

Extração Automática de Atributos de Sinais de Emissão Acústica com Redes Neurais Autocodificadoras para Predição de Integridade em Tubulações

Bernardo Dutra¹, Antônio Neves¹, Marcos Carvalho¹, Sérgio Daniel Carvalho Canuto²,
Jorge Wanderley Ribeiro³, Rodrigo Pires³,
Douglas Soares dos Santos⁴, André Lopes Gama da Fonseca⁵,
Jussara Almeida¹, Marcos André Gonçalves¹

¹Departamento de Ciência da Computação
Universidade Federal de Minas Gerais (UFMG) 31270-901 - Belo Horizonte - Brasil

²Instituto Federal de Goiás.

³ISQ Brasil

⁴Petrogal Brasil, Av. República do Chile, 330 - Centro, Rio de Janeiro - RJ, Brasil - 20031-170

⁵Galp, Avenida da Índia 8, 1349-065 Lisboa, Portugal

{bernardolemos, antonioneto, marcoscarvalho, jussara, mgoncalv}@dcc.ufmg.br

sergio.canuto@ifg.edu.br, {jwribeiro, rropires}@isqbrasil.com.br

douglas.santos@petrogalbrasil.com, andre.gama@galp.com

Abstract. *Acoustic Emission (AE) signal analysis is a widely used technique for monitoring structural degradation, such as wall thinning in pipelines. However, extracting informative features from raw AE signals remains challenging due to noise and the complexity of signal propagation. This work proposes a novel feature extraction approach using convolutional Denoising Autoencoders to reduce noise and generate compact, informative latent representations. The proposed method outperforms traditional techniques in thickness prediction tasks, particularly in high-noise scenarios with both synthetic and real data.*

Resumo. *A análise de sinais de Emissão Acústica (EA) é uma técnica amplamente utilizada para monitorar processos de degradação estrutural, como a perda de espessura em tubulações. Contudo, a extração de atributos informativos a partir desses sinais é dificultada por ruídos e pela complexidade física da propagação acústica. Este trabalho propõe uma nova abordagem baseada em Denoising Autoencoders com redes neurais convolucionais para atenuação de ruídos e geração de representações latentes compactas e relevantes. A técnica proposta supera métodos tradicionais de extração de atributos na predição de espessura, especialmente em cenários com alto nível de ruído, utilizando dados sintéticos e reais.*

1. Introdução

O monitoramento contínuo da integridade estrutural de tubulações industriais e oleodutos é crucial para prevenir impactos ambientais, riscos à segurança operacional

e prejuízos econômicos, o que hoje ultrapassa 2.5 trilhões de dólares por ano [Bender et al. 2022]. Sinais de Emissão Acústica (EA) fornecem informações sensíveis associadas à degradação do material [Zang et al. 2023]. A predição da integridade de tubulações via EA é realizada com modelos de aprendizado de máquina treinados a partir de instâncias extraídas dos sinais brutos e suas respectivas rotulações (e.g., espessura da tubulação). No entanto, fatores como ruídos ambientais e efeitos de dispersão, reflexão e espalhamento dos sinais podem comprometer a generalização dos modelos [Zang et al. 2023].

Diversas abordagens têm sido utilizadas com o objetivo de extrair atributos relevantes e capazes de representar as principais características desses sinais. Tais estratégias usualmente envolvem estatísticas simples, como média, desvio padrão e mediana dos valores do sinal de EA [Aldosari et al. 2020], e atributos extraídos por meio das Transformadas Wavelet ou Fourier [Song et al. 2024, Xiao et al. 2019]. Essas transformadas decompõem os sinais em componentes de frequência e tempo-frequência, sendo especialmente úteis para capturar padrões transientes ou localizados ao longo do tempo.

Técnicas no estado da arte utilizam a matriz de coeficientes da Transformada Wavelet Contínua (CWT) como atributos extraídos de sinais de EA, explorando simultaneamente informações no tempo e na frequência [Song et al. 2024, Heuel and Friederich 2022]. Os sinais são inicialmente normalizados com *Z-score* e reduzidos por *Piecewise Aggregate Approximation* [Keogh et al. 2001], que segmenta o sinal e calcula a média de cada segmento. Em seguida, são convertidos em matrizes tempo-frequência via CWT, utilizadas como entrada para redes neurais convolucionais (CNN), cujos filtros extraem padrões espaciais relevantes nas matrizes de coeficientes.

