

Detecção de Falhas em Bombas Industriais baseada em Análise Cicloestacionária de Sinais Acústicos, Decomposição em Modos Empíricos e 1D-CNN

Ayrton Silva, Agostinho Brito Júnior, Luiz Silveira Júnior, Luiz Felipe Silveira

Resumo—A necessidade de se detectar precocemente falhas em máquinas industriais cresceu significativamente devido à demanda por confiabilidade, disponibilidade e redução do custo de manutenção desses sistemas. Neste trabalho, propõe-se uma abordagem para detecção de falhas em bombas industriais, a partir da análise de sinais acústicos, que envolve Decomposição em Modos Empíricos, extração de características cicloestacionárias e 1D-CNN, buscando aumentar a robustez e a sensibilidade da detecção de falhas em condições de ruído elevado. A solução proposta mostra-se robusta, atingindo acurácias de classificação superiores a 90% no conjunto de sinais utilizado.

Palavras-Chave—Detecção de falhas em máquinas, EMD, 1D-CNN, Análise cicloestacionária, Aprendizagem de máquina.

Abstract—The need for early detection of faults in industrial machines has grown significantly due to the demand for reliability, availability, and reduced maintenance costs of systems. This work proposes an approach for fault detection in industrial pumps using acoustic signals that involves decomposition into empirical modes, extraction of cyclostationary features, and 1D-CNN, aiming to increase the robustness and sensitivity of fault detection under high noise conditions. The proposed solution proves to be robust, achieving classification accuracies greater than 90% in the set of signals used.

Keywords—Fault detection in machinery, EMD, 1D-CNN, Cyclostationary analysis, Machine learning.

I. INTRODUÇÃO

A confiabilidade operacional de máquinas industriais rotativas é crucial na indústria moderna, pois falhas podem causar paradas inesperadas. Nesse cenário, a detecção precoce de degradações nos sistemas torna-se essencial para reduzir custos de manutenção, evitar interrupções não planejadas e ampliar a vida útil dos equipamentos [1].

Técnicas de aprendizado de máquina têm impulsionado abordagens modernas orientadas por dados para tornar a detecção de falhas mais rápida, precisa e automatizada [1]. Em [2], por exemplo, os autores incorporaram a análise *wavelet* multirresolução diretamente na arquitetura de uma Rede Neural Convolucional (CNN) para analisar sinais de vibração unidimensionais de caixas de engrenagens, enquanto que em [3], redes convolucionais em grafos foram construídas a partir da decomposição *wavelet* de sinais de vibração.

A Decomposição em Modos Empíricos (EMD - *Empirical Mode Decomposition*), conforme discutido por Zhang et al.

[4], também tem sido empregada tanto para extração de características quanto para filtragem adaptativa, na qual determinadas Funções de Modo Intrínseco (IMFs) são removidas do sinal original com o objetivo de reduzir componentes ruidosos. Técnicas de pós-processamento das IMFs com representações bidimensionais vêm sendo exploradas a fim de usá-las como entradas em redes neurais convolucionais com bastante sucesso.

A maioria dos trabalhos presentes na literatura, a exemplo de [2] e [3], utilizam sinais de vibração para a detecção automática de falhas em máquinas industriais. Neste trabalho, propomos uma abordagem de detecção de falhas em bombas industriais baseada na análise de sinais acústicos reais adquiridos em ambiente industrial por meio de equipamento de aquisição de baixo custo [11].

Em comparação com os sinais de vibração, os sinais acústicos apresentam vantagens práticas relevantes, como aquisição não intrusiva, menor custo de instrumentação e maior simplicidade de instalação. Entretanto, por serem captados no ar, esses sinais são mais suscetíveis à interferência de ruídos presentes no ambiente industrial [5]. Essa limitação pode reduzir sua capacidade discriminativa, especialmente quando as assinaturas associadas às falhas se manifestam como modulações de baixa intensidade, facilmente mascaradas por ruído ou por variações normais de operação.

