A New Approach for Bearing Condition Monitoring and Damage Detection Using Machine Learning Algorithms ⋆

André da Silva Barcelos * **Fábio Muniz Mazzoni** **
Antonio J. Marques Cardoso *

* *CISE-Electromechatronic Systems Research Centre, Universidade da Beira Interior, Calçada Fonte do Lameiro, 6200-358, Covilhã, Portugal (e-mail: andre.s.barcelos@ubi.pt; ajmcardoso@ieee.org)*
** *Instituto de Ciência e Tecnologia, Universidade Federal Fluminense, Rua Recife, Jardim Bela Vista, 28895-532, Rio das Ostras, Brasil (e-mail: fabiomuma@hotmail.com)*

Abstract:
Bearing condition monitoring (BCM) and damage identification are usually performed by vibration-based signals and supervised learning algorithms. However, this approach is impractical in many industrial facilities because some industrial motors are unable to provide access to vibration-based signals or they are prevented from performing their functions under damaged conditions. In this context, this work employs a density-based and fractal approach to extract features from current-based signals. These features create an unlabelled database that feeds a fuzzy c-means algorithm to perform BCM and the support vector machines to classify bearing damages in an unsupervised learning approach. Tests with several bearing damages under various load and speed conditions are reported, presenting promising results.

Keywords: Bearing fault diagnosis, Fuzzy c-means, Unlabelled learning.

1. INTRODUCTION

Electric motors are used in many industrial processes due to their flexibility for a wide range of applications, efficiency, and reliability even in severe environments. According to recent estimates, these machines consume 70% of the electric energy demanded by the European Union's industrial sector (Cardoso, 2018).

Moreover, owing to the intrinsic characteristics, electrical faults and mechanical damages occur on a regular basis, resulting in reduced performance or interruption of production processes. Indeed, according to studies conducted by the Institute of Electrical and Electronics Engineers, bearings damages are responsible for 40% of the defects caused in electric motors (Cerrada et al., 2018).

The two categories of bearing damages are punctual and distributed damages. Punctual damages arise on a delimited bearing surface, presenting cracks, holes, scratches, particles, pitting, material removal, impact points, or others forms that cause impulsive mechanical vibrations (Barcelos and Cardoso, 2021). Otherwise, distributed damages are exemplified by flushing, encrustations, material degradation, corrosion, wear, brinelling, or other types of damages that spread a defect along the entire length of the bearing's raceways, causing continuous mechanical vibrations (Cardoso, 1991).

The most common approach for conducting bearing condition monitoring (BCM) and damage detection is to input the vibration-based data into a signal processing technique. After that, one may perform a feature extraction method to build a labelled database, and applying a supervised learning algorithm to classify the bearing damages (Moshrefzadeh, 2021).

However, most industrial facilities are unable to provide a labelled database, because their machines are prevented to operate under damaged conditions. Consequently, it is also necessary to develop artificial intelligence approaches that perform bearing damage classification with unlabelled databases (Liu and Gryllias, 2020).

Furthermore, many industrial motors are unable to supply vibration-based signals because they are located in inaccessible sites or their housings are inadequate to accommodate new devices. Besides, the vibration data acquisition is costly, requiring the use of new sensors (e.g. accelerometers) and devices to transduce, transmit, and process the information (Neupane and Seok, 2020).

Otherwise, current-based data acquisition is remotely available in most industrial electric motors because the stator current is monitored for control, supply, and protection purposes. Also, the stator phases of electric motors are typical multiple data sources, allowing information fusion

⋆ This work was supported by the European Regional Development Fund (ERDF) through the Operational Programme for Competitiveness and Internationalization (COMPETE 2020) under Project POCI-01-0145-FEDER-029494 and by National Funds through the FCT, Portuguese Foun-dation for Science and Technology under Projects PTDC/EEI-EEE/29494/2017, UIDB/04131/2020 and UIDP/04131/2020.

approaches. Indeed, these advantages has attract more attention for researches on BCM through current-based signals in recent years (Leite et al., 2014).

The main disadvantage for current-based BCM are the poor signal-to-noise ratio (SNR), information losses in the magnetic field, saturation harmonics, electrical faults, and interference (Barcelos et al., 2021). Moreover, conventional signal processing techniques for denoising and extracting information from vibration-based signals may work inadequately in current-based signals (Wang et al., 2021). In summary, for the same signal processing technique, the current-based BCM has less accessible information (e.g., indirect measure) and more feature extraction complexity (e.g., poor SNR).

In this context, recent researches employs the α-stable probability density function (α-SPDF) to extract features from vibration-based signals (Hebda-Sobkowicz et al., 2020). However, the most common approach to extract density-based features from current-based signals is to fit them into a Gaussian or other quasi-normal parametric distributions, avoiding the computational efforts required to fit the α-SPDF (Wodecki et al., 2021).

Indeed, due to the poor SNR, the current-based distributions from healthy and damaged bearings are similar to the Gaussian distributions, within a log-likelihood test. As a result, the BCM and feature extraction methods, employing α-SPDF into current-based signals is practically absent in the literature. However, the non-Gaussian parameters for the α-SPDF may fully represent the healthy and the bearing damaged behavior for current-based signals, which have dense and elongated tails (Puchalski and Komorska, 2018).

Therefore, this study fit the α-SPDF into current-based distributions to extract non-Gaussian parameters. Furthermore, this paper introduces a change of coordinates in time domain, creating orbits to extract the fractal dimension (FD), which is a measure for the capacity to cover a multidimensional space (Puchalski and Komorska, 2018). A similar procedure that changes the time coordinates can project the Fourier transform into delayed coordinates (FTDC) to extract several non-conventional features.

