

Diagnóstico Inteligente de Falhas em Equipamentos Dinâmicos Através da Transformada Wavelet e Dados Estatísticos

Diego Assereuy Lobão* Luiz Alberto Pinto*

* Grupo de Aprendizado de Máquina e Automação - Gama
Programa de Pós-Graduação em Engenharia de Controle e Automação
- Propecaut, Instituto Federal do Espírito Santo, Serra, Brasil
(e-mails: diego.lobao@gmail.com, pintoluizalberto@gmail.com)

Abstract: This article addresses the topic of fault diagnosis in dynamics equipments applying Machine Learning techniques using descriptors extracted by means of Wavelet Transform and statistical descriptors calculated in time domain. The performances obtained by classification algorithm k -Nearest Neighbor, Support Vector Machine, Random Forest and Artificial Neural Networks Perceptron are compared. In addition, the work brings a brief approach on dimensionality reduction through Principal Component Analysis. The dataset used in the experiments was obtained in the rotating machine simulator Mafaulda, and the best result was obtained by Random Forest (98.46 %), with the data in the wavelet domain.

Resumo: Este artigo aborda o tema do diagnóstico de falhas em equipamentos dinâmicos aplicando técnicas de aprendizado de máquinas utilizando descritores extraídos por meio da Transformada *Wavelet*, e descritores estatísticos calculados no domínio do tempo. Os desempenhos dos modelos de classificação obtidos pelos algoritmos k -Nearest Neighbour, Support Vector Machine, Floresta Aleatória e Redes Neurais Artificiais *Perceptron* Multicamadas são comparados. Além disso, o trabalho traz uma breve abordagem sobre redução de dimensionalidade através da Análise de Componentes Principais. O conjunto de dados utilizado nos experimentos foi obtido no simulador de máquinas rotativas *Mafaulda*, e o melhor resultado foi alcançado com a Floresta Aleatória (98.46%), com os dados no domínio *wavelet*.

Keywords: Fault diagnosis; machine learning; wavelet transform; feature extraction; vibration analysis; mafaulda; pattern recognition.

Palavras-chaves: Diagnóstico de falhas; aprendizado de máquinas; transformada wavelet; extração de características; análise de vibração; mafaulda; reconhecimento de padrões.

1. INTRODUÇÃO

As máquinas rotativas são comumente utilizadas em sistemas eletro-mecânicos e desempenham um papel muito importante no ramo industrial (Lei et al., 2013). Devido as condições adversas no ambiente de operação, esses equipamentos estão sujeitos a diversos tipos de falhas que podem causar danos materiais envolvendo o próprio equipamento ou instalação, danos ao meio ambiente devido a fuga de insumos da produção, ou até mesmo danos pessoais envolvendo a integridade física das pessoas que trabalham ou que vivem em torno das instalações.

Atualmente existem diferentes tipos de máquinas rotativas nos mais diversificados ramos industriais, como por exemplo os aeroderivados no setor aeroespacial e as turbinas a gás e eólicas no ramo de geração de energia (Lei, 2016). A construção desses equipamentos possuem componentes em comuns como rotores, rolamentos e engrenagens. Dessa forma, o projeto de fabricação desses equipamentos requer a existência de folgas mecânicas muito pequenas entre partes, que de forma geral, operam em altas velocidades de rotação. Devido as condições rígidas de operação, uma

falha incipiente local pode se propagar por outras partes do equipamento podendo resultar em falhas catastróficas no sistema de produção. Diante deste possível cenário, uma manutenção eficiente torna-se extremamente importante para garantir a disponibilidade, a confiabilidade e a segurança do equipamento e de toda instalação.

A gestão de manutenção nesses equipamentos vem ganhando importância cada vez maior devido à alta exigência de disponibilidade e confiabilidade. Estudos recentes mostram que o tempo de inatividade de equipamentos custa aos fabricantes britânicos £180 bilhões por ano (Ford, 2017). Assim, é evidente que uma boa gestão é necessária para manter a confiabilidade dos sistemas mecânicos e reduzir custos de manutenção.

Baseado na estratégia da manutenção preditiva e no desenvolvimento de sistemas inteligentes, um novo conceito chamado *Prognostics and Health Management (PHM)* foi desenvolvido recentemente (Lei, 2016). A técnica se baseia em cinco etapas: (i) aquisição dos dados através de sistemas de monitoramento, (ii) processamento dos sinais extraíndo características discriminativas, no domínio do tempo, ou no domínio da frequência ou ainda no domí-

nio tempo/frequência, (iii) diagnóstico utilizando métodos tradicionais ou sistemas inteligentes, (iv) prognóstico utilizando métodos baseados em modelos ou dados estatísticos e (v) tomada de decisão, no qual é definido o plano de intervenção do equipamento.

Nesse contexto, o presente trabalho apresenta um estudo sobre a aplicação de *machine learning* ao desenvolvimento de sistemas de diagnósticos de falhas num simulador de equipamentos dinâmicos utilizando uma fusão de descritores obtidos através da Transformada *Wavelet* (TW) e parâmetros estatísticos extraídos no domínio do tempo. Para a realização dos testes, foram utilizados diversos algoritmos de classificação como o *k-Nearest Neighbour* (*k-NN*), *Support Vector Machine* (*SVM*), Floresta Aleatória (FA) e Redes Neurais Artificiais *Perceptron* Multicamadas (RNA/*MLP*).

Na sequência este trabalho está constituído pelas seguintes seções, a Seção 2 contém uma revisão bibliográfica sobre o diagnóstico de falhas; a Seção 3 faz uma descrição sucinta dos algoritmos de classificação empregados e uma breve descrição da Transformada *Wavelet*; a Seção 4 apresenta a bancada experimental utilizada nesta pesquisa; a Seção 5 aborda a metodologia utilizada; a Seção 6 apresenta os resultados de classificação e a Seção 7 traz as conclusões finais.

2. REVISÃO BIBLIOGRÁFICA

Pesquisas envolvendo diagnóstico de falhas através do processamento de sinais de vibração já tem sido realizadas ao longo das últimas décadas.