Este trabalho propõe uma metodologia que integra EMD e descritores cicloestacionários como etapas de pré-processamento, seguidas por um classificador 1D-CNN (Rede Neural Convolucional Unidimensional). A hipótese primária é que a extração cicloestacionária aplicada às IMFs produza representações mais robustas e discriminativas, explorando tanto a decomposição adaptativa do sinal quanto a sensibilidade da análise cicloestacionária a periodicidades estatísticas e modulações frequentemente induzidas por falhas em máquinas rotativas, mesmo sob ruído elevado.

Este trabalho é organizado em mais quatro seções. A Seção II sumariza fundamentos teóricos das ferramentas usadas. A Seção III apresenta detalhes importantes na aquisição e preparação dos dados, bem como da arquitetura do modelo 1D-CNN e das etapas de pré-processamento empregadas. Seguindo, as Seções IV e V apresentam e discutem, respectivamente, os resultados obtidos, e as conclusões da proposta.

II. FUNDAMENTAÇÃO TEÓRICA

Nesta seção são discutidos os fundamentos teóricos essenciais ao entendimento do trabalho, apresentando técnicas

Ayrton Silva, Programa de Pós-Graduação em Eng. Elétrica e de Computação, UFRN, ayrton.pinheiro.701@ufrn.edu.br; Agostinho Brito Júnior, DCA, UFRN, Natal, ambj@dca.ufrn.br; Luiz Silveira Júnior, DCO, UFRN, Natal, luiz.silveira@ufrn.br; Luiz Felipe Silveira, DCA, UFRN, Natal, felipe.silveira@ufrn.br.

capazes de decompor os sinais brutos de áudio em componentes oscilatórios e, destas, extrair padrões cicloestacionários para treinar o modelo de inteligência artificial utilizado para a detecção de falhas no sistema.

A. Decomposição em Modos Empíricos

A EMD foi proposta por Huang et al. [6] como um método adaptativo para decompor um sinal em componentes oscilatórios denominados Funções de Modo Intrínseco (IMF - *Intrinsic Mode Function*) e um resíduo final. Cada IMF pode ser entendida como uma oscilação que pode variar tanto em amplitude quanto em frequência ao longo do tempo. Isso significa que as IMFs não são componentes fixos, mas sim modos oscilatórios que se adaptam às características locais do sinal. Ou seja, diferente de outras técnicas de decomposição, as IMFs geradas pela EMD podem se ajustar dinamicamente ao comportamento do sinal.

Cada IMF é uma função que satisfaz duas condições: (1) em todo o conjunto de dados, o número de extremos e o número de cruzamentos por zero devem ser iguais ou diferir no máximo em um; (2) em qualquer ponto, o valor médio entre o envelope definido pelos máximos locais e o envelope definido pelos mínimos locais é zero ou aproximadamente zero [6].

A EMD é adequada para sinais não lineares e não estacionários, como sinais fisiológicos, mecânicos ou geofísicos, pois realiza a decomposição de forma adaptativa e empírica no domínio temporal, sem impor bases fixas como ocorre em métodos tradicionais. Além disso, as IMFs preservam características do sinal original e sua soma permite reconstruí-lo integralmente, possibilitando tanto o isolamento de componentes associados a falhas quanto a recombinação seletiva de IMFs para atenuar ruídos ou destacar padrões relevantes [6].

Esse processo, ilustrado na Figura 1, é descrito pelo algoritmo proposto por Huang [6]:

- 1) Identifique todos os máximos e mínimos;
- 2) Interpole os máximos e mínimos para formar o envelope superior e o envelope inferior (geralmente utiliza-se *cubic spline line* para interpolar). Observação: Os envelopes superior e inferior devem cobrir todos os dados entre eles;
- 3) Calcule a média dos envelopes e subtraia-a do sinal inicial;
- 4) Verifique se os passos anteriores geraram uma IMF (satisfazendo as duas condições das IMFs). Caso não seja IMF, repita os passos do 1 ao 4 neste sinal gerado;
- 5) Armazene a IMF extraída e a subtraia-a do sinal original, gerando um resíduo;
- 6) Repita todo o processo para o resíduo gerado até que este se torne um sinal monotônico (sinal que apresenta menos de 2 mínimos ou máximos locais).

