



INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA DA BAHIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE MATERIAIS

Luiz Claudio de Souza da Cunha

**COMBINAÇÃO DE ENSAIOS ULTRASSÔNICOS E APRENDIZADO
DE MÁQUINA NA DETECÇÃO DE DEFEITOS EM REVESTIMENTO
DE TUBULAÇÕES INDUSTRIAIS**

Salvador

2025

Luiz Claudio de Souza da Cunha

**COMBINAÇÃO DE ENSAIOS ULTRASSÔNICOS E APRENDIZADO
DE MÁQUINA NA DETECÇÃO DE DEFEITOS EM REVESTIMENTO
DE TUBULAÇÕES INDUSTRIAIS**

Dissertação apresentada ao Programa de Pós-Graduação Stricto Sensu – Mestrado Profissional em Engenharia de Materiais (PPGEM) – do Instituto Federal da Bahia (IFBA), como parte dos requisitos necessários à obtenção do título de Mestre em Engenharia dos Materiais.

Orientador:

Prof. Dr. Ivan Costa da Silva

Coorientadora:

Profa. Dra. Maria Cléa Soares de
Albuquerque

Salvador

2025

FICHA CATALOGRÁFICA ELABORADA PELO SISTEMA DE BIBLIOTECAS DO IFBA, COM OS
DADOS FORNECIDOS PELO(A) AUTOR(A)

C972c Cunha, Luiz Claudio de Souza da

Combinação de ensaios ultrassônicos e aprendizado de máquina na detecção de defeitos em revestimento de tubulações industriais / Luiz Claudio de Souza da Cunha; orientador Ivan Costa da Silva; coorientadora Maria Cléa Soares de Albuquerque -- Salvador, 2025.

93 p.

Dissertação (Programa de Pós-Graduação em Engenharia de Materiais) -- Instituto Federal da Bahia, 2025.

1. Técnica ultrassônica pulso-eco. 2. Processamento de sinais. 3. SVM. 4. Revestimento interno em argamassa. 5. Tubulações industriais. I. Silva, Ivan Costa da, orient. II. Albuquerque, Maria Cléa Soares de, coorient. III. TÍTULO.

CDU 621.644

INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA DA BAHIA
R. Emídio dos Santos - Bairro Barbalho - CEP 40301-015 - Salvador - BA - www.portal.ifba.edu.br

PRÓ-REITORIA DE PESQUISA, PÓS-GRADUAÇÃO E INOVAÇÃO

PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE MATERIAIS - PPGEM

**COMBINAÇÃO DE ENSAIOS ULTRASSÔNICOS E APRENDIZADO DE MÁQUINA NA
DETECÇÃO DE DEFEITOS EM REVESTIMENTO DE TUBULAÇÕES INDUSTRIAIS**

LUIZ CLÁUDIO SOUZA CUNHA

Produto(s) Gerado(s): () Dissertação (X) Relatório Técnico () Publicação em Periódico
() Artigo Técnico () Patente () Registro de software () Outro: especificar

Orientador: Prof. Dr. Ivan Costa da Silva – PPGEM/IFBA

Coorientadora: Profa. Dra. Maria Cléa Soares de Albuquerque – PPGEM/IFBA

Banca examinadora:

Prof. Dr. Ivan Costa da Silva

Orientador do PPGEM – Instituto Federal da Bahia – IFBA

Profa. Dra. Maria Cléa Soares de Albuquerque

Coorientadora do PPGEM – Instituto Federal da Bahia – IFBA

Prof. Dr. Eduardo Telmo Fonseca Santos

Examinador Externo ao PPGEM – Instituto Federal da Bahia – IFBA

Prof. Dr. Lucas Cruz da Silva

Examinador Externo – SENAI / CIMATEC

Prof. Dr. Manoel Messias Silva Júnior

Examinador Externo – UFOB



Documento assinado eletronicamente por **Manoel Messias Silva Júnior, Usuário Externo**, em 04/04/2025, às 12:34, conforme decreto nº 8.539/2015.



Documento assinado eletronicamente por **Lucas Cruz da Silva, Usuário Externo**, em 07/04/2025, às 07:56, conforme decreto nº 8.539/2015.



Documento assinado eletronicamente por **MARIA CLEA SOARES DE ALBUQUERQUE, Professor Efetivo**, em 08/04/2025, às 20:23, conforme decreto nº 8.539/2015.



Documento assinado eletronicamente por **IVAN COSTA DA SILVA, Professor(a) do Ensino Básico, Técnico e Tecnológico - EBTT**, em 11/04/2025, às 14:52, conforme decreto nº 8.539/2015.



Documento assinado eletronicamente por **EDUARDO TELMO FONSECA SANTOS, Docente Permanente**, em 12/04/2025, às 10:12, conforme decreto nº 8.539/2015.



A autenticidade do documento pode ser conferida no site http://sei.ifba.edu.br/sei/controlador_externo.php?acao=documento_conferir&acao_origem=documento_conferir&id_orgao_acesso_externo=0 informando o código verificador **4094203** e o código CRC **4BCDF146**.

AGRADECIMENTOS

A realização deste trabalho só foi possível graças ao apoio e incentivo de inúmeras pessoas e instituições, às quais expresso minha mais profunda gratidão.

Ao Prof. Dr. Ivan Costa da Silva, por sua orientação precisa, paciência e, sobretudo, disponibilidade, sempre disposto a me atender, mesmo em horários e dias muitas vezes inconvenientes, contribuindo de forma inestimável para o avanço desta pesquisa.

À minha coorientadora, Profa. Dra. Maria Cléa Soares de Albuquerque, cujo olhar crítico e busca incessante pela excelência me desafiaram a aprimorar constantemente este trabalho.

O compromisso de ambos com o PPGEM e sua dedicação ao ensino são fontes genuínas de inspiração.

À minha família, que esteve ao meu lado em cada desafio desta jornada. Aos meus pais, pelo amor, pelos valores que me transmitiram e por acreditarem em mim.

À minha esposa, Aline, pelo suporte incansável, pela compreensão diante das minhas ausências, e por assumir tantas responsabilidades para que eu pudesse me dedicar a este projeto.

Aos meus filhos, João Pedro e Lara, e ao meu enteado, Ruan, dedico este trabalho com todo o meu amor. Que esta conquista sirva de exemplo para que nunca desistam de seus sonhos e acreditem na importância da perseverança e do conhecimento.

À Transpetro e ao IFBA, pelo suporte direto e indireto à realização deste estudo.

Aos meus colegas de trabalho e do PPGEM, pelo compartilhamento de conhecimento, pelas valiosas contribuições e pelas sugestões enriquecedoras ao longo dessa trajetória.

Um agradecimento especial às graduandas Ana Eliza Melo e Beatriz Gomes de Oliveira, que, como integrantes da iniciação científica, contribuíram significativamente no desenvolvimento dos experimentos nos laboratórios de Inspeção Não Destrutiva do IFBA

Por fim, a todos que, de alguma forma, contribuíram para que este trabalho se tornasse realidade, meu mais sincero agradecimento.

O sucesso é a soma de pequenos
esforços repetidos dia após dia.

Robert Collier

RESUMO

A inspeção de tubulações industriais com revestimento interno de argamassa representa um desafio, especialmente em dutos que não permitem inspeção interna com equipamentos automatizados, como pigs instrumentados. Este estudo avaliou como o classificador SVM Linear, alimentado por diferentes técnicas de processamento de sinais ultrassônicos, distingue entre regiões íntegras e defeituosas nesses revestimentos. Ensaios ultrassônicos Pulso-Eco foram realizados em corpos de prova fabricados de tubos de aço carbono, e os sinais adquiridos foram analisados nos domínios do tempo e da frequência. A partir do sinal *A-scan*, no domínio do tempo, foram extraídos parâmetros como energia do sinal, coeficiente de atenuação e estatísticas descritivas. Já no domínio da frequência, o processamento envolveu a aplicação das Transformadas Discreta de Fourier (DFT) e do Cosseno (DCT), tanto isoladamente quanto de forma combinada. Para redução de dimensionalidade, foram avaliadas a Análise de Componentes Principais (PCA), a Análise Discriminante Linear (LDA) e sua combinação PCA-LDA. Os experimentos foram conduzidos no MATLAB®, utilizando um corpo de prova para treinamento e validação cruzada em cinco divisões, além de um conjunto de teste independente, composto por 20% das amostras deste mesmo corpo de prova. Na sequência, o modelo foi testado em um segundo corpo de prova para avaliar sua capacidade de generalização em um conjunto de dados não visto previamente. O estudo confirmou a validade da metodologia proposta na classificação da integridade de revestimentos internos de tubulações, destacando a combinação das transformadas DFT e DCT com LDA como a combinação mais eficiente.

Palavras-chave: Técnica ultrassônica Pulso-eco; processamento de sinais; SVM; revestimento interno em argamassa; tubulações industriais.

ABSTRACT

The inspection of industrial pipelines with internal mortar linings poses a challenge, especially for pipelines that do not allow internal inspection with automated equipment such as in-line inspection (ILI) tools. This study evaluates the capability of a Linear SVM classifier, trained with different ultrasonic signal processing techniques, to distinguish between intact and defective regions in these linings. Pulse-echo ultrasonic testing was conducted on carbon steel pipe specimens, and the acquired signals were analyzed in both time and frequency domains. From the time-domain A-scan signals, features such as signal energy, attenuation coefficient, and statistical descriptors were extracted. In the frequency domain, the processing involved the application of the Discrete Fourier Transform (DFT) and the Discrete Cosine Transform (DCT), both individually and in combination. Dimensionality reduction techniques, including Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA), and their combination (PCA-LDA), were assessed. Experiments were conducted in MATLAB®, using one specimen for training with five-fold cross-validation and an independent test set comprising 20% of the samples. The trained model was subsequently tested on a second specimen to evaluate its generalization capability on previously unseen data. The study confirmed the effectiveness of the proposed methodology in classifying the integrity of internal coatings in pipelines, with the combination of DFT and DCT with LDA emerging as the most efficient approach.

Keywords: Pulse-echo ultrasonic technique; signal processing; SVM; internal mortar lining; industrial pipelines.

