Methodology for the Detection of Contamination and Gradual Outer Race Faults in Bearings by Fusion of Statistical Vibration–Current Features and SVM Classifier

Geovanni Díaz-Saldaña 1,2, Jonathan Cureño-Osornio 1, Israel Zamudio-Ramírez 1,2, Roque A. Osornio-Ríos 1, Larisa Dunai 3, Lilia Sava 4 and Jose A. Antonino-Daviu 2,*

1 Cuerpo Académico Mecatrónica, Facultad de Ingeniería, Universidad Autónoma de Querétaro, Campus San Juan del Río, Av. Río Moctezuma 249, San Juan del Río 76807, Querétaro, Mexico; gdiaz17@alumnos.uaq.mx (G.D.-S.); jcureno08@alumnos.uaq.mx (J.C.-O.); iszara@doctor.upv.es (I.Z.-R.); roosornio@hspdigital.org (R.A.O.-R.)
2 Instituto Tecnológico de la Energía, Universitat Politècnica de València (UPV), Camino de Vera s/n, 46022 Valencia, Spain
3 Department of Graphic Engineering, Universitat Politècnica de València (UPV), Camino de Vera s/n, 46022 Valencia, Spain; ladu@upv.es
4 Faculty of Electronics and Telecommunications, Technical University of Moldova (UTM), MD-2004 Chisinau, Moldova; lilia.sava@fet.utm.md
* Correspondence: joanda@die.upv.es; Tel.: +34-96387-7592

Abstract: Bearings are one of the main components of induction motors, machines widely employed in today’s industries, making their monitoring a primordial task; however, most systems focus on measuring one physical magnitude to detect one kind of fault at a time. This research tackles the combination of two common faults, grease contamination and outer race damage, as lubricant contamination significantly impacts the life of the bearing and the emergence of other defects; as a contribution, this paper proposes a methodology for the diagnosis of this combination of faults based on a proprietary data acquisition system measuring vibration and current signals, from which time domain statistical and fractal features are computed and then fused using LDA for dimensionality reduction, ending with an SVM model for classification, achieving 97.1% accuracy, correctly diagnosing the combination of the contamination with different severities of the outer race damage, improving the classification results achieved when using vibration and current signals individually by 7.8% and 27.2%, respectively.

Keywords: bearing; fault diagnosis; grease contamination; data fusion; support vector machine

1. Introduction

Induction motors are essential elements in industries and are used in a wide range of applications. Despite being robust, these machines are prone to faults in different parts and components, with the bearings being responsible for up to 45% percent of all the faults in electric motors [1]. Bearings are one of the most important parts of electric motors, as they allow the movement of the rotor and the transmission of the torque to the kinematic chain. They are subjected to continuous loads, stresses, impacts, environmental conditions, and inadequate manipulation. All these factors play a part in the various mechanisms that cause bearing failure, making it imperative to identify early signs of deterioration to prevent downtime and extended damage to the machines. For this reason, several researchers have studied the mechanisms causing the bearings to fail, classifying the causes as fatigue, wear, corrosion, deformation, fracture, and incorrect handling and maintenance [2–8]; improper handling and maintenance is one of the most important and preventable mechanisms, as it impacts the lubrication of the bearing and up to 52% of their faults can be due to lubricant contamination [4]. At the same time, one of the biggest challenges for the detection of
these faults is the correct diagnosis, as it involves the accurate measurement of physical variables that provide significant information regarding the faults and the development of methodologies considering the correct manipulation and analysis of the data to discern between conditions involving the presence of different severities of a particular defect as well as the combination of faults in one bearing.

Over the years, several methodologies have been developed to mainly diagnose defects in the inner and outer races of bearings, measuring a variety of physical magnitudes that are analyzed by various processing tools and algorithms to determine the presence of the faults. In this regard, temperature is a novel approach to bearing failure diagnosis that shows promising results; the authors of [9] developed a methodology to diagnose between five different one-severity damaged bearings and a healthy one using infrared thermography and a Convolutional Neural Network (CNN), achieving 99.8% effectiveness in the diagnosis with the noninvasive method; however, they noted the drawbacks of the sensitivity to the installation of the equipment and electromagnetic interference. On the other hand, current signal analysis is a methodology that has proven efficiency in bearing diagnostics; in reference [10], the authors focused on characteristic fault frequency components in a stator current to detect different severities of outer race (OR) damage using the Fast Fourier Transform (FFT), being able to detect the defects through the analysis of the magnitudes across all of the current signal spectrum to relate the increase in magnitude of specific components to the severity of the fault. On the other hand, the authors of [11] employed the Discrete Wavelet Transform (DWT) to diagnose one graduality of inner race (IR) and OR defects by separating the measured current signals into frequency bands and calculating the energy in the bands to relate the results to the damage. Similarly, stray flux analysis has grown in popularity in recent years in the condition monitoring field; the authors of [12] focused on the detection of two levels of damage in the OR of a bearing by the analysis of one stray flux signal using thresholds on the kurtosis values to distinguish between conditions, developing a system for online use on a Field Programmable Gate Array (FPGA) platform.