O trabalho realizado por Prabhakar et al. (2002) apresenta um estudo sobre diagnóstico de falhas em mancais de rolamentos de esferas utilizando a Transformada *Wavelet* Discreta (TWD) para análise dos sinais de vibração. O experimento aborda falhas nas pistas interna e externa de cinco mancais. O trabalho se baseia na análise multiresolução e realiza a decomposição dos sinais de vibração utilizando a *Wavelet* da família *Daubechies* (*db4*). Esse procedimento possibilita a análise da sensibilidade da detecção das condições de falhas presentes nos diversos níveis de detalhe. Os autores investigaram também a utilização de parâmetros estatísticos obtidos no domínio do tempo, e verificaram que a curtose e o valor *rms* do sinal são sensíveis às diferentes falhas, contudo não possuem a capacidade de detectar falhas em estágios iniciais. Por último, os autores investigaram a utilização da curtose calculada a partir dos sinais de detalhe em diversos níveis de decomposição, e verificaram a grande sensibilidade desse parâmetro à detecção de falhas incipientes. O experimento concluiu que o uso de técnicas envolvendo a TWD combinadas com extração de parâmetros estatísticos favorecem no diagnóstico de falhas em mancais de rolamentos.

Uma metodologia de diagnóstico de falhas em máquinas rotativas é apresentada no trabalho de Santiago et al. (2004). O autor utiliza a Transformada *Wavelet* Contínua (TWC) da família *Morlet* para análise da resposta a sinais de vibração em regimes transientes e estacionários. Para a classificação das condições de falhas foi utilizada a energia contida nas bandas de frequências extraídas através do algoritmo *Wavelet Packet Transform* (*WPT*) em

conjunto com uma RNA, técnica conhecida como *Wavelet Packet Neural Network* (*WPNN*). Os resultados obtidos mostraram que o *WPNN* como técnica de classificação e diagnóstico de falhas em máquinas rotativas é eficiente, no qual foi possível obter acurácia próxima de 100%.

Pesquisas sobre diagnóstico inteligente de falhas em rolamentos de esferas aplicando RNA e *SVM* foram apresentadas no trabalho de Kankar et al. (2011). Através da análise e extração de descritores de sinais de vibração, os autores investigaram diversas falhas nos rolamentos de esferas, tais como, defeitos nas pistas externa e interna, defeitos nos elementos rolantes e uma combinação de todos eles. Métodos estatísticos são utilizados para extração de descritores no domínio do tempo como a média, o desvio padrão, a assimetria, a amplitude, a curtose e o fator de crista. Técnicas de seleção de atributos utilizando algoritmos de aprendizado de máquinas foram utilizados para reduzir a dimensionalidade do conjunto. Apesar da baixa acurácia devido a pequena amostragem contida no conjunto de dados, 71,23% para o modelo utilizando RNA e 73,97% para o modelo utilizando *SVM*, os autores consideraram que os resultados da classificação se mostraram promissores e com grande potencial de aplicação em sistemas de diagnósticos inteligentes para prevenção de falhas.

Pesquisas que utilizaram métodos de similaridades foram realizadas por Ribeiro et al. (2017). Nessa técnica, a utilização do modelo de similaridade auxilia na geração de descritores para um classificador e pode ser utilizada também como um classificador independente. O modelo baseado em similaridade é um algoritmo não paramétrico, capaz de procurar pelo menor número de amostras representativas que sejam suficientes e eficazes para descrever cada classe. Os autores utilizam dados estatísticos no domínio do tempo e da frequência para extração dos descritores, e como classificador é utilizado um modelo baseado em Floresta Aleatória. Resultados experimentais utilizando os conjuntos de dados *Mafaulda* e *CWRU* (*Case Western Reserve University*) mostram o poder de generalização do modelo, no qual foi possível obter taxas de acerto acima de 96%.

Uma combinação de Rede Neural Convolutiva (RNC) e *Extreme Learning Machine* (*ELM*) foi utilizada no trabalho de Chen et al. (2019). A pesquisa foi realizada utilizando uma bancada experimental composta por um motor acionador, um transdutor de torque e uma caixa multiplicadora, na qual foram introduzidas falhas nos mancais de rolamentos e nas engrenagens. O experimento utiliza sinais de vibração através de acelerômetros instalados no equipamento, esses sinais foram processados utilizando a TWC, a partir dos quais foram gerados escalogramas. Na segunda etapa foi utilizado uma RNC para a extração do vetor de características das imagens, que foram classificados por uma rede *ELM*. Durante os testes, foram reproduzidos sete tipos de falhas num total de 2100 amostras, sendo 300 amostras para cada classe. Segundo os autores, o método proposto superou técnicas tradicionais de *machine learning* que utilizaram descritores baseados na transformada rápida de *Fourier*. O melhor resultado obtido com o método proposto, que utilizou RNC-*ELM*, obteve uma acurácia acima de 99%.

3. REFERENCIAL TEÓRICO

Esta seção descreve os princípios que fundamentam os algoritmos de classificação, bem como uma breve descrição da Transformada *Wavelet*, aqui utilizada para extração de descritores no domínio tempo/frequência.

3.1 Algoritmos de Classificação

A etapa de classificação no presente trabalho foi realizada com a aplicação dos algoritmos *k-NN*, *SVM*, Floresta Aleatória e *RNA/MLP*. No que segue é feita uma breve descrição de cada um.

k-NN: é um classificador não paramétrico que atribui à amostra desconhecida a mesma classe das *k* amostras mais próximas, segundo um critério de decisão. Sua simplicidade e eficácia fazem com que esse algoritmo seja amplamente utilizado em problemas de classificação (Fayed and Atiya, 2009). O princípio sobre o qual o *k-NN* se baseia, é de que amostras similares tendem a se concentrar em uma mesma região do espaço. Apesar de sua implementação ser simples, o *k-NN* não é muito robusto quando os dados apresentam *outliers* (Sugiyama, 2015). Para a classificação, o modelo necessita de uma regra de decisão, que geralmente consiste em atribuir à classe desconhecida a classe da maioria dos *k* vizinhos mais próximos. Além disso, um valor prévio de *k*, assim como a métrica de distância a ser utilizada, devem ser definidos previamente. Nesta pesquisa foram utilizados os valores de *k=1, 3 e 5*, e a distância euclidiana foi utilizada como métrica de distância.

