

# Detecção de Falhas em Bombas Industriais usando Frequências Cíclicas com *Random Forest*

Daniela M. da Silva e Luiz F. Q. Silveira

**Resumo**— Máquinas industriais desempenham um papel fundamental no setor produtivo, e a detecção de anomalias por meio de aprendizado de máquina é essencial para melhorar a confiabilidade. Este estudo propõe um método de detecção de falhas em bombas industriais, baseado no cálculo da Densidade Espectral Cíclica (SCD) e na classificação utilizando o *Random Forest*. A SCD extrai características dos sinais sensorizados, objetivando aprimorar a acurácia da classificação. O método foi avaliado com dados reais, alcançando até 96,5% de acurácia, F1-score médio de 0,80 e 100% de precisão, demonstrando seu potencial para o diagnóstico de falhas industriais.

**Palavras-Chave**— Detecção de falhas, Bombas industriais, Sinais cicloestacionários, Densidade espectral cíclica, Aprendizado de máquina, *Random Forest*.

**Abstract**— Industrial machines play a fundamental role in the productive sector, and anomaly detection through machine learning is essential to improve reliability. This study proposes a fault detection method for industrial pumps, based on the computation of Spectral Correlation Density (SCD) and classification using the *Random Forest* algorithm. The SCD extracts cyclostationary features from sensor-acquired signals to enhance classification accuracy. The method was evaluated using real data from the MIMII dataset, achieving up to 96.5% accuracy, an average F1-score of 0.80, and 100% precision, demonstrating its potential for industrial fault diagnosis applications.

**Keywords**— Fault detection, Industrial pumps, Cyclostationary signals, Spectral correlation density, Machine learning, *Random Forest*.

## I. INTRODUÇÃO

O monitoramento automático de máquinas industriais é fundamental para aumentar a confiabilidade operacional, reduzir custos com manutenção corretiva e evitar paradas não planejadas. Tradicionalmente, sensores de vibração são amplamente utilizados nesse contexto, por captarem diretamente as oscilações mecânicas associadas ao funcionamento de motores, engrenagens e rolamentos. No entanto, sensores acústicos surgem como uma alternativa promissora por sua natureza não invasiva, menor custo e facilidade de instalação, especialmente em locais de difícil acesso. Embora ambos os tipos de sensores possam ser utilizados para diagnóstico de falhas, há diferenças importantes: sinais de vibração geralmente apresentam maior fidelidade em componentes mecânicos rígidos, enquanto sinais acústicos podem capturar informações propagadas pelo ar, sendo mais sensíveis a ruídos ambientais. Mesmo assim, diversas pesquisas indicam que técnicas avançadas de processamento, como análise cicloestacionária, podem extrair padrões

relevantes de ambos os tipos de sinais, tornando os sensores acústicos uma alternativa viável em muitos cenários industriais [1].

Em ambientes industriais, é comum que os sinais acústicos e de vibração apresentem comportamento não-estacionário, especialmente quando gerados por máquinas rotativas em operação contínua. Nesses casos, os sinais frequentemente exibem características cicloestacionárias, com repetitividade estatística associada ao movimento rotativo ou oscilatório dos componentes mecânicos. Técnicas de processamento que exploram essas propriedades não-estacionárias tendem a alcançar melhor desempenho na detecção de falhas, especialmente em cenários com alto nível de ruído. Nesse sentido, a Densidade Espectral Cíclica (SCD) tem se mostrado uma ferramenta poderosa para extrair componentes espectrais moduladas ciclicamente, revelando padrões que escapam a abordagens espectrais convencionais [2].

Estudos recentes têm explorado estratégias diversas de análise cicloestacionária no diagnóstico de falhas. Zhao *et al.* [3] propuseram um método com sinais acústicos captados ao longo de trilhos ferroviários, utilizando análise cíclica para identificar falhas em rolamentos. Por outro lado, a maioria dos trabalhos na literatura emprega sinais de vibração. Michalak *et al.* [4] aplicaram análise cicloestacionária com testes estatísticos para detectar falhas em mancais de motores elétricos. Javorskyj *et al.* [5] modelaram sinais de vibração como processos bi-periodicamente correlacionados em turbinas eólicas, enquanto Roshanmanesh *et al.* [6] combinaram decomposição em modos empíricos por conjunto (EEMD) com análise cíclica para isolar padrões modulares em ambientes ruidosos.