B. Análise Cicloestacionária

Este trabalho propõe a análise de características cicloestacionárias das IMFs para a caracterização de modos de operação normal e em falha. Funções de análise cicloestacionária permitem caracterizar sinais cujas propriedades estatísticas variam

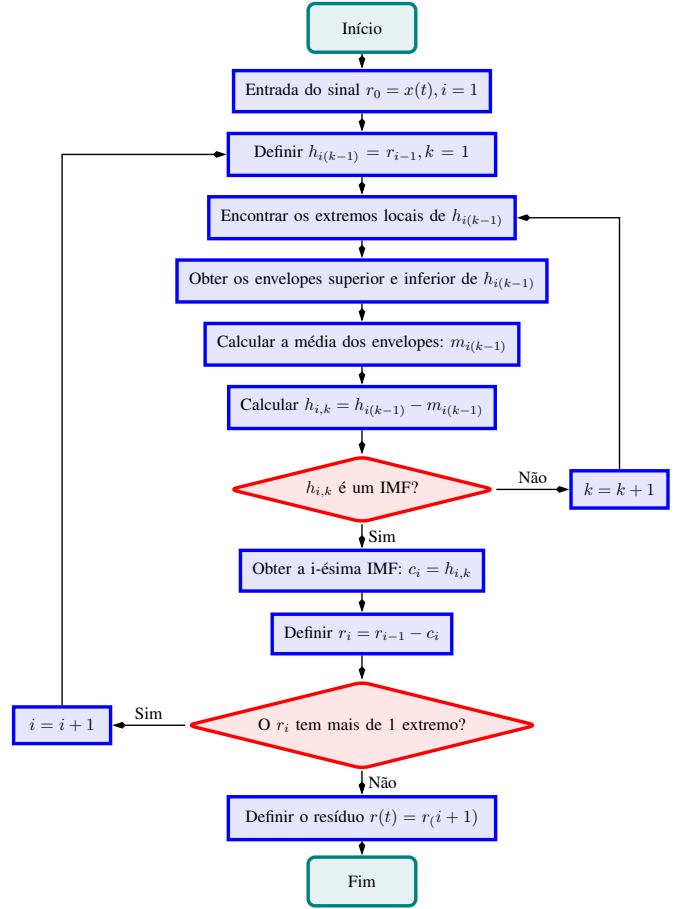


Fig. 1. Procedimento para obter as IMFs utilizando a técnica EMD. Adaptado de [7].

periodicamente no tempo, mesmo quando o sinal não é estritamente periódico [8]. Essa propriedade é útil na detecção de falhas em máquinas rotativas, pois defeitos mecânicos podem introduzir impulsos repetitivos e modulações em amplitude ou fase que nem sempre são evidentes no domínio do tempo ou na análise espectral convencional [9].

Neste trabalho, considera-se a cicloestacionariedade de segunda ordem, baseada na Função de Autocorrelação Cíclica (CAF - *Cyclic Autocorrelation Function*) e na Densidade Espectral Cíclica (SCD - *Spectral Cyclic Density*). Para um processo aleatório $x(t)$, a CAF [8] é definida por

$$R_x^\alpha(\tau) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} x\left(t + \frac{\tau}{2}\right) x^*\left(t - \frac{\tau}{2}\right) e^{-j2\pi\alpha t} dt$$

$$= \left\langle x\left(t + \frac{\tau}{2}\right) x^*\left(t - \frac{\tau}{2}\right) e^{-j2\pi\alpha t} \right\rangle_t \quad (1)$$

em que τ representa o atraso temporal, α é a frequência cíclica e $\langle \cdot \rangle_t$ indica média temporal. Valores significativos de $R_x^\alpha(\tau)$ para $\alpha \neq 0$ indicam a presença de periodicidade estatística de segunda ordem.