Floresta Aleatória: é uma coleção de árvores de decisão (Breiman et al., 1984). A árvore de decisão é um classificador estruturado na forma de um grafo com nós de decisão e arcos, aos quais estão associadas às regras que serão aplicadas na classificação (Fürnkranz et al., 2012). A classificação de amostras desconhecidas resulta do conjunto de decisões, segundo critérios específicos, que são tomadas ao se percorrer à árvore. Critérios de decisão que podem ser utilizados são, por exemplo, a entropia e o índice *gini*. De forma geral, a árvore de decisão apresenta bons desempenhos em tarefas de classificação. Contudo, para problemas que envolvam grandes conjuntos de dados e alta variância, a FA pode ser mais adequada. Uma das formas de se implementar a FA é a utilização do *bagging*. O *bagging* particiona o conjunto de treino em diversos subconjuntos. Para cada subconjunto, *features* são selecionadas e um modelo é treinado formando diversas árvores cujos resultados são combinados. Ao estimar uma entrada, cada árvore irá classificar o subconjunto de forma independente, sendo que a decisão a respeito da classe que a amostra desconhecida pertence será resultado de um processo de votação, e a classe atribuída será aquela com maior número de votos. Nesta pesquisa foi utilizada uma coleção de 100 árvores.

SVM: são classificadores lineares binários não probabilísticos que se baseiam na teoria do aprendizado estatístico (Cortes and Vapnik, 1995). Para a classificação, o algoritmo busca um hiperplano de separação entre duas classes, de forma a maximizar a distância entre dois objetos mais próximos. Em problemas não-lineares, o *SVM* utiliza funções *kernel*, as quais, realizam a transformação do

espaço dimensional dos dados não linearmente separáveis para uma dimensão superior, onde pode ser encontrado um limite de decisão linear. Por se tratar de um classificador binário, problemas de classificação que envolvam mais de duas classes requerem a aplicação de métodos multiclases. Os métodos multiclases mais utilizados são *one-vs-one* e *one-vs-all*. No primeiro, também conhecido como todos-contra-todos, dadas *n* classes são gerados $n(n-1)/2$ classificadores binários, sendo cada um deles responsável por determinar apenas um par de classes. O resultado final é definido pelo método de votação, a classe mais votada é atribuída à amostra. No segundo método, conhecido também como um-contra-todos, para um problema com *n* classes são gerados *n* classificadores binários, cada um especializado na sua classe. Essa técnica pode apresentar desvantagem quando ocorre um desbalanceamento considerável no conjunto de dados (Faceli et al., 2011). Como decisão de projeto, foi utilizado o método multiclasse *one-vs-one* e *kernel* do tipo *Linear*.

RNA/MLP: são compostas por uma ou mais camadas intermediárias de neurônios e uma camada de saída. Na topologia *MLP*, o fluxo de informações se inicia na camada de entrada e, na sequência, percorre as camadas intermediárias até alcançarem a camada de saída (Haykin, 2010). Para o ajuste dos pesos, essa arquitetura de rede utiliza o algoritmo *backpropagation*, que consiste em duas etapas. Esse ajuste ocorre em dois estágios: *forward* e *backward*. Na primeira, *forward*, cada neurônio da primeira camada recebe a amostra e aplica a função de ativação. A saída dessa camada é utilizada como entrada para a camada seguinte, assim sucessivamente até que seja obtida a saída da rede. Na segunda etapa, *backward*, a saída da rede é comparada com o valor desejado e o valor do erro é calculado. Os valores dos pesos sinápticos e dos limiares dos neurônios são ajustados num caminho que percorre a camada de saída até a primeira camada intermediária. Existe uma grande variação do algoritmo *backpropagation*, neste projeto foi utilizado o *Levenberg-Marquardt*, (Hagan and Menhaj, 1994). A rede neural utilizada nesse trabalho contém duas camadas ocultas, sendo a primeira com quarenta neurônios e a segunda com vinte, e o número de épocas no treinamento foi limitado a vinte.

3.2 Transformada Wavelet

A TW é uma ferramenta matemática eficiente para processamento de sinais, notadamente quando aplicada a sinais não-estacionários. Uma das vantagens de sua utilização é a possibilidade de, ao mesmo tempo, analisar um trecho específico, ou um intervalo do sinal, retendo as informações espectrais importantes e não-visíveis no domínio original, característica que pode ser vantajosa para aplicações de detecção de falhas (Al-Badour et al., 2011).

A TWC (Strang and Nguyen, 1996) pode ser definida por

$$TWC(a, b) = \frac{1}{\sqrt{|a|}} \int_{-\infty}^{+\infty} x(t) \psi \left(\frac{t-b}{a} \right) dt, \quad (1)$$

com $a \neq 0$, onde $x(t)$ é o sinal a ser analisado e a função $\psi(t)$ é denominada wavelet-mãe. A Equação 1 pode ser interpretada como a convolução do sinal no domínio do

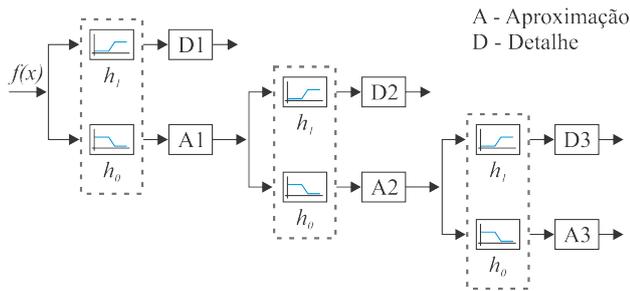


Figura 1. Decomposição do sinal $f(x)$ em análise multiresolução.

tempo. A função $\psi(t - b/a)$ corresponde a wavelet-mãe dilatada/contrainda pela escala a e transladada pelo fator b (Strang and Nguyen, 1996). O termo $1/\sqrt{|a|}$ é utilizado para assegurar que os valores da TWC em todas as escalas possam ser diretamente comparados, mantendo constante a energia (Strang and Nguyen, 1996).