SUMÁRIO

| | |
|--|-----------|
| Lista de figuras | iii |
| Lista de tabelas..... | vi |
| Lista de abreviaturas, siglas e símbolos..... | vii |
| 1 INTRODUÇÃO..... | 8 |
| 1.1 Objetivo geral..... | 12 |
| 1.2 Objetivos específicos | 12 |
| 2 REVISÃO DA LITERATURA | 13 |
| 2.1 Ultrassom | 13 |
| 2.1.1 Técnica de Pulso-Eco..... | 14 |
| 2.1.2 Propriedades das Ondas Ultrassônicas | 16 |
| 2.1.2.1 Velocidade de Propagação | 17 |
| 2.1.2.2 Impedância Acústica..... | 17 |
| 2.1.2.3 Energia do Sinal..... | 18 |
| 2.1.2.4 Parâmetros Estatísticos no Domínio do Tempo | 19 |
| 2.1.3 Reflexão e Transmissão | 21 |
| 2.1.4 Atenuação | 22 |
| 2.2 Processamento de Sinais | 24 |
| 2.2.1 Transformada Discreta de Fourier (DFT) | 24 |
| 2.2.2 Transformada Discreta do Cosseno (DCT) | 26 |
| 2.2.3 Análise de Componentes Principais (PCA)..... | 28 |
| 2.2.4 Análise Discriminante Linear (LDA) | 32 |
| 2.3 Aprendizado de Máquina | 34 |
| 2.3.1 Fundamentos de Aprendizado de Máquina..... | 35 |
| 2.3.2 <i>Support Vector Machine</i> (SVM) | 38 |
| 3 MATERIAL E MÉTODOS | 41 |
| 3.1 Material..... | 41 |
| 3.2 Métodos..... | 42 |
| 3.2.1 Preparação dos Corpos de Prova..... | 43 |
| 3.2.2 Inspeção Ultrassônica Pulso-Eco | 45 |
| 3.2.3 Processamento dos Sinais..... | 48 |
| 3.2.3.1 Pré-processamento dos Sinais no Domínio do Tempo..... | 48 |
| 3.2.3.2 Pré-processamento dos Sinais no Domínio da Frequência..... | 50 |
| 3.2.3.3 Análise dos Componentes Principais (PCA) | 51 |
| 3.2.3.4 Análise Discriminante Linear (LDA)..... | 51 |

| | |
|--|-----------|
| 3.2.4 Aprendizado de Máquina | 52 |
| 4 RESULTADOS E DISCUSSÃO | 55 |
| 4.1 Avaliação das Características dos Sinais no Domínio do Tempo..... | 55 |
| 4.1.1 Coeficiente de Atenuação | 55 |
| 4.1.2 Energia do Sinal | 56 |
| 4.1.3 Média do Sinal | 57 |
| 4.1.4 RMS (<i>Root Mean Square</i>) do Sinal..... | 58 |
| 4.1.5 Desvio Padrão do Sinal | 59 |
| 4.1.6 Curtose do Sinal | 60 |
| 4.1.7 Análise Consolidada e Limitações dos Parâmetros Ultrassônicos | 61 |
| 4.1.8 Desempenho do SVM Linear com as Características Isoladas | 62 |
| 4.2 Análise Espectral e Extração de Parâmetros no Domínio da Frequência..... | 63 |
| 4.3 SVM Linear com Diferentes Métodos de Processamento no CP2..... | 65 |
| 4.4 Generalização do Modelo com Dados do CP1..... | 69 |
| 5 CONCLUSÕES..... | 75 |
| 6 PERSPECTIVAS PARA TRABALHOS FUTUROS | 77 |
| REFERÊNCIAS BIBLIOGRÁFICAS..... | 78 |

Lista de figuras

| | |
|---|----|
| Figura 1. Foto aérea do terminal aquaviário de Madre de Deus e suas instalações..... | 8 |
| Figura 2. Matriz de Tolerabilidade de Riscos. | 9 |
| Figura 3. Técnicas de Inspeção: a) Pig instrumentado; b) ondas guiadas digital; c) MFL. ... | 10 |
| Figura 4. Revestimento interno de argamassa em tubulação: (a) aplicação do revestimento; (b) tubulação substituída devido à degradação severa do revestimento interno. | 11 |
| Figura 5. Medidor de espessura por ultrassom. | 13 |
| Figura 6. Circuito básico de ultrassom pulso-eco. | 14 |
| Figura 7. Sinal típico de <i>A-scan</i> com o pulso inicial e os ecos refletidos da superfície oposta. | 15 |
| Figura 8. Deformações causadas por ondas planares volumétricas ao longo do eixo x..... | 16 |
| Figura 9. Reflexão e transmissão da onda sônica..... | 21 |
| Figura 10. Sinal típico de <i>A-scan</i> com curva de atenuação ajustada. | 23 |
| Figura 11. Transformada de sinal do domínio do tempo para o domínio da frequência via DFT. | 24 |
| Figura 12. Diferença entre a DCT e DFT. | 27 |
| Figura 13. Análise de Componentes Principais..... | 29 |
| Figura 14. Diferença entre PCA e LDA. | 33 |
| Figura 15. Validação cruzada | 36 |
| Figura 16. Complexidade ideal do modelo em aprendizado de máquina. | 37 |
| Figura 17. Funcionamento do algoritmo SVM. | 38 |
| Figura 18. Vista frontal (a) e posterior (a) do tubo usado para fabricação do corpo de prova. | 41 |
| Figura 19. Fluxograma de Metodologia para Análise de Sinais Ultrassônicos..... | 42 |

| | |
|---|----|
| Figura 20. Corpos de prova CP1 e CP2 com indicação das regiões de interesse..... | 44 |
| Figura 21. Configuração experimental e equipamentos utilizados para a coleta dos sinais ultrassônicos..... | 45 |
| Figura 22. Sinais típicos <i>A-scan</i> de onda completa nas três regiões de interesse (S1, S2 e S3) dos corpos de prova CP1 e CP2. | 46 |
| Figura 23. Sinais típicos <i>A-scan</i> do segundo eco de fundo nas três regiões de interesse (S1, S2 e S3) dos corpos de prova CP1 e CP2. | 47 |
| Figura 24 - Gráfico <i>boxplot</i> - Coeficiente de atenuação das diferentes regiões dos CP1 e CP2. | 55 |
| Figura 25. Gráfico <i>boxplot</i> – Energia do sinal das diferentes regiões dos CP1 e CP2..... | 56 |
| Figura 26. Gráfico <i>boxplot</i> – Média do sinal das diferentes regiões dos CP1 e CP2. | 57 |
| Figura 27. Gráfico <i>boxplot</i> – RMS do sinal das diferentes regiões dos CP1 e CP2..... | 58 |
| Figura 28. Gráfico <i>boxplot</i> – Desvio padrão do sinal das diferentes regiões dos CP1 e CP2. | 59 |
| Figura 29. Gráfico <i>boxplot</i> – Curtose do sinal das diferentes regiões dos CP1 e CP2. | 60 |
| Figura 30. Espectros Médios de Frequência nas Regiões de CP1 e CP2..... | 63 |
| Figura 31. Matrizes de confusão do SVM Linear com a cadeia DFT-DCT-LDA nas etapas de (a) validação e (b) teste com dados do CP2. | 66 |
| Figura 32. Matrizes de confusão do SVM Linear com a cadeias DFT-DCT-PCA-LDA nas etapas de (a) validação e (b) teste com dados do CP2..... | 67 |
| Figura 33. Gráfico de dispersão (<i>Scatter Plot</i>) DFT-PCA. | 68 |
| Figura 34. Gráfico de dispersão (<i>Scatter Plot</i>) DFT-LDA. | 68 |
| Figura 35. Matrizes de confusão do SVM Linear com as cadeias (a) DFT-LDA, (b) DCT-LDA e (c) DFT-DCT-LDA no teste com dados do CP1. | 70 |
| Figura 36. Gráficos de dispersão dos discriminantes lineares gerados pelas cadeias (a) DFT-DCT- LDA sem normalização para o (a) CP1 e (b) CP2. | 71 |

| | |
|--|----|
| Figura 37. Gráficos de dispersão dos discriminantes lineares gerados pelas cadeias DFT-DCT- LDA com normalização dos dados para o (a) CP1 e (b) CP2..... | 71 |
| Figura 38. Gráficos de dispersão dos discriminantes lineares gerados pelas cadeias DFT-DCT- PCA-LDA para (a) CP1 e (b) CP2. | 72 |
| Figura 39. Matrizes de confusão do SVM Linear com as cadeias (a) DFT-PCA-LDA, (b) DCT-PCA-LDA e (c) DFT-DCT-PCA-LDA no teste com dados do CP1..... | 72 |
| Figura 40. Efeito da redução de dimensionalidade via PCA sobre a acurácia do modelo SVM Linear para a cadeia DFT-DCT-PCA-LDA. | 73 |

Lista de tabelas

| | |
|--|----|
| Tabela 1. Propriedades acústicas dos materiais. | 18 |
| Tabela 2. Descrição das regiões de interesse nos corpos de prova CP1 e CP2. | 44 |
| Tabela 3. Cadeias de processamento avaliadas em combinação com o algoritmo SVM Linear | 52 |
| Tabela 4. Resultados consolidados dos parâmetros ultrassônicos | 61 |
| Tabela 5. Desempenho SVM Linear com as Característica no Domínio do Tempo no CP2. | 62 |
| Tabela 6. Parâmetros Extraídos dos Espectros de Frequência para CP1 e CP2. | 64 |
| Tabela 7. Desempenho SVM Linear com diferentes métodos de processamento dos sinais CP2. | 65 |
| Tabela 8. Desempenho SVM Linear com Dados do CP1..... | 69 |

Lista de abreviaturas, siglas e símbolos

- **ANP:** Agência Nacional do Petróleo, Gás Natural e Biocombustíveis
- **API:** *American Petroleum Institute*
- **AWWA:** *American Water Works Association*
- **CP:** Corpo de Prova
- **DA:** *Direct Assessment* (Avaliação Direta)
- **DFT:** *Discrete Fourier Transform* (Transformada Discreta de Fourier)
- **DCT:** *Discrete Cosine Transform* (Transformada Discreta do Cosseno)
- **END:** Ensaio Não Destrutivo
- **FFT:** *Fast Fourier Transform* (Transformada Rápida de Fourier)
- **LDA:** *Linear Discriminant Analysis* (Análise Discriminante Linear)
- **MFL:** *Magnetic Flux Leakage* (Vazamento de Fluxo Magnético)
- **PCA:** *Principal Component Analysis* (Análise de Componente Principal)
- **RMS:** *Root Mean Square* (Valor Efetivo Quadrático)
- **RTT:** Regulamento Técnico de Terminais
- **SVM:** *Support Vector Machine* (Máquina de Vetores de Suporte)

1 INTRODUÇÃO

A gestão da integridade de dutos portuários é essencial para mitigar riscos à segurança das pessoas e proteger o meio ambiente, particularmente em terminais aquaviários destinados ao transporte de petróleo e seus derivados, gás natural e biocombustíveis. Esses dutos desempenham um papel estratégico na logística de transporte entre terminais e embarcações, atravessando frequentemente áreas ecologicamente sensíveis, onde falhas podem gerar impactos ambientais severos e custos financeiros elevados. Estudos demonstram que a instalação e operação dessas infraestruturas frequentemente resultam em conflitos socioambientais, afetando comunidades tradicionais e gerando desafios na mitigação de danos ambientais (SILVA e FURTADO, 2021).

A Figura 1 apresenta uma vista aérea do terminal aquaviário de Madre de Deus, situado na Baía de Todos os Santos, Bahia. Sua infraestrutura, composta por tanques de armazenamento, dutos e instalações portuárias, evidencia a complexidade operacional e a importância da gestão da integridade dos dutos para garantir a segurança das operações e preservação do meio ambiente.

Figura 1. Foto aérea do terminal aquaviário de Madre de Deus e suas instalações.



Fonte: Autoria Própria

Nesse contexto, o Regulamento Técnico de Terminais (RTT), instituído pela Resolução nº 810/2020 da Agência Nacional do Petróleo, Gás Natural e

Biocombustíveis (ANP), estabelece diretrizes para a segurança operacional e a importância de programas de inspeção e manutenção como parte da gestão de integridade estrutural destas instalações (BRASIL, 2020).