A SCD é obtida pela transformada de Fourier da CAF em relação ao atraso τ , conforme a Relação Cíclica de Wiener-Khinchin [10]:

$$S_x^\alpha(f) = \int_{-\infty}^{\infty} R_x^\alpha(\tau) e^{-j2\pi f\tau} d\tau \quad (2)$$

A SCD opera no domínio bidimensional formado pela frequência espectral f e pela frequência cíclica α , permitindo evidenciar correlações espectrais associadas a modulações periódicas. Assim, falhas em componentes rotativos podem se manifestar como componentes cíclicas específicas no plano (f, α) , mesmo quando mascaradas por ruído [7].

Para utilizar essa representação em uma CNN unidimensional, uma das formas é extrair, para cada frequência cíclica α , o valor máximo da SCD ao longo do eixo das frequências f . O vetor resultante, denominado perfil- α , descreve a intensidade das componentes cíclicas do sinal e é utilizado como descritor cicloestacionário neste trabalho.

C. Rede Neural Convolutacional Unidimensional

Ao contrário das CNNs tradicionais, que operam sobre imagens bidimensionais, as 1D-CNNs aplicam filtros diretamente sobre sinais unidimensionais. Dessa forma, ao trabalhar com um sinal de uma única dimensão, como no presente estudo, é mais conveniente fazer uso de uma 1D-CNN.

A operação básica de convolução para uma entrada unidimensional x e um kernel k de tamanho n é definida como

$$(x * k)(i) = \sum_{j=1}^n x(i+j) \cdot k(j), \quad (3)$$

em que $*$ representa a convolução e i é o índice da saída. De fato, esta operação equivale a uma correlação cruzada, ou seja, difere de uma convolução tradicional por usar o kernel na ordem direta [4].

III. METODOLOGIA

Esta seção descreve todo o fluxo de processamento proposto neste trabalho para a detecção de falhas em bombas industriais.

A. Aquisição e preparação dos dados

Para o desenvolvimento e validação da metodologia proposta, foi utilizado o conjunto de dados MIMII (*Malfunctioning Industrial Machine Investigation and Inspection*) [11], um banco público composto por gravações de sons de máquinas industriais operando em condições normais e anômalas. O MIMII contém sinais acústicos reais obtidos de quatro tipos de máquinas: válvulas, bombas, ventiladores e esteiras.

Cada sinal do banco possui duração de 10 segundos, com amostragem de 16 kHz. Para simular ambientes industriais sob diferentes condições de relação sinal-ruído (SNR), especificamente 6 dB, 0 dB e -6 dB, o banco de dados coletou ruídos reais de fábricas e os adicionou artificialmente às gravações originais das máquinas. Neste trabalho, foram usados apenas dados de bombas, considerando operações de sucção e descarga de água. As condições anômalas analisadas incluem problemas como vazamentos, contaminação e obstruções.

Os dados disponíveis estão organizados em quatro seções distintas: id_00, id_02, id_04 e id_06, que representam modelos diferentes de bomba. Cada seção contém amostras correspondentes às condições normal e anômala. Entretanto, observa-se um desbalanceamento significativo entre as classes, com aproximadamente 1000 segmentos normais e 100 segmentos anômalos por seção.

Devido a limitações computacionais e com o objetivo de manter o balanceamento entre as condições analisadas, foram selecionados os 60 primeiros segmentos de áudio para cada condição operacional, em cada seção e para cada nível de SNR. Dessa forma, obteve-se um total de 60 segmentos $\times$ 2 condições $\times$ 4 seções $\times$ 3 níveis de SNR, resultando em 1440 segmentos de áudio, dos quais 720 correspondem à condição normal e 720 à condição anômala. Para representar cenários mais próximos de aplicações reais, ainda que com maior complexidade de classificação, os dados provenientes de diferentes tipos de bombas e múltiplos níveis de SNR foram combinados de forma proporcional entre as classes normal e anômala. Em seguida, foram distribuídos de maneira estratificada entre os conjuntos de treinamento, validação e teste, nas proporções de 70%, 15% e 15%, respectivamente.