A versão discreta da transformada pode ser diretamente obtida da TWC pela discretização dos parâmetros a e b , fazendo $a = a_0^j$ e $b = kb_0a_0^j$, sendo j, k valores inteiros e $a_0 > 1, b_0 > 0$. Fazendo $a_0 = 2$ e $b_0 = 1$, obtém-se um caso particular de discretização denominado discretização diádica (Strang and Nguyen, 1996). Neste caso, pode-se obter a Transformada *Wavelet* Discreta de forma computacionalmente eficiente através do uso de bancos de filtros digitais passa-altas e passa-baixas. A Figura 1 apresenta o esquema dos bancos de filtros de decomposição para cálculo da TWD. Nesse exemplo, foi considerado três níveis de resolução para a decomposição do sinal de entrada $f(x)$.

A estrutura do banco de filtros de decomposição consiste de um par de filtros, sendo um passa-baixas (h_0) e um passa-altas (h_1), seguidos por operações de subamostragem. As saídas subamostradas dos filtros passa-baixas e passa-altas são denominadas coeficientes de aproximação e detalhe, respectivamente. As operações de filtragem/subamostragem podem ser reaplicadas, recursivamente, aos coeficientes de aproximação até o número de níveis de resolução especificado pelo analista. Os filtros h_0 e h_1 são de comprimento finito e, portanto, cada coeficiente de aproximação e detalhe corresponde a um trecho do sinal original. Como já anteriormente destacado esta característica de localização espacial é uma das principais vantagens da transformada *Wavelet* sobre a transformada de Fourier.

4. O CONJUNTO DE DADOS MAFAULDA

Para a realização dos experimentos utilizou-se o conjunto de dados de propriedade do *Signals, Multimedia, and Telecommunications Laboratory - COPPE/UFRJ*, que pode ser obtido livremente em <http://www02.smt.ufrj.br/~offshore/mfs/index.html#T0C1>. O conjunto de dados é constituído por 1951 sinais de vibração, distribuídos entre dez classes, colhidos na bancada experimental *Machinery Fault Database - Mafaulda*, para análise e diagnóstico das falhas em máquinas rotativas. A bancada é composta por um motor CC de 0.25 CV, e a faixa de velocidade utilizada nos testes variou de 700 a 3600 RPM. Para aquisição

dos sinais de vibração, foram utilizados sensores do tipo acelerômetro fixados sobre os mancais do equipamento, configurados para realizar medidas nas direções radiais e axiais. A coleta dos dados foi realizada com uma taxa de amostragem de 50 kHz durante 5 segundos, totalizando 250 mil pontos por amostra. O experimento contém dez condições de operação do equipamento sendo (i) operação normal (NR), (ii) desalinhamento paralelo (DP), (iii) desalinhamento angular (DA), (iv) desbalanceamento (DB), (v) falha no elemento rolante do mancal 1 (FRM1), (vi) falha na gaiola do mancal 1 (FGM1), (vii) falha na pista externa do mancal 1 (FPM1), (viii) falha no elemento rolante do mancal 2 (FRM2), (ix) falha na gaiola do mancal 2 (FGM2) e (x) falha na pista externa do mancal 2 (FPM2). A Tabela 1 mostra a quantidade de amostras para cada condição de operação.

Tabela 1. Número de amostras por classe do conjunto de dados *Mafaulda*.

Classe	Condição de operação	Nº de amostras
C1	Normal	49
C2	Desalinhamento paralelo	197
C3	Desalinhamento angular	301
C4	Desbalanceamento	333
C5	Falha no elemento rolante do mancal 1	186
C6	Falha na gaiola do mancal 1	188
C7	Falha na pista externa do mancal 1	184
C8	Falha no elemento rolante do mancal 2	137
C9	Falha na gaiola do mancal 2	188
C10	Falha na pista externa do mancal 2	188
Total		1951

As falhas por desalinhamentos foram simuladas com a instalação de calços nos suportes dos mancais (variando de 0mm a 2mm). Para simular as falhas por desbalanceamento, massas variando de 6 a 35 gramas foram acrescentadas no rotor. Devido a complexidade das falhas em mancais de rolamento, o fabricante da bancada forneceu rolamentos defeituosos com três tipos de falhas distintas sendo (i) falha na gaiola, (ii) falha na pista externa e (iii) falha no elemento rolante. Cada rolamento defeituoso foi colocado, um por vez, em posições diferentes entre o rotor e o motor (mancal 1), e na posição externa (mancal 2). A Figura 2 ilustra a bancada experimental e a posição de instalação dos acelerômetros. Os pontos de coleta dos sinais de vibração estão localizados nos mancais 1 e 2 e são representados por 1A, 1H e 1V referentes ao mancal 1, e os pontos 2A, 2H e 2V referentes ao mancal 2. Os sensores radiais são identificados como vertical (V) e horizontal (H), e permitem identificar falhas geradas por forças no sentido perpendicular ao eixo central do equipamento. As medições no sentido axial (A) permitem identificar vibrações por movimentos paralelos ao eixo (Mobley, 1999).

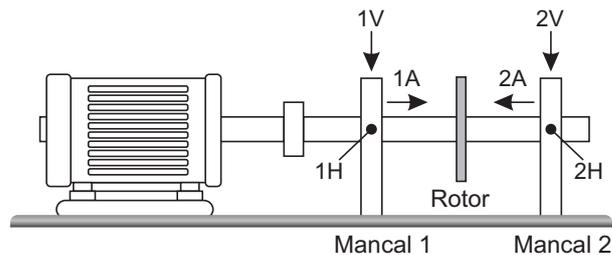


Figura 2. Bancada experimental *Mafaulda*.

5. METODOLOGIA

Esta seção apresenta as etapas executadas para a realização dos experimentos. São descritas as etapas de extração de características, pré-processamento, normalização dos dados, redução de dimensionalidade e treinamento dos classificadores.