Apesar da eficácia das técnicas tradicionais de extração de atributos, estatísticas e transformadas podem falhar no tratamento de nuances relevantes, especialmente em cenários com variações de níveis de ruído. Propomos como principal contribuição desse artigo, uma estratégia de extração de atributos automática baseada em *denoising autoencoders*, que realiza a atenuação adaptativa de ruídos nos sinais de EA e gera automaticamente um conjunto compacto de atributos latentes que preservam informações discriminativas essenciais dos sinais.

Experimentos realizados com dados reais e sintéticos refletindo cenários com diferentes espessuras, frequências de sinal e níveis de ruído demonstraram que a nossa proposta supera as abordagens existentes. O método se destacou especialmente na predição da espessura de dano em cenários com alta variabilidade espectral e ruído, alcançando erros até três vezes menores que o segundo melhor método e 220 vezes inferiores aos métodos tradicionais. A análise também evidenciou o impacto da espessura e da frequência do sinal no desempenho dos modelos. Por fim, a avaliação com um conjunto experimental real revelou limitações relacionadas ao baixo volume de dados, destacando a necessidade de estratégias de ampliação de dados para aplicações futuras.

2. Trabalhos Relacionados

A predição automática da integridade estrutural em tubulações geralmente se baseia na extração de atributos estatísticos dos sinais de EA, seguidos por modelos de aprendizado de máquina [Aldosari et al. 2020]. Embora simples, essas estatísticas globais têm baixa capacidade de capturar variações locais e fenômenos transitórios. Transformadas como

Wavelet e Fourier oferecem representações mais informativas ao decompor os sinais em faixas de frequência. [Xiao et al. 2019] utilizaram a Transformada Discreta de Wavelet (DWT) para atenuar ruídos e extrair padrões localizados, combinando coeficientes transformados e métricas estatísticas como atributos. A técnica RELIEF foi empregada para reduzir a dimensionalidade, aprimorando a generalização de modelos SVM. De forma semelhante, [May et al. 2021] aplicaram a Transformada por Pacotes Wavelet (PWT) e a Transformada Rápida de Fourier (FFT), utilizando os coeficientes e estatísticas extraídas para classificar a severidade dos danos.

Transformadas discretas representam padrões locais com baixa resolução e perda de informação devido à amostragem. Abordagens baseadas em CWT [Song et al. 2024] têm ganhado destaque por fornecerem representação tempo-frequência densa dos sinais de EA. Os coeficientes CWT são organizados em matrizes bidimensionais e usados diretamente como atributos para redes neurais convolucionais, que exploram sua capacidade de capturar padrões espaciais complexos na estrutura do sinal. Embora os coeficientes CWT utilizados como atributos evitem a perda de informação típica das transformadas discretas, as matrizes CWT apresentam desafios importantes: (1) incorporam ruídos do sinal original, (2) têm alta dimensionalidade, e (3) exigem modelos de predição capazes de extrair padrões locais, como filtros convolucionais. Este trabalho contribui ao propor uma nova estratégia baseada em *denoising autoencoders*, que realiza atenuação adaptativa de ruídos nos sinais de EA, gerando automaticamente um conjunto compacto de variáveis latentes que preservam informações discriminativas essenciais, superando as limitações das abordagens anteriores. Descrevemos nossa proposta a seguir.