A segurança e a confiabilidade das instalações industriais dependem diretamente da identificação de perigos e da análise de riscos em todas as fases do seu ciclo de vida. Como exemplo, a Figura 2 mostra a Matriz de Tolerabilidade de Risco, conforme definida na norma Petrobras N-2782, que estabelece uma abordagem sistemática para a classificação de riscos industriais, considerando a severidade das consequências e a probabilidade de ocorrência de eventos indesejáveis (PETROBRAS, 2015). Neste contexto, a inspeção de equipamentos desempenha um papel relevante, fornecendo informações importantes sobre a integridade e permitindo a mitigação de riscos operacionais.

Figura 2. Matriz de Tolerabilidade de Riscos.

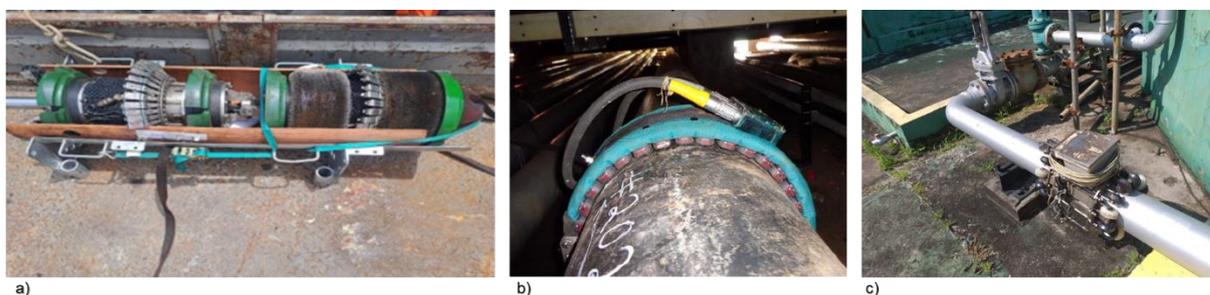
| Descrição / características | | | | | Categorias de frequência | | | | | | |
|--|-----|--------------|--|--|--------------------------|---|---|---|---|--|----|
| | | | | | A | B | C | D | E | | |
| | | | | | Extremamente remota | Remota | Pouco provável | Provável | Frequente | | |
| | | Pessoas | Patrimônio / continuidade operacional | Meio ambiente (ver Nota 1) | Imagem | Possível mas sem referências na indústria | Não esperado ocorrer, apesar de haver referências em instalações similares na indústria | Pouco provável de ocorrer durante a vida útil de um conjunto de instalações similares | Possível de ocorrer uma vez durante a vida útil da instalação | Possível de ocorrer muitas vezes durante a vida útil da instalação | |
| Categorias de Severidade das Consequências | V | Catastrófica | Múltiplas fatalidades intramuros ou fatalidade extramuros (ver Nota 2) | Danos catastróficos podendo levar à perda da instalação industrial | Danos catastróficos | Repercussão internacional | M | M | NT | NT | NT |
| | IV | Crítica | Fatalidade intramuros ou lesões graves extramuros (ver Nota 3) | Danos severos a sistemas / equipamentos (reparação lenta) | Danos severos | Repercussão nacional | T | M | M | NT | NT |
| | III | Média | Lesões graves intramuros ou lesões leves extramuros | Danos moderados a sistemas / equipamentos | Danos moderados | Repercussão regional | T | T | M | M | NT |
| | II | Marginal | Lesões leves | Danos leves a sistemas / equipamentos | Danos leves | Repercussão local | T | T | T | M | M |
| | I | Desprezível | Sem lesões ou no máximo casos de primeiros socorros | Danos leves a equipamentos sem comprometimento da continuidade operacional | Danos insignificantes | Repercussão insignificante | T | T | T | T | M |

Fonte: (PETROBRAS, 2015)

Embora métodos de inspeção, como pigs instrumentados (Figura 3a) sejam capazes de avaliar a integridade de dutos com alta confiabilidade, estes apresentam

limitações em tubulações com geometrias complexas, diâmetros reduzidos ou submetidas a condições adversas (DELGADILLO, 2020). Outro método de inspeção comumente realizado é o teste hidrostático. Contudo, além de ser considerado um ensaio destrutivo, pois pode propagar defeitos subcríticos na estrutura, traz impactos operacionais, como a necessidade de interrupção do fluxo, consumo elevado de água e a necessidade do descarte adequado do fluido (WANG, LAMBORN e CHEN, 2022). Para superar essas limitações, técnicas de avaliação direta (DA, do inglês *Direct Assessment*) têm sido empregadas, como ultrassom de ondas guiadas (Figura 3b) e MFL (Figura 3c) (ONUAHA, 2024).

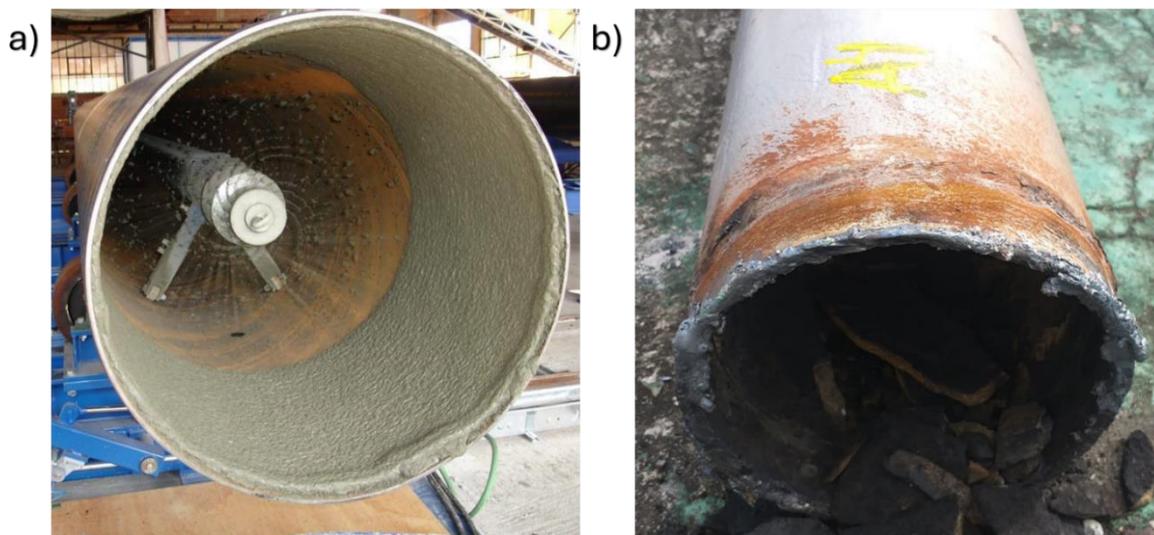
Figura 3. Técnicas de Inspeção: a) Pig instrumentado; b) ondas guiadas digital; c) MFL.



Fonte: Autoria Própria

Essas técnicas, embora amplamente utilizadas na inspeção, raramente têm como foco a avaliação da integridade do revestimento interno das tubulações, como os de argamassa, usados para proteção contra corrosão e danos mecânicos. A Figura 4 ilustra esse tipo de revestimento, desde sua aplicação (Figura 4a) até sua degradação ao longo do tempo (Figura 4b). Esse tipo de revestimento é geralmente aplicado em sistemas que transportam água potável, água salgada e espuma (PETROBRAS, 2020; BONDS, 2005). Em aplicações de maior responsabilidade, como o transporte de água oleosa, os desafios relacionados à degradação desses revestimentos tornam-se ainda mais críticos. A degradação do revestimento pode manifestar-se na forma de fissuras, descolamentos e exposição do aço ao meio, favorecendo a corrosão interna (SONG, 2016). Essas falhas aumentam os riscos de vazamentos, os quais podem resultar em impactos ambientais significativos, prejuízos financeiros consideráveis e penalidades regulatórias severas. Na Bahia, por exemplo, o Decreto nº 14.024/2012 prevê multas de até R\$ 50 milhões, dependendo da gravidade do dano ambiental (BAHIA, 2012).

Figura 4. Revestimento interno de argamassa em tubulação: (a) aplicação do revestimento; (b) tubulação substituída devido à degradação severa do revestimento interno.



Fonte: a) (SELMERS, 2025); b) Autoria Própria

Nesse cenário, a aplicação de técnicas de inspeção ultrassônica combinadas com algoritmos de aprendizado de máquina apresenta-se como uma solução promissora para a detecção precoce de defeitos em tubulações. O uso de ferramentas de triagem (*screening*), que identificam preliminarmente regiões potencialmente comprometidas e direcionam inspeções mais detalhadas para essas áreas, tem se mostrado relevantes para otimizar a avaliação da integridade de dutos, permitindo um monitoramento mais eficiente e reduzindo a dependência de inspeções extensivas (KRISHNAMURTHY, MILLIGAN e SLUDER, 2023). A combinação dessas tecnologias pode otimizar o tempo e os custos de inspeção, além de aumentar a eficiência do monitoramento da integridade. Com sua aplicabilidade em cenários industriais, a metodologia permitirá aumentar a confiabilidade operacional, reduzindo o escopo de investigação para regiões críticas, prevenindo impactos ambientais e falhas em dutos portuários.

Ensaio ultrassônicos são amplamente utilizados na detecção de descontinuidades em materiais metálicos e não metálicos, sendo potencializados pelo aprendizado de máquina na classificação de falhas. Yuan et al. (2024) e Lari et al. (2024) aplicaram a DFT na identificação de defeitos em soldas e materiais estruturais, enquanto Cruz et al. (2017) e Xu et al. (2023) demonstraram a eficácia da DCT na extração e compressão de características ultrassônicas. Além disso, Smagulova et al. (2024) utilizaram a LDA para aprimorar a separação de classes na identificação de

defeitos em adesivos estruturais. Apesar do avanço dessas metodologias, sua aplicação integrada na inspeção de revestimentos internos de argamassa em tubulações de aço carbono ainda é pouco explorada, evidenciando a relevância do presente estudo.

É sob estes aspectos que este trabalho visa contribuir ao desenvolver e validar uma metodologia integrada baseada em ensaios ultrassônicos e aprendizado de máquina para detecção de defeitos em revestimentos internos em argamassa de tubulações de aço carbono, contribuindo para a inspeção não destrutiva em aplicações industriais.

1.1 Objetivo geral

Investigar a aplicação de ensaios ultrassônicos e aprendizado de máquina na inspeção de revestimentos internos feitos de argamassa em tubulações de aço carbono, com foco na identificação e classificação de defeitos que possam comprometer a integridade estrutural e a segurança operacional destes sistemas.

1.2 Objetivos específicos

Projetar e preparar corpos de prova representativos, incluindo defeitos simulados em revestimentos de argamassa aplicados a tubulações de aço carbono.

Coletar e processar sinais ultrassônicos utilizando a técnica Pulso-Eco, avaliando parâmetros nos domínios do tempo e da frequência.

Aplicar transformadas espectrais, como DFT e DCT, para análise de sinais ultrassônicos e extração de características relevantes.

Implementar técnicas de estatísticas, como PCA e LDA, para otimizar a representação dos dados e facilitar a classificação dos defeitos.

Treinar modelos de aprendizado supervisionado, avaliando o classificador quanto à acurácia e capacidade de distinguir regiões íntegras e defeituosas.