These parameters, features, and measures are used to construct an unlabelled database to input machine learning algorithms that may perform classification in this context. Some recurrent examples are the fuzzy c-means (FCM), support vector data description, one-class neural networks, one-class support vector classifiers, generative adversarial networks, among others, that may perform positive unlabelled learning (Barcelos and Cardoso, 2021). The intrinsic characteristics of the FCM algorithm, such as adaptive fuzzifiers, overlapping clusters, ellipsoidal shape, outlier rejection, density-based approach, adaptive membership functions, and many others, make the FCM suitable for bearing damage identification (Li et al., 2018b).

As a result, this research's contributions are stated as follows: i) Introduce a novel feature extraction approach from current-based signals to construct an unlabelled database with the α-SPDF, FD, and FTDC measures; ii) Develops a fuzzy c-means (FCM) algorithm to perform BCM, while a support vector machine (SVM) perform damage detection. The remainder of this paper start with section 2, presenting the coordinate changes in time and frequency domain, the FCM and SVM algorithms, and α-SPDF distribution. Section 3 describes a database for experimental tests, while section 4 explains the modeling procedure and setup. Section 5 presents the results of the SVM performance, and the conclusion is in section 6.

2. MATHEMATICAL FORMULATION

2.1 Embedding dimension

The time series $x(t)$ of current-based signals can be converted into circular trajectories, creating an orbit with a fractal dimension (FD), while filling a space with integer dimension D>FD. With this premise, the signal $x(t)$ is sampled in the discrete signal $x(n)$, to build a D-dimensional vector $z(n)$ with time delay τ, as follows:

$$z(n,\tau) = [x(n), x(n-\tau), ..., x(n(-(D-1)\tau)] \quad (1)$$

This procedure changes the coordinates representation of $x(n)$, unfolding all projections (Strogatz, 2018). In these orbits, one can identify the outliers, the path of signals, the eigenvalues, peaks, volume, and other measures that offer insight about geometric characteristics. This research develops the generalization of the Eq. 1 to minimize the orbit overlapping. The parameters a_i and exponents b_i, scales and provide a nonlinear transformation, as follows:

$$\begin{aligned} z(n,\tau,a,b) = [a_0 x(n)^{b_0}, a_1 x(n-\tau)^{b_1}, \\ a_{(D-1)} x(n-(D-1)\tau)^{b_{(D-1)}}] \end{aligned} \quad (2)$$

Different values of τ define a distinct nonlinear transformation between each pair of vectors, increasing computational efforts. Thus, the orbits in this research use the same value of τ for each pair of vectors in the same time series. The procedure to find the value of τ that unfolds the projections of Eq. 2 is described in the following sections.

2.2 Mutual Information

The entropy of a random variable is an information theory measure that quantifies the uncertainties associated with variables. To calculate the entropy $H(x)$, it is necessary to consider the probability $P_X(x_i)$, as follows:

$$H(x) = -\sum_i P_X(x_i) \log P_X(x_i) \quad (3)$$

Furthermore, the Mutual Information MI can be defined using joint probability functions, expressing the MI for each pair of vectors as follows:

$$MI(x_i; x_j, \tau) = \sum_{i=1}^{D-1} \sum_{j=i+1}^{D} P(x_i, x_j) \log \left[\frac{P(x_i, x_j)}{P(x_i)P(x_j)} \right] \quad (4)$$

An average value of τ that minimizes the MI of the pairs of vectors in Eq. 2, generates symmetric orbits.

2.3 Renyi Dimension

The Eq. 3 was adapted by Renyi, using the exponents of the probability to obtain the Eq. 5 that generalizes $H(x)$.

$$R_q(x) = \frac{1}{1-q} log\left(\sum_{i=1}^{n} P_X(x_i^q)\right) \tag{5}$$

When $q \to 1$, $R_q(x) = H(x)$, assuring that the entropy $H(x)$ is a particular case of Reny's entropy. The Eq. 6 calculates the fractal dimension (FD) based on Renyi's entropy, as follows:

$$D_q = \lim_{\Delta x \to 0} \frac{1}{q-1} \frac{log \sum_{i=1}^{n} P_X(x_i^q)}{log(\Delta x)} \tag{6}$$

In Maragos and Sun (1993), the authors explain how to construct an algorithm to compute the Hausdorff fractal dimension (D_H) with the Eq. 6.

2.4 Box-Count Dimension

The Minkowski-Bouligand fractal dimension (D_M) employs the principle of Minkowski covers, which consists of covering a boundary region with disks of area $A(\epsilon)$ and radius ϵ. The Box-Count is a similar approach for calculating the fractal dimension (D_B), that uses a cover of n-dimensional boxes with sides ϵ and a counter $B(\epsilon)$ to sum the number of boxes as follows:

$$D_B = 2 - \lim_{\epsilon \to 0} \frac{log\ \ B(\epsilon)}{log\ \ 1/\epsilon} \tag{7}$$

In continuous systems, D_H, D_B e D_M are equivalent; but, in discrete systems these dimensions depends on computational issues and numerical approximations (So et al., 2017). In this work, the D_M is chosen as feature from the orbits, however these three dimensions are computed and contrasted to ensure computational convergence.