Apenas após essa divisão, os sinais de cada conjunto foram segmentados independentemente (evitando vazamento de dados entre os diferentes conjuntos) em janelas de 2048 amostras, com uma sobreposição de 80%. Assim, cada segmento de áudio original, que tem 160 mil amostras, produziu 387 novos segmentos menores, totalizando $387 \times 1440 = 557280$ blocos de 2048 amostras.

Todos os conjuntos passaram por uma normalização Z-score da forma

$$x' = \frac{x - \mu_{treino}}{\sigma_{treino}}, \quad (4)$$

em que x' é o conjunto de dados após a normalização, x é o conjunto de dados, μ_{treino} e σ_{treino} são, respectivamente, a média e o desvio padrão das amplitudes dos dados no conjunto de treino. Ou seja, todos os conjuntos de dados são normalizados usando os parâmetros do conjunto de treino.

B. Pré-processamento

Na etapa de pré-processamento, cada bloco de 2048 amostras é decomposto por EMD, produzindo um conjunto de IMFs. Como o número de IMFs varia entre os blocos, apenas as quatro primeiras foram mantidas, com base em análises preliminares com a arquitetura CNN. Os dados resultantes possuem dimensão $(557280, 4, 2048)$, correspondente ao número de blocos, IMFs e amostras por IMF, respectivamente.

Para cada segmento de áudio original, considera-se o conjunto de 387 blocos obtidos após a segmentação dos sinais. Em seguida, para cada IMF correspondente a esses blocos, as matrizes de SCD são obtidas por meio de uma adaptação do algoritmo para o cálculo da CAF descrito por Farias et al. [12]. Inicialmente, cada bloco de 2048 amostras foi segmentado em janelas de tamanho $N = 1024$, com sobreposição de 75%, resultando em $L = 5$ subjanelas por bloco.

Seja $x_{b,l}^{(m)}[n]$ a l -ésima subjanela extraída do bloco b , associada à m -ésima IMF, com $b = 0, 1, \dots, 386$, $l = 0, 1, \dots, L-1$, $m = 1, 2, 3, 4$ e $n = 0, 1, \dots, N-1$. Para cada subjanela, calcula-se

$$V_{b,l}^{(m)}[n, \tau] = x_{b,l}^{(m)}[n + \tau] \left(x_{b,l}^{(m)}[n - \tau] \right)^*, \quad (5)$$

em que τ representa o atraso discreto. Em seguida, para cada bloco b , obtém-se uma matriz média a partir das $L = 5$

TABELA I
 ARQUITETURA DA 1D-CNN UTILIZADA PARA CLASSIFICAÇÃO DOS PERFIS- α .

Camada	Composição da camada	Canais de entrada	Canais de saída	Tamanho do kernel	Stride	Padding	Tamanho da saída (canal $\times$ comprimento)
L1	Conv1d + (BatchNorm1d + LeakyReLU)	4	32	5	1	same	32×512
	MaxPool1d + (Dropout)	/	/	4	4	0	32×128
L2	Conv1d + (BatchNorm1d + LeakyReLU)	32	64	5	1	2	64×128
	MaxPool1d + (Dropout)	/	/	4	4	0	64×32
L3	Flatten	/	/	/	/	/	1×2048
L4	Totalmente conectada	/	/	/	/	/	1×2

subjanelas:

$$V_b^{(m)}[n, \tau] = \frac{1}{L} \sum_{l=0}^{L-1} V_{b,l}^{(m)}[n, \tau]. \quad (6)$$