Validar a metodologia proposta por meio da análise dos resultados obtidos, comparando diferentes cadeias de processamento para identificar a abordagem mais adequada com base em métricas de desempenho, como acurácia, sensibilidade e especificidade.

2 REVISÃO DA LITERATURA

2.1 Ultrassom

O ensaio por ultrassom é uma técnica de ensaio não destrutivo (END) reconhecida por sua versatilidade, sendo aplicável a uma ampla gama de materiais, tanto metálicos quanto não metálicos. Esse método permite a identificação de descontinuidades superficiais e internas, podendo ser utilizado acessando apenas de um lado do objeto inspecionado. Os sistemas ultrassônicos utilizam transdutores piezoelétricos e dispositivos de aquisição de dados, como pulsadores-receptores, para analisar sinais retornados (RAJ e AL, 2007). A Figura 5 apresenta um exemplo de equipamento de ultrassom muito utilizado na indústria para a medição de espessuras de equipamentos. Ele utiliza um transdutor que é acoplado à superfície do material, os resultados são exibidos numa tela digital por meio de gráficos no modo amplitude *A-scan*, permitindo uma avaliação rápida da integridade do material e da espessura do componente inspecionado.

Figura 5. Medidor de espessura por ultrassom.



Fonte: (OLYMPUS, 2025)

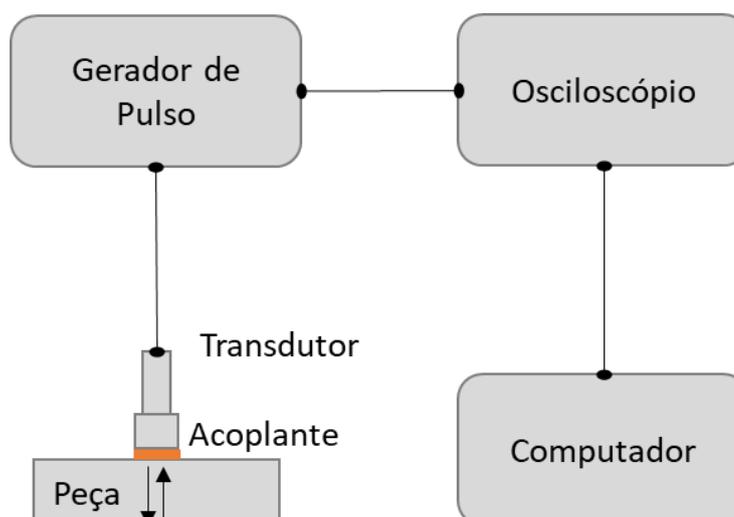
2.1.1 Técnica de Pulso-Eco

A técnica de pulso-eco é uma das metodologias mais amplamente empregadas na inspeção não destrutiva por ultrassom devido à sua versatilidade e capacidade de fornecer informações detalhadas sobre descontinuidades internas e a espessura de materiais. Esse método baseia-se na geração e recepção de pulsos ultrassônicos em um único transdutor ou em transdutores separados, operando em sincronia. A técnica utiliza pulsos ultrassônicos de alta frequência que são introduzidos no material por meio de um transdutor acoplado à superfície do objeto de teste. Um gel de acoplamento, geralmente composto por óleo ou glicerina, é utilizado para garantir a transmissão eficiente da energia sonora entre o transdutor e a superfície do material (RAJ e AL, 2007).

Quando a onda atravessa o material, esta interage com interfaces internas, como descontinuidades ou superfícies opostas. Parte da energia da onda é refletida de volta ao transdutor, enquanto a outra parte é transmitida ou refratada. A energia refletida é captada pelo transdutor, convertida em sinais elétricos e exibida em um osciloscópio ou dispositivo similar, permitindo a análise do tempo decorrido entre a emissão e a recepção (ROSE, 2014).

A Figura 6 ilustra o circuito básico da técnica de Pulso-Eco. O sistema pode ser integrado a um computador para processamento avançado.

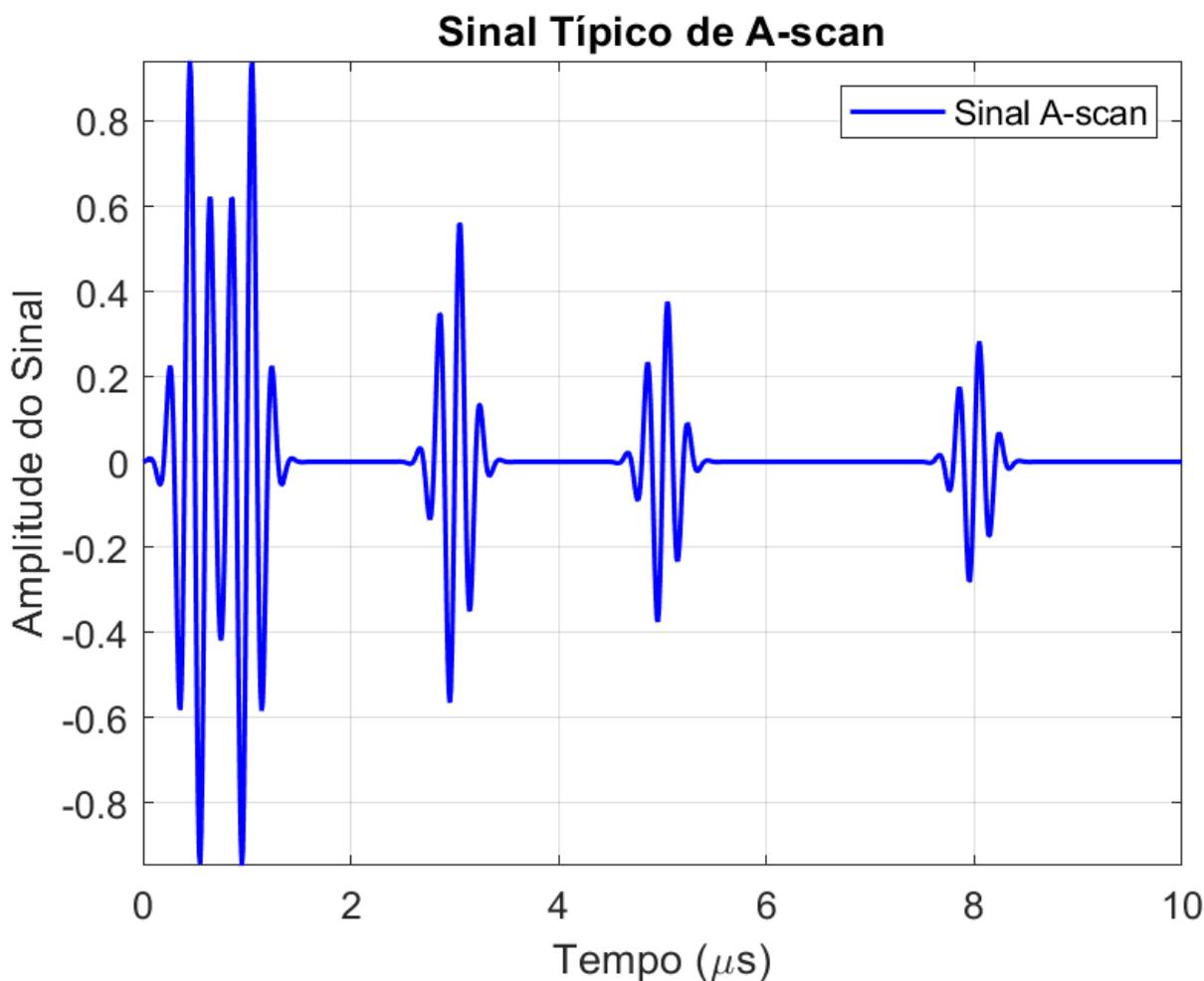
Figura 6. Circuito básico de ultrassom pulso-eco.



Fonte: Autoria própria

A visualização dos sinais em ensaios ultrassônicos é essencial para caracterizar discontinuidades e indicar suas dimensões. O *A-scan* é o modo de amplitude mais utilizado, apresentando os ecos em um gráfico onde o eixo X indica o tempo, relacionado à distância percorrida pela onda, e o eixo Y representa a amplitude do eco, correspondente à intensidade do sinal refletido. A Figura 7 mostra um sinal típico de *A-scan*. Essa exibição, geralmente realizada em osciloscópios ou em equipamentos portáteis como o da Figura 5, permite avaliar profundidade de discontinuidades e espessuras do material. Embora seja um método unidimensional, avanços como controle automatizado de sondas e análise computacional aprimoram sua aplicação em cenários industriais. Outros modos, como o *B-scan* e o *C-scan*, são utilizados para análises específicas, mas o *A-scan* permanece amplamente empregado pela sua simplicidade e eficiência (RAJ e AL, 2007).

Figura 7. Sinal típico de *A-scan* com o pulso inicial e os ecos refletidos da superfície oposta.



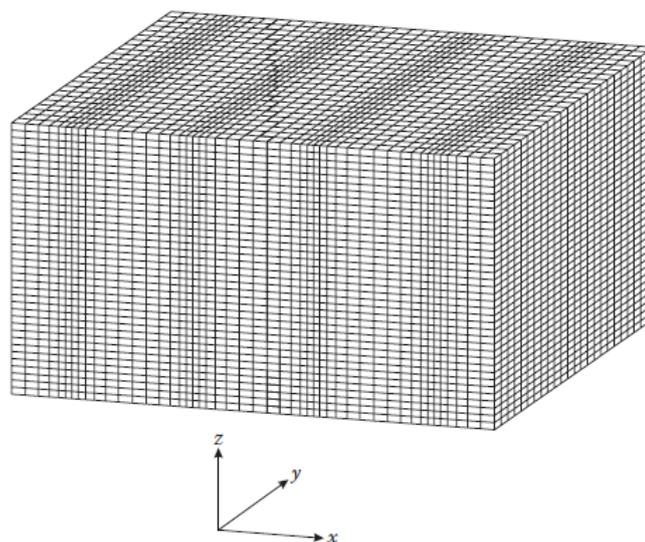
Fonte: Autoria própria

2.1.2 Propriedades das Ondas Ultrassônicas

Ondas sonoras, sejam elas audíveis ou ultrassônicas, são vibrações mecânicas que se propagam por meio do movimento das partículas do material em que estão inseridas. As ondas sonoras podem ser transmitidas em qualquer material que tenha comportamento elástico e são classificadas segundo o modo como as partículas do meio vibram em relação à direção de propagação, sendo divididas basicamente em ondas longitudinais (compressão ou "P") e transversais (cisalhamento ou "S"). Existindo também outras formas, como as ondas de superfície (*Rayleigh*) e guiadas (*Lamb*) (RAJ e AL, 2007).

As ondas longitudinais ou ondas de compressão, distinguem-se por sua propagação no mesmo sentido do movimento das partículas no meio, criando zonas alternadas de compressão e rarefação, conforme ilustrado na Figura 8:

Figura 8. Deformações causadas por ondas planares volumétricas ao longo do eixo x.



Fonte: (CHEEKE, 2012)

Essa característica, impulsionada por forças compressivas e expansivas, torna-as eficazes em sólidos devido à alta velocidade proporcionada pela rigidez compressiva do material. (CHEEKE, 2012). Sua alta penetração e precisão tornam as ondas longitudinais ferramentas amplamente empregadas em ensaios não destrutivos, sendo o tipo de onda usada neste trabalho.