2.5 Alpha-stable Distribution

The α-SPDF describes a class of sub-Gaussian distributions of impulsive nature, with dense and elongated tail. This distribution is parameterized in Eq. 8, where Cauchy, Lévy, and Gauss distributions are particular cases.

$$\vartheta(t) = exp\{j\delta t - \gamma|t|^{\alpha}[1 + j\beta sgn(t)w(t,\alpha)]\} \tag{8}$$

where

$$w(t,\alpha) = \begin{cases} \tan\dfrac{\alpha\pi}{2} & \text{for } \alpha \neq 1 \\ \dfrac{2}{\pi} log|t| & \text{for } \alpha = 1 \end{cases}$$

The α-SPDF has four parameters: i) Location $\delta \in [-\infty, \infty]$; ii) Dispersion $\gamma > 0$; iii) Asymmetry $\beta \in [-1, 1]$; and iv) Exponent $\alpha \in (0, 2]$. The parameter estimation may be achieved by the maximum log-likelihood method (Mittnik et al., 1999).

2.6 Peak Identification

A peak identification algorithm consists of monitoring the moving mean of the time series and updating the standard deviation to verify the positions that the signal exceeds an adjustable threshold. In this work, after the peak identification, the prominence I_p, width I_w, prominence kurtosis I_{pk}, and width kurtosis I_{wk} are used as features.

2.7 Fuzzy C-Means

The fuzzy c-means (FCM) is a clustering algorithm that allows a data point to belongs to several clusters (Li et al., 2018a). The most recurrent approach is based on the minimization of the objective function, as follows:

$$f(u_{ij}, c_j) = \sum_{i=1}^{N}\sum_{j=1}^{m} u_{ij}^{\zeta}||x_i - c_j||^2$$

under the constrains

$$\sum_{j=1}^{m} u_{ij} = 1, \quad 0 \leq \sum_{k=1}^{N} u_{ij} \leq N$$

where N is a sample of feature vectors x_i, m is the number of clusters, u_{ij} is the grade of membership functions, c_j is the centroid, and ζ is a nonnegative fuzzifier that determines the fuzziness between clusters. The Lagrange approach minimizes the optimization problem as follows:

$$L(u_{ij}, c_j, \lambda_i) = \sum_{i=1}^{N}\sum_{j=1}^{m} u_{ij}^{\zeta}||x_i - c_j||^2 + \sum_{i=1}^{N} \lambda_i \left(\sum_{j=1}^{m} u_{ij} - 1\right)$$

where λ is a Lagrange multiplier. Whether $\zeta = 2$, the algorithm is the FCM, while $\zeta \to 1$, the behavior is close to K-means algorithm. Moreover, this work proposes that the parameter ζ change to ζ_i to compute adaptive fuzzifiers. The functions to update the centers, Lagrange multipliers and the membership functions are omitted for the sake of brevity, but can be found in Liu et al. (2018).

Assuming that the data moves from the healthy cluster to the bearing damaged cluster, it is possible to monitor the healthy cluster center (c_1), using moving means (MM) and moving standard deviations (MSTD). Therefore, this work takes the intervals of 10% (fast), 15% (regular), and 20% (slow) for the total amount of data to measure the MM and MSTD. Fig. 1 presents 250 sequential samples of a healthy bearing within 80 seconds.

On the left side, it is presented scatter plot of $\gamma \times$ FD, where the center and the standard deviation are available. On the right side, the FD series and γ parameter monitor the center behavior.

The faster MM has the sensitivity to detect intermittent outliers during transient changes. In Fig. 2, after 250 data points, the healthy bearing changes to a damaged condition gradually. The MM detects anomalies in the FD and γ parameter, while the persistence of the outliers alters the slower MM permanently.

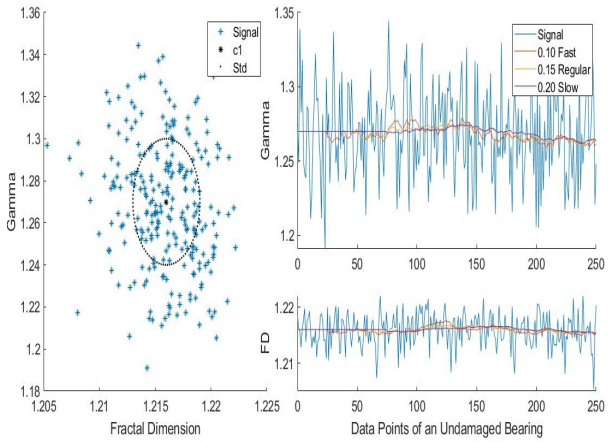

Figure 1. Clustering and MM behavior of γ and FD.

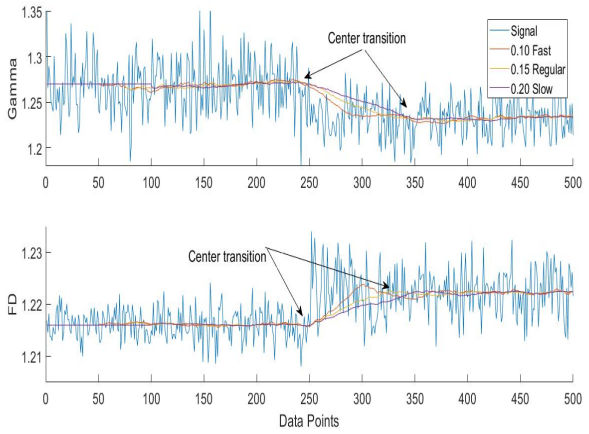

Figure 2. The center transition monitored by γ and FD.

The transition from healthy signals to damaged bearing signals creates a path from one cluster to another cluster. Fig. 3 presents the same transition of Fig. 2 by the viewpoint of each clusters.