5.1 Extração de características

- (1) Domínio do Tempo: na ocorrência de uma falha, as estruturas mecânicas rígidas no entorno do equipamento são afetadas produzindo sinais de vibração. De acordo com Rauber et al. (2015), Nayana and Geethanjali (2017) e Tahir et al. (2017), o comportamento da falha representado pelo sinal de vibração pode ser adequadamente descrito por parâmetros estatísticos tais como o valor médio (X_m), a amplitude da raiz (X_{root}), a raiz quadrada média (X_{rms}), o valor de pico (X_{peak}) e o desvio padrão (X_{std}). Esses atributos refletem as variações na intensidade de uma falha consolidada, porém, não são sensíveis à detecção de falhas em estágios iniciais. Para compensar a baixa sensibilidade dos parâmetros anteriormente mencionados, outras grandezas estatísticas podem ser utilizadas, entre elas, a assimetria ($X_{skewness}$), a curtose ($X_{kurtosis}$), o fator de crista (X_{crest}), o fator de folga ($X_{clearance}$), o fator de forma (X_{shape}) e o fator de impulso ($X_{impulse}$). A utilização dos referidos parâmetros, para representar os sinais de vibração no domínio do tempo, pode resultar em sistemas de diagnósticos sensíveis a falhas precoces, e que ainda mantenham o bom desempenho com o aumento da severidade destas (Lei, 2016). A Tabela 2 apresenta a descrição matemática dos parâmetros estatísticos extraídos do sinal no domínio do tempo. Para cada amostra de sinal, obtida por cada um dos seis sensores, onze parâmetros foram extraídos, constituindo assim um conjunto de dados com 1951 amostras e 66 descritores.

Tabela 2. Parâmetros estatísticos extraídos do sinal no domínio do tempo (Lei, 2016).

Parâmetro	Descrição Matemática
Valor médio	$X_m = \frac{\sum_{n=1}^N x(n)}{N}$
Valor de pico	$X_{peak} = \max x(n) $
Fator de folga	$X_{clearance} = \frac{X_{peak}}{X_{root}}$
Amplitude da raiz	$X_{root} = \left(\frac{\sum_{n=1}^N \sqrt{ x(n) }}{N} \right)^2$
Curtose	$X_{kurtosis} = \frac{\sum_{n=1}^N (x(n) - X_m)^4}{(N-1)X_{std}^4}$
Fator de impulso	$X_{impulse} = \frac{X_{peak}}{\frac{1}{N} \sum_{n=1}^N x(n) }$
Desvio padrão	$X_{std} = \sqrt{\frac{\sum_{n=1}^N (x(n) - X_m)^2}{N-1}}$
Skewness	$X_{skewness} = \frac{\sum_{n=1}^N (x(n) - X_m)^3}{(N-1)X_{std}^3}$
Fator de forma	$X_{shape} = \frac{X_{rms}}{\frac{1}{N} \sum_{n=1}^N x(n) }$
Raiz quadrada média	$X_{rms} = \sqrt{\frac{\sum_{n=1}^N (x(n))^2}{N}}$
Fator de crista	$X_{crest} = \frac{X_{peak}}{X_{rms}}$

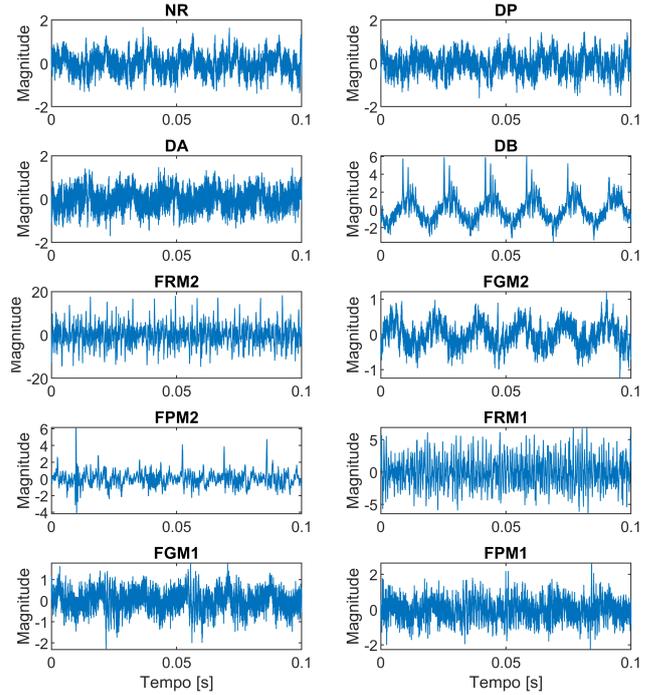


Figura 3. Sinais de vibração no domínio do tempo para cada condição de operação.

A Figura 3 apresenta o comportamento do sinal de vibração capturado pelo sensor 1V para cada condição de operação no período de 0,1 segundos, numa velocidade próxima de 60 Hz. Pode-se notar que as diferentes condições de operação são refletidas nas diferenças entre as magnitudes em cada caso. Dessa forma, descritores baseados em parâmetros estatísticos no domínio do tempo podem representar diferentes padrões de comportamentos para cada natureza de falha.

- (2) Domínio tempo/frequência: para a extração de descritores no domínio tempo/frequência foi aplicada a análise multiresolução, utilizando a TW como ferramenta para sua implementação. Considerando sua utilização em trabalhos anteriores similares, nesse trabalho, para a decomposição dos sinais de vibração, foi aplicada a função *Wavelet db10* da família *Daubchies*, e os sinais foram decompostos em 10 níveis de resolução. A Figura 4 apresenta a árvore de decomposição dos sinais, assim como as frequências em cada nível de aproximação e detalhe após aplicação dos filtros passa-baixa e passa-alta. Nessa figura, S representa o sinal original, Ai e Di representam os sinais de aproximação e detalhe no i-ésimo nível. Para a representação dos sinais de vibração no domínio tempo/frequência, foi calculado o valor da entropia de *Shannon*. Este indicador reflete a distribuição de energia contida num sinal, sendo definida matematicamente em (2) (Correa and Guzman, 2020).

$$H_E = - \sum_{j=1}^n p_j \log p_j. \quad (2)$$

A Figura 5 apresenta o valor médio da entropia de *Shannon* para cada nível de decomposição em cada condição de operação. Nos gráficos, a primeira barra