2.1.2.1 Velocidade de Propagação

A velocidade de propagação de ondas longitudinais é influenciada pela densidade do material e pelas suas propriedades elásticas. A relação fundamental da velocidade longitudinal (V_L) com as propriedades elásticas do material é dada pela Equação 2.1 (BUNGEY e GRANTHAM, 2006):

$$V_L = \sqrt{\frac{E_d(1 - \nu)}{\rho(1 + \nu)(1 - 2\nu)}} \quad (2.1)$$

onde,

- E_d é o módulo dinâmico de elasticidade;
- ν o coeficiente de Poisson;
- ρ densidade do material.

A análise da velocidade de propagação das ondas ultrassônicas é amplamente utilizada para a avaliação não destrutiva de materiais, permitindo investigar sua integridade física e suas propriedades. Hofmann (2015) mostrou que alterações na velocidade das ondas estão relacionadas à presença de fissuras no concreto, permitindo monitorar danos progressivos durante ensaios de flexão. Freitas (2009) utilizou as velocidades longitudinal e transversal para correlacionar as propriedades elásticas e microestruturais de aços AISI e duplex, evidenciando a influência de fases como ferrita e martensita. Carvalho Jr et al. (2024) aplicaram a análise da velocidade para caracterizar hastes de aterramento, destacando sua sensibilidade a mudanças nas propriedades após ataque químico. Esses exemplos ressaltam a importância da velocidade ultrassônica na inspeção e caracterização de materiais.

2.1.2.2 Impedância Acústica

A impedância acústica é definida como a resistência que um meio oferece à passagem de uma onda sonora. Representada por (Z) e expressa em unidades de rayl (kg/m^2), a impedância é calculada pelo produto da densidade do material (ρ) pela velocidade de propagação da onda (V_L) no meio, conforme a Equação 2.2:

$$Z = \rho \cdot V_L \quad (2.2)$$

Quando há uma grande diferença de impedância acústica entre materiais (Tabela 1), como entre um transdutor ultrassônico e o objeto a ser inspecionado, ocorre reflexão significativa, reduzindo a energia transmitida. Para mitigar esse efeito, líquidos acoplantes com impedâncias intermediárias são frequentemente utilizados em ENDs para minimizar as perdas por reflexão (SCHMERR, 1998).

Tabela 1. Propriedades acústicas dos materiais.

| Material | Peso Específico (kg/m ³) | Velocidade Transversal (m/s) | Velocidade Longitudinal (m/s) | Impedância Acústica (10 ⁵ kg/m ² s) |
|-------------|--------------------------------------|------------------------------|-------------------------------|---|
| Ar | 1 | - | 330 | 0,04 |
| Água (20°C) | 1.000 | - | 1.480 | 1,48 |
| Aço Carbono | 7.850 | 3.250 | 5.920 | 46,47 |
| Concreto | 2.000 | - | 4.600 | 9,20 |

Fonte: (SANTIN, 2003)

2.1.2.3 Energia do Sinal

No domínio do tempo, a energia de um sinal E pode ser definida como a integral do quadrado do valor absoluto do sinal, garantindo que todas as contribuições, positivas ou negativas, sejam contabilizadas (LATHI, 2008).

Para cálculos específicos, uma adaptação utilizada envolve determinar a energia pela área sob o sinal retificado. Essa abordagem, descrita por Carelli (2014) pode ser expressa pela Equação 2.3:

$$E = \int_0^{t_i} |A(t)| dt \quad (2.3)$$

onde,

- $|A(t)|$ = módulo da amplitude do sinal no tempo t ;
- t_i = fim da janela de tempo fixada nos experimentos.

O valor absoluto da energia não possui significância direta devido à influência do acoplamento entre os transdutores e a amostra. Dessa forma, é comum realizar uma análise comparativa desse parâmetro entre uma amostra ou estrutura íntegra e outra com danos (CARELLI, 2014).

2.1.2.4 Parâmetros Estatísticos no Domínio do Tempo

Estudos recentes, como o de Smagulova et al. (2024), indicam que o uso de múltiplos parâmetros extraídos do domínio do tempo têm demonstrado eficácia na detecção de defeitos, indo além da amplitude do sinal e incorporando métricas como valor médio, variância, desvio padrão, valor RMS (*root mean square*) e curtose. Esses parâmetros fornecem uma descrição estatística abrangente dos sinais capturados, permitindo a distinção entre regiões íntegras e áreas comprometidas.

A seguir, apresentam-se as fórmulas matemáticas de cada parâmetro, acompanhadas de uma análise de sua aplicação prática em ensaios ultrassônicos (MATHWORKS, 2024):

Média (μ): A média representa a tendência geral do sinal em uma região analisada. Em materiais homogêneos, os valores se mantêm consistentes, apresentada pela Equação 2.4:

$$\mu = \frac{1}{N} \sum_{i=1}^N x_i \quad (2.4)$$

onde,

- x_i é o valor do sinal na i -ésima amostra;
- N é o número de amostras.

Variância (σ^2): A variância mede a dispersão dos valores do sinal em relação à média, apresentada pela Equação 2.5 para um conjunto de dados amostral:

$$\sigma^2 = \frac{1}{N-1} \sum_{i=1}^N (x_i - \mu)^2 \quad (2.5)$$

onde,

- x_i é o valor do sinal na i -ésima amostra;
- μ é a média;
- N é o número de amostras.

Desvio Padrão (σ): O desvio padrão é diretamente relacionado à dispersão dos valores individuais do sinal em relação à média, apresentado pela Equação 2.6:

$$\sigma = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (x_i - \mu)^2} \quad (2.6)$$

onde,

- x_i é o valor do sinal na i -ésima amostra;
- μ é a média;
- N é o número de amostras.

Valor RMS (x_{RMS}): O valor RMS fornece a energia total associada ao sinal propagado, sendo uma métrica utilizada para avaliar mudanças na integridade em materiais sujeitos à inspeção ultrassônica, apresentado pela Equação 2.7:

$$x_{RMS} = \sqrt{\frac{1}{N} \sum_{i=1}^N x_i^2} \quad (2.7)$$

onde,

- x_i é o valor do sinal na i -ésima amostra;
- N é o número de amostras.

Curtose (K): A curtose avalia a proeminência de picos no sinal, podendo indicar falhas pontuais ou inclusões estruturais, apresentada pela Equação 2.8:

$$K = \frac{\frac{1}{N} \sum_{i=1}^N (x_i - \mu)^4}{\sigma^4} \quad (2.8)$$

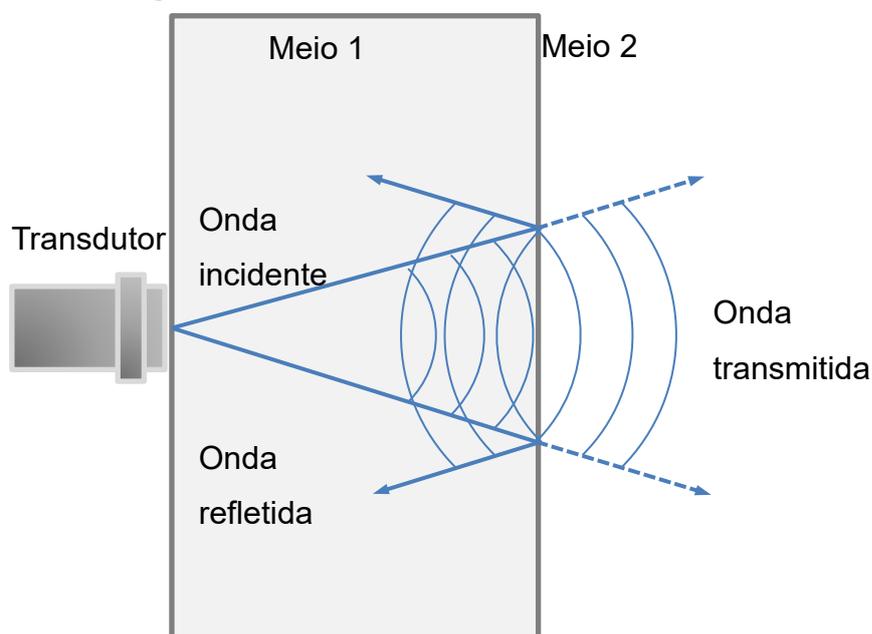
onde,

- x_i é o valor do sinal na i -ésima amostra;
- μ é a média;
- σ é o desvio padrão;
- N é o número de amostras.

2.1.3 Reflexão e Transmissão

Quando ondas ultrassônicas incidem em uma interface entre dois materiais com propriedades acústicas distintas, ocorre a divisão da energia em componentes refletidas e transmitidas (Figura 9). Esse fenômeno é consequência direta das diferenças entre as impedâncias acústicas dos materiais.

Figura 9. Reflexão e transmissão da onda sônica.



Fonte: Adaptado de Santin (2003)

No caso de incidência normal, os coeficientes de reflexão (R) e transmissão (T) são dados pelas equações 2.9 e 2.10: (RAJ e AL, 2007):

$$R = \left(\frac{Z_2 - Z_1}{Z_2 + Z_1} \right)^2 \quad (2.9)$$

$$T = \frac{4 \cdot Z_2 \cdot Z_1}{(Z_2 + Z_1)^2} \quad (2.10)$$

onde,

- R é o coeficiente de reflexão, fração da energia refletida,
- T é coeficiente de transmissão, fração da energia transmitida,
- Z_2 e Z_1 são as impedâncias acústicas dos meios 1 e 2, respectivamente.

2.1.4 Atenuação

A atenuação das ondas sonoras em materiais ocorre por dois processos principais: espalhamento e absorção. O espalhamento é causado por heterogeneidades no material, como poros ou inclusões, que desviam as ondas sonoras, especialmente em materiais não isotrópicos ou com grãos grosseiros. Para grãos com dimensões entre 1/1000 e 1/10 do comprimento de onda, o espalhamento é insignificante, mas torna-se significativo à medida que o tamanho dos grãos aumenta, dificultando ou impossibilitando testes, principalmente em materiais anisotrópicos. A absorção, por sua vez, converte a energia sonora em calor, sendo influenciada por fenômenos como fricção interna, histerese elástica e condução térmica, aumentando com a frequência. Essas perdas limitam os testes ultrassônicos. Enquanto a absorção enfraquece a energia transmitida, o espalhamento é mais crítico no método de eco, pois reduz os sinais úteis e gera ruído, dificultando a identificação de ecos reais. Esse ruído não pode ser compensado aumentando a energia inicial, pois também amplifica o efeito. A única solução prática é reduzir a frequência, embora isso limite a capacidade de detectar pequenas falhas (RAJ e AL, 2007).