Apesar dos avanços, a aplicação de técnicas baseadas em SCD a sinais acústicos industriais ainda é pouco explorada, assim como a combinação entre descritores espectrais cíclicos e classificadores de baixo custo computacional, como o *Random Forest*. Essa lacuna é especialmente relevante em contextos nos quais se busca alta acurácia com baixo tempo de resposta, utilizando apenas sinais acústicos de fácil aquisição.

Este artigo propõe uma abordagem para detecção de falhas em bombas industriais com base na extração de descritores por meio da SCD, seguida de classificação com o algoritmo *Random Forest*. A proposta é avaliada com dados reais do *MIMII Dataset*, explorando múltiplas seções de operação, demonstrando sua eficácia em diferentes condições operacionais e sua superioridade em relação a métodos baseados em espectrogramas convencionais.

O restante deste artigo está organizado da seguinte forma: a Seção II detalha a metodologia, incluindo a extração de características e a classificação; a Seção III descreve o conjunto de dados; a Seção IV apresenta e discute os resultados; e a

Daniela M. da Silva, Programa de Pós-Graduação em Engenharia Elétrica e de Computação, Universidade Federal do Rio Grande do Norte, Natal-RN, e-mail: daniela.maia.064@ufrn.edu.br; Luiz Felipe de Queiroz Silveira, Departamento de Engenharia de Computação e Automação, Universidade Federal do Rio Grande do Norte, Natal-RN, e-mail: felipe.silveira@ufrn.br

Seção V traz as conclusões e perspectivas futuras.

## II. MÉTODO PROPOSTO

Esta seção descreve a metodologia proposta para detecção de falhas em bombas industriais, que combina a análise espectral de sinais acústicos via Densidade Espectral Cíclica (SCD) com a classificação por meio do algoritmo *Random Forest*, incluindo suas motivações e fundamentos teóricos.

### A. Visão Geral do Método

O método proposto visa a detecção automática e robusta de falhas em bombas industriais por meio da análise cicloestacionária dos sinais acústicos adquiridos. A abordagem adotada combina a extração de características espectrais via Densidade Espectral Cíclica (SCD - *Spectral Correlation Density*) com um classificador baseado em *Random Forest*. O fluxo de processamento consiste nas seguintes etapas principais:

- 1) Aquisição dos sinais acústicos das máquinas em condições normais e anômalas;
- 2) Cálculo da Densidade Espectral Cíclica (SCD) dos sinais adquiridos;
- 3) Seleção criteriosa das frequências cíclicas mais relevantes;
- 4) Classificação automática utilizando *Random Forest* para diagnóstico de falhas.

### B. Cálculo da SCD

Em processamento digital de sinais, é comum adotar a hipótese de que os sinais são estacionários, o que permite a aplicação de técnicas clássicas e computacionalmente eficientes, como a Transformada Rápida de Fourier (FFT). No entanto, muitos sinais encontrados em aplicações reais, incluindo os gerados por sistemas mecânicos industriais, são, na prática, não-estacionários. Dentro dessa categoria, destaca-se a classe dos sinais cicloestacionários, cuja principal característica é a periodicidade dos seus momentos estatísticos. Formalmente, um sinal  $x(t)$  é dito cicloestacionário de segunda ordem quando sua média  $m_x(t)$  e sua função de autocorrelação  $R_x(t, \tau)$  são funções periódicas no tempo  $t$ , com período  $T$ . Isso implica que:

$$m_x(t) = E[x(t)] = m_x(t + T). \quad (1)$$

$$R_x(t, \tau) = E \left[ x \left( t + \frac{\tau}{2} \right) x \left( t - \frac{\tau}{2} \right) \right] = R_x(t + T, \tau). \quad (2)$$

Essas propriedades permitem representar  $R_x(t, \tau)$  como uma série de Fourier em  $t$ , cujos coeficientes correspondem à autocorrelação cíclica do sinal para diferentes frequências cíclicas  $\alpha$ .