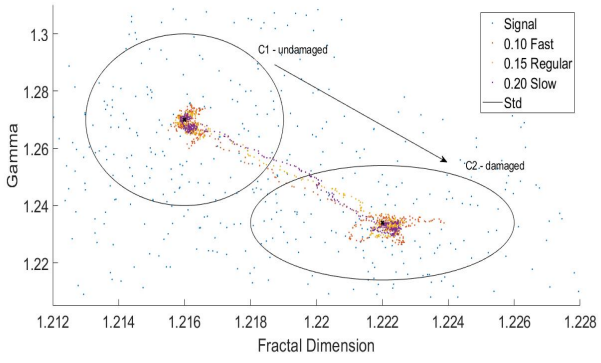

Figure 3. The data transition from c_1 to c_2.

The first cluster (c_1) is the healthy bearing condition, while the c_2 is the damaged bearing condition. The standard deviation of each cluster limits the transient region, while the persistence of outliers from a c_1 viewpoint represents the data migration from the healthy center. Fig. 4 presents the MM, MSTD, and the data migration.

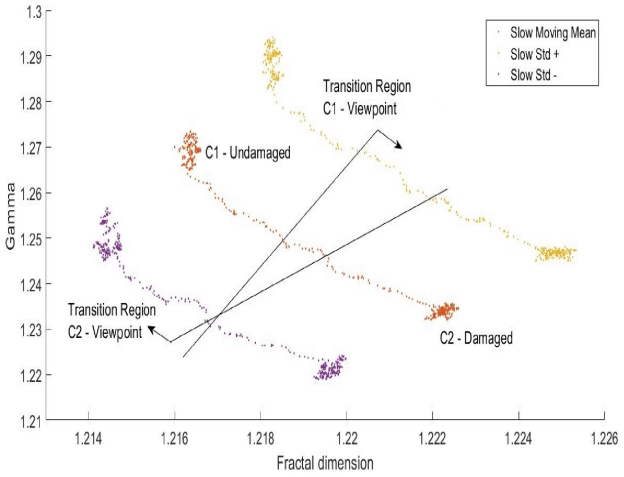

Figure 4. Data migration boundaries from c_1 and c_2 viewpoints.

From the c_1 perspective, the cluster standard deviation defines a border for the transition region, while the MSTD (Slow Std + and Slow Std -) limits the region. From c_2 perspective, there is a different border for the transition region, represented by another straight line enclosing the transition data points.

The slow MM and MSTD (yellow and purple) tracking the data migration, creating a hyperplane on which a SVM may distinguish clusters. In summary, the MM allows the initialization of the clusters (c_1) and (c_2), while the MSTD defines the data migration path. Furthermore, the nuclei migration velocity and trajectory determine the bearing damage trend.

2.8 Support Vector Machines

The SVM builds classification hyperplanes based on optimum weight vector $\mathbf{w_o}$, establishing an optimization problem with Lagrange multipliers (λ_i) limited by a constant V, as follows:

$$\max \sum_{i=1}^{N} \lambda_i - \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \lambda_i \lambda_j d_i d_j {x_i}^T x_j$$

$$\text{subject to } \sum_{i=1}^{N} \lambda_i d_i = 0 \qquad 0 \leq \lambda_i \leq V$$

where N is the size of an labelled database with variables $\mathbf{x_i}$ and label $\mathbf{d_i}$. The optimal weight vectors $\mathbf{w_o}$ and bias $\mathbf{b_o}$ are according to Eqs. 9 and 10:

$$\mathbf{w_o} = \sum_{i=1}^{N} \lambda_i d_i x_i \tag{9}$$

$$b_o = 1 - w_o^T x_i \quad \text{iff} \quad d_i = 1 \tag{10}$$

The kernel $x_i^T x_j$ can be replaced by a product functions $k(x, x')$. It is necessary that $k(x, x')$ be symmetrical, continuous, have eigenfunctions $\phi(x)$ and $\phi(x')$ with positive eigenvalues, and satisfy the Mercer's condition (Haykin,

2010). Assuming that k_1 and k_2 are kernels, the following properties of Eq. 11 and 12 hold.

$$k(x, x') = \eta k_1(x, x') + k_2(x, x') \quad \text{for} \quad \eta \geq 0 \tag{11}$$

$$k_s(x, x') = exp\Big(\frac{-||x - x'||^2}{2\sigma^2}\Big) \tag{12}$$

In this case, the SVM is a machine learning algorithm that provide separation by kernels.

2.9 Uninorms

Uninorm is a set operator, which generalizes the T-norm and S-norm of fuzzy set theory (Souza et al., 2018). A T-norm is a mapping defined in $T : [0, 1]^n \rightarrow [0, 1]$ which must satisfy the following conditions:

- Comutativity: $T(a, b) = T(b, a)$
- Associativity: $T(a, T(b, c)) = T(T(a, b), c)$
- Monocity: $T(a, b) \leq T(c, d)$ if $a \leq c$ and $b \leq d$
- Identity: $T(a, 1) = a$

The dual form of T-norms are the S-norms, obtained by the reversal operation, as follows:

$$S(a, b) = 1 - T(1 - a, 1 - b) \tag{13}$$

In Pedrycz (2006), the author explains that Commutativity, Associativity, and Monocity of T-norm persist for S-norm, however, the Identity is $S(a, 0) = a$. Thus, the Uninorm is an operator with neutral element g, as follows:

$$g \in [0, 1] \rightarrow \mathbb{R} \begin{cases} g = 1 \text{ then } U \rightarrow \text{T-Norm} \\ g = 0 \text{ then } U \rightarrow \text{S-Norm} \end{cases}$$