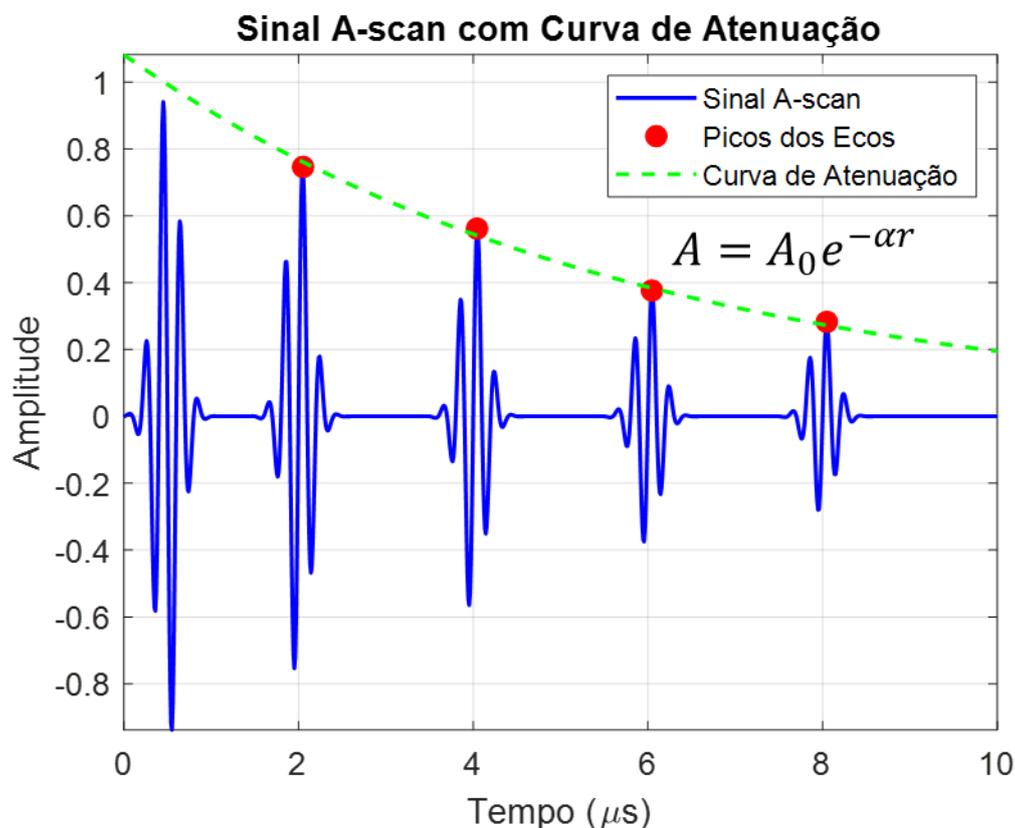
Sua descrição matemática é dada pela Equação 2.11 e seu comportamento é mostrado na Figura 10 que relaciona a intensidade da onda com a distância percorrida e o coeficiente de atenuação (SCHMERR, 1998):

$$A = A_0 e^{-\alpha r} \quad (2.11)$$

onde,

- A_0 é a amplitude incidente;
- A é a amplitude após percorrer uma distância r ;
- α é o coeficiente de atenuação;
- r é a distância percorrida.

Figura 10. Sinal típico de A-scan com curva de atenuação ajustada.



Fonte: Autoria própria

Emery (2024) analisou o impacto da porosidade e das interações nas interfaces materiais sobre a atenuação de ondas ultrassônicas em amostras de concreto. Utilizando altas frequências, o autor demonstrou que a atenuação aumenta em regiões com maior presença de porosidade e interfaces heterogêneas. Essa abordagem permitiu uma análise quantitativa da distribuição de poros em diferentes misturas de concreto, destacando a relevância desse parâmetro para caracterizar microestruturas críticas. Souza e Pinto (2020) investigaram a eficiência do reparo de fissuras em concreto com resinas epóxi, utilizando parâmetros relacionados à energia das ondas ultrassônicas influenciados pela atenuação. Os resultados mostraram que fissuras parcialmente preenchidas apresentaram redução significativa na energia do sinal, devido à dispersão e absorção nas interfaces descontínuas. Hofmann (2015) demonstrou que a atenuação é um parâmetro eficaz para monitorar fissuração em concreto, correlacionando-a ao comportamento mecânico durante ensaios de flexão. O aumento da atenuação foi associado ao desenvolvimento de micro e microfissuras, indicando a degradação progressiva da estrutura.

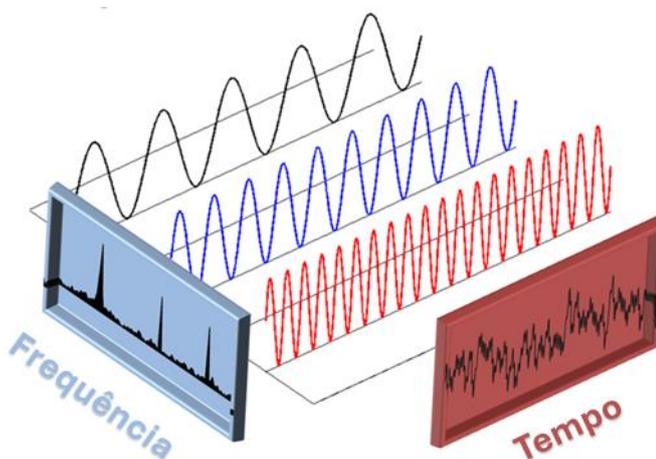
2.2 Processamento de Sinais

O processamento de sinais é um campo multidisciplinar que abrange técnicas matemáticas, algoritmos computacionais e implementações práticas voltadas para a análise, transformação e extração de informações de sinais. Seja no domínio do tempo ou da frequência, o objetivo principal do processamento de sinais é melhorar a interpretação dos dados e habilitar o desenvolvimento de sistemas avançados de medição, comunicação e controle. Esta parte do capítulo explora as bases teóricas e aplicações práticas de ferramentas transformadoras, destacando seu papel em aplicações como inspeção ultrassônica, compressão de dados e análise espectral.

2.2.1 Transformada Discreta de Fourier (DFT)

A Transformada Discreta de Fourier (DFT, do inglês *Discrete Fourier Transform*) é uma ferramenta matemática no processamento de sinais, com a finalidade de converter sinais do domínio do tempo para o domínio da frequência (Figura 11). Diferentemente da Transformada de Fourier contínua, que analisa sinais contínuos e infinitos, a DFT é projetada para processar sinais discretos e de comprimento finito, sendo especialmente relevante para aplicações computacionais. Seu uso permite a decomposição de sinais discretos em componentes sinusoidais básicas, caracterizadas por frequências, amplitudes e fases distintas, proporcionando uma análise detalhada das propriedades frequenciais do sinal (SMITH, 1999).

Figura 11. Transformada de sinal do domínio do tempo para o domínio da frequência via DFT.



Fonte: Adaptado de Kalhara (2017)

A DFT é particularmente útil na análise de sinais digitais. Sua capacidade de representar sinais como combinações lineares de funções cossenoidais e senoidais, denominadas funções base, torna-a relevante em tarefas como filtragem de ruídos, compressão de dados e diagnóstico de sistemas. Essas funções base, descritas matematicamente pelas equações 2.12 e 2.13, visam capturar as oscilações harmônicas que compõem o sinal original, permitindo uma visão abrangente das suas características espectrais (SMITH, 1999).

Função Cossenoidal – Equação 2.12:

$$c_k[i] = \cos\left(\frac{2\pi ki}{N}\right) \quad (2.12)$$

Função Senoidal – Equação 2.13 :

$$s_k[i] = \sin\left(\frac{2\pi ki}{N}\right) \quad (2.13)$$

onde:

- $c_k[i]$ é a componente cossenoidal da base.;
- $s_k[i]$ é a componente senoidal da base;
- N é o número total de amostras no sinal;
- i é o índice da amostra no domínio do tempo;
- k é o índice da componente harmônica, que corresponde à $f_k = \frac{k}{N} f_s$, sendo a frequência real associada ao índice k e a frequência de amostragem do sinal, respectivamente.

Essas funções base representam as oscilações harmônicas que constituem os blocos construtivos do sinal. A DFT calcula como cada uma dessas harmônicas contribui para a composição do sinal original. A fórmula completa da DFT é dada pela Equação 2.14:

$$X[k] = \sum_{n=0}^{N-1} x[n] \cdot e^{-i \cdot 2\pi kn/N} \quad (2.14)$$

onde:

- $X[k]$ é o coeficiente no domínio da frequência associado ao índice da componente harmônica k ;
- $x[n]$ é o sinal discreto no domínio do tempo;
- n é o índice de amostra no domínio do tempo;
- N é o número total de amostras no sinal;
- $e^{-i \cdot 2\pi kn/N}$ é o termo exponencial complexo.

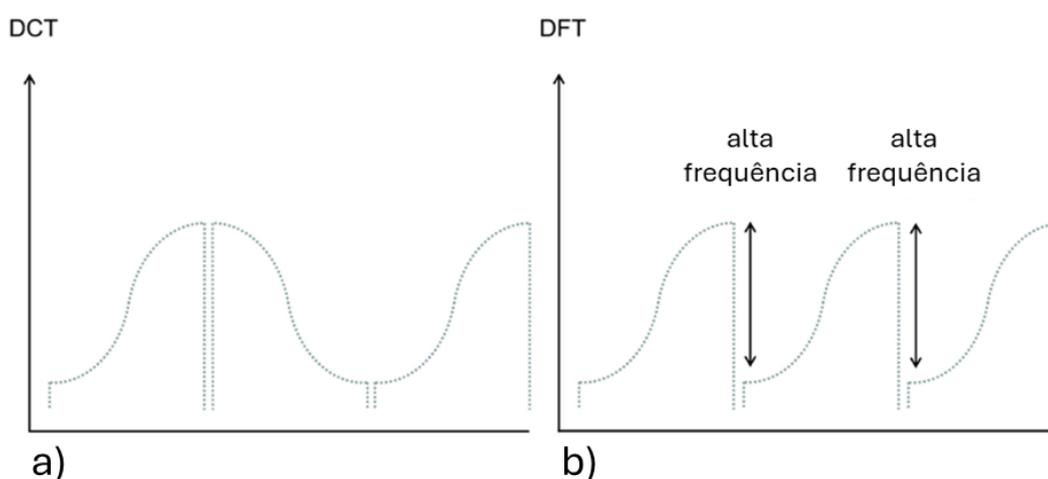
A DFT tem sido amplamente empregada em aplicações de ultrassom para análise de sinais e detecção de falhas estruturais. Alguns exemplos destacados em estudos recentes incluem a avaliação de defeitos em juntas soldadas no trabalho apresentado por Yuan et al. (2024), que utilizaram a DFT para analisar sinais de ondas ultrassônicas orientadas, possibilitando a detecção de defeitos em juntas soldadas de estruturas tubulares, melhorando a relação sinal-ruído e a capacidade de identificar imperfeições em regiões críticas de solda. Lari et al. (2024), no trabalho intitulado Caracterização de Geometrias de Falhas, demonstraram a aplicação da DFT para extrair características de amplitude e frequência de sinais ultrassônicos em testes não destrutivos. Essas informações foram usadas para treinar redes neurais artificiais na caracterização geométrica de falhas em materiais, destacando sua utilidade na identificação de defeitos internos com alta precisão. Outro estudo por Yuan et al. (2024), intitulado Avaliação de Corrosão em Dobras L, aplicou a DFT em conjunto com simulações de elementos finitos para identificar e quantificar afinamentos na parede de estruturas em forma de L causados por corrosão. Este método comprovou ser eficaz para correlacionar profundidade de defeitos com coeficientes de reflexão obtidos através da transformada.