No contexto industrial, sinais acústicos oriundos de máquinas rotativas e outros equipamentos com ciclos operacionais regulares frequentemente apresentam comportamento cicloestacionário, refletindo a repetitividade intrínseca das operações mecânicas. Nesses casos, técnicas convencionais de análise espectral podem ser insuficientes para evidenciar componentes periódicas sutis, sobretudo em ambientes ruidosos. A análise

cicloestacionária, e em especial o cálculo da Densidade Espectral Cíclica (SCD), mostra-se particularmente eficaz ao revelar essas periodicidades ocultas e ao correlacionar diferentes bandas espectrais, possibilitando uma caracterização mais robusta e informativa do sinal [7].

Formalmente, a Densidade Espectral Cíclica é definida como a transformada de Fourier da Função de Autocorrelação Cíclica (CAF), dada por:

$$S_x^\alpha(f) = \int_{-\infty}^{+\infty} R_x^\alpha(\tau) e^{-j2\pi f\tau} d\tau. \quad (3)$$

em que  $R_x^\alpha(\tau)$  é o coeficiente da série de Fourier da função de autocorrelação periódica, calculada por:

$$R_x(t, \tau) = \sum_{\alpha} R_x^\alpha(\tau) e^{-j2\pi\alpha t}. \quad (4)$$

Na função  $S_x^\alpha(f)$ , a variável  $f$  representa a frequência espectral, enquanto  $\alpha$  é a frequência cíclica, associada à periodicidade estatística do sinal. A Densidade Espectral Cíclica quantifica a correlação entre componentes espectrais separadas por  $\alpha$ , ou seja, localizadas em  $f + \alpha/2$  e  $f - \alpha/2$ , revelando estruturas de modulação ou periodicidade no sinal que não seriam capturadas por métodos espectrais convencionais [8]. Essa propriedade é especialmente útil para identificar componentes moduladas ciclicamente em sinais industriais, cujas periodicidades são associadas ao funcionamento rotativo ou oscilatório das máquinas.

Na prática, a SCD é computada em um grid discreto de frequências espectrais  $f$  e frequências cíclicas  $\alpha$ , resultando em uma matriz complexa bidimensional:

$$S_x^\alpha(f) \in C^{N_\alpha \times N_f}. \quad (5)$$

em que  $N_\alpha$  representa o número de amostras discretas da frequência cíclica  $\alpha$ , e  $N_f$  representa o número de amostras da frequência espectral  $f$ . Cada linha da matriz corresponde a um valor fixo de  $\alpha$ , enquanto cada coluna representa uma frequência  $f$ . Para viabilizar sua estimativa eficiente, o cálculo da SCD é implementado por meio de sucessivas transformadas rápidas de Fourier (FFTs), seguindo abordagens consagradas na literatura [7].

### C. Perfil- $\alpha$

Após o cálculo das matrizes SCD, cada clipe de áudio é representado por um vetor unidimensional que caracteriza o perfil de magnitude em função da frequência cíclica, denominado *perfil- $\alpha$* . Esse vetor, denotado  $v(\alpha)$ , é construído ao selecionar, para cada valor positivo de  $\alpha$ , o maior valor da magnitude espectral ao longo de  $f$ , ou seja:

$$v(\alpha) = \max_f |S_x^\alpha(f)|. \quad (6)$$

Como apenas os valores positivos de  $\alpha$  foram considerados, o vetor  $v(\alpha) \in R^{N_\alpha/2}$  representa a assinatura espectral cíclica do clipe de áudio. O *perfil- $\alpha$*  resulta em uma redução significativa da dimensionalidade, ao condensar a matriz complexa bidimensional  $S_x^\alpha(f)$  em um vetor real unidimensional  $v(\alpha)$ . A consideração apenas dos valores positivos de  $\alpha$  contribui adicionalmente para essa redução.