3. Proposta: CWT-CNN Denoising Autoencoder (CWT-CNNDA)

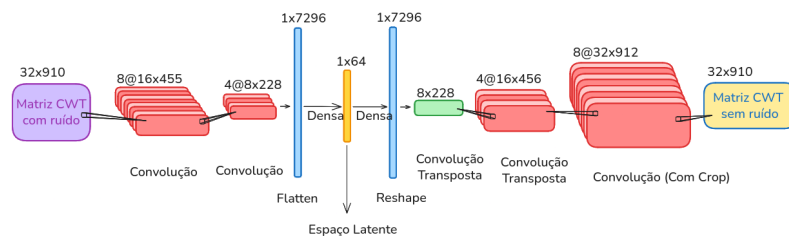


Figure 1. Arquitetura do CWT-CNNDA

A estratégia proposta, denominada CWT-CNNDA, estende a representação de atributos da matriz de coeficientes CWT [Song et al. 2024] visando a construção de um novo conjunto de atributos por meio da aplicação de redes neurais do tipo *Denoising Autoencoder* [Li et al. 2023]. Essas redes foram concebidas, de forma geral, para atender a dois objetivos principais: (1) a remoção automática de ruídos presentes nos dados de entrada e (2) a construção de um espaço latente capaz de capturar as informações discriminativas das instâncias de dados. A Figura 1 ilustra a arquitetura da solução proposta, baseada na codificação da entrada em um espaço latente e compacto, de 64 atributos, e decodificação para reconstrução do sinal original sem ruído.

Inicialmente, os dados de EA são convertidos em uma matriz de coeficientes CWT, de dimensão 32x910, representada no lado direito da figura. Durante o treinamento, para cada instância de EA, foi gerada uma versão artificialmente corrompida, com a adição

de ruído branco ao sinal original [Xu et al. 2021], resultando um par do tipo ⟨sinal ruidoso, sinal original⟩. Com isso, é gerada a Matriz CWT com ruído, representada no lado esquerdo da figura, cuja dimensão é 32×910. A arquitetura do CWT-CNNDA segue o padrão de um *Convolutional Denoising Autoencoder*, com blocos convolucionais no codificador e no decodificador. O processo inicia-se com duas camadas convolucionais que processam a matriz de entrada (dimensão 32×910), gerando representações intermediárias com formatos 8×16×455 e 4×8×228, respectivamente. Após as convoluções, a saída é achatada (*flatten*) e passa por uma camada densa de 64 neurônios, formando o espaço latente, que representa a codificação comprimida da instância de entrada. Na etapa de decodificação, a saída latente é expandida por meio de uma nova camada densa e reformatada para 8×228. Em seguida, são aplicadas duas camadas de convolução transposta que recuperam as dimensões espaciais, passando pelas representações 4×16×456 e 8×32×912. Por fim, uma última camada convolucional (com *crop* para ajuste dimensional) reconstrói o sinal final, resultando na Matriz CWT sem ruído, com as mesmas dimensões da entrada original (32×910). Essa arquitetura foi adotada de forma a capturar de forma eficiente os padrões locais presentes nos escalogramas CWT, utilizando filtros espaciais das camadas convolucionais [Kaji et al. 2020].

4. Resultados

4.1. Configuração Experimental

As estratégias de extração de atributos foram avaliadas em dois conjuntos de dados para a tarefa de regressão (estimativa de perda de espessura) e um conjunto para a tarefa de classificação (nível de corrosão). O primeiro conjunto, denominado como *Coleção de Dados 1*, foi gerado por simulação com a ferramenta COMSOL [Multiphysics C 1998], modelando uma tubulação de 0,42m. A simulação acústica simulou quatro cenários de equipamentos que utilizam frequências de 200kHz e 240kHz, e duas espessuras de tubulação (50mm e 100mm). Com isso, criou-se quatro subconjuntos com 2.548.000 pontos no total. A simulação considerou 14 variações de perda de espessura, de 9,5 mm a 3 mm em passos de 0,5 mm, com adição de ruído gaussiano (desvio padrão 0,2) aos sinais acústicos para representar condições reais [Xu et al. 2021].

O segundo conjunto, denominado como *Coleção de Dados 2*, foi disponibilizado pelos autores em [May et al. 2021]. Nele, a corrosão foi induzida pela imersão de um tubo de metal em uma solução de ácido clorídrico, sendo 457 instâncias de EA coletadas ao longo de 12 horas de experimento em uma tubulação. Cerca de 50 instâncias foram registradas na primeira hora, aproximadamente 300 entre a sexta e a sétima hora, e o restante distribuído de forma semelhante nas demais horas. Por fim, o terceiro conjunto de dados, denominado como *Coleção de Dados 3* foi disponibilizado por [Sung et al. 2024], com 60.000 instâncias de sinais de EA e seis categorias de níveis de corrosão correspondentes à concentração de óxido de ferro (1, 3, 5, 200, 300 e 15ml/L) mensurada em laboratório.