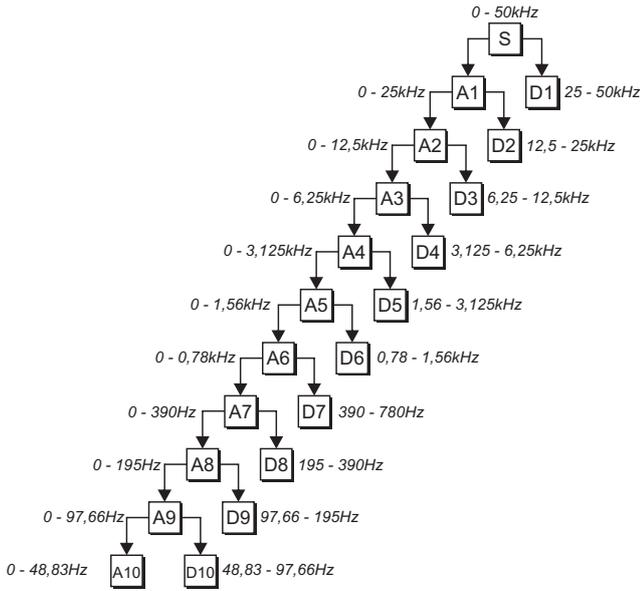


Figura 4. Árvore de decomposição utilizando a *Wavelet db10*.

representa o valor da entropia de *Shannon* do sinal de aproximação (A10) e as demais barras representam, na ordem, as entropias nos níveis de detalhe de D1 a D10. Para cada sensor, onze valores de entropia foram extraídos, totalizando sessenta e seis descritores.

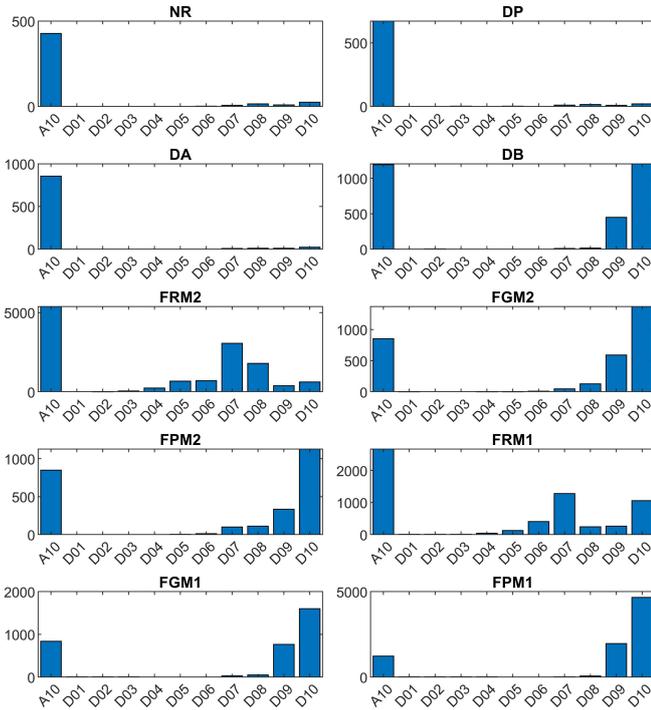


Figura 5. Entropia nos níveis de decomposição para cada condição de operação.

5.2 Normalização dos Dados

A normalização é uma transformação numérica utilizada quando o conjunto de dados possui atributos com escalas e ordens de grandeza diferentes. Sua aplicação tem por

objetivo evitar que o valor de um atributo predomine sobre os outros, e dessa forma exerça maior influência no modelo, ainda que não seja o mais relevante (Faceli et al., 2011). A técnica de normalização utilizada nessa pesquisa foi a denominada normalização por reescala, definida em (3), no qual os valores dos atributos estarão dentro do intervalo de zero a um.

$$X_{novo} = \frac{X_{atual} - \min}{\max - \min}. \quad (3)$$

5.3 Análise de Componentes Principais

Principal Component Analysis (PCA) é uma técnica de extração de características, na qual novos descritores são encontrados através de combinações lineares dos descritores originais (Pearson, 1901). A *PCA* busca obter uma representação do espaço amostral em uma região com menor dimensionalidade, através da realização de projeções ortogonais do espaço original, de maior dimensão, sobre um espaço com menor dimensão.

A Figura 6 ilustra a representação dos dados em três dimensões, e ao lado a representação bidimensional das duas primeiras componentes principais. O objetivo é encontrar uma representação espacial de menor dimensão em um conjunto de dados maximizando a variância do conjunto.

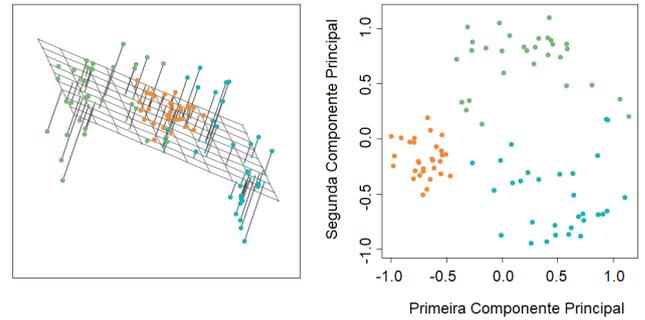


Figura 6. Redução da dimensão do espaço amostral utilizando a *PCA*.

Nesta pesquisa, a *PCA* foi utilizada para a redução da quantidade de descritores no domínio do tempo e no domínio tempo/frequência. Em ambos os casos, a variância retida foi 99% da variância total contida no conjunto original. A Figura 7 apresenta a relação entre o número de componentes principais e a variância explicada quando a *PCA* foi aplicada aos descritores no domínio do tempo e no domínio tempo/frequência. Observa-se que, quando aplicada no domínio do tempo, a *PCA* reduz a quantidade de descritores de 66 para 27, o que corresponde a uma redução de 59,09%. Já para os descritores no domínio tempo/frequência, obtidos pela decomposição *Wavelet* dos sinais, houve uma redução de 66 para 37 descritores, o que equivale a uma redução de 43,94% do total.