A family of Uninorms is characterized by:

$$T_g \equiv T(\frac{a}{g}, \frac{b}{g}) \quad \text{and} \quad S_g \equiv S\Big(\frac{a-g}{1-g}, \frac{b-g}{1-g}\Big)$$

$$U(a, b, g) = \begin{cases} gT_g & \text{if } a, b \in [0, g) \\ g + (1 - g)S_g & \text{if } a, b \in [g, 1] \\ \text{min or max (a,b)} & \text{if } otherwise \end{cases}$$

With the definitions of Uninorm, it is possible to construct the operator *Uni*, as follows:

$$Uni(x, x', g) = \begin{cases} gT_g & \text{if } x_i \in [0,\ g) \\ g + (1 - g)S_g & \text{if } x_i \in [g,\ 1] \\ max(x, x') & \text{if } g = [0,\ 0.5) \\ min(x, x') & \text{if } g = [0.5,\ 1] \end{cases}$$

The parameter g can change continuously, then the Uni(x, x', g) operator can assume intermediate values between the T-norm and S-norm. This work uses the product T-norm and the respective S-norm, as follows:

$$Product \begin{cases} T_g(a, b) = a \times b \\ S_g(a, b) = 1 - (1 - a) \times (1 - b) \end{cases}$$

The Uni(x, x', g) operator can provide a scalar, and therefore, the properties described in Eq. 11 and 12 enable this operator to scale a Gaussian kernel $k_s(x, x')$ according to:

$$k(x, x', g) = \frac{1}{Uni(x, x', g)} \times k_s(x, x') \tag{14}$$

In summary, the variation of the parameter g scale $k_s(x, x')$ between the S-norm and T-norm of the input variables. This work uses a SVM with a gaussian kernel to perform classification, and the Uni operator (Uni-SVM) to scale the gaussian kernel, improving the convergence in a non-Boolean frame.

3. DATA DESCRIPTION

The validations of this work use the dataset developed by the Chair of Design and Drive Technology, from the University of Paderborn in Germany, which contains the current-based signals from an electric motor with damaged and healthy bearings (Barcelos et al., 2021). Table 1 describes this data set with the bearing damage location (Inner Raceway - IR and Outer Raceway - OR) in the second column.

The third column presents the damages, where an electric drilling machine (EDM) provides artificial damages, a drilling machine produces the holes, and an electric engraver (EE) makes the scratches. The others two types of damages are pitting and indentations by plastic deformations.

Table 1. The characteristics of the damages in the bearings of the Paderborn time series

Serie	Local	Damage	Severity	Characteristic
KA01	OR	EDM	1	single point
KA07	OR	Drilling	1	single point
KA08	OR	Drilling	2	single point
KA05	OR	EE	1	single point
KA04	OR	Pitting	1	single point
KA16	IR & OR	Pitting	2	single point
KA22	OR	Pitting	1	single point
KI16	IR	Pitting	3	single point
KI17	IR	Pitting	1	single point
KI18	IR	Pitting	2	single point
KA30	OR	Indentations	1	distributed
KB27	IR & OR	Indentations	1	distributed
KB24	IR & OR	Indentations	3	distributed
KI04	IR & OR	Indentations	1	single point

The severity express the extension of each damage, represented on an increasing scale of the degree of damage (1-low, 2-medium, 3-high). And finally, the fifth column is the characteristic of the damages, as described in section 1. These time series and the healthy bearing K002, are available at speeds of 900 rpm and 1500 rpm (N09 and N15), with loading of conditions 0.1 Nm and 0.7 Nm (M01 and M07).

4. MODELING PROCEDURE

4.1 Feature extraction

In this work, the α-SPDF parameters for each time series in Paderborn database are determined by the log-maximum likelihood fitting according to Mittnik et al.

(1999). The $\mathrm{MI}(x_i, x_j, (\tau+1))$ and $\mathrm{MI}(x_i, x_j, \tau)$ ratios, finds a local minimum to project each time series into orbits (Eq. 2). Fig. 5 presents the K002 and KA16 orbits with $\tau = 400$ bytes in different load and speed conditions (i.e. 1.56 ms).

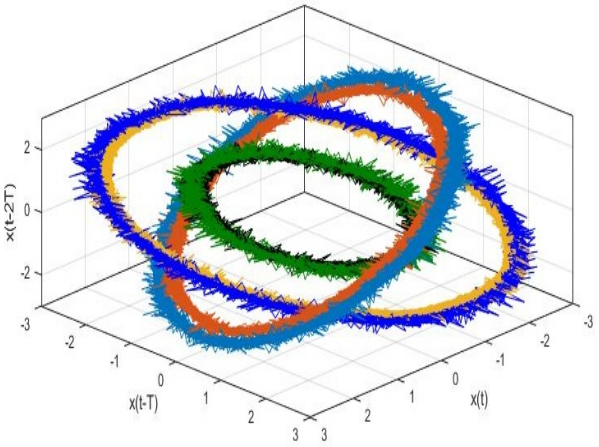

Figure 5. The K002 versus KA16 orbits in N15M01 (black and green), N09M07 (yellow and blue), and N15M07 (orange and light blue)

In these orbits, the peaks are evident, the projections are unfolded, the angular difference in N09M07 condition is caused by the motor speed, while the lower magnitude of N15M01 is caused by the load conditions. Consequently, the fractal dimension D_M is determined for each orbit using the algorithm established by Maragos and Sun (1993).