Nas comparações de efetividade, foram utilizados atributos extraídos de abordagens previamente propostas na literatura [Song et al. 2024, Xiao et al. 2019, May et al. 2021], além de três grupos de atributos: estatísticos, Wavelet e Fourier, anteriormente adotados de forma combinada, sem a análise de cada grupo individual [May et al. 2021]. Os atributos estatísticos consistem em cinco medidas simples extraídas diretamente da sequência de pontos do sinal acústico: média, desvio padrão,

RMS, mediana e média dos valores absolutos. Para os atributos Wavelet (DWT) e Fourier (FFT), as mesmas cinco estatísticas são extraídas dos sinais decompostos.

Para a tarefa de classificação foi utilizado o método Support Vector Machine (SVM), amplamente empregado no contexto de predição da integridade de tubulações, devido à sua reconhecida capacidade de generalização [Xiao et al. 2019, May et al. 2021]. No SVM, o parâmetro de regularização C foi ajustado com validação cruzada para obtenção do melhor de 9 valores entre 10^{-4} à 10^4 . Já para tarefas de regressão, foi utilizado o método Extreme Gradient Boosting (XGBoost), que emprega comitês de árvores de regressão, sendo amplamente reconhecido como um dos modelos estado da arte para tarefas de regressão [Shwartz-Ziv and Armon 2022]. Neste trabalho, foram adotados os valores padrão de parametrização do XGBoost. Todos os modelos foram treinados utilizando validação cruzada com 5 *folds*, e os resultados foram reportados com intervalos de confiança de 95% para as métricas de acurácia e erro quadrático médio (RMSE), correspondentes às tarefas de classificação e regressão, respectivamente.

4.2. Efetividade para Tarefa de Predição de Espessura

A Tabela 1 apresenta os resultados da tarefa de regressão, mostrando os valores médios de RMSE na predição da espessura de dano. Primeiramente, são apresentados os valores de RMSE para a *Coleção de Dados 1*, que considera dois níveis de espessura e duas frequências. Em seguida, é mostrado o valor médio do RMSE para a *Coleção de Dados 2* ([May et al. 2021]). O CWT-CNNDA obteve o melhor desempenho nos dados sintéticos, com erros significativamente menores que os demais métodos. Superou em pelo menos 3x a abordagem de [Song et al. 2024], que também usa CWT e CNN, mas sem denoising autoencoders. Em comparação com métodos baseados em estatísticas, Wavelet ou Fourier, os erros foram até 220x menores, evidenciando a superioridade da extração robusta com remoção de ruído.

Atributos	Coleção de Dados 1				Coleção de Dados 2
	Espessura da Tubulação		Frequência do Sinal		
	50mm	100mm	240KHz	200KHz	
CWT-CNNDA	0.0007(0.0001)	0.003(0.002)	0.02(0.004)	0.07(0.005)	1.8(0.5)
[Song et al. 2024]	0.0183(0.0028)	0.024(0.004)	0.07(0.010)	0.19(0.012)	1.8(0.4)
Estatística	0.3474(0.0291)	1.263(0.035)	1.42(0.039)	1.68(0.032)	1.4(0.2)
Wavelet	1.8562(0.0783)	1.657(0.059)	1.74(0.029)	1.75(0.032)	1.2(0.4)
Fourier	0.1908(0.0145)	0.673(0.038)	1.35(0.039)	1.62(0.041)	2.0(0.6)
[Xiao et al. 2019]	0.0024(0.0010)	0.009(0.005)	0.14(0.008)	0.34(0.017)	1.1(0.1)
[May et al. 2021]	0.5102(0.0358)	0.592(0.019)	1.22(0.039)	1.62(0.038)	0.9(0.2)

Table 1. Erro RMSE médio (e ICs com 95% de confiança) dos métodos de extração de features para predição da espessura em diferentes conjuntos de dados simulando espessura e frequência do sinal - dados de [May et al. 2021].