Para cada IMF m , obtém-se um conjunto de 387 matrizes $V_b^{(m)}[n, \tau]$, associadas aos blocos do segmento de áudio original. Essas matrizes são então divididas em seis grupos temporais. Essa estratégia foi adotada para aumentar o número de amostras de treinamento, preservando a estabilidade das médias da SCD. Cinco grupos contêm 65 matrizes, enquanto o grupo restante contém 62 matrizes. Denotando por $\mathcal{B}_g$ o conjunto de índices das matrizes pertencentes ao grupo g , com $g = 1, 2, \dots, 6$, calcula-se a matriz média de cada grupo por

$$\bar{V}_g^{(m)}[n, \tau] = \frac{1}{|\mathcal{B}_g|} \sum_{b \in \mathcal{B}_g} V_b^{(m)}[n, \tau]. \quad (7)$$

Como cada matriz $V_b^{(m)}[n, \tau]$ resulta da média de $L = 5$ subjanelas, cada grupo com 65 matrizes incorpora informação de $65 \times 5 = 325$ subjanelas, enquanto o grupo com 62 matrizes incorpora informação de $62 \times 5 = 310$ subjanelas. Por fim, para cada grupo g e para cada IMF m , aplica-se a Transformada Rápida de Fourier bidimensional em $\bar{V}_g^{(m)}[n, \tau]$, sendo a transformada ao longo de n associada à frequência cíclica α e a transformada ao longo de τ associada à frequência espectral f . Dessa forma, são obtidas seis matrizes SCD por IMF para cada segmento de áudio original.

Para utilizar essa representação em uma CNN unidimensional, foi extraído o perfil- α , definido como o vetor formado pelos valores máximos da SCD ao longo do eixo de frequências f para cada frequência cíclica α . Considerando os 1440 segmentos de áudio e a extração de 6 perfis- α por segmento, o pré-processamento resultou em um total de 8640 blocos.

Como o sinal de áudio é real e apresenta simetria no domínio da frequência, utilizou-se apenas a metade positiva do perfil- α , resultando em vetores de tamanho 512. Assim, a estrutura final dos dados possui dimensão $(8640, 4, 512)$. O pré-processamento foi realizado de forma a evitar vazamento de dados entre os conjuntos de treinamento, validação e teste.

C. Modelo 1D-CNN

A Tabela I apresenta a arquitetura do modelo utilizado neste trabalho, tendo como referência inicial a arquitetura utilizada por Zhang et al. [4].

A arquitetura adotada é composta por duas camadas convolucionais seguidas por operações de normalização, ativação e redução dimensional. As camadas convolucionais são responsáveis pela extração automática de características relevantes

dos perfis- α , permitindo identificar padrões relacionadas às modulações periódicas associadas às falhas.

Após a etapa convolucional, os mapas de características são convertidos em um vetor unidimensional por meio do *Flatten*, sendo então encaminhados para uma camada totalmente conectada responsável pela classificação final.

Os parâmetros de treinamento (Tabela II) foram definidos empiricamente de forma a proporcionar estabilidade e boa capacidade de generalização ao modelo. O número de épocas escolhido foi 60, mas a técnica de parada antecipada interrompe o treinamento quando o modelo deixa de melhorar no conjunto de validação, evitando *overfitting*.

 TABELA II
 PARÂMETROS DE TREINAMENTO DO MODELO 1D-CNN.

Nome	Configuração
Número de épocas	60
Tamanho do <i>batch</i>	32
Taxa de aprendizado inicial	0.001
Método de ajuste da taxa de aprendizado	<i>ReduceLRonPlateau</i> (<i>patience</i> = 5)
Função de perda	Entropia cruzada
Otimizador	Adam com decaimento de 0.001
Inclinação negativa do <i>Leaky ReLU</i>	0.01 (padrão)
Parada antecipada	10 épocas sem melhora
Taxa de <i>dropout</i>	0.3

IV. RESULTADOS

Os resultados apresentados nesta seção foram obtidos considerando a combinação proporcional entre diferentes modelos de bombas e níveis de SNR, conforme descrito na metodologia. Esse cenário torna a tarefa de classificação mais desafiadora e mais próxima de aplicações reais, nas quais o nível de ruído e as características dos equipamentos não são previamente controlados. Dessa forma, torna-se possível avaliar a robustez das técnicas de pré-processamento propostas em condições mais realistas.