The same approach that unfold orbits from time series can be extended to Fourier transforms. In this case, the parameter τ from Eq. 2 is replaced by a frequency delay (θ) while the MI minimization project the Fourier transform into degenerate orbits with delayed coordinates (FTDC). Fig. 6 presents the FTDC for K002, KA04, KA16, and KA22 signals with $a_i = [1.00, 1.00, 1.00]$, $b_i = [1.10, 1.10, 1.10]$, and $\theta = 120$ bytes.

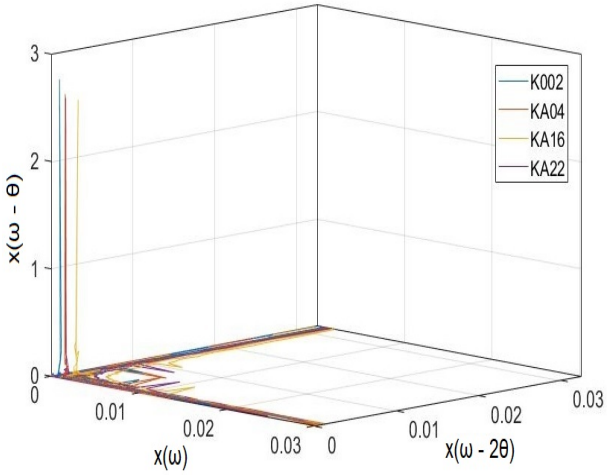

Figure 6. The FTDC projection for K002, KA04, KA16, and KA22 under N15M07 conditions.

The axes $x(\omega)$ and $x(\omega - 2\theta)$ captures the general behavior, dynamics and noises, while the plane $x(\omega) \times x(\omega - \theta)$ overlap the harmonics associated with the damages, improving intra-class separation. Fig. 7 presents the FTDC for all time series in N15M07 condition.

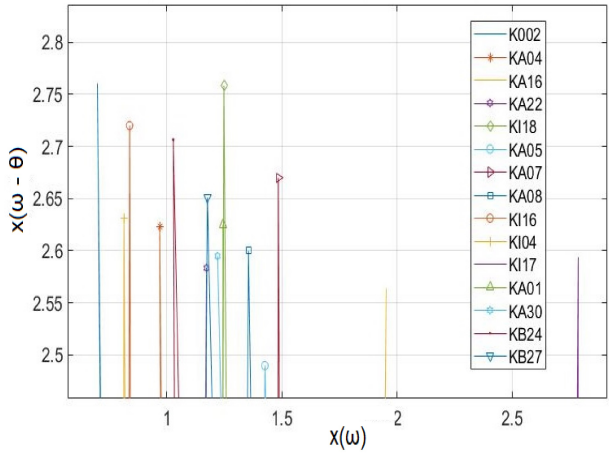

Figure 7. The FTDC for all signals under N15M07 conditions

The trio KI16, KI17, and KI18 have different severity for the pitting damage, producing almost the same harmonics with different magnitudes, which are mapped on distant points over the support $x(w)$. These patterns also happen with the trio KA30, KB24, and KB27 for distributed damages, which are scattered over the support, improving intra-class separation.

Another relevant aspect for FTDC is that the K002 signal, therefore without harmonics from damages, is mapped close to the origin over the $x(w)$ support. Otherwise, the α-SPDF parameters and D_M can perform extra-class separation, improving the classifiers accuracy. Fig. 8 presents the parameter γ and the D_M behavior for each motor condition.

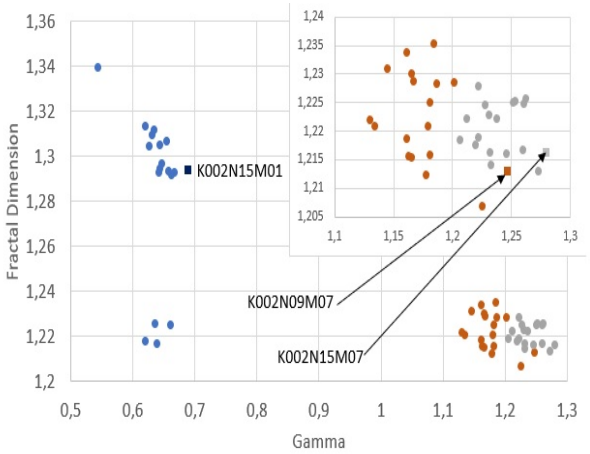

Figure 8. The D_M and the γ parameter of N15M01 (blue), N09M07 (orange), and N15M07 (gray) conditions.

The extra-class separation is evident in the $\gamma \times D_M$ frame. Indeed, the γ parameter alone can identify the damage condition, while the D_M can be associated with the load.

4.2 Setup and configurations

The α-SPDF parameters, D_M, the FTDC measures $\{x(\omega), max(x(\omega-\theta))\}$, and the features from the peaks are used to construct the database. The FCM has five Gaussian membership functions to perform the BCM and damage detection, while the Uni-SVM classifies the damages. In this case, the parameters g, b_i, and ζ_i start at 0.500, [1.00, 1.00, 1.00], and 2.000 respectively, and they are updated in 0.001 steps per epoch.

Since the b_i parameters influence the orbit's convexity, they are updated only after the g and ζ_i parameters have been set, reducing computational efforts and avoiding non-convex optimization algorithms. In the supervised learning approach, the training set contains 70% of data, while the validation set contains 30% of data, with a stop criterion of 0.005%. In the unlabelled learning, these parameters are adjusted only after the cluster migration, and therefore the training set is absent.

5. RESULTS

The parameters and features for the healthy bearing K002 are presented in Table 2. The FTDC measurements are consistent over all load and speed conditions.