It must also be noted that vibration signal analysis is the preferred method to study bearing faults; researchers in [13] utilized a CNN to analyze vibration data to detect OR and IR defects with different loads and speeds, achieving 93.97% in their classification efficiency. In a different approach, reference [14] shows the use of statistical features to identify between OR, IR, and ball defects in bearings by using a Random Forest Algorithm (RFA), obtaining results of 99.57% in the classification for the one-severity faults. These methods can detect the presence of faults, and some can differentiate between a couple of faulty states with only one severity; however, most systems require significant amounts of information to discriminate between a more extensive set of faults or severities. In that regard, data fusion from different sensors is a trend in condition monitoring as the measurement of multiple variables allows the information of one sensor or physical magnitude to be complemented with other sources to facilitate the diagnosis, sometimes with the requirement of more robust analysis techniques to manage the considerable amounts of data produced [15,16]; as an example, the authors of [17] implemented a multi-sensor system using an accelerometer and a load cell to diagnose different bearing conditions using time and frequency indices, fusing the decisions of a k-Nearest Neighbors (kNN) classifier, showing an improvement in diagnosis when combining the decisions of the classifier for each sensor. A different approach was published in [18], where the researchers utilized Transfer Learning to generate a complex network capable of the fusion of information from different domains, being capable of identifying OR, IR, and ball defects with 95.97% to 98.31% accuracy. The focus of [19] was the analysis and comparison of features in time, frequency, and time–frequency domains to identify the best option to diagnose OR defects in a ball bearing through the measurement of current and vibration signals, showing better results with the application of the Hilbert–Huang Transform (HHT). Similarly, in reference [20], acoustic emission (AE) and vibration signals are used to identify the presence of an OR defect under various speeds and load levels, showing better results in
fault identification when combining the Root Mean Square (RMS) values of the vibration signals and the amplitude of the AE signals.

The diagnosis of multiple simultaneous faults has also been studied, showing interesting alternatives to the detection of combined defects. In this area, works like [21] implemented a vibration-based system to detect sixteen defect conditions in bearings consisting of the combinations of OR, IR, ball, and cage damages simulated from single fault signals using DWT, Hilbert Transform (HT), statistical features, metaheuristic algorithms and comparing the results of a Feed-Forward Back-Propagation (FFBP) Artificial Neural Network (ANN), a Support Vector Machine (SVM) with a Gaussian kernel, and a kNN classifier, resulting in a maximum outcome of 91.6% effectiveness from the ANN. In reference [22], a wavelet-based moving average control chart is studied to diagnose the combination of OR, IR, and rolling element faults using vibration signals and comparing the results to those obtained with Empirical Mode Decomposition (EMD) and Hilbert envelope spectrum analysis, showing positive results with the method and also revealing the difficulty of identifying the faults if the signals have similar energy dispersion patterns.

As previously established, lubricant contamination is a primary cause of defects in bearings; the first studies to investigate the effect of contaminants were carried out in the 1960s, focusing on the analysis of their concentrations and effects, finding relations for their hardness, sizes, and the impact on bearing life [23]. In reference [24], the authors compared SVM models with a radial basis function kernel fed with statistical features against a CNN structure to diagnose two levels of sand dust contamination on bearings using vibration signals, showing 100% effectiveness in the classification of the healthy, 50, and 100 mg of contamination classes using RMS on the SVM model and 98.33% with the CNN. On the other hand, in [25], a system using vibration, current, AE, and shock pulses is developed to diagnose different levels of Silica and Ferric Oxide contamination using various contaminant particle sizes to determine the best magnitude for the detection, concluding on the effectiveness of the peak amplitude on AE signals to differentiate between the conditions. Similarly, the authors of [26] studied the effect of contamination with steel and corundum particles on the wear and fatigue life of bearings, reporting the use of time and frequency indicators on vibration signals to identify the damage of the particles, concluding on the effect of the hardness of the particles over their size related to the damage produced inside the bearings.

As shown in the literature, bearing failure is a critical phenomenon as it presents in different ways, with the defects frequently being related to lubricant contamination. Most of the methods presented focus on one type of fault at the time, mainly defects on the outer and inner race of the bearings, while few works study the presence of multiple simultaneous defects; in addition, most of the proposals use vibration signals as they are captured directly from the motor, while other sources provide information that could relate to different conditions and require more advanced processing or the fusion of data from various physical magnitudes to complement the information from the measured sources. Considering these facts, this research presents a bearing fault diagnosis method based on the fusion of time domain statistical and fractal features from raw vibration and current signals, capable of detecting the presence of black iron oxide in the lubrication of bearings with OR damage. The features fused using Linear Discriminant Analysis (LDA) allow an improvement in classification with an SVM model, as the fusion creates clusters that maximize the separation between conditions while compacting them, in comparison to the analysis of the individual signals, showing the method’s capability to differentiate between the conditions.

2. Theoretical Background
2.1. Bearing Faults and Lubrication
2.1.1. Bearing Faults

The study of bearing faults has been an important research area for decades due to the significance of these elements in most mechanisms in all industries. Many researchers have
analyzed the causes and mechanisms that originate the different types of faults. The leading causes of bearing failure can be summarized as improper selection or design, improper material selection and quality, common defects in the bearings, and incorrect assembly, use, and maintenance [2–7], with the last being responsible for up to 49% of bearing failures [2]. For the mechanisms of occurrence, the classification considers deformation, fracture, corrosion, and wear caused by fatigue, erosion, cavitation, and stress [2–7].

Bearings can present defects in different locations, and numerous researchers have categorized these faults into OR, IR, rolling element, and cage defects, with all being important due to the vibrations created with each kind of fault, which affect the kinematic chain and impact the state of other elements being capable of producing more significant damage. The most common way to diagnose these defects is by searching for frequency components in vibration signals from the electric motors that have been well reported in the literature, where data from the bearing and the speed of the motor are used to identify characteristic frequencies related to OR, IR, cage, and rolling element faults. Similarly, for the analysis of current signals in the frequency domain, the literature focuses on the search for frequencies that are related to the characteristic frequencies in the vibration spectrum in relation to the supply frequency to identify the presence of faults according to the amplitude of these components [1,3,7]. These frequency components are approximated using Equations (1)–(4) in Table 1 and considering \( N \) as the speed of the shaft in revolutions per minute (RPM), \( n \) as the number of balls in the bearing, \( \alpha \) as the contact angle of the bearing due to load from the radial plane, \( d \) as the diameter of the ball, and \( D \) as the mean diameter of bearing [1,3,7].