2.2.2 Transformada Discreta do Cosseno (DCT)

A Transformada Discreta do Cosseno (DCT, do inglês *Discrete Cosine Transform*) é uma ferramenta matemática utilizada no processamento de sinais para converter sinais do domínio do tempo para o domínio da frequência. Diferentemente da DFT, a DCT opera exclusivamente com funções reais, o que a torna especialmente eficaz em aplicações voltadas para compressão de dados, como no processamento de imagens e vídeos (KHAYAM, 2003).

A Figura 12 compara a DCT e a DFT. A DCT (Figura 12a) converte o sinal original em um sinal par real, evitando a introdução de componentes de alta frequência artificiais nas bordas, concentrando a energia em baixas frequências. Já a DFT (Figura 12b) não aplica simetria, resultando em maior presença de componentes de alta frequência. Isso torna a DCT ideal para compressão de dados, enquanto a DFT é mais adequada para análise espectral detalhada.

Figura 12. Diferença entre a DCT e DFT.



Fonte: Adaptado de Peng (2024)

A DCT possui quatro variantes principais — DCT-I, DCT-II, DCT-III e DCT-IV —, cada uma projetada para atender a necessidades específicas no processamento de sinais (MATHWORKS, 2024). A DCT-II é bastante empregada em algoritmos de compressão de dados, como JPEG e MPEG, por sua capacidade de concentrar a maior parte da energia do sinal nos primeiros coeficientes. Essa característica reduz redundâncias, otimizando a compactação de dados (KHAYAM, 2003).

A DCT-III é utilizada como a transformada inversa da DCT-II, possibilitando a reconstrução do sinal original. As variantes DCT-I e DCT-IV, apesar de menos frequentes, têm aplicações específicas em áreas como análise matemática e solução de problemas especializados. No contexto deste trabalho, será adotado o MATLAB® como ferramenta de desenvolvimento e análise, a DCT será aplicada por meio da função integrada “*dct*”, que implementa a DCT-II, que é matematicamente apresentada na Equação 2.15 (MATHWORKS, 2024).

$$y(k) = \sqrt{\frac{2}{N}} \sum_{n=1}^N x(n) \frac{1}{\sqrt{1 + \delta_{k1}}} \cos \left(\frac{\pi}{2N} (2n - 1)(k - 1) \right) \quad (2.15)$$

onde:

- n é o índice das amostras do sinal no domínio do tempo;
- $x(n)$ é a amplitude da n -ésima amostra do sinal no domínio do tempo;
- N é Tamanho do vetor de entrada $x(n)$, que indica o n° total de amostras;
- k é o índice do coeficiente de frequência, variando de 1 a N ;
- δ_{k1} é o Delta de *Kronecker*, que vale 1 se $k=1$ e 0 caso contrário;
- $\cos \left(\frac{\pi}{2N} (2n - 1)(k - 1) \right)$ é a função cosseno responsável por mapear o sinal para o domínio da frequência.

A DCT tem sido aplicada em várias áreas de ultrassom para análise de sinais e detecção de falhas estruturais. Alguns exemplos destacados de trabalhos incluem o trabalho sobre detecção de falhas em estruturas de aço soldado de Cruz et al. (2017), que utilizaram a DCT para extrair características relevantes de sinais ultrassônicos na avaliação de juntas soldadas. Essa técnica foi combinada com redes neurais artificiais para classificar defeitos em solda. Outro trabalho relevante foi sobre a compressão de dados em inspeção de soldas em bicos de Xu et al. (2023), que exploraram o uso da DCT no contexto da compressão sensorial de dados de array ultrassônico capturados em soldas de bicos. O trabalho demonstrou que a DCT pode ser usada como base de transformação esparsa, mostrando precisão na reconstrução dos sinais ultrassônicos comprimidos e contribuindo para a redução do espaço de armazenamento e tempo de processamento em aplicações industriais.

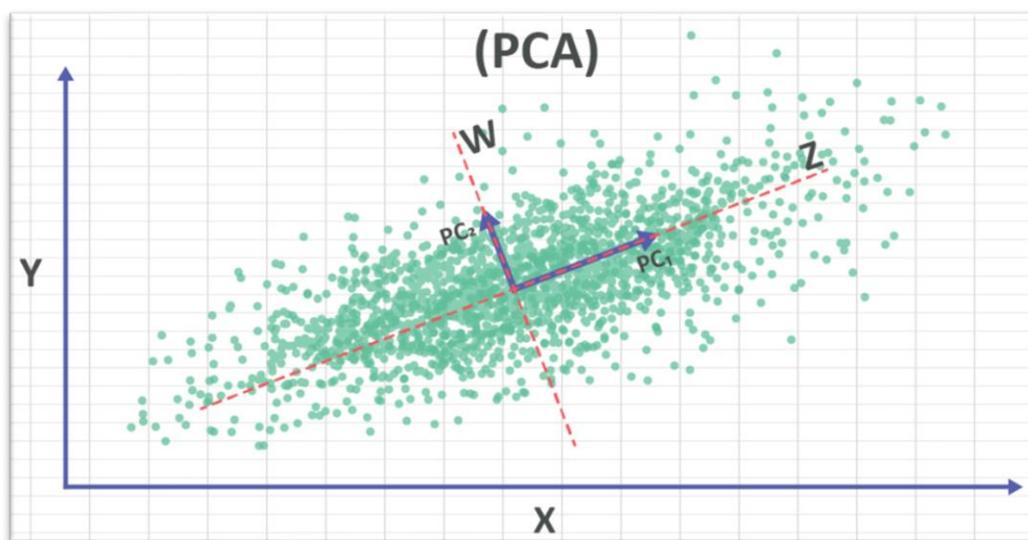
2.2.3 Análise de Componentes Principais (PCA)

A Análise de Componentes Principais (PCA, do inglês *Principal Component Analysis*) é uma técnica estatística utilizada para reduzir a dimensionalidade de um conjunto de dados composto por múltiplas variáveis inter-relacionadas. Seu objetivo é preservar o máximo possível da variabilidade dos dados enquanto os transforma em um novo sistema de coordenadas. Nesse processo, são geradas novas variáveis, chamadas componentes principais, que são mutuamente ortogonais e ordenadas de

modo que as primeiras componentes retêm a maior parte da informação presente nos dados originais (JOLLIFFE, 2002).

Na Figura 13, os eixos originais dos dados são rotacionados para alinhar-se às novas direções principais, evidenciando as dimensões mais relevantes e reduzindo redundâncias (OMARZAI, 2025).

Figura 13. Análise de Componentes Principais.



Fonte: (OMARZAI, 2025)

O processo de desenvolvimento da PCA envolve centralizar os dados, calcular a matriz de covariância ou correlação e determinar autovalores e autovetores. Os autovalores representam a variância explicada, e os autovetores fornecem as direções dos componentes, ordenados de forma decrescente para priorizar os mais relevantes e envolve as seguintes etapas (JOLLIFFE, 2002):

Os dados são organizados em uma matriz \mathbf{X} de dimensão $M \times N$, essa matriz é expressa como:

$$\mathbf{X} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1N} \\ x_{21} & x_{22} & \cdots & x_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ x_{M1} & x_{M2} & \cdots & x_{MN} \end{bmatrix}$$

onde:

- M representa o número de observações;
- N o número de variáveis.

A PCA é modelada matematicamente como a solução de um problema de autovalores e autovetores, conforme apresentado na Equação 2.16:

$$\Sigma \mathbf{a}_k = \lambda_k \mathbf{a}_k \quad (2.16)$$

onde:

- Σ matriz de covariância dos dados;
- \mathbf{a}_k autovetor associado ao k-ésimo componente principal;
- λ_k autovalor correspondente, representando a variância explicada pelo k-ésimo componente.

Os dados são centralizados para garantir que cada variável tenha média zero, conforme descrito pela Equação 2.17:

$$\mathbf{X}_c = \mathbf{X} - \boldsymbol{\mu} \quad (2.17)$$

onde:

- \mathbf{X}_c é a matriz de dados centralizados;
- \mathbf{X} é a matriz de dados originais;
- $\boldsymbol{\mu}$ é o vetor de médias de cada variável (de dimensão).

Para reduzir a dimensionalidade, calcula-se a matriz de covariância Σ a partir dos dados centralizados \mathbf{X}_c , conforme Equação 2.18:

$$\Sigma = \frac{1}{M-1} \mathbf{X}_c^T \mathbf{X}_c \quad (2.18)$$

onde:

- Σ : matriz de covariância dos dados;
- \mathbf{X}_c : matriz de dados centralizados;
- M : número de amostras.

Após encontrar os autovalores λ_k e seus autovetores correspondentes \mathbf{a}_k , os dados são projetados no espaço definido pelos m primeiros autovetores, formando os componentes principais mais relevantes, segundo a Equação 2.19:

$$\mathbf{Z} = \mathbf{X}_c \mathbf{A} \quad (2.19)$$

onde:

- \mathbf{Z} é matriz projetada no espaço dos componentes principais ($M \times m$);
- \mathbf{A} é a matriz formada pelos m autovetores correspondentes aos maiores autovalores ($N \times m$).

O número m de componentes selecionados é determinado com base na variância explicada acumulada, que é calculada conforme Equação 2.20:

$$\text{Variância acumulada} = \frac{\sum_{k=1}^m \lambda_k}{\sum_{k=1}^N \lambda_k} \quad (2.20)$$

onde:

- λ_k é o k -ésimo autovalor, representando a variância explicada pelo respectivo componente principal
- m é o número de componentes principais selecionados;
- N é o número total de componentes principais.

A PCA tem encontrado aplicações diversas, especialmente em áreas que lidam com grandes volumes de dados e necessidade de otimização de processamento, como análise de sinais ultrassônicos e diagnóstico de falhas industriais.

Huang et al. (2023) empregaram a PCA na detecção de delaminações em compósitos de fibra de carbono reforçada. Neste estudo, a técnica foi utilizada para melhorar a relação sinal-ruído em dados térmicos, permitindo uma detecção mais precisa de delaminações, mesmo em camadas profundas dos materiais, evidenciando sua eficácia na extração de informações relevantes em cenários complexos.

Cruz et al. (2017) também destacaram a importância da PCA ao aplicá-la na seleção de características para redes neurais na detecção de falhas em juntas soldadas de aço. A PCA foi empregada para reduzir um conjunto de 100 características para apenas 20, mantendo alta precisão na classificação de falhas como porosidade e inclusões, demonstrando sua utilidade na otimização de modelos de aprendizado de máquina.

De forma complementar, Moshrefi e Nabki (2024) aplicaram a PCA na simplificação de conjuntos de dados de alta dimensionalidade em sistemas industriais, combinando-a com classificadores de conjunto para detecção de falhas em tempo real. Sun et al. (2023), em uma revisão sobre aprendizado de máquina aplicado à inspeção não destrutiva de soldas, ressaltaram o uso da PCA para seleção de

características, evidenciando sua relevância na classificação de falhas em soldas e zonas afetadas pelo calor.

Esses trabalhos demonstram a versatilidade e relevância da PCA em diferentes aplicações, destacando sua contribuição para a simplificação de dados, otimização de modelos e melhoria para análise de dados.

