the angular or order domains compared to the time or frequency domains. Some hypotheses are suggested to justify these results, but further investigations are needed to validate them. Firstly, it is suggested that the angular and order domains may be more advantageous in traditional machine learning methodologies, particularly in the feature extraction step. In such cases, these domains may offer an advantage by extracting signal information in scenarios with rotational variation due to their standardization characteristics. In the case of using a deep learning model, as employed in this work, it is assumed that the model was able to extract useful information for classification directly from the frequency domain, even with rotational variation. Additionally, the use of a sequence of techniques for signal manipulation, as is the case with the angular and order domains, may lead to a loss of information, resulting in a decrease in performance.

Another possible hypothesis is the occurrence of information leakage in the proposed data split for the multiclass approach. In this case, signals from the same bearing are present in both the training and test sets, as the split considers only the rotational condition. With this leakage, there is a possibility that the model is simply memorizing the intrinsic characteristics of the fault signals. Since order analysis normalizes the signals, the model may better memorize the particularities of each class when the data is in the frequency domain. Although there is no leakage in the proposed split for the binary approach, based on the results of the third experiment, a significant amount of associated randomness was observed, suggesting that the model was not able to efficiently perform the proposed task, regardless of the use of order analysis.

6. CONCLUSIONS

We propose the evaluation of a framework for bearing fault detection using vibration signals, considering variable rotational conditions over time. A deep learning model was employed to classify faults across different signal representation domains (time, frequency, angular, and order) to assess whether the angular and order domains could provide additional benefits. To this end, three approaches were tested using a dataset of vibration signals from bearings under five conditions (healthy, inner race fault, outer race fault, rolling element fault, and combined fault). In the first approach, a multiclass task was defined, maintaining the original acquisition setup (200 kHz sampling frequency) and 2-second segments. In the second and third approaches, the signals were resampled to 2 kHz, and the segments were extended to 4 seconds. The second approach involved a multiclass task, while the third approach proposed a binary task focusing only on differentiating between healthy and faulty bearings.

When conducting the evaluation experiments of the three approaches, it was observed that the frequency domain representation achieved the best performance. In the first approach, the obtained values were high for all tested representations, likely due to the signal acquisition configuration, which, with a high sampling rate, captured data with more information, making it easier to distinguish between the classes. In the second approach, there was a drop in values, especially in the time and angle domains, possibly caused by the loss of information resulting from the filtering process. Despite this limitation, the order and, particularly, the frequency domains maintained satisfactory performance. The third approach showed the worst performance, with high variability in the results, regardless of the representation used. Therefore, it is concluded that it is not possible to extract a generalized representation for bearing defects in this application.

The analyses conducted in this work indicate that the use of angular and order domains does not offer significant advantages over the frequency domain. However, these domains may provide benefits in feature extraction when traditional machine learning algorithms are employed, due to their ability to standardize signals under rotational variations. To explore this issue in more detail, future studies should focus on evaluating their impact within different machine learning frameworks. Moreover, it is essential to investigate potential issues, such as information leakage during data splitting, to determine whether the observed results stem from inherent limitations of order analysis application or from external factors, such as memorizing specific fault patterns. Therefore, a more comprehensive analysis, including alternative strategies to mitigate information leakage and the use of other datasets, is recommended to fully understand the implications of using these domains.

ACKNOWLEDGMENT

We thank the Laboratory of Vibrations and Acoustics at UFSC and Dynamox S/A, supported by FEESC, for making this work possible.

REFERENCES

Borghesani, P., Pennacchi, P., Chatterton, S., & Ricci, R. (2014). The velocity synchronous discrete fourier transform for order tracking in the field of rotating machinery. *Mechanical Systems and Signal Processing*, *44*(1-2), 118–133.

Domingues, R. K. (2023). *Avaliação de metodologia de detecção de falhas em mancais de rolamento utilizando análise de ordem e média síncrona no tempo* (Masters dissertation). Federal University of Santa Catarina, Florianópolis.

Géron, A. (2022). *Hands-on machine learning with scikit-learn, keras, and tensorflow*. " O'Reilly Media, Inc.".

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT Press. (`http://www.deeplearningbook.org`)

Guyon, I., & Elisseeff, A. (2006). An introduction to feature extraction. In I. Guyon, M. Nikravesh, S. Gunn, & L. A. Zadeh (Eds.), *Feature extraction: Foundations and applications* (pp. 1–25). Berlin, Heidelberg: Springer Berlin Heidelberg. doi: 10.1007/978-3-540-35488-8_1

He, C., Cao, Y., Yang, Y., Liu, Y., Liu, X., & Cao, Z. (2023). Fault diagnosis of rotating machinery based on the improved multidimensional normalization resnet. *IEEE Transactions on Instrumentation and Measurement*, *72*. doi: 10.1109/TIM.2023.3293554

Hoang, D.-T., & Kang, H.-J. (2019). A survey on deep learning based bearing fault diagnosis. *Neurocomputing*, *335*, 327–335.

Huang, H., & Baddour, N. (2019). *Bearing vibration data under time-varying rotational speed conditions.* Mende-

ley Data. doi: 10.17632/v43hmbwxpm.2

Huang, X., Zhang, J., Xu, Z., & Xu, S. (2024). Fault diagnosis of motor bearings with multiple time-frequency extraction method under variable speed conditions. In *2024 10th international power electronics and motion control conference (ipemc 2024-ecce asia)*. Nanjing, China: IEEE. doi: 10.1109/IPEMC-ECCEAsia60879.2024.10567995

Lei, Y. (2016). *Intelligent fault diagnosis and remaining useful life prediction of rotating machinery*. Butterworth-Heinemann.