### Table 1. Characteristic frequencies for the identification of bearing defects using vibration signals.

<table>
<thead>
<tr>
<th>Characteristic Frequency</th>
<th>Equation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inner race defect frequency</td>
<td>( \frac{n}{2} \left( N - \frac{d}{D} \cos \alpha \right) )</td>
</tr>
<tr>
<td>Outer race defect frequency</td>
<td>( \frac{n}{2} \left( N + \frac{d}{D} \cos \alpha \right) )</td>
</tr>
<tr>
<td>Ball defect frequency</td>
<td>( \frac{D}{2} \left( 1 - \frac{d}{D} \cos \alpha \right) )</td>
</tr>
<tr>
<td>Cage defect frequency</td>
<td>( \frac{1}{2} \left( N - \frac{d}{D} \cos \alpha \right) )</td>
</tr>
</tbody>
</table>

Another approach to the diagnosis of faulty bearings is using statistical techniques, using large datasets of time domain information, finding trends and relations in the data when calculating different statistical indicators, such as RMS or Kurtosis [3,4], and finding thresholds related to the severity of the defects that allow the diagnosis. However, today’s industry requires more advanced techniques as the availability of information from multiple sensors requires novel methodologies to deal with larger amounts of data, making the fusion of the information a promising approach for better diagnostic technology [15,16].

#### 2.1.2. Bearing Grease Lubrication

Grease is a common lubrication method used in rolling element bearings. Its primary function is to prevent contact between the moving parts of the bearing, which is achieved by the liberation of a small amount of oil contained in the grease that will adhere to the surfaces; the oil degrades gradually by oxidation and is lost because of evaporation, centrifugal forces, and other phenomena [27].

Grease lubrication has advantages when compared to liquid lubricants, as it reduces wear and friction, helps to dissipate heat from friction, prevents rusting, and provides some sealing between particles and the bearing as well as reduced maintenance as the proper quantity of lubricant is confined in the housing and no oil level needs to be maintained [28].

This method is generally used for low-speed operations and relubrication depends mainly on the specifications from the manufacturer, as well as bearing and grease type, size, speed, operating temperature, and ambient conditions. Also, when changing the grease, the
operator must apply the correct amount to prevent corrosion and seal the bearing correctly to prevent contamination [27].

Bearing failure due to lubrication is classified as inadequate lubrication, including the incorrect selection or amount of lubricant, the incorrect interval for lubrication, and contamination, caused mostly by seal failure and improper handling [5]. Grease condition is challenging to monitor, as its distribution and the movement of the parts degrade the grease unequally, making it complex to use online techniques, with vibration and AE signals being the most used, and meaning laboratory tests with samples from the grease are more accurate [29]. Vibration and AE are primarily analyzed with statistical indicators in the time domain and characteristic frequencies in the frequency domain as the particles interact with the surface of the elements inside the bearing, creating changes in contact pressure that generate vibration and noise.

2.2. Statistical Features

In the condition monitoring field, statistical features in different domains allow the identification of characteristics within signals with a low computational load. Time domain indicators analyze a time series signal based on the relative position, separation, distribution, or other data characteristics, giving information helpful for the discrimination of the different conditions under analysis. As mentioned in Section 1, the use of statistical indicators is common for condition monitoring and fault detection in bearings as they provide information related to the waveforms captured, allowing the characterization of the defects and avoiding the problem of similar amplitudes or the closeness of the components when using techniques such as the FFT. Table 2 presents a set of 15 indicators reported in the literature for fault identification with satisfactory results [30], which were employed in this research.

Table 2. Statistical features employed in the analysis.

<table>
<thead>
<tr>
<th>Statistical Features</th>
<th>Equation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$ (5)</td>
</tr>
<tr>
<td>Maximum value</td>
<td>$\hat{x} = \max(x)$ (6)</td>
</tr>
<tr>
<td>Root Mean Square</td>
<td>$x_{RMS} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (x_i)^2}$ (7)</td>
</tr>
<tr>
<td>Square Root Mean</td>
<td>$x_{SRM} = \left(\frac{1}{n} \sum_{i=1}^{n} \sqrt{</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>$\sigma = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2}$ (9)</td>
</tr>
<tr>
<td>Variance</td>
<td>$\sigma^2 = \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2$ (10)</td>
</tr>
<tr>
<td>RMS Shape Factor</td>
<td>$SF_{RMS} = \frac{\hat{x}}{\frac{1}{n} \sum_{i=1}^{n}</td>
</tr>
<tr>
<td>SRM Shape Factor</td>
<td>$SF_{SRM} = \frac{x_{RMS}}{\frac{1}{n} \sum_{i=1}^{n}</td>
</tr>
<tr>
<td>Crest Factor</td>
<td>$CF = \frac{\hat{x}}{x_{RMS}}$ (13)</td>
</tr>
<tr>
<td>Latitude Factor</td>
<td>$LF = \frac{\hat{x}}{x_{SRM}}$ (14)</td>
</tr>
<tr>
<td>Impulse Factor</td>
<td>$IF = \frac{\hat{k}}{\frac{1}{n} \sum_{i=1}^{n}</td>
</tr>
<tr>
<td>Skewness</td>
<td>$S_k = \frac{\Sigma[(x-\mu)^3]}{\sigma^3}$ (16)</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>$k = \frac{\Sigma[(x-\mu)^4]}{\sigma^4}$ (17)</td>
</tr>
<tr>
<td>Fifth Moment</td>
<td>$5thM = \frac{\Sigma[(x-\mu)^5]}{\sigma^5}$ (18)</td>
</tr>
<tr>
<td>Sixth Moment</td>
<td>$6thM = \frac{\Sigma[(x-\mu)^6]}{\sigma^6}$ (19)</td>
</tr>
</tbody>
</table>
2.3. Fractal Dimension