5.4 Treinamento do Classificador

Para a modelagem, o conjunto de dados foi particionado em treino e teste, com 70% e 30% das amostras, respectivamente. Os modelos foram obtidos utilizando os algoritmos *k-NN*, *SVM*, *FA* e *RNA/MLP*, que foram treinados com

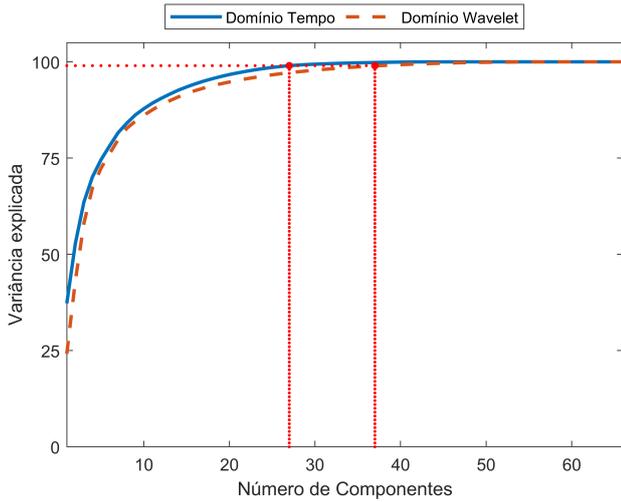


Figura 7. Variância explicada através da *PCA*.

validação cruzada utilizando 10 partições. Como métrica de avaliação de desempenho foi utilizada a acurácia. Todos os testes foram realizados no Matlab®, R2019a.

6. RESULTADOS

A Tabela 3 apresenta os resultados dos testes no domínio do tempo. Para os testes, duas configurações do conjunto de dados foram utilizadas sendo (i) utilização de todos os sessenta e seis descritores do conjunto de dados (66-D) e (ii) utilização das vinte e sete componentes principais (27-CP) obtidas com a aplicação da *PCA*. Em todos os casos foi verificado o desempenho do classificador na fase de treino e de teste.

Tabela 3. Desempenho dos classificadores utilizando descritores estatísticos no domínio do tempo.

Classificadores	Acurácia (%)			
	66-D		27-CP	
	Treino	Teste	Treino	Teste
1-NN	91.80	79.83	91.80	80.51
3-NN	90.34	81.54	90.48	80.85
5-NN	88.87	81.20	88.95	81.71
SVM	95.31	28.55	92.75	75.38
RNA/MLP	97.00	73.68	98.75	51.97
FA	96.56	90.43	91.15	81.20

Como pode ser observado na Tabela 3, de forma geral, todos os classificadores apresentaram bons desempenhos. Utilizando o conjunto original de variáveis (66-D), o melhor resultado foi obtido com o modelo FA com uma taxa de acerto de 90,43%. Os resultados utilizando o conjunto reduzido através da *PCA* (27-CP) obtiveram desempenhos inferiores, nesta configuração, o melhor resultado foi obtido com o modelo 5-NN com 81,71% de acurácia.

Testes utilizando os descritores baseados no domínio tempo/frequência, com a aplicação da análise multiresolução utilizando a TW, foram realizados utilizando a mesma abordagem adotada para os descritores no domínio do tempo. Foram realizados testes utilizando (i) todo o conjunto original de descritores (66-D) e (ii) as trinta e sete componentes principais (37-CP) obtidas pela *PCA*, os resultados estão apresentados na Tabela 4.

Tabela 4. Desempenho dos classificadores utilizando descritores extraídos no domínio tempo/frequência.

Classificadores	Acurácia (%)			
	66-D		37-CP	
	Treino	Teste	Treino	Teste
1-NN	88.80	86.67	87.41	85.98
3-NN	84.41	84.44	83.60	82.74
5-NN	84.11	84.10	82.43	82.91
SVM	98.98	79.83	90.41	77.09
RNA/MLP	95.24	83.59	94.65	76.92
FA	98.98	98.46	93.48	88.03

Conforme pode ser visto na Tabela 4, todos os modelos apresentaram bons resultados de classificação. O modelo baseado em FA obteve os melhores resultados em todos os testes, utilizando o conjunto original de descritores (66-D) foi obtido uma acurácia de 98,46%, e utilizando o conjunto reduzido (37-CP) a taxa de acerto foi de 88,03%.

Quando analisado o desempenho dos classificadores, verifica-se que a utilização de um único descritor (entropia de Shannon), no domínio Wavelet, supera os resultados dos classificadores que utilizaram descritores extraídos no domínio do tempo, ou seja, em todos os testes, os resultados dos modelos utilizando descritores baseados no domínio tempo/frequência foram superiores. Pode-se notar também que o modelo utilizando FA obteve o melhor resultado com acurácia de 98,46% utilizando descritores no domínio tempo/frequência. A redução da dimensionalidade através da *PCA*, mantendo 99% de variância, reduziu o desempenho dos classificadores nos dois domínios em que os testes foram realizados, o que significa que informações relevantes são perdidas.

Para fins de comparação dos desempenhos nos domínios do tempo e tempo/frequência, a Figura 8 apresenta os resultados de forma gráfica, para a modelagem com todos os descritores (66-D) e utilizando o conjunto reduzido através da *PCA*. A metodologia através da decomposição utilizando a TW obteve, no geral, desempenhos superiores se comparado ao método utilizando descritores no domínio do tempo.

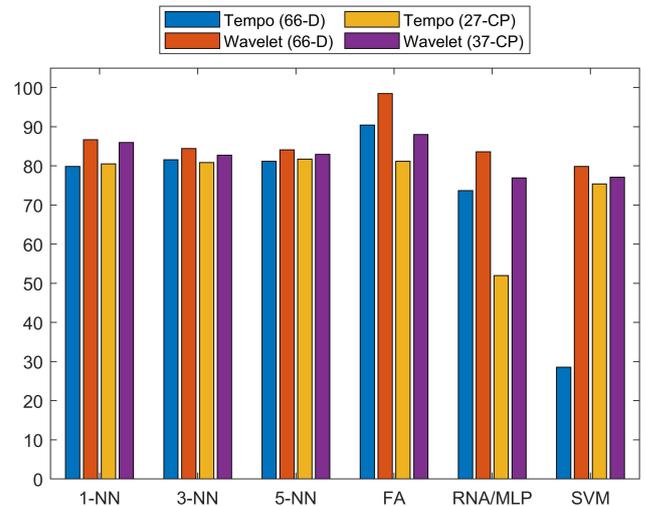


Figura 8. Desempenho entre os métodos no domínio do tempo e Wavelet.