#### D. Classificação Automática com Random Forest

A escolha do algoritmo *Random Forest* foi motivada por sua capacidade comprovada de lidar com dados de alta dimensionalidade e presença de ruído, além de sua robustez frente ao *overfitting* e facilidade de interpretação dos resultados [9].

O *Random Forest* é um método de aprendizado supervisionado baseado em um conjunto de árvores de decisão, cada uma construída a partir de subconjuntos aleatórios de amostras e atributos. O treinamento de cada árvore é realizado de forma independente por meio da técnica de *Bootstrap Aggregation* (bagging), que consiste em gerar subconjuntos por amostragem com reposição a partir do conjunto original [10]. Durante a construção da árvore, a divisão de cada nó é determinada com base em um subconjunto aleatório de atributos, selecionando-se aquele que melhor separa as classes naquele ponto.

Essa diversidade entre amostras e atributos gera árvores complementares, cuja combinação por votação majoritária contribui para reduzir a variância do modelo e melhorar sua capacidade de generalização. O funcionamento do algoritmo é ilustrado na Fig. 1.

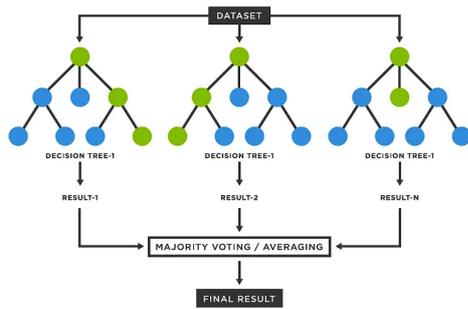


Fig. 1: Algoritmo *Random Forest*.

### III. CONJUNTO DE DADOS

Esta seção apresenta o conjunto de dados utilizado na avaliação da abordagem proposta, abordando suas principais características.

#### A. Descrição Geral

Neste estudo, foram utilizados dados do *MIMII Dataset* [11], amplamente empregado em pesquisas sobre detecção de falhas industriais. O banco de dados contém 32.157 clipes de áudio com 10 segundos de duração cada, amostrados a 16 kHz, dos quais 26.092 correspondem a condições normais de operação. Os dados foram coletados a partir de quatro tipos distintos de máquinas industriais: ventiladores, válvulas, bombas e trilhos deslizantes.

Cada tipo de máquina é subdividido em sete seções (de 00 a 06), que representam variações estruturais ou diferentes modelos do mesmo equipamento. Todas as seções incluem gravações em estado normal e sob condições anômalas simuladas, permitindo avaliar a generalização dos métodos aplicados em diferentes contextos operacionais.

#### B. Dados Considerados

Este trabalho utiliza exclusivamente os clipes de áudio referentes ao tipo de máquina *bomba*, sob condição de relação sinal-ruído (SNR) de +6 dB, considerando as seções 00, 02 e 04. As bombas operam em um sistema de sucção e descarga conectado a um reservatório de água, reproduzindo condições típicas de uso industrial. As anomalias presentes nas gravações incluem falhas como vazamentos, entupimentos, contaminações e obstruções no sistema hidráulico [12].

### IV. RESULTADOS E DISCUSSÃO

Esta seção apresenta os resultados experimentais obtidos com a metodologia proposta, incluindo exemplos dos perfis espectrais gerados, bem como a avaliação quantitativa do desempenho do classificador *Random Forest*.

#### A. Particionamento dos Dados e Pré-processamento

Os experimentos foram conduzidos com dados acústicos de bombas industriais das seções 00, 02 e 04 do MIMII Dataset, sob condição de SNR de +6 dB. A Tabela I resume o número de segmentos normais e anômalos por seção.

TABELA I: Composição do conjunto de dados por seção.

Tipo de máquina	ID do modelo	Segmentos em condição normal	Segmentos em condição anômala
Bomba	00	1006	143
	02	1005	111
	04	702	100
Total		2713	354

Os dados foram divididos em conjuntos de treino (80%) e teste (20%) de maneira aleatória e estratificada, preservando a proporção entre amostras normais e anômalas. A aleatoriedade foi controlada por meio da definição de uma semente fixa (*random state*), assegurando a reprodutibilidade dos resultados.