A fractal is a geometric shape in which, when segmented, each section approximates the original shape on a smaller scale. The concept of fractal dimension corresponds to how a fractal fills a space while approaching smaller scales [31], as well as an indication of the self-similarity or repeatability of patterns in time domain signals [32]. Fractal dimensions range from 1.0 for a straight line to 2.0 for a plane, with several formulations to calculate this quantity, with the following being used in this paper.

2.3.1. Higuchi’s Fractal Dimension (HFD)

The process of calculating HFD consists of four steps [33]:

1. The signal $X$ in the time domain, with $N$ samples, gets decomposed into new series or sequences, $x_k^m$, using Equation (20), with $m$ as the initial time and $k$ as the time interval.

$$x_k^m = x_m, x_{m+k}, x_{m+2k}, \ldots, x_{m+[N-m]k} \quad (m = 1, 2, \ldots, k)$$

2. The average normalized length, $L_m$, of each $x_k^m$ is calculated with Equation (21).

$$L_m(k) = \left( \frac{N-1}{k} \right) \sum_{i}^{\left[ \frac{N-m}{k} \right]} |x_{m+ik} - x_{m+(i-1)k}|$$

3. The total length, $L(k)$, is calculated for every $k$ using Equation (22).

$$L(k) = \sum_{m=1}^{k} L_m(k)$$

4. The value $k$ is updated with $k = k + 1$. Steps 1 to 3 are repeated while $k < k_{\text{max}}$, with $k_{\text{max}}$ being a selected value at which the slope of the best-fit line of the plotted diagram in the plane ($\ln[L(k)]$ versus $\ln[1/k]$) is constant, with the slope being the value of HFD.

2.3.2. Katz’ Fractal Dimension (KFD)

To obtain the KFD value of a signal, the process is as follows [34]:

1. Calculate the maximum Euclidean distance, $d$, between the first sample, $x_1$, and the sample $x_k$ of a signal $X$ with $N$ samples, with $k = 1, \ldots, N$.

2. Obtain the sum of the Euclidean distances between consecutive samples, $L$, of the signal $X$ and its average, $a$, as shown in Equations (23) and (24).

$$L = \sum_{k=2}^{N} \text{dist}(x_k - x_{k-1})$$

$$a = \frac{L}{N-1}$$

3. The value of KFD is defined by Equation (25).

$$\text{KFD} = \frac{\log(L/a)}{\log(d/a)}$$

2.3.3. Petrosian’s Fractal Dimension (PFD)

To calculate the PFD value, the process consists of three steps [35]:

1. Take a signal, $X$, with $N$ samples, and binarize it according to the criteria in Equation (26).

$$z_i = \begin{cases} 
1, & x_i > \text{mean}(X) \\
-1, & x_i \leq \text{mean}(X) 
\end{cases}, \quad i = 1, 2, \ldots, N$$
2. Find the total number of sign changes in the binarized signal, $Z$, using Equation (27).

\[
N_\Delta = \sum_{i=1}^{N-2} \left| \frac{z_{i+1} - z_i}{2} \right|
\]  

(27)

3. Calculate the PFD as described in Equation (28).

\[
PFD = \frac{\log_{10} N}{\log_{10} N + \log_{10} \left( \frac{N}{N + \Delta N} \right)}
\]  

(28)

2.3.4. Sevcik’s Fractal Dimension (SFD)

SFD can be obtained following three steps [36]:

1. Considering the signal to be analyzed as a series of points $(x_i, y_i)$ with length $N$, normalize the signals as proposed in Equation (29), with $x_{\text{min}}$ and $y_{\text{min}}$ as the minimum values in the series while $x_{\text{max}}$ and $y_{\text{max}}$ are the maximum values:

\[
x_i^* = \frac{x_i - x_{\text{min}}}{x_{\text{max}} - x_{\text{min}}}; \quad y_i^* = \frac{y_i - y_{\text{min}}}{y_{\text{max}} - y_{\text{min}}}
\]  

(29)

2. Obtain the length of the waveform, $L$, using Equation (30):

\[
L = \sum_{i=0}^{N-2} \sqrt{(y_{i+1}^* - y_i^*)^2 + (x_{i+1}^* - x_i^*)^2}
\]  

(30)

3. Calculate SFD as shown in Equation (31):

\[
L = \sum_{i=0}^{N-2} \sqrt{(y_{i+1}^* - y_i^*)^2 + (x_{i+1}^* - x_i^*)^2}
\]  

(31)

2.4. Linear Discriminant Analysis (LDA)

When working with large, high-dimensional sets of data, it is necessary to reduce the dimension of the dataset to interpret the information and manipulate it for the classification or clustering tasks at hand. In this context, different techniques have been developed for supervised, when there is previous knowledge about the classes within the data, and unsupervised methods, with no prior knowledge of the classes, with LDA being one of the main techniques for supervised dimension reduction. LDA focuses on linear combinations of the variables of features to maximize the between-class scatter matrix ($S_b$) to separate the classes as much as possible while minimizing the within-class scatter matrix ($S_w$) to make the clusters more compact [37]. Equations (32) and (33) present the procedure to calculate $S_b$ and $S_w$, where $k$ is the dimension of the subspace defined as $k = C - 1$, $C$ is the number of classes, $m_k$ is the mean value of the class $C_k$, $m$ is the global mean value, $n_k$ is the number of samples, and $x_i$ is the $i$-th sample of class $C_k$.