A extração de atributos via transformadas, aliada à remoção de ruído pela DWT [Xiao et al. 2019], mostrou-se eficaz na redução de erros em comparação a métodos mais simples (e.g., estatísticas básicas, Fourier, Wavelet isoladas). No entanto, na *Coleção de Dados 1*, os erros ainda foram cerca de três vezes maiores do que os obtidos com o CWT-CNNDA. Por outro lado, na *Coleção de Dados 2*, CWT-CNNDA teve desempenho inferior ao de abordagens mais simples, como a de [May et al. 2021], devido à limitação de dados. A eficácia de autoencoders depende de um volume representativo de exemplos para aprender representações latentes robustas, o que compromete sua performance em cenários com poucos dados ou alta variabilidade, como no EXP. O erro de

predição aumentou com a espessura da tubulação, especialmente em 100mm, devido à maior atenuação e distorção dos sinais acústicos em materiais espessos, dificultando a extração de atributos informativos. Sinais com frequência de 240kHz apresentaram melhor resolução espacial que os de 200kHz, sendo mais eficazes na detecção de defeitos sutis e variações finas de espessura.

Atributos	Nível de Ruído (<i>Coleção de Dados 3</i>)				
	0	0.05	0.10	0.20	0.40
CWT-CNNDA	1.00(0.00)	0.99(0.001)	0.95(0.003)	0.86(0.010)	0.74(0.006)
[Song et al. 2024]	1.00(0.00)	0.96(0.007)	0.87(0.005)	0.78(0.009)	0.67(0.031)
Estatísticas	0.90(0.003)	0.79(0.003)	0.65(0.003)	0.52(0.003)	0.37(0.006)
Wavelet	0.89(0.003)	0.74(0.041)	0.65(0.004)	0.52(0.003)	0.39(0.006)
Fourier	0.88(0.005)	0.70(0.004)	0.59(0.002)	0.48(0.03)	0.35(0.003)
[Xiao et al. 2019]	0.99(0.001)	0.89(0.006)	0.79(0.011)	0.62(0.006)	0.33(0.017)
[May et al. 2021]	0.99(0.002)	0.91(0.010)	0.79(0.005)	0.61(0.009)	0.42(0.005)

Table 2. Acurácia média (e ICs com 95% de confiança) dos métodos de extração de features em diferentes níveis de ruído para classificação da corrosão da tubulação (presença de óxido de ferro) nos dados de [Sung et al. 2024].

A Tabela 2 apresenta os resultados da tarefa de classificação, especificamente a acurácia média obtida pelas diferentes estratégias de extração de atributos sob distintos níveis de ruído na *Coleção de Dados 3*. CWT-CNNDA obteve desempenho consistentemente superior em todos os níveis, alcançando acurácia de 0.74 mesmo com ruído elevado, com ganhos de 10% a 55% em relação às demais técnicas nesse cenário adverso. Tal resultado se deve à capacidade dos *Denoising Autoencoders* de eliminar componentes ruidosos e gerar atributos latentes robustos, a partir do treinamento com mais de 40 mil instâncias rotuladas. A abordagem de [Song et al. 2024] apresentou, em geral, o segundo melhor desempenho, destacando-se pela combinação de representações multiescalares via matriz CWT e redes convolucionais, eficazes na captura de padrões recorrentes em grandes volumes de dados. Em contraste, métodos baseados em estatísticas simples, Fourier e Wavelet tiveram quedas expressivas de desempenho com o aumento do ruído, sugerindo que seus atributos se confundem com padrões ruidosos em cenários degradados. Tendência similar foi observada nos métodos de [Xiao et al. 2019] e [May et al. 2021], também sensíveis ao ruído, evidenciando a limitação de abordagens sem mecanismos de atenuação adaptativa.