Table 2. Parameters and features of K002

	N15M01	N09M07	N15M07
α	2.000	1.999	1.999
β	0.933 ± 0.024	0.031 ± 0.081	0.300 ± 0.051
γ	0.688 ± 0.011	1.246 ± 0.068	1.270 ± 0.030
δ	-0.019 ± 0.020	-0.017 ± 0.016	-0.024 ± 0.019
I_p	0.112 ± 0.2276	0.116 ± 0.303	0.117 ± 0.389
I_{pk}	144.841 ± 8.325	265.950 ± 6.121	171.547 ± 3.127
I_w	3.181 ± 22.038	3.324 ± 29.718	2.993 ± 22.251
I_{wk}	191.554 ± 1.890	304.880 ± 1.587	192.523 ± 1.796
D_M	1.293 ± 0.030	1.213 ± 0.010	1.216 ± 0.003
$FTDC$	(0.248, 2.731)	(0.246, 2.748)	(0.244, 2.752)

A relevant observation is that the magnitude of $\alpha < 2.000$ and $\beta \neq 0$ or 1 defines a sub-Gaussian PDF. Indeed, only the N15M01 condition can be properly fitted into a Gaussian distribution with a reduced relative error, while the others loading and speed conditions properly typify sub-Gaussian distributions. The average of the parameters and features from the damaged bearings are presented in Table 3.

Table 3. Parameters and features from the damaged bearings

	N15M01	N09M07	N15M07
α	1.999	1.999	1.999
β	0.064 ± 0.807	-0.038 ± 0.314	0.086 ± 0.463
γ	0.642 ± 0.028	1.171 ± 0.028	1.234 ± 0.020
δ	-0.021 ± 0.023	-0.031 ± 0.026	-0.035 ± 0.029
I_p	0.114 ± 0.210	0.118 ± 0.289	0.116 ± 0.375
I_{pk}	145.789 ± 8.379	259.349 ± 6.005	174.514 ± 3.219
I_w	3.193 ± 22.089	3.342 ± 29.632	3.019 ± 22.426
I_{wk}	191.249 ± 1.905	304.039 ± 1.645	191.640 ± 1.702
D_M	1.294 ± 0.036	1.222 ± 0.007	1.222 ± 0.004

Comparing the results from Table 2 and Table 3, the parameter β is sensitive to load and speed, the features I_{pk} and γ may discriminate the classes, while the D_M increases when the damages occur. Since the FTDC output an intra-class measure for several types of bearings, the average of these outputs is an equilibrium point in the support (e.g., a mass center), without physical meaning. In this context, an adequate description of FTDC measurements in N15M07 condition is introduced in Fig 7. Moreover, the N15M01 and N09M07 conditions produce a similar pattern and therefore are omitted for the sake of brevity.

In summary, the FTDC measurements improve the classifiers accuracy, increasing the number of quasi-orthogonal vectors available in this database. Consequently, Fig. 9 present the Uni-SVM classifications in supervised learning context to verify the FTDC behavior, considering random signals (i.e., number of clusters) from the Paderborn data set.

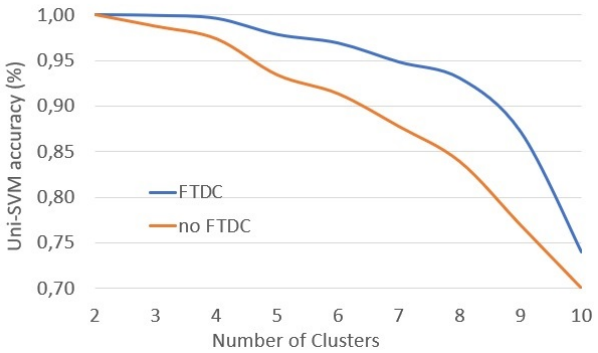

Figure 9. The Uni-SVM accuracy with FTDC (blue) and without FTDC as feature (orange).

In this case, when the number of clusters remains less than four (i.e., three damages), the Uni-SVM algorithm achieves 100% of accuracy. Furthermore, the empirical boundary for this approach is eight clusters, and therefore the following experiments are performed in this condition. The Uni-SVM accuracy for labelled and unlabelled databases are presented in Table 4.

Table 4. The Uni-SVM peformance with labelled and unlabelled databases.

	Labelled	Unlabelled	Uni(x, x',g)
N09M07	68.37%	65.05%	1.000
	92.99%	84.45%	0.235
	96.37%	91.89%	0.159
N15M01	74.53%	70.20%	1.000
	92.32%	85.52%	0.210
	94.24%	92.12%	0.120
N15M07	82.45%	78.86%	1.00
	95.00%	90.35%	0.282
	96.39%	93.55%	0.162

When the kernel Uni(x, x', g) is set to 1.000, the Uni-SVM behaves like a Gaussian SVM. However, scaling the Uni-kernel to 1/0.16 provide the highest accuracy for N09M07 and N15M07 conditions, while the N15M01 condition is better classified at 1/0.12. Moreover, the identification of eight clusters is an unusual situation, since most studies with current-based signals will distinguish three or four bearing damages.

6. CONCLUSION

This work proposes an innovative method for bearing condition monitoring, using a density-based approach, fractal dimension, and the Fourier transform into delayed coordinates from current-based signals. The FCM algorithm create clusters to identify bearing damages, while the Uni-SVM perform classification. The FTDC measures improve intra-class separation, while the fractal dimension and the α-SPDF parameters improve extra-class classification.