\[
S_b = \sum_k n_k (m_k - m)(m_k - m)^T
\]  

(32)

\[
S_w = \sum_k \sum_{i \in C_k} (x_i - m_k)(x_i - m_k)^T
\]  

(33)

The optimal subspace $G = (g_1, \ldots, g_k)$ is obtained by optimizing Equation (34), with the solution being given by the $k$ eigenvectors of $S_w^{-1}S_b$ associated with the highest eigenvalues.

\[
\max \text{Tr} G^T S_b G = \frac{\text{Tr} G^T S_b G}{\text{Tr} G^T S_w G}
\]  

(34)
Once the subspace $G$ is obtained, the original data, $X$, are projected onto $G$ to transform it into the new space, $Y$, as shown in Equation (35).

$$Y = G^T X$$  \hspace{1cm} (35)

2.5. Support Vector Machines (SVMs)

SVMs are supervised learning techniques used for classification, regression, novelty detection, and other purposes [38]. SVMs were conceived as linear models, capable of classifying data by creating a hyperplane to achieve maximum separation between the classes in the feature space that depends on the datapoints near the hyperplane, called support vectors [39]. Considering that real-world data tend to present nonlinear characteristics, SVMs have evolved to implement different kernel functions to deal with data distribution, achieving great accuracy, flexibility, robustness, and efficiency with high generalization from the training data [39].

For a linearly separable case, the starting point is a data set $D$, expressed as (36), formed by a $n$-dimensional vector of samples, $\vec{x}$, and their labels, $y$. The SVM classification function, or hyperplane, is defined with Equation (37), where $\vec{w}$ is a weight vector, and $b$ is the bias; this function separates the data as shown in Equation (38).

$$D = \left\{ (x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n) \right\} \hspace{1cm} (36)$$

$$F(\vec{x}) = \vec{w} \cdot \vec{x} - b \hspace{1cm} (37)$$

$$y_i (\vec{w} \cdot \vec{x}_i - b) > 0, \ \forall (\vec{x}_i, y_i) \in D \hspace{1cm} (38)$$

The hyperplane needs to provide the greatest separation from both classes, called the margin; this can be achieved by modifying (38) by scaling $\vec{w}$ and $b$ as shown in Equation (39).

$$y_i (\vec{w} \cdot \vec{x}_i - b) \geq 1, \ \forall (\vec{x}_i, y_i) \in D \hspace{1cm} (39)$$

The margin from the closest vector $\vec{x}$ to the hyperplane is expressed in Equation (40). The vectors that satisfy (39) with an equality are the support vectors. To maximize the margin, the denominator of (40) needs to be minimized using Equations (41) and (42).

$$\text{margin} = \frac{|F(\vec{x}_i)|}{\|\vec{w}\|} = \frac{1}{\|\vec{w}\|} \hspace{1cm} (40)$$

$$\text{minimize : } \frac{1}{2} \|\vec{w}\|^2 \hspace{1cm} (41)$$

$$\text{subject to : } y_i (\vec{w} \cdot \vec{x}_i - b) \geq 1, \ \forall (\vec{x}_i, y_i) \in D \hspace{1cm} (42)$$

3. Materials and Methods

This section describes the methodology developed in this research for diagnosing five different conditions of metallic ball bearings. Figure 1 presents the block diagram of the steps followed, consisting of five stages described in the following subsections: the Testbench, the Data Acquisition stage, the Feature Extraction stage, High Dimensional Feature Matrices, and the Dimensionality Reduction and Classification stage.
3.1. Testbench

The experiments were conducted using a three-phase, 1 hp, WEG induction motor with one pair of poles, with a nominal speed of 3355 rpm, fed via a Variable Frequency Drive (VFD) with 220 Vac at 60 Hz, and programmed with a starting ramp of 10 s. The load was an automotive alternator coupled to the motor through a belt and pulleys system, as shown in Figure 2.

The metallic bearings used for the experiments were from SKF, model 6203 2RS, with an external diameter of 40 mm, an internal diameter of 17 mm, a 12 mm width, and eight caged balls, as presented in Figure 3a. To study the OR defects, holes with a 1 mm, 3 mm, and 5 mm diameter were drilled, simulating pitting [40] as a standard method to induce this kind of fault [9,11,41,42]. As the contaminant, black iron oxide, Fe₃O₄, with an approximate particle size of 30 µm, was used; to pollute the lubricant, 1 g of oxide was added to the BATat-3 grease used for the bearings. Figure 3c shows the measurement of the contaminant, and in Figure 3b it is possible to observe one of the bearings with the contaminated grease.
Figure 2. Testbench used for the experiments based on the induction motor, the coupled load, and the data acquisition system (DAS).

The metallic bearings used for the experiments were from SKF, model 6203 2RS, with an external diameter of 40 mm, an internal diameter of 17 mm, a 12 mm width, and eight caged balls, as presented in Figure 3a. To study the OR defects, holes with a 1 mm, 3 mm, and 5 mm diameter were drilled, simulating pitting [40] as a standard method to induce this kind of fault [9,11,41,42]. As the contaminant, black iron oxide, Fe3O4, with an approximate particle size of 30 µm, was used; to pollute the lubricant, 1 g of oxide was added to the BATat-3 grease used for the bearings. Figure 3c shows the measurement of the contaminant, and in Figure 3b it is possible to observe one of the bearings with the contaminated grease.