5. Conclusões e Trabalhos Futuros

Nosso método CWT-CNNDA obteve desempenho superior nas coleções sintéticas, superando amplamente as demais abordagens. Seus erros foram até três vezes menores que o segundo melhor método. Foi ainda cerca de 220x mais eficaz que métodos baseados em estatísticas simples, Fourier ou Wavelet. Mesmo sob ruído, o CWT-CNNDA manteve alta acurácia, com ganhos de até 55%. Isso se deve à capacidade dos *denoising autoencoders* de filtrar ruído e gerar representações latentes robustas. Como trabalhos futuros, propõem-se testes com sinais reais, análise de eficiência e uso de atributos adicionais à matriz CWT.

Acknowledgments

Este trabalho foi apoiado por Petrogal Brasil S.A. (Joint Venture Galp — Sinopec), Agência Nacional do Petróleo, Gás Natural e Biocombustíveis (ANP)(Nº ANP: 23490-6), CNPq, Capes, Fapemig, Fapesp, AWS, NVIDIA, CIIA-Saúde e Instituto Nacional de Ciência e Tecnologia em Inteligência Artificial Responsável para Linguística Computacional, Tratamento e Disseminação de Informação (INCT-TILD-IAR; 408490/2024-1).

References

- Aldosari, H., Elfouly, R., and Ammar, R. (2020). Evaluation of machine learning-based regression techniques for prediction of oil and gas pipelines defect. In *2020 International Conference on Computational Science and Computational Intelligence (CSCI)*, pages 1452–1456.
- Bender, R., Féron, D., Mills, D., Ritter, S., Bäßler, R., Bettge, D., De Graeve, I., Dugstad, A., Grassini, S., Hack, T., et al. (2022). Corrosion challenges towards a sustainable society. *Materials and corrosion*, 73(11):1730–1751.
- Heuel, J. and Friederich, W. (2022). Suppression of wind turbine noise from seismological data using nonlinear thresholding and denoising autoencoder. *Journal of Seismology*, 26(5):913–934.
- Kaji, M., Parvizian, J., and van de Venn, H. W. (2020). Constructing a reliable health indicator for bearings using convolutional autoencoder and continuous wavelet transform. *Applied Sciences*, 10(24).
- Keogh, E., Chakrabarti, K., Pazzani, M., and Mehrotra, S. (2001). Dimensionality reduction for fast similarity search in large time series databases. *Knowledge and Information Systems*, 3(3):263–286.
- Li, P., Pei, Y., and Li, J. (2023). A comprehensive survey on design and application of autoencoder in deep learning. *Applied Soft Computing*, 138:110176.
- May, Z., Alam, M., Nayan, N. A., Rahman, N. A. A., and Mahmud, M. S. (2021). Acoustic emission corrosion feature extraction and severity prediction using hybrid wavelet packet transform and linear support vector classifier. *Plos one*, 16(12):e0261040.
- Multiphysics C (1998). *Introduction to COMSOL Multiphysics®*. COMSOL Multiphysics, Burlington, MA. Accessed February 9, 2018.
- Shwartz-Ziv, R. and Armon, A. (2022). Tabular data: Deep learning is not all you need. *Information Fusion*, 81:84–90.
- Song, L., Cui, X., Han, X., Gao, Y., Liu, F., Yu, Y., and Yuan, Y. (2024). A non-metallic pipeline leak size recognition method based on cwt acoustic image transformation and cnn. *Applied Acoustics*, 225:110180.
- Sung, Y., Jeon, H. J., Kim, D., et al. (2024). Internal pipe corrosion assessment method in water distribution system using ultrasound and convolutional neural networks. *npj Clean Water*, 7:63.
- Xiao, R., Hu, Q., and Li, J. (2019). Leak detection of gas pipelines using acoustic signals based on wavelet transform and support vector machine. *Measurement*, 146:479–489.
- Xu, Z.-D., Zhu, C., and Shao, L.-W. (2021). Damage identification of pipeline based on ultrasonic guided wave and wavelet denoising. *Journal of Pipeline Systems Engineering and Practice*, 12(4):04021051.
- Zang, X., Xu, Z.-D., Lu, H., Zhu, C., and Zhang, Z. (2023). Ultrasonic guided wave techniques and applications in pipeline defect detection: A review. *International Journal of Pressure Vessels and Piping*, 206:105033.