2.2.4 Análise Discriminante Linear (LDA)

A Análise Discriminante Linear (LDA, do inglês *Linear Discriminant Analysis*) é uma técnica estatística amplamente utilizada em tarefas de classificação e redução de dimensionalidade. Desenvolvida inicialmente por Fisher (1936), o método busca encontrar uma transformação linear que maximize a separação entre as médias das classes, enquanto minimiza a variabilidade dentro de cada classe (LI, 2023). A LDA é baseada na suposição de que os dados de cada classe seguem uma distribuição normal multivariada, com matrizes de covariância homogêneas entre as classes (GRAF, ZELDOVICH e FRIEDRICH, 2024).

Uma das principais aplicações da LDA está na projeção dos dados originais em um espaço de menor dimensão, facilitando a visualização e interpretação dos padrões de separação entre as classes. Essa projeção é realizada otimizando o critério de Fisher, que maximiza a razão entre a variância entre classes e a variância dentro das classes, conforme descrito por LI et al. (2023):

$$W = \operatorname{argmax}_W \frac{W^T S_b W}{W^T S_w W} \quad (2.21)$$

onde:

- W é o vetor ou matriz de transformação linear que maximiza a separabilidade entre as classes;
- S_b é a matriz de dispersão entre as classes, calculada como:

$$S_b = \sum_{c=1}^c N_c (\mu_c - \mu)(\mu_c - \mu)^T \quad (2.22)$$

onde μ_c é a média dos dados da classe c , μ é a média global e N_c é o número de amostras na classe c ;

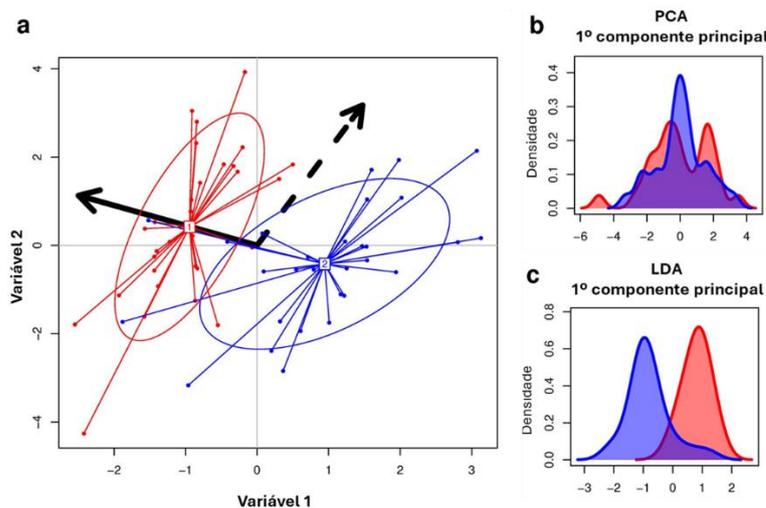
S_w é a matriz de dispersão dentro das classes, definida por:

$$S_w = \sum_{c=1}^C \sum_{x \in \mathcal{D}_c} (x - \mu_c)(x - \mu_c)^T \quad (2.23)$$

onde x representa os dados individuais da classe C , \mathcal{D}_c é o conjunto de amostras da classe C , e μ_c é a média dos dados dessa classe; C é o número total de classes.

A Figura 14 ilustra a diferença fundamental entre a PCA e a LDA, destacando como cada método opera em um espaço de características com duas variáveis. Na Figura 14a, os indivíduos (pontos) e os grupos (representados por cores e elipses) estão dispostos em um plano bidimensional. O PCA busca identificar a direção que maximiza a variância total dos dados (seta tracejada), sem levar em conta as diferenças entre os grupos. Em contrapartida, a LDA identifica a direção que maximiza a separação entre os grupos (seta contínua), enquanto minimiza a variabilidade dentro de cada grupo. Como resultado, na Figura 14b, a PCA falha em discriminar adequadamente os grupos, enquanto na Figura 14c, a LDA exibe de forma clara as diferenças entre eles. Essa característica faz da LDA uma abordagem mais apropriada para problemas de classificação, pois otimiza a separabilidade entre as classes enquanto reduz a variância intra-classe (JOMBART, DEVILLARD e BALLOUX, 2010).

Figura 14. Diferença entre PCA e LDA.



Fonte: Adaptado de Jombart, Devillard e Balloux (2010)

No entanto, a aplicação da LDA é limitada por alguns fatores. Entre eles, destaca-se o problema da matriz de covariância singular, que ocorre quando o número de variáveis supera o número de amostras, e pela sensibilidade do método a ruídos e dados não gaussianos (LI, 2023).

Para contornar essas limitações, diversas extensões da LDA foram desenvolvidas, incluindo métodos que integram a LDA com outras técnicas, como a própria técnica de PCA, para reduzir a dimensionalidade antes da aplicação do discriminante (JOMBART, DEVILLARD e BALLOUX, 2010).

No contexto do trabalho de Jombart, Devillard e Balloux (2010), a integração do PCA antes do LDA é uma abordagem sustentada pela necessidade de mitigar problemas inerentes ao uso isolado da LDA, como a singularidade da matriz de covariância e a sensibilidade a dados não gaussianos. A PCA reduz dimensionalidade ao preservar as componentes com maior variância global, proporcionando uma estrutura mais compacta e menos ruidosa para a aplicação subsequente da LDA. Esse processo atende ao princípio discutido no estudo, de que o LDA é mais eficiente quando operando em um espaço de características previamente reduzido e bem definido.

Smagulova et al. (2024) utilizaram a LDA como parte de um sistema de detecção automática de defeitos em juntas adesivas, analisando dados ultrassônicos. O método foi empregado para reduzir a dimensionalidade dos dados e maximizar a separação entre classes de defeitos, facilitando a classificação. A integração da LDA com técnicas de extração de características relevantes do sinal ultrassônico garantiu maior precisão e eficiência no processo.

2.3 Aprendizado de Máquina

O aprendizado de máquina (ML, do inglês *Machine Learning*) é uma subárea da inteligência artificial que busca desenvolver sistemas capazes de aprender e tomar decisões com base em dados, sem a necessidade de programação explícita para cada tarefa. Esse campo representa uma transformação significativa nos paradigmas computacionais tradicionais, permitindo que os sistemas processem informações complexas e identifiquem padrões de maneira autônoma (JOHNSON, 2024).

Diversos estudos destacam a integração de aprendizado de máquina com ensaios ultrassônicos em aplicações industriais. Xiao et al. (2020) propuseram um método híbrido combinando o modelo ultrassônico aprimorado e *Support Vector Machine* (SVM) para identificar defeitos em materiais metálicos, com resultados que demonstram uma melhora significativa na precisão da classificação de inclusões e cavidades, utilizando parâmetros como amplitude e frequência de eco.

Outro exemplo relevante é o trabalho de Sun et al. (2023) revisaram sistematicamente o uso de aprendizado de máquina na análise de dados ultrassônicos para inspeção de soldas, explorando métodos como SVM, com aplicações na indústria nuclear.

No contexto de adesivos estruturais, Smagulova et al. (2024) aplicaram aprendizado de máquina para detecção automática de defeitos em juntas adesivas com base em dados ultrassônicos, utilizando SVM para classificação e predição de profundidade de descontinuidades com alta precisão, destacando a eficácia de técnicas de redução de dimensionalidade na melhora do desempenho dos modelos.

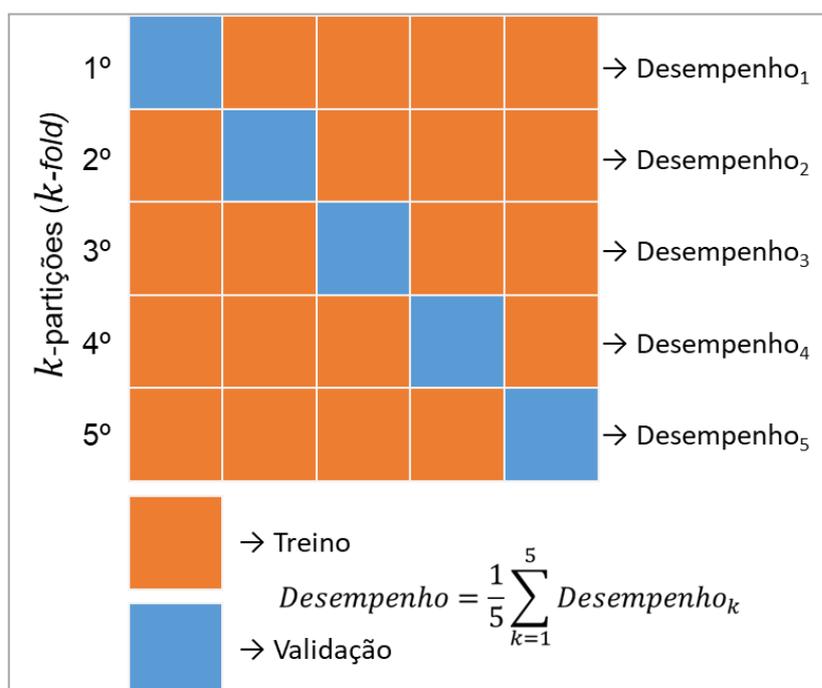
2.3.1 Fundamentos de Aprendizado de Máquina

Os métodos de aprendizado de máquina podem ser categorizados em diferentes abordagens de acordo com a disponibilidade de rótulos nos dados e o objetivo da tarefa. Entre as principais abordagens estão o aprendizado supervisionado, o aprendizado não supervisionado e o aprendizado por reforço, cada um com características e aplicações específicas (ALPAYDIN, 2010).

No aprendizado supervisionado, os algoritmos treinam com dados rotulados, associando entradas específicas a saídas desejadas. Esse método é amplamente utilizado em tarefas como classificação e regressão. Wu et al. (2023) exploraram essa abordagem ao aplicá-la à detecção de defeitos em estruturas de concreto utilizando ultrassom, demonstrando a capacidade de modelos supervisionados em resolver problemas de alta criticidade. Já no aprendizado não supervisionado, os dados não possuem rótulos, e o algoritmo busca identificar padrões ou agrupamentos com base em características compartilhadas entre os exemplos. Essa abordagem é aplicada em problemas de agrupamento (*clustering*) e segmentação de dados (ALPAYDIN, 2010).

O aprendizado por reforço, diferente dos métodos anteriores, envolve interações com o ambiente, nas quais o sistema recebe recompensas ou penalizações. Esse processo é particularmente útil em áreas como robótica e jogos, onde a interação contínua e o ajuste progressivo de estratégias são fundamentais (JOHNSON, 2024). A avaliação de desempenho é uma etapa importante no desenvolvimento de modelos de aprendizado de máquina, e a validação cruzada (*cross-validation*) é amplamente utilizada para essa finalidade. O método mais comum, conhecido como validação cruzada em k -partições (*k-fold*), divide os dados em k subconjuntos de tamanhos semelhantes. Em cada iteração, um subconjunto é reservado para teste, enquanto os demais são usados para treinamento. Após k iterações, todos os subconjuntos terão sido utilizados como teste, e a média das métricas obtidas fornece uma estimativa consolidada de desempenho do modelo, conforme Figura 15. Essa abordagem é particularmente útil em cenários com dados limitados, pois aproveita ao máximo as informações disponíveis e reduz o risco de sobreajuste (*overfitting*). Além disso, variantes como a validação estratificada preservam a proporção das classes nos subconjuntos, garantindo uma avaliação mais adequada em problemas de classificação (ALPAYDIN, 2010).