Figure 3. Bearings and black iron oxide used for the experiments: (a) bearing without contaminant; (b) bearing with polluted grease; and (c) measuring the oxide.

The tests considered five conditions: healthy (HLT), contaminated (CONT), contaminated with 1 mm OR defect (C1OR), contaminated with 3 mm OR defect (C3OR), and contaminated with 5 mm OR defect (C5OR). Figure 4 shows the HLT, C3OR, and C5OR bearings used in the experiments, allowing the comparison of the OR damage considered for the tests.

3.2. Data Acquisition Stage

The data capture was carried out by using a proprietary FPGA-based system with a Spartan 6 XC6SLX45, a 12-bit Analog-to-Digital Converter (ADC), and an on-board three-axis accelerometer sensor model LSM6DS3 with a ±2 g acceleration scale and a 0.061 mg/LSB sensitivity, mounted on the motor to measure the vibrations, as shown in
Figure 5a; for current sensing, a SCT013-030 current transformer with a rated 30 A/1 V conversion factor was used, shown in Figure 5b. As shown in Figure 2, the DAS was mounted over the connection box since the accelerometer was integrated into the board, and it was the most adequate location due to the limitations in the available space and the size of the board.

![Figure 4. Bearings used for the experiments; a comparison of the OR defects considered for the tests.](image1)

![Figure 5. DAS used in the experiments: (a) FPGA-based DAS system, with the accelerometer used for measuring the vibrations highlighted; (b) current sensor measuring one of the phases of the motor from the VFD.](image2)

For every condition, 15 experiments were run, obtaining 75 tests in total; each experiment consisted of a 10 s start ramp and 30 s steady state, measuring 40 s signals from the triaxial accelerometer and the current sensor; the vibration signals were sampled at a frequency of 1k samples/s, while current signals were sampled at 8k samples/s.

3.3. Feature Extraction Stage

The processing and classification of the conditions were performed using MATLAB R2023b. Because of the different sampling frequencies, a digital resampling of the vibration signals was performed to obtain vectors of the same size as the current signals. Once the signals were resampled, the 30 s steady-state segments were normalized in a $-1$ to $1$ range by applying Equation (43), where $X$ is the original signal, $X_n$ is the normalized signal, and $a$ and $b$ are the intervals for the normalized range. The normalized signals were divided into 1 s windows, and, for each window, the statistical indicators described...
with Equations (5)–(19) were calculated, as well as the four fractal dimensions presented in Section 2.3., obtaining a feature matrix of 19 features and 30 samples per signal.

\[ X_n = a + \left( \frac{X - \min(X)}{\max(X) - \min(X)} \right) (b - a) \]  

(43)

3.4. High-Dimensional Feature Matrices

The process described in Section 3.3 was repeated for the 15 test per bearing condition to generate a 450 × 19 feature matrix for each signal, concatenated for each of the five conditions, resulting in a 450 × 19 × 5 matrix per signal. For the current data, a 450 × 19 × 5 matrix was generated, and the 450 × 19 × 5 matrices generated from each of the accelerometer axes measuring the vibration signals were concatenated to obtain a 450 × 57 × 5 feature matrix.

3.5. Dimensionality Reduction and Classification Stage

The resulting matrices were then fused using LDA and projected into a 3D space to better separate the five classes by maximizing the \( S_b \) matrix and minimizing the \( S_w \) matrix of the features to cluster the data in groups as compact and as separated as possible, obtaining an optimal subspace, \( G \), from the eigenvectors of \( S_w^{-1}S_b \) associated with the highest eigenvalues, as explained in Section 2.4. The projected data were then fed to train and test an SVM model with a linear kernel, a box constraint of 1, a one vs. one coding design, an Error-Correcting Output Codes (ECOCs) model, and a Hinge loss function, using 80% of the data for the training process and 20% for validation, to classify and differentiate the condition of the bearings.

4. Results

In this section, the results obtained using statistical features as proof of the complexity of the identification of the combined contamination and OR faults and their fusion and classification using an SVM model are reported.

4.1. Data Preparation

The first step was preparing the signals, as there were different sampling rates and amplitudes. In Figure 6a, the resampled signals corresponding to the current and vibration HLT class are shown to illustrate the acquired signals; the amplitudes of the signals correspond to the values obtained with the 12-bit ADC. In contrast, Figure 6b–e present the steady-state, normalized samples to be analyzed, corresponding to the CONT, C1OR, C3OR, and C5OR classes. This figure shows that the changes in the signals due to the faults are not graphically appreciable.

4.2. Feature-Based Separation

As a first approach to the analysis using statistical features, the 15 statistical indicators and the four fractal dimensions were calculated for the three vibration signals (x, y, and z axes) and the current signal, and their distribution was observed using boxplots, with the outliers indicated with red plus signs, to identify if any of the features allowed for the discrimination of the classes on their own. Figure 7 shows the box plots generated for the use of RMS; in these graphs, it is possible to observe that the feature can separate the HLT class from the faulty ones with vibration signals from the y and z axes (Figure 7b,c) and current signals (Figure 7d), while not being able to discriminate between the damaged conditions adequately on any of the sources.

As mentioned in Section 2.1.1, kurtosis is another indicator commonly used for condition monitoring; Figure 8 presents the boxplots for this indicator on each source captured in the experiments. The results show no separation between conditions in current signals (Figure 8d) and some trends in the change in value for the kurtosis as the severity increases when analyzing vibration signals, but no clear separation that could lead to a correct diagnosis, as in Figure 8b,c, where the results from the vibration signals show some increment
in the value with the severity of the faults but not all conditions follow the trend, like the C1OR condition in both figures.