7. CONCLUSÃO

Os resultados obtidos nos testes mostram a viabilidade de se implementar sistemas inteligentes de monitoramento em equipamentos dinâmicos. As amostras investigadas possuem características de separabilidade, o que facilita no bom desempenho dos algoritmos de classificação, no qual alguns modelos atingiram acurácia superior a 90%. No geral, o melhor desempenho foi obtido com o modelo FA (90,43% e 98,46%) obtidos pelos testes nos domínios do tempo e tempo/frequência, respectivamente, o que demonstra a possibilidade da elaboração diagnósticos precisos e confiáveis.

Nas estratégias envolvendo a redução da dimensionalidade através da *PCA* foram obtidas taxas de acerto de no máximo 88,03% utilizando o modelo FA (37-CP). Os resultados comprovam que informações relevantes são perdidas durante a transformação dos dados.

De forma geral, os resultados demonstram que técnicas de reconhecimento de padrões ganham cada vez mais atenção em aplicações que envolvem o monitoramento de equipamentos de grande porte, ou até mesmo plantas de processos. No contexto atual, o monitoramento das variáveis de processos, combinadas com dados históricos dos equipamentos, juntamente com a utilização de técnicas de *machine learning* formam uma ferramenta poderosa no diagnóstico de falhas. Totalmente alinhada ao conceito de *Prognostics and Health Management*, podem estimar a ocorrência de falhas programando intervenções quando estas ainda se encontram em fase inicial, o que reduz os custos de manutenção. A intervenção sobre o processo, ainda quando as falhas não estão totalmente consolidadas, evita ocorrências emergenciais que, de forma geral, são mais prejudiciais pois resultam em perdas de produção, impactos ambientais e danos aos equipamentos ou pessoas.

REFERÊNCIAS

- Al-Badour, F., Sunar, M., and Cheded, L. (2011). Vibration analysis of rotating machinery using time-frequency analysis and wavelet techniques. *Mechanical Systems and Signal Processing*, 25(6), 2083–2101.
- Breiman, L., Friedman, J., Stone, C.J., and Olshen, R.A. (1984). *Classification and Regression Trees*. CRC press.
- Chen, Z., Gryllias, K., and Li, W. (2019). Mechanical fault diagnosis using convolutional neural networks and extreme learning machine. *Mechanical Systems and Signal Processing*, 133, 106272.
- Correa, J.C.J. and Guzman, A.L. (2020). *Mechanical vibrations and condition monitoring*. Elsevier.
- Cortes, C. and Vapnik, V. (1995). Support-vector networks. *Machine learning*, 20(3), 273–297.
- Faceli, K., Lorena, A.C., Gama, J., and de Carvalho, A.C. (2011). *Inteligência Artificial: Uma Abordagem de Aprendizagem de Máquina*, volume 1. LTC, Rio de Janeiro.
- Fayed, H.A. and Atiya, A.F. (2009). A novel template reduction approach for the k -nearest neighbor method. *IEEE Transactions on Neural Networks*, 20(5), 890–896.
- Ford, J. (2017). Machine downtime costs uk manufacturers £180bn a year. the engineer. URL <https://www.theengineer.co.uk/faulty-machinery-machine-manufacturers>.
- Fürnkranz, J., Gamberger, D., and Lavrač, N. (2012). *Foundations of rule learning*. Springer Science & Business Media.
- Hagan, M.T. and Menhaj, M.B. (1994). Training feed-forward networks with the marquardt algorithm. *IEEE transactions on Neural Networks*, 5(6), 989–993.
- Haykin, S. (2010). *Neural Networks and Learning Machines*, 3ed. Pearson Education India.
- Kankar, P., Sharma, S.C., and Harsha, S. (2011). Fault diagnosis of ball bearings using machine learning methods. *Expert Systems with Applications*, 38(3), 1876 – 1886. doi:<https://doi.org/10.1016/j.eswa.2010.07.119>. URL <http://www.sciencedirect.com/science/article/pii/S0957417410007475>.
- Lei, Y. (2016). *Intelligent fault diagnosis and remaining useful life prediction of rotating machinery*. Butterworth-Heinemann.
- Lei, Y., Lin, J., He, Z., and Zuo, M.J. (2013). A review on empirical mode decomposition in fault diagnosis of rotating machinery. *Mechanical systems and signal processing*, 35(1-2), 108–126.
- Mobley, R.K. (1999). *Vibration fundamentals*. Elsevier.
- Nayana, B.R. and Geethanjali, P. (2017). Analysis of statistical time-domain features effectiveness in identification of bearing faults from vibration signal. *IEEE Sensors Journal*, 17(17), 5618–5625.
- Pearson, K. (1901). Liii. on lines and planes of closest fit to systems of points in space. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 2(11), 559–572.
- Prabhakar, S., Mohanty, A., and Sekhar, A. (2002). Application of discrete wavelet transform for detection of ball bearing race faults. *Tribology International*, 35(12), 793 – 800. doi:[https://doi.org/10.1016/S0301-679X\(02\)00063-4](https://doi.org/10.1016/S0301-679X(02)00063-4). URL <http://www.sciencedirect.com/science/article/pii/S0301679X02000634>.
- Rauber, T.W., de Assis Boldt, F., and Varejão, F.M. (2015). Heterogeneous feature models and feature selection applied to bearing fault diagnosis. *IEEE Transactions on Industrial Electronics*, 62(1), 637–646.
- Ribeiro, F., Marins, M., Netto, S., and da Silva, E. (2017). Rotating machinery fault diagnosis using similarity-based models. doi:10.14209/sbrt.2017.133.
- Santiago, D.F.d.A. et al. (2004). Diagnóstico de falhas em máquinas rotativas utilizando transformada de wavelet e redes neurais artificiais.
- Strang, G. and Nguyen, T. (1996). *Wavelet and Filter Banks*. Wellesley: Cambridge Press.
- Sugiyama, M. (2015). *Introduction to statistical machine learning*. Morgan Kaufmann.
- Tahir, M.M., Khan, A.Q., Iqbal, N., Hussain, A., and Badshah, S. (2017). Enhancing fault classification accuracy of ball bearing using central tendency based time domain features. *IEEE Access*, 5, 72–83.