#### B. Perfis Cíclicos Ilustrativos

As Figuras 2 e 3 ilustram a SCD para exemplos de sinais acústicos capturados em condição normal e anômala para a seção 00, enquanto que as Figs 4 e 5 ilustram os respectivos perfis- $\alpha$ ,  $v(\alpha)$ , desses sinais, evidenciando diferenças nas amplitudes associadas às frequências cíclicas.

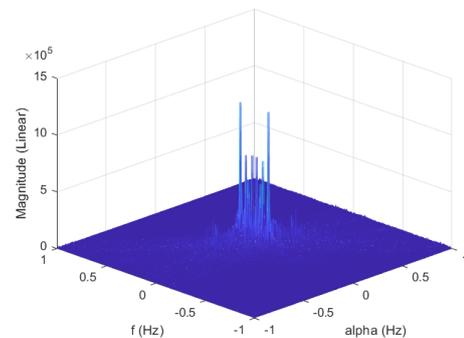


Fig. 2: Função SCD do sinal acústico em condição normal (seção 00).

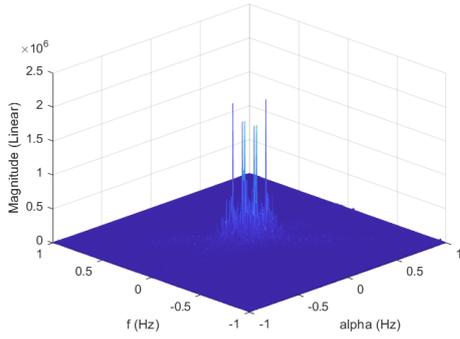


Fig. 3: Função SCD do sinal acústico com anomalia (seção 00).

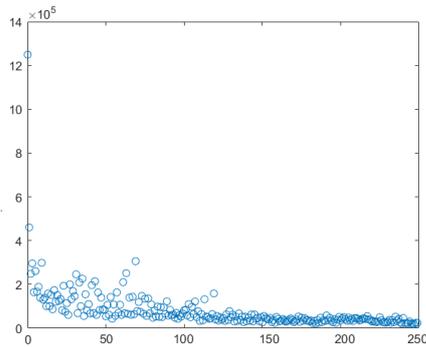


Fig. 4: Perfil- $\alpha$  para o sinal acústico em condição normal (seção 00).

### C. Treinamento do Classificador

O classificador *Random Forest* foi configurado com 100 estimadores (árvores), critério de divisão *log\_loss* e máximo de atributos avaliados por nó definido como *log2*. Esses hiperparâmetros foram definidos com base em testes empíricos, com o objetivo de maximizar o desempenho em termos de precisão e robustez.

Antes do treinamento, os dados foram organizados em um *dataframe* estruturado, associando cada instância ao seu respectivo rótulo (normal ou anômalo), conforme ilustrado na Tabela II.

O modelo foi então treinado utilizando validação cruzada repetida, por meio do método *Repeated Stratified K-Fold*

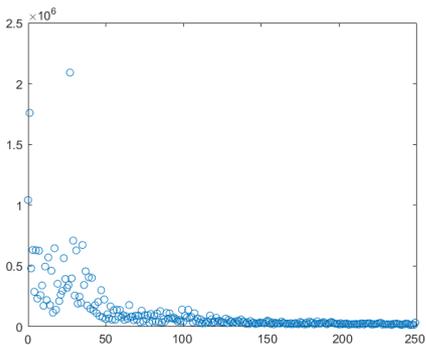


Fig. 5: Perfil- $\alpha$  para o sinal acústico com anomalia (seção 00).

TABELA II: Exemplo de estrutura do dataframe.

Áudio	$\alpha_1$	$\alpha_2$	...	$\alpha_{N_\alpha/2}$	Classe
1	valor	valor	...	valor	0
2	valor	valor	...	valor	1
...	...	...	...	...	...
N	valor	valor	...	valor	0

*Cross-Validation*, com 5 divisões (*folds*) e 10 repetições independentes. A estratificação preserva a proporção entre classes em cada partição, o que é especialmente importante devido ao desequilíbrio entre amostras normais e anômalas. A semente do gerador aleatório foi fixada em 42 para assegurar a reprodutibilidade dos experimentos.

### D. Avaliação do Modelo

Os vetores  $v(\alpha)$  extraídos de cada clipe do conjunto de teste, após a etapa de normalização, foram utilizados como entrada no classificador *Random Forest* treinado conforme descrito anteriormente. O desempenho do modelo foi avaliado com base na matriz de confusão e nas métricas de acurácia, precisão, *recall* e *F1-score*, calculadas a partir da validação cruzada repetida.

A Figura 6 apresenta a matriz de confusão referente ao modelo treinado com os dados da seção 00. Observa-se que o classificador alcançou uma acurácia de 96,52%, com desvio padrão de 0,86%, demonstrando alto desempenho na distinção entre sinais normais e anômalos. O modelo cometeu apenas oito falsos negativos e nenhum falso positivo, evidenciando sua elevada precisão na detecção de falhas e sua robustez contra a emissão de alarmes indevidos.

Para a seção 02, o modelo alcançou uma acurácia média de 95,54% com desvio padrão de 1,23%. Já na seção 04, a acurácia média foi de 95,65%, com desvio padrão de 1,16%. Esses resultados reforçam a eficácia e a estabilidade do método proposto nas diferentes condições operacionais avaliadas.

Matriz de confusão

True Labels	Normal	199	0
	Anomaly	8	23
		Normal	Anomaly
		Predicted Labels	

Fig. 6: Matriz de confusão para a seção 00.

A Tabela III resume as métricas de desempenho obtidas para os modelos treinados nas seções analisadas, incluindo os valores médios correspondentes. Esses resultados permitem uma comparação direta da eficácia do classificador em diferentes subconjuntos do conjunto de dados. Cada seção representa um modelo de máquina distinto, e as variações no desempenho

entre elas podem ser atribuídas às diferenças nas características espectrais e operacionais dos respectivos sinais acústicos.

TABELA III: Métricas de desempenho por seção.

Tipo de máquina	ID do modelo	AUC (%)	F1-Score	Precisão	Recall
Bomba	00	96,52	0,85	1,00	0,74
	02	95,54	0,75	1,00	0,60
	04	95,65	0,82	1,00	0,71
Média		95,90	0,80	1,00	0,70

### E. Análise Comparativa

No estudo original que introduziu o *MIMII Dataset*, foram utilizados autoencoders treinados com espectrogramas em escala Mel como descritores. Os resultados relataram uma grande variação de desempenho entre as seções: 84% para a seção 00, 45% para a seção 02 e 99% para a seção 04 [12].

A abordagem proposta neste trabalho apresentou desempenho superior ao do estudo original, especialmente na seção 02, onde a acurácia média obtida superou em mais de 50 pontos percentuais o valor anteriormente reportado. A análise dos vetores  $v(\alpha)$  indicou maior separabilidade entre as classes, evidenciando que os descritores extraídos via Densidade Espectral Cíclica capturam com mais eficácia as assinaturas características das anomalias do que os espectrogramas convencionais.

Esses resultados sugerem que a combinação da SCD com classificadores *Random Forest* constitui uma estratégia eficaz para o diagnóstico de falhas em sinais acústicos, mesmo em ambientes industriais ruidosos. Além da alta precisão, o método demonstrou boa capacidade de generalização entre diferentes modelos de bombas industriais.

## V. CONCLUSÕES

A utilização da Densidade Espectral Cíclica (SCD) mostrou-se uma estratégia eficaz para evidenciar padrões cíclicos em sinais acústicos industriais, capturando informações relevantes associadas ao comportamento cicloestacionário das máquinas. A extração dos valores máximos de frequência para cada valor de  $\alpha$  contribuiu para reduzir a dimensionalidade dos dados, mantendo características discriminativas e viabilizando um processamento mais eficiente na etapa de classificação.