Figure 6. Acquired signals: (a) pre-normalization and -segmentation HLT signals; (b) processed CONT signals; (c) C1OR condition signals; (d) signals corresponding to the C3OR class; and (e) prepared C5OR signals.
for the use of RMS; in these graphs, it is possible to observe that the feature can separate the HLT class from the faulty ones with vibration signals from the y and z axes (Figure 7b,c) and current signals (Figure 7d), while not being able to discriminate between the damaged conditions adequately on any of the sources.

Figure 7. Boxplots for the RMS feature of the measured signals: (a) results for the x-axis vibration signal; (b) plot showing the distribution for the y-axis vibrations; (c) distribution of the z-axis vibration signal; and (d) distribution of the RMS for the current signal.

As mentioned in Section 2.1.1, kurtosis is another indicator commonly used for condition monitoring; Figure 8 presents the boxplots for this indicator on each source captured in the experiments. The results show no separation between conditions in current signals (Figure 8d) and some trends in the change in value for the kurtosis as the severity increases when analyzing vibration signals, but no clear separation that could lead to a correct diagnosis, as in Figure 8b,c, where the results from the vibration signals show some increment in the value with the severity of the faults but not all conditions follow the trend, like the C1OR condition in both figures.

Figure 8. Boxplots for the analysis of the conditions using kurtosis feature: (a) plot for the vibrations on the x-axis; (b) results for the y-axis vibrations; (c) boxplot for the z-axis vibration signals; and (d) current signal boxplots.

Similarly, the HFD boxplots are shown in Figure 9, displaying that each class is more compact than for RMS and kurtosis. When using this indicator, a clear separation between the HLT condition and the rest of the classes is appreciated across the vibration and current signals. In contrast, the classes corresponding to the faults are closer in value, and no trend in their distribution is shown, making it difficult to differentiate between the faults.
Similarly, the HFD boxplots are shown in Figure 9, displaying that each class is more compact than for RMS and kurtosis. When using this indicator, a clear separation between the HLT condition and the rest of the classes is appreciated across the vibration and current signals. In contrast, the classes corresponding to the faults are closer in value, and no trend in their distribution is shown, making it difficult to differentiate between the faults.

![Boxplots](image)

**Figure 9.** Boxplots for the use of HFD feature: (a) results for the analysis of x-axis vibrations; (b) plot for the vibrations on the y-axis; (c) z-axis vibrations results; and (d) current signal boxplots.

### 4.3. Proposed Method

As can be observed in the results in Figures 7–9, none of the features calculated from the current and vibration signals could distinguish between all the conditions under study. This led to the fusion of the features using a dimensionality reduction process to obtain sets of features that allowed better separation between the classes.

Starting with the $450 \times 19 \times 5$ matrix of current features, a dimensionality reduction was performed to cluster the data according to the conditions and project them in a 3D space for better differentiation. Figure 10 presents the results obtained when analyzing the current signals and Figure 10a shows the projection of the reduced feature matrix with the LDA process and the hyperplanes of the SVM, where it is possible to observe the separation of the HLT class from the faulty conditions; however, the defect-related data are overlapped, making the classification difficult. In the Figure 10b, the confusion matrix for the validation of the SVM model corroborates the complexity of separating the faults correctly when using current signals, as it obtains 66.9% accuracy.
Figure 10. Analysis results for the current data: (a) 3D projection of the fused data after the LDA process and the hyperplanes created by the SVM model; (b) confusion matrix for the validation of the SVM model, achieving 66.9% accuracy.

Similarly, the results of the vibration data analysis are presented in Figure 11. For the vibration signals, the $450 \times 19 \times 5$ feature matrix was reduced to a three-dimensional representation with LDA, displayed in Figure 11a, as well as the hyperplanes from the SVM model; it is possible to observe a better separation than with the current data, but the classes are not as compact, leading to some overlap between conditions. Figure 11b corresponds to the confusion matrix for the model validation, achieving 89.3% accuracy, a 20% improvement compared to the current analysis.
Figure 11. Analysis results for the vibration data: (a) 3D projection of the fused data after the LDA process and the hyperplanes created by the SVM model; (b) confusion matrix for the validation of the SVM model, achieving 89.3% accuracy.

Considering the previous results, the proposed methodology, focused on the fusion of current and vibration signals, allowed a better separation of the classes, as can be observed in Figure 12a, where the hyperplanes generated by the SVM model are also shown; note the separation between the HLT condition and the faults, as they are far apart, as well as some overlapping of the C3OR and C5OR classes, which correspond to those with the more significant defects, and more scattered CONT data. In Figure 12b, the confusion matrix for the validation of the model is shown, achieving 97.1% accuracy for the classification of the
five conditions, a notable improvement when compared to the results from the analysis of current and vibration signals individually.

![Confusion Matrix](image)

**Figure 12.** Analysis results for the fused vibration–current data: (a) 3D projection of the fused data after the LDA process and the hyperplanes created by the SVM model; (b) confusion matrix for validating the SVM model, achieving 97.1% accuracy.