Para avaliar a estabilidade dos resultados, o treinamento foi repetido vinte vezes utilizando diferentes valores de *seed*, de 0 a 19. Também foram avaliadas todas as combinações possíveis das quatro primeiras IMFs. Os valores médios de acurácia e *F1-score* obtidos para as classes normal (N) e anômala (A) são apresentados na Tabela III. A combinação contendo as quatro IMFs apresentou o melhor desempenho, com uma acurácia média de 92.61%. As combinações com acurácia inferior a 90% foram desprezadas nesta análise.

Para avaliar o ganho proporcionado pelo uso da SCD, o treinamento foi repetido utilizando apenas os dados obtidos após o pré-processamento por EMD, ou seja, considerando as IMFs no domínio do tempo. Para esse caso, foi necessária

TABELA III
RESULTADOS MÉDIOS NOS DADOS COM EMD E SCD

IMFs	Acurácia	F1-score N	F1-score A	Seeds
(1, 2, 3, 4)	0.9261 ± 0.0053	0.9289 ± 0.0048	0.9232 ± 0.0060	20
(1, 2, 4)	0.9249 ± 0.0050	0.9274 ± 0.0047	0.9221 ± 0.0053	20
(1, 2, 3)	0.9226 ± 0.0070	0.9251 ± 0.0065	0.9199 ± 0.0075	20
(1, 2)	0.9181 ± 0.0061	0.9202 ± 0.0056	0.9160 ± 0.0066	20
(1, 3, 4)	0.9129 ± 0.0048	0.9152 ± 0.0048	0.9105 ± 0.0049	20
(1, 3)	0.9039 ± 0.0051	0.9058 ± 0.0050	0.9018 ± 0.0054	20

uma pequena adaptação na arquitetura da rede neural, uma vez que o comprimento das IMFs é quatro vezes maior que o dos perfis- α obtidos pela SCD, correspondendo a 2048 e 512 amostras, respectivamente. Assim, o tamanho da entrada da camada totalmente conectada foi ajustado de 2048 (valor descrito na saída da camada *Flatten* da Tabela I) para 8192. Além disso, o tamanho do *batch* foi alterado de 32 para 128 devido à quantidade significativamente maior de sinais utilizados nos treinamentos com EMD (390096 blocos de IMFs, contra 6048 blocos de perfis- α).

Nesses treinamentos, foi considerada apenas a combinação de IMFs com melhor desempenho nos experimentos anteriores. Os resultados obtidos são apresentados na Tabela IV.

TABELA IV
RESULTADOS MÉDIOS NOS DADOS COM EMD

IMFs	Acurácia	F1-score N	F1-score A	Seeds
(1, 2, 3, 4)	0.8643 ± 0.0043	0.8643 ± 0.0051	0.8644 ± 0.0041	20

Com isso, a hipótese proposta neste trabalho mostrou-se válida, indicando que a extração de características cicloestacionárias aplicada às IMFs é capaz de produzir representações mais robustas e discriminativas em cenários com diferentes níveis de ruído, uma vez que os resultados obtidos utilizando apenas a técnica EMD apresentaram desempenho consideravelmente inferior (acurácia média de 86.43%).

V. CONCLUSÕES

O presente trabalho propôs uma metodologia para detecção de falhas em bombas industriais baseada na análise de sinais acústicos. A abordagem combina decomposição em modos empíricos e análise cicloestacionária para extrair descritores robustos, posteriormente utilizados como entrada de uma rede neural convolucional unidimensional. Os resultados demonstram que a metodologia proposta foi capaz de classificar, com elevada acurácia, bombas operando em condições normais e anômalas, mesmo considerando sinais reais ruidosos. Além disso, o desempenho obtido é particularmente promissor por ter sido alcançado em um cenário mais desafiador, no qual sinais provenientes de diferentes bombas e diferentes condi-

ções de ruído foram combinados em um mesmo processo de treinamento e avaliação.