Lu, S., Yan, R., Liu, Y., & Wang, Q. (2019). Tacholess speed estimation in order tracking: A review with application to rotating machine fault diagnosis. *IEEE Transactions on Instrumentation and Measurement*, *68*(7), 2315–2332.

Mushtaq, S., Islam, M. M., & Sohaib, M. (2021). Deep learning aided data-driven fault diagnosis of rotatory machine: A comprehensive review. *Energies*, *14*(16), 5150.

Peng, B., Bi, Y., Xue, B., Zhang, M., & Wan, S. (2022). A survey on fault diagnosis of rolling bearings. *Algorithms*, *15*(10), 347.

Randall, R. B. (2021). *Vibration-based condition monitoring: industrial, automotive and aerospace applications*. John Wiley & Sons.

Rémond, D., Antoni, J., & Randall, R. (2014). *Editorial for the special issue on instantaneous angular speed (ias) processing and angular applications* (Vol. 44) (No. 1-2). Elsevier.

Rosa, R. K., Borges, V. K., de S. Braga, D., & Silva, D. (2023). Diagnóstico de falhas em rolamentos usando redes convolucionais: Otimização da representação de sinais e uma nova metodologia de avaliação. In *Xli simpósio brasileiro de telecomunicações e processamento de sinais (sbrt2023)*. doi: 10.14209/sbrt.2023.1570915812

Shin, K., & Hammond, J. (2008). *Fundamentals of signal processing for sound and vibration engineers*. John Wiley & Sons.

Wang, K., Huang, Y., Zhang, B., Luo, H., Yu, X., Chen, D., & Zhang, Z. (2024). Improved synchronous sampling and its application in high-speed railway bearing damage detection. *Machines*, *12*(2), 101. doi: 10.3390/machines12020101

Zhang, S., Zhang, S., Wang, B., & Habetler, T. G. (2020). Deep learning algorithms for bearing fault diagnostics—a comprehensive review. *IEEE Access*, *8*, 29857-29881. doi: 10.1109/ACCESS.2020.2972859

Zhang, W., Peng, G., Li, C., Chen, Y., & 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), 425.

Zhang, X., Zhao, B., & Lin, Y. (2021). Machine learning based bearing fault diagnosis using the case western reserve university data: A review. *Ieee Access*, *9*, 155598–155608.

BIOGRAPHIES

M.Sc Racquel K. Domingues is a researcher and doctoral student in the field of Vibrations and Acoustics in the Mechanical Engineering postgraduate program at the Federal University of Santa Catarina in Brazil, where she also obtained a master's degree. She holds a bachelor's degree in Control and Automation Engineering from the Federal University of Pelotas, also in Brazil. Racquel has worked in various areas, gaining experience in mathematical modeling of vibrational systems, numerical modeling of vibroacoustic systems, and digital signal processing techniques. Currently, her research focus is on predictive maintenance through vibration analysis, using both traditional signal processing techniques and machine learning algorithms.

Prof. Julio A. Cordioli graduated in Mechanical Engineering from Federal University of Santa Catarina (UFSC), Brazil, in 2000 and received his PhD in 2006 also from UFSC, with part of his PhD research developed at the University of Cambridge, UK. He has previously worked for EMBRAER (Brazil) on the design of noise control treatments for aircraft, and for the ESI Group (USA) as an acoustic scientist on the development of vibroacoustic numerical methods. Since 2011, he has been an Associate Professor at the Mechanical Engineering Department, UFSC. His main research interests include numerical methods in vibroacoustics, aeroacoustics, and vibration monitoring.

Prof. Danilo Silva received the B.Sc. degree from the Federal University of Pernambuco (UFPE), Recife, Brazil, in 2002, the M.Sc. degree from the Pontifical Catholic University of Rio de Janeiro (PUC-Rio), Rio de Janeiro, Brazil, in 2005, and the Ph.D. degree from the University of Toronto, Toronto, Canada, in 2009, all in electrical engineering. In 2010, he joined the Department of Electrical and Electronic Engineering, Federal University of Santa Catarina (UFSC), Florianópolis, Brazil, where he is currently an Associate Professor. His current research interests include machine learning, computer vision and signal processing.

Dr. Danilo S. Braga is a distinguished mechanical engineer and vibration specialist, currently serving as a Specialist at Dynamox SA. With a strong application background and a Ph.D. degree in noise and vibration, Danilo possesses extensive expertise in vibration monitoring and analysis, wireless sensor technologies, and signal processing. Additionally, Danilo has played a pivotal role in software development for vibration analysis (Dynamox's platform).

M.Sc Guilherme C. Miron is a Mechanical Engineer with a master's degree in Noise and Vibration, possessing in-depth knowledge in MEMS technology, piezoelectricity, and wireless sensors. His expertise centers on vibration monitoring and signal processing, where he pioneers innovative approaches to machinery fault detection and reliability enhancement. As a key member of Dynamox's Research and Innovation Department, Guilherme continues to drive advancements in industrial machinery monitoring, aiming to integrate theoretical insights with practical applications.