5. Discussion

The results for the three cases show an improvement when the features are fused compared to when each sensor is analyzed independently, complementing the measurements of each kind of signal and making the system more robust and capable of correctly classifying the conditions. In all cases, the bearing with no damage (HLT) is well separated from the rest of the conditions. This could indicate a significant influence of the contamination, as
the faulty condition clusters are closer together. For the different OR defects, the projection of the features from the current signals presents considerable overlapping (Figure 10a), and the SVM model was not capable of classifying the conditions, as shown in Figure 10b; on the other hand, vibration data (Figure 11a) separate the OR damages better than when using current data, but the classification carried out with the SVM presents problems in correctly identifying the 3 mm and 5 mm holes, while recognizing the 1 mm hole better. As mentioned, the fusion data allow better distinction between the 3 and 5 mm defects in the OR, as seen in Figure 12b. This research contributes to the detection of common conditions that affect bearings, OR defects, and lubricant contamination, presenting the possibility of discriminating between the defects and their combinations by considering different severities of OR damage.

Additionally, Table 3 compares the proposed model to other approaches previously exposed in the literature, presenting an accuracy level similar to those works. It is possible to appreciate the primary use of vibration signals and neural networks, as well as the focus on defects located on the parts of the bearing, and a lack of studies on lubricant contamination, even though it is a leading cause of different faults [4], and the fusion of different sources of information. This comparison allows the identification of the possibilities and less explored areas in bearing condition monitoring to focus future works on new techniques and different faulty conditions.

Table 3. Comparison of the proposed methodology to proposals reported in the literature.

<table>
<thead>
<tr>
<th>Proposal</th>
<th>Classification Model</th>
<th>Faults</th>
<th>Number of Severities</th>
<th>Signals</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>[8]</td>
<td>CNN</td>
<td>OR, IR, ball, cage, lack of lubrication</td>
<td>6</td>
<td>Thermal images</td>
<td>99.8%</td>
</tr>
<tr>
<td>[12]</td>
<td>CNN</td>
<td>OR and IR</td>
<td>3</td>
<td>Vibration</td>
<td>93.97%</td>
</tr>
<tr>
<td>[13]</td>
<td>RFA</td>
<td>OR, IR, ball</td>
<td>3</td>
<td>Vibration</td>
<td>99.57%</td>
</tr>
<tr>
<td>[17]</td>
<td>Multi-source domain information fusion network (MDIFN)</td>
<td>OR, IR, ball</td>
<td>13</td>
<td>Vibration</td>
<td>98.31%</td>
</tr>
<tr>
<td>[20]</td>
<td>ANN</td>
<td>OR, IR, ball, cage</td>
<td>16</td>
<td>Vibration</td>
<td>91.6%</td>
</tr>
<tr>
<td>[23]</td>
<td>SVM and CNN</td>
<td>Contamination</td>
<td>3</td>
<td>Vibration</td>
<td>100% and 98.33%</td>
</tr>
<tr>
<td>Proposed Method</td>
<td>Linear SVM</td>
<td>OR and Contamination</td>
<td>5</td>
<td>Vibration and current</td>
<td>97.1%</td>
</tr>
</tbody>
</table>

6. Conclusions

This research contributes to the detection of common defects that affect bearings, OR damage and lubricant contamination, with the development of a methodology for measuring and processing vibration and current signals using time domain features and data fusion. The sensors and acquisition system allowed the correct measurement of the physical magnitudes, capturing low-noise signals processed directly with low conditioning.

The results show the advantage of data fusion from different sensors for diagnosing combined faults, as individual indicators present difficulties for the correct separation of the classes. Similarly, the fusion and projection of the data using LDA allowed the classes to be compact and well distributed in a 3D space for better diagnosis.

The linear SVM classifier proposed in the methodology presents a low efficiency when working with current signals, achieving 69.9%; for the fusion of the vibration signals measured with a three-axis accelerometer, the results increased to 89.3% efficiency, while when fusing the four signals (x-, y-, and z-axis vibration and one-phase current signals), the SVM reached 97.1% efficiency, being able to discriminate correctly across the five conditions under study, resulting in a system capable of diagnosing the presence of contaminants in the lubricating grease and three different levels of OR damage in bearings used in a VFD-fed induction motor. These results are contrasted with other works in the literature,
presenting an approach that considers the presence of contamination in the bearings and proposing a different classification model, achieving results on par with those studies.

Improvements can be made to the methodology to increase the performance and capability to detect the defects under study. These prospective improvements include the use of other noninvasive sensors for the measurement of other physical magnitudes, such as stray magnetic flux, the testing of different locations of the sensors to study the effects on the methodology, analysis of more conditions, including different contaminant concentrations, the introduction of individual OR damages, variations of the motor speed or the frequency of the VFD and load levels to generate a more extensive database, as well as the selection of the most significant features to improve class separation, and the implementation of different kernels for the SVM and other classifiers, such as ANNs or k-NN.


Funding: This research was funded by the Spanish ‘Ministerio de Ciencia e Innovación’, Agencia Estatal de Investigación and FEDER program in the framework of the ‘Proyectos de Generación de Conocimiento 2021’ of the “Programa Estatal para Impulsar la Investigación Científico-Técnica y su Transferencia”, belonging to the “Plan Estatal de Investigación Científica, Técnica y de Innovación 2021–2023”. (ref: PID2021-122343OB-I00).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to privacy.

Acknowledgments: The authors would like to thank Consejo Nacional de Humanidades, Ciencias y Tecnologías, and CONAHCYT, with scholarships registered under CVU numbers 1184202 and 1083233, respectively.

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References
42. Zamudio-Ramirez, I.; Osornio-Rios, R.A.; Antonino-Daviu, J.A.; Cureño-Osornio, J.; Saucedo-Dorantes, J.J. Gradual wear diagnosis of outer-race rolling bearing faults through artificial intelligence methods and stray flux signals. *Electronics* 2021, 10, 1486. [CrossRef]

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.