The fuzzy c-means algorithm allows a fast convergence of the SVM, selecting the most appropriate data points to perform classification. Indeed, with these approaches and techniques, the classification accuracy achieves 100% for four clusters. Moreover, tests performed with eight clusters in labelled and unlabelled databases reach 96% and 93% of accuracy in N15M07 condition, respectively. The main advantage of this method is the capability to perform current-based bearing condition monitoring and damage detection without any prior knowledge of the bearing dimensions, damage location, severity, or damage type.

As a suggestion for future work, is the application of another method of extracting non-conventional features, for comparison purposes with the proposed in this work.

ACKNOWLEDGEMENT

This work is dedicated to the author André da Silva Barcelos (in memoriam), for all his efforts in the development of this research.

REFERENCES

Barcelos, A.S. and Cardoso, A.J.M. (2021). Current-based bearing fault diagnosis using deep learning algorithms. *Energies*, 14(9), 2509.

Barcelos, A.S., Mazzoni, F.M., and Cardoso, A.J.M. (2021). Análise de avarias em rolamentos, utilizando algoritmos de inteligência artificial. *Brazilian Journal of Development*, 7(3), 29080–29093.

Cardoso, A.J.M. (1991). *Fault diagnosis in three-phase induction motors*. Coimbra, Portugal: Coimbra Editora.

Cardoso, A.J.M. (2018). *Diagnosis and Fault Tolerance of Electrical Machines, Power Electronics, and Drives*. IET - The Institution of Engineering and Technology, London, UK.

Cerrada, M., Sánchez, R.V., Li, C., Pacheco, F., Cabrera, D., de Oliveira, J.V., and Vásquez, R.E. (2018). A review on data-driven fault severity assessment in rolling bearings. *Mechanical Systems and Signal Processing*, 99, 169–196.

Haykin, S. (2010). *Neural networks and learning machines, 3/E*. Pearson Education India.

Hebda-Sobkowicz, J., Zimroz, R., Pitera, M., and Wyłomańska, A. (2020). Informative frequency band selection in the presence of non-gaussian noise. *Mechanical Systems and Signal Processing*, 145, 106971.

Leite, V.C., da Silva, J.G.B., Veloso, G.F.C., da Silva, L.E.B., Lambert-Torres, G., Bonaldi, E.L., and de Oliveira, L.E.d.L. (2014). Detection of localized bearing faults in induction machines by spectral kurtosis and envelope analysis of stator current. *IEEE Transactions on Industrial Electronics*, 62(3), 1855–1865.

Li, C., Cerrada, M., Cabrera, D., Sanchez, R.V., Pacheco, F., Ulutagay, G., and Valente de Oliveira, J. (2018a). A comparison of fuzzy clustering algorithms for bearing fault diagnosis. *Journal of Intelligent & Fuzzy Systems*, 34(6), 3565–3580.

Li, Y., Yang, Y., Wang, X., Liu, B., and Liang, X. (2018b). Early fault diagnosis of rolling bearings based on hierarchical symbol dynamic entropy and binary tree support vector machine. *Journal of Sound and Vibration*, 428, 72–86.

Liu, C. and Gryllias, K. (2020). A semi-supervised support vector data description-based fault detection method for rolling element bearings based on cyclic spectral analysis. *Mechanical Systems and Signal Processing*, 140, 106682.

Liu, R., Yang, B., Zio, E., and Chen, X. (2018). Artificial intelligence for fault diagnosis of rotating machinery: A review. *Mechanical Systems and Signal Processing*, 108, 33–47.

Maragos, P. and Sun, F.K. (1993). Measuring the fractal dimension of signals: morphological covers and iterative optimization. *IEEE Transactions on signal Processing*, 41(1), 108.

Mittnik, S., Doganoglu, T., Chenyao, D., et al. (1999). Maximum likelihood estimation of stable paretian models. *Mathematical and Computer Modelling*, 29(10-12), 275–293.

Moshrefzadeh, A. (2021). Condition monitoring and intelligent diagnosis of rolling element bearings under constant/variable load and speed conditions. *Mechanical Systems and Signal Processing*, 149, 107153.

Neupane, D. and Seok, J. (2020). Bearing fault detection and diagnosis using case western reserve university dataset with deep learning approaches: A review. *IEEE Access*, 8, 93155–93178.

Pedrycz, W. (2006). Logic-based fuzzy neurocomputing with unineurons. *IEEE Transactions on Fuzzy Systems*, 14(6), 860–873.

Puchalski, A. and Komorska, I. (2018). Stable distributions and fractal diagnostic models of vibration signals of rotating systems. In *Advances in Condition Monitoring of Machinery in Non-Stationary Operations*, 91–101. Springer.

So, G.B., So, H.R., and Jin, G.G. (2017). Enhancement of the box-counting algorithm for fractal dimension estimation. *Pattern Recognition Letters*, 98, 53–58.

Souza, P.V.C., Silva, G.R.L., and Torres, L.C.B. (2018). Uninorm based regularized fuzzy neural networks. In *2018 IEEE conference on evolving and adaptive intelligent systems (EAIS)*, 1–8. IEEE.

Strogatz, S.H. (2018). *Nonlinear dynamics and chaos with applications to physics, biology, chemistry, and engineering*. CRC press.

Wang, C., Wang, M., Yang, B., Song, K., Zhang, Y., and Liu, L. (2021). A novel methodology for fault size estimation of ball bearings using stator current signal. *Measurement*, 171, 108723.

Wodecki, J., Michalak, A., and Zimroz, R. (2021). Local damage detection based on vibration data analysis in the presence of gaussian and heavy-tailed impulsive noise. *Measurement*, 169, 108400.