Como continuidade desta pesquisa, propõe-se investigar o uso da abordagem desenvolvida em conjunto com autoencoders, no contexto de uma estratégia semisupervisionada. Essa investigação permitirá avaliar a robustez do método proposto diante de mudanças no contexto operacional, bem como em cenários nos quais há predominância de dados normais e disponibilidade limitada de amostras rotuladas de falha.

AGRADECIMENTOS

Este trabalho foi realizado com o apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001, e do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), no âmbito dos processos 445158/2024-7 e 421235/2025-0.

DECLARAÇÃO DE USO DE INTELIGÊNCIA ARTIFICIAL GENERATIVA

Os autores declaram que utilizaram a ferramenta ChatGPT, da OpenAI, exclusivamente como apoio à revisão gramatical e ao aprimoramento da clareza textual em alguns trechos do manuscrito.

REFERÊNCIAS

- [1] Neupane, D., M. R. Bouadjenek, R. Dazeley, S. Aryal, “Data-driven machinery fault diagnosis: A comprehensive review”. *Neurocomputing*, v. 627, 2025. DOI: 10.1016/j.neucom.2025.129588.
- [2] H. Wang, Z. Liu, D. Peng, M. J. Zuo, “Interpretable convolutional neural network with multilayer wavelet for noise-robust machinery fault diagnosis”. *Mechanical Systems and Signal Processing*, vol. 195, 2023.
- [3] X. Yu, B. Tang, K. Zhang, “Fault Diagnosis of Wind Turbine Gearbox Using a Novel Method of Fast Deep Graph Convolutional Networks”. *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1-14, 2021. DOI: 10.1109/TIM.2020.3048799.
- [4] L. Zhang, Q. Fan, J. Lin, Z. Zhang, X. Yan, C. Li, “A nearly end-to-end deep learning approach to fault diagnosis of wind turbine gearboxes under nonstationary conditions”. *Engineering Applications of Artificial Intelligence*, vol. 119, 2023. DOI: 10.1016/j.engappai.2022.105735.
- [5] G. Jombo and Y. Zhang, “Acoustic-Based Machine Condition Monitoring—Methods and Challenges”. *Eng* 4(1), pp. 47–79, 2023.
- [6] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N.-C. Yen, C. C. Tung, H. H. Liu, “The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis”. *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences* 454(1971), pp. 903–995, 1998. DOI: 10.1098/rspa.1998.0193.
- [7] S. Roshanmanesh, F. Hayati, M. Papaelias, “Utilisation of ensemble empirical mode decomposition in conjunction with cyclostationary technique for wind turbine gearbox fault detection”. *Applied Sciences* 10(9), 2020. DOI: 10.3390/app10093334.
- [8] W. A. Gardner, “An introduction to cyclostationary signals, in: W.A. Gardner (Ed.), *Cyclostationarity in Communications and Signal Processing*, pp. 1–90, IEEE Press, New York, 1994.
- [9] S. Tang, S. Yuan, Y. Zhu, “Cyclostationary analysis towards fault diagnosis of rotating machinery”. *Processes*, 8(10), 2020.
- [10] K. Kim, I. Akbar, K. Bae, J.-S. Um, C. Spooner, J. Reed, “Cyclostationary approaches to signal detection and classification in cognitive radio”. *IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks*, pp. 212–215, 2007.
- [11] H. Purohit, R. Tanabe, K. Ichige, T. Endo, Y. Nikaido, K. Suefusa and Y. Kawaguchi, “MIMII Dataset: Sound Dataset for Malfunctioning Industrial Machine Investigation and Inspection”. *CoRR*, 2019.
- [12] A. Farias; C. E. N. Melo; R. V. Dias. “Uma abordagem a métodos de sensoriamento do espectro aplicados a rádio cognitivo”. 102 f. Monografia (Bacharelado)-Universidade de Brasília, Brasília, 2010.