O modelo *Random Forest*, utilizado para a detecção de anomalias, demonstrou robustez frente à variabilidade entre seções e à presença de ruído, obtendo resultados superiores aos do estudo original do *MIMII Dataset*, especialmente na seção 02. Esses resultados destacam o potencial da abordagem proposta para aplicações práticas de diagnóstico de falhas em ambientes industriais com sinais acústicos.

Como trabalhos futuros, pretende-se investigar o uso de abordagens baseadas em aprendizado não supervisionado combinados à SCD, com o objetivo de detectar anomalias raras ou em sistemas onde a rotulação não é viável. Essa direção visa ampliar a aplicabilidade do método proposto para cenários reais.

## AGRADECIMENTOS

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

Os autores agradecem ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pelo apoio financeiro a este trabalho, por meio do projeto fomentado sob processo nº 445158/2024-7.

## REFERÊNCIAS

- [1] Xiaoyang Bi, Shuqian Cao, and Daming Zhang. A variety of engine faults detection based on optimized variational mode decomposition-robust independent component analysis and fuzzy c-mean clustering. *IEEE Access*, 7:27756–27768, 2019.
- [2] Zrar Khalid Abdul, Abdulbasit K. Al-Talabani, and Dlair O. Ramadan. A hybrid temporal feature for gear fault diagnosis using the long short term memory. *IEEE Sensors Journal*, 20(23):14444–14452, December 2020.
- [3] Xing Zhao, Yiming Lu, Baoxian Chang, and Liqun Chen. Study on the extraction method for track-side acoustic features based on cyclic stationary analysis. *Machines*, 11(10):957, October 2023.
- [4] Anna Michalak, Justyna Hebda-Sobkowicz, Jacek Wodecki, Krzysztof Szabat, Marcin Wolkiewicz, Sebastien Weisse, Jerome Valire, Radosław Zimroz, and Agnieszka Wylomańska. Enhancement of cyclic spectral coherence map by statistical testing approach—application to bearing faults diagnosis in electric motors. *Measurement Science and Technology*, 36(1):016169, November 2024.
- [5] Ihor Javorskyj, Ivan Matsko, Roman Yuzefovych, Oleh Lychak, and Roman Lys. Methods of hidden periodicity discovering for gearbox fault detection. *Sensors*, 21(18):6138, September 2021.
- [6] Sanaz Roshanmanesh, Farzad Hayati, and Mayorkinos Papaelias. Utilisation of ensemble empirical mode decomposition in conjunction with cyclostationary technique for wind turbine gearbox fault detection. *Applied Sciences*, 10(9):3334, May 2020.
- [7] W. Gardner. Measurement of spectral correlation. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 34(5):1111–1123, October 1986.
- [8] Steven Hong, Eric Like, Zhiqiang Wu, and Cem Tekin. Multi-user signal classification via spectral correlation. In *2010 7th IEEE Consumer Communications and Networking Conference*, page 1–5. IEEE, January 2010.
- [9] Leo Breiman. *Machine Learning*, 45(1):5–32, 2001.
- [10] Tin Kam Ho. Random decision forests. In *Proceedings of 3rd International Conference on Document Analysis and Recognition*, volume 1, pages 278–282 vol.1, 1995.
- [11] Harsh Purohit, Ryo Tanabe, Takeshi Ichige, Takashi Endo, Yuki Nikaido, Kaori Suefusa, and Yohei Kawaguchi. MIMII Dataset: Sound dataset for malfunctioning industrial machine investigation and inspection. In *Proceedings of the Detection and Classification of Acoustic Scenes and Events 2019 Workshop (DCASE2019)*, pages 209–213, November 2019.
- [12] Harsh Purohit, Ryo Tanabe, Kenji Ichige, Takashi Endo, Yuki Nikaido, Kaori Suefusa, and Yohei Kawaguchi. Mii dataset: Sound dataset for malfunctioning industrial machine investigation and inspection, 2019.