

# Detecção de Anomalias por Classificação de Áudio em Ventiladores Industriais

Andriel de Figueiredo Furtado  
Instituto de Informática  
Universidade Federal de Goiás  
Goiânia, Brasil  
andriel@discente.ufg.br

Anna Pietra Vitória L. Bastos Moreira  
Instituto de Informática  
Universidade Federal de Goiás  
Goiânia, Brasil  
annapietra@discente.ufg.br

Maria Carolina Xavier de Almeida  
Instituto de Informática  
Universidade Federal de Goiás  
Goiânia, Brasil  
mariaalmeida2@discente.ufg.br

Mariana Nogueira Carvalho da Silveira  
Instituto de Informática  
Universidade Federal de Goiás  
Goiânia, Brasil  
mariananogueira@discente.ufg.br

Rosimeire Alves da Silva  
Instituto de Informática  
Universidade Federal de Goiás  
Goiânia, Brasil  
rosimeire\_alves@discente.ufg.br

**Resumo**— No contexto da Indústria 4.0, a detecção precoce de anomalias em ventiladores industriais é importante para evitar falhas e otimizar operações. Este trabalho expande estudos anteriores ao desenvolver uma ferramenta que analisa diferentes níveis de ruído (6dB, 0dB e -6dB) para identificar anomalias sonoras, utilizando processamento de sinais e aprendizado de máquina. A validação é realizada com o dataset MIMII, permitindo a classificação do áudio em diferentes condições de ruído em ambiente industrial.

**Palavras-chave** — Detecção de anomalias, Classificação de áudio, Processamento de sinais, Aprendizado de máquina.

## I. INTRODUÇÃO

Fábricas inteligentes utilizam a detecção de anomalias em equipamentos para melhorar a qualidade do produto, reduzir perdas de insumos, identificar deterioração de máquinas e aumentar a confiabilidade e eficiência dos processos industriais, além de reduzir custos e melhorar a segurança dos trabalhadores. O ambiente industrial é caracterizado por ruídos diversos, tornando desafiadora a identificação de eventos sonoros específicos. Técnicas baseadas em aprendizado de máquina, como autoencoders e classificadores de sinais, têm sido amplamente empregadas para detectar anomalias em sinais sonoros [6]. Estudos demonstram que redes neurais com filtros adaptativos apresentam melhor desempenho na classificação de sinais sonoros em comparação com abordagens tradicionais [3]. Além disso, métodos como MFCC (Mel Frequency Cepstral Coefficients) e PCA (Principal Component Analysis) são utilizados para caracterizar e clusterizar eventos sonoros [4].

### A. Revisão bibliográfica

Os avanços em aprendizado profundo (DL) têm transformado a área de detecção de anomalias. Ciaburro et al. [5] propuseram uma metodologia para diagnóstico de falhas em ventiladores industriais utilizando redes neurais convolucionais (CNNs) e transferência de aprendizado, alcançando uma precisão de 95% na classificação de condições operacionais. Tang et al. [6] destaca a eficácia de redes neurais recorrentes (RNNs) em extrair características complexas em condições de trabalho não estacionárias. Shubita et al. [7] desenvolveu uma solução otimizada para execução em dispositivos de baixo custo, mostrando que técnicas simplificadas podem ser eficazes. Barreto et al. [8] aplicou aprendizado de máquina (ML) para detecção de falhas

em transportadores de correia, atingindo uma ROC AUC de 91% ao empregar o modelo XGBoost. Por fim, Kawaguchi et al. [6] introduziu benchmarks para autoencoders aplicados à detecção de anomalias em ambientes fabris, utilizando o dataset MIMII.

### B. Conjunto de Dados

O dataset MIMII (Malfunctioning Industrial Machine Investigation and Inspection) [6] disponibiliza sons de quatro tipos de máquinas industriais: válvulas, bombas, ventiladores e trilhos deslizantes. Para este trabalho, utilizamos exclusivamente os dados relacionados a ventiladores, considerando níveis de SNR (relação sinal-ruído) de 6 dB, 0 dB e -6 dB. O dataset foi estendido para incluir modelos a diferentes níveis de SNR, expandindo o trabalho prévio que analisava apenas ventiladores a 0 dB. Os três arquivos de áudio Wav de ventiladores domésticos baixados do Kaggle [6] foram gravados usando o Audacity e exportados como PCM assinado de 16 bits, volume de gravação 0,70.

### C. Métodos e Avaliação

Utilizamos redes neurais convolucionais (CNNs) para a detecção de anomalias em áudios de ventiladores industriais do dataset MIMII, considerando níveis de ruído de 6 dB, 0 dB e -6 dB. As características sonoras foram extraídas com coeficientes MFCCs, e o modelo foi avaliado com métricas como AUC, precisão e matriz de confusão. Expandimos abordagens anteriores ao incluir diferentes níveis de ruído, permitindo uma análise comparativa da robustez do modelo em condições variadas.

## II. FUNDAMENTOS TEÓRICOS

A detecção de anomalias em áudio industrial combina processamento de sinais com aprendizado profundo, visando identificar desvios sonoros que indiquem falhas em máquinas.

### A. Processamento de Sinais de Áudio:

Os sinais sonoros são representados por coeficientes cepstrais na frequência Mel (MFCCs), que capturam características acústicas essenciais. Esses coeficientes são calculados aplicando a Transformada Rápida de Fourier (FFT), seguida de filtros na escala Mel e pela Transformada Discreta do Cosseno (DCT), garantindo uma representação compacta e informativa do áudio.

### B. Arquitetura de Rede Neural:

A solução utiliza Redes Neurais Convolucionais (CNNs), que são eficazes na extração de padrões locais de dados temporais. A arquitetura inclui camadas convolucionais e de pooling para extração de características, e camadas densas para realizar a classificação binária.

### C. Avaliação de Desempenho:

O desempenho do modelo é avaliado com métricas como precisão, recall, F1-score, AUC e matriz de confusão, possibilitando a análise detalhada da capacidade do modelo em distinguir entre condições normais e anômalas.

Esses fundamentos teóricos orientam a aplicação das técnicas e métricas que serão detalhadas na metodologia, fornecendo uma base consistente para a abordagem implementada.

## III. METODOLOGIA

A metodologia para o desenvolvimento do detector de anomalias em áudio industrial foi estruturada em etapas, desde a preparação dos dados até a aplicação do modelo em novos cenários

### A. Aquisição e Organização dos Dados

Os dados foram obtidos do MIMII Dataset, com foco exclusivo no tipo de máquina "fan", abrangendo áudios em três níveis de ruído (6 dB, 0 dB e -6 dB). Um sistema automatizado foi implementado para verificar, criar diretórios necessários e organizar os arquivos em classes ("normal" e "anormal"). O método load\_dataset encapsulou a organização e carregamento dos arquivos de áudio, armazenando metadados relevantes em um DataFrame para análise eficiente.

### B. Extração de Características

Os áudios foram convertidos em vetores de características utilizando coeficientes cepstrais na frequência Mel (MFCCs). O método extract\_features calculou os MFCCs para cada arquivo de áudio, extraíndo a média de suas transpostas para capturar as características espectrais mais relevantes. O método prepare\_data consolidou as entradas (X) e os rótulos (y) para alimentar o modelo.

### C. Construção do Modelo

Foi projetada uma rede neural convolucional (CNN) com camadas Conv1D, pooling (MaxPooling1D), normalização por lotes (BatchNormalization) e dropout para evitar overfitting. A arquitetura incluiu uma camada densa final com ativação sigmoide para classificação binária. A função de perda foi a entropia cruzada binária, e o otimizador Adam foi utilizado para ajuste eficiente dos pesos do modelo.

### D. Treinamento e Validação

Os dados foram divididos em conjuntos de treinamento, validação e teste, mantendo a proporção das classes através de divisão estratificada. Callbacks foram aplicados durante o treinamento:

- ModelCheckpoint: Para salvar os melhores pesos com base na AUC do conjunto de validação.
- EarlyStopping: Para interromper o treinamento em caso de estagnação na melhoria do modelo.
- ReduceLROnPlateau: Para ajustar a taxa de aprendizado quando o desempenho parava de melhorar.

### E. Avaliação do Modelo

O desempenho do modelo foi avaliado utilizando métricas como precisão, recall, F1-Score, AUC e matriz de confusão. Essas métricas forneceram uma análise detalhada da capacidade do modelo de distinguir entre áudios normais e anômalos.

### F. Predição em Novos Dados

Implementamos o método predict\_audio, que permite ao modelo salvo realizar previsões em novos áudios, identificando se o som é "normal" ou representa uma "anomalia". Essa funcionalidade viabiliza a aplicação prática em ambientes industriais.

## IV. RESULTADOS E CONCLUSÕES

Os experimentos realizados avaliaram o desempenho de um modelo de redes neurais convolucionais (CNNs) para detecção de anomalias em áudios de ventiladores industriais do dataset MIMII. Modelos foram treinados e testados em três níveis de ruído distintos: 6 dB, 0 dB e -6 dB.

### A. Resultados

Os resultados para cada nível de ruído são apresentados na Tabela 1 a seguir, com base nas métricas de acurácia, precisão, recall, F1-Score e AUC, complementados pela matriz de confusão.

TABELA I. RESULTADOS PARA CADA NÍVEL DE RUÍDO

Métrica	6 dB	0 dB	-6 dB
Acurácia	99,82%	99,09%	93,66%
Precisão	100,00%	99,58%	92,86%
Recall	99,30%	97,90%	81,82%
F1-Score	99,65%	98,25%	86,90%
AUC	1,0000	0,9998	0,9769

O modelo apresentou desempenho excepcional no nível de ruído 6 dB, com AUC máxima e mínima ocorrência de falsos positivos e negativos.

Embora ligeiramente inferior ao cenário de 6 dB, o desempenho geral manteve-se consistente no nível de ruído 0 dB, evidenciando alta precisão.

Em condições mais adversas de ruído (nível -6 dB), o desempenho apresentou uma queda significativa no recall, indicando maior dificuldade do modelo em identificar corretamente as anomalias.

A seguir, os resultados estão ilustrados por meio das matrizes de confusão para cada nível de ruído (vide Figuras 1, 2 e 3).

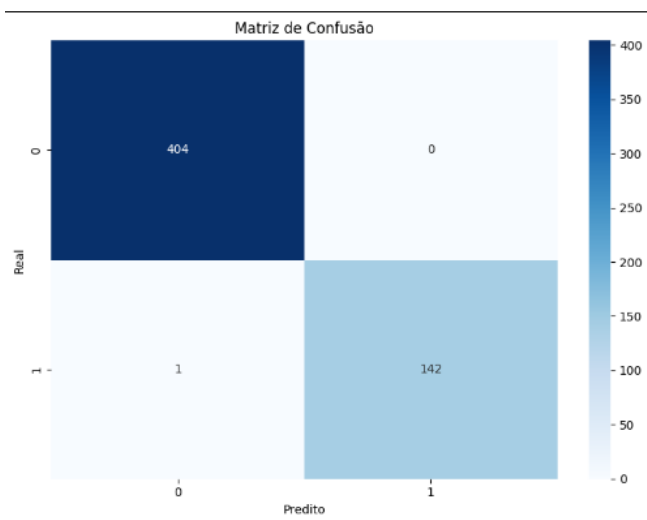


Figura 1. Matriz de confusão para nível de ruído 6 dB

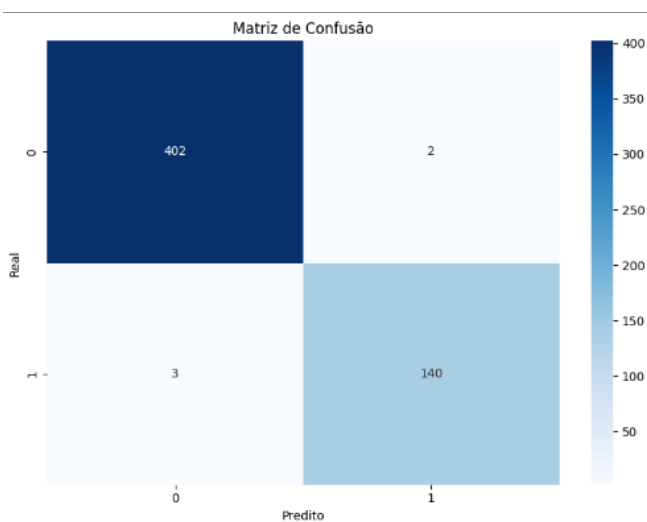


Figura 2. Matriz de confusão para nível de ruído 0 dB

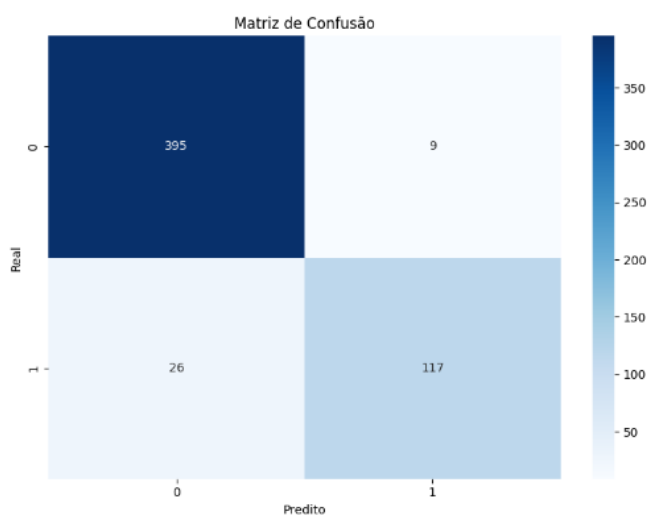


Figura 3. Matriz de confusão para nível de ruído -6 dB

## Discussão dos Resultados

Em termos práticos, os resultados obtidos reforçam a viabilidade do uso de aprendizado profundo para detecção de anomalias em áudios industriais, especialmente em cenários onde o ruído é minimizado. A queda de desempenho em -6 dB, entretanto, destaca a necessidade de investigar técnicas adicionais, como aumento de dados, treinamento com ruído adversarial ou modelos mais robustos, como redes neurais recorrentes.

## B. Conclusões

Os resultados obtidos validaram a eficácia da arquitetura de Redes Neurais Convolucionais (CNNs) na extração de padrões locais em dados temporais, demonstrando desempenho elevado em diferentes níveis de ruído. Entretanto, observou-se que o ruído afeta significativamente o desempenho do modelo, especialmente na identificação de anomalias (recall), evidenciando a necessidade de reforçar a robustez da solução em cenários adversos.

A análise comparativa entre os diferentes níveis de ruído foi fundamental para avaliar a capacidade de generalização do modelo e identificar oportunidades de melhoria. Este trabalho ampliou estudos anteriores ao incluir condições variadas de ruído, destacando a importância de abordagens especializadas para manutenção preditiva em ambientes industriais.

Para trabalhos futuros, considera-se a incorporação de técnicas de regularização avançada e a exploração de arquiteturas híbridas que combinem CNNs com redes recorrentes, visando maior robustez e precisão em cenários desafiadores.

## REFERÊNCIAS

- [1] Duman, T.B., Bayram, B., İnce, G. (2020). Acoustic Anomaly Detection Using Convolutional Autoencoders in Industrial Processes. In: Martínez Álvarez, F., Troncoso Lora, A., Sáez Muñoz, J., Quintián, H., Corchado, E. (eds) 14th International Conference on Soft Computing Models in Industrial and Environmental Applications (SOCO 2019). SOCO 2019. Advances in Intelligent Systems and Computing, vol 950. Springer, Cham. [\[https://doi.org/10.1007/978-3-030-20055-8\\_41\]](https://doi.org/10.1007/978-3-030-20055-8_41)
- [2] Harsh, P., Ryo, T., Kenji, I., Takashi, E., Yuki, N., Kaori, S., Yohei, K. "MIMII Dataset: Sound Dataset for Malfunctioning Industrial Machine Investigation and Inspection". September 2019. ZENODO. [\[https://zenodo.org/records/3384388\]](https://zenodo.org/records/3384388) (Accessed: Nov., 06, 2024)
- [3] Tagawa, Y., Maskeliūnas, R., Damasevicius, R. "Acoustic Anomaly Detection of Mechanical Failures in Noisy Real-Life Factory Environments". MDPI. Journal Electronics, vol. 10, Issue 19. 10.3390/electronics10192329. [\[https://www.mdpi.com/2079-9292/10/19/2329\]](https://www.mdpi.com/2079-9292/10/19/2329) (Accessed: Nov., 06, 2024.)
- [4] A. Abbasi, A. R. R. Javed, A. Yasin, Z. Jalil, N. Kryvinska and U. Tariq, "A Large-Scale Benchmark Dataset for Anomaly Detection and Rare Event Classification for Audio Forensics," in IEEE Access, vol. 10, pp. 38885-38894, 2022, doi: 10.1109/ACCESS.2022.3166602. [\[https://ieeexplore.ieee.org/abstract/document/9755147\]](https://ieeexplore.ieee.org/abstract/document/9755147) (Accessed: Nov., 06, 2024)
- [5] Ciaburro, G., Padmanabhan, S., Maleh, Y., & Puyana-Romero, V. (2023). Fan Fault Diagnosis Using Acoustic Emission and Deep Learning Methods. Informatics, 10(1). <https://doi.org/10.3390/informatics10010024>
- [6] Tang, S., Yuan, S., & Zhu, Y. (2020). Deep learning-based intelligent fault diagnosis methods toward rotating machinery. IEEE Access, 8, 9335–9346. <https://doi.org/10.1109/ACCESS.2019.2963092>
- [7] Shubita, R. R., Alsadeh, A. S., & Khater, I. M. (2023). Fault Detection in Rotating Machinery Based on Sound Signal Using Edge

Machine Learning. IEEE Access, 11, 6665–6672. <https://doi.org/10.1109/ACCESS.2023.3237074>

[8] Barreto, L. D. T., Rosa, R. K., Silva, D., & Braga, D. (n.d.). Fault detection for rotating machinery based on vibration data using machine learning.

[9] 我校在第九届国际权威声学场景和事件检测及分类竞赛中荣获国际第一名. 发布时间：2023年08月14日.

Disponível em: <https://www.nwpu.edu.cn/info/1198/69658.htm> > Acesso: 27/11/2024. [Tradução: Northwestern Polytechnical University. Nossa escola conquistou o primeiro lugar no mundo na 9ª Competição Internacional Autorizada de Cena Acústica e Detecção e Classificação de Eventos. Agosto. 14. 2023.]

[10] Dohi, Kota; Nishida, Tomoya; Purohit, Harsh; Tanabe, Ryo; Endo, Takashi; Yamamoto, Masaaki; Nikaido, Yuri & Kawaguchi, Yonei; MIMII DG: Sound Dataset for malfunctioning Industrial

Machine Investigation for Domain Generalization Task. Maio. 9. 2022. Version 1.0. Disponível em:

<<https://zenodo.org/records/6529888> > Acessado: 21/11/2024

[11] Durrani, Muhammad Mubashirullah. Laptop Fan Noise. Disponível em:

[<https://www.kaggle.com/datasets/muhammadmdurrani/laptop-fan-noise?resource=download>] (Acessed: Nov., 21, 2024)

[12] Koizumi, Yuma; Kawaguchi, Yohei; Imoto, Keisuke; Nakamura, Toshiaki; Nikaido, Yuki; Tanabe, Ryo; Purohit, Harsh; Suefusa, Kaori; Endo, Takashi; Yasuda, Masahito & Harada, Noburu. Dcase 2020 Challenge Task 2 Development Dataset. Published Mar, 1, 2022. Version 1.0. Disponível em: [<https://zenodo.org/records/3678171>] (Acessed: Nov., 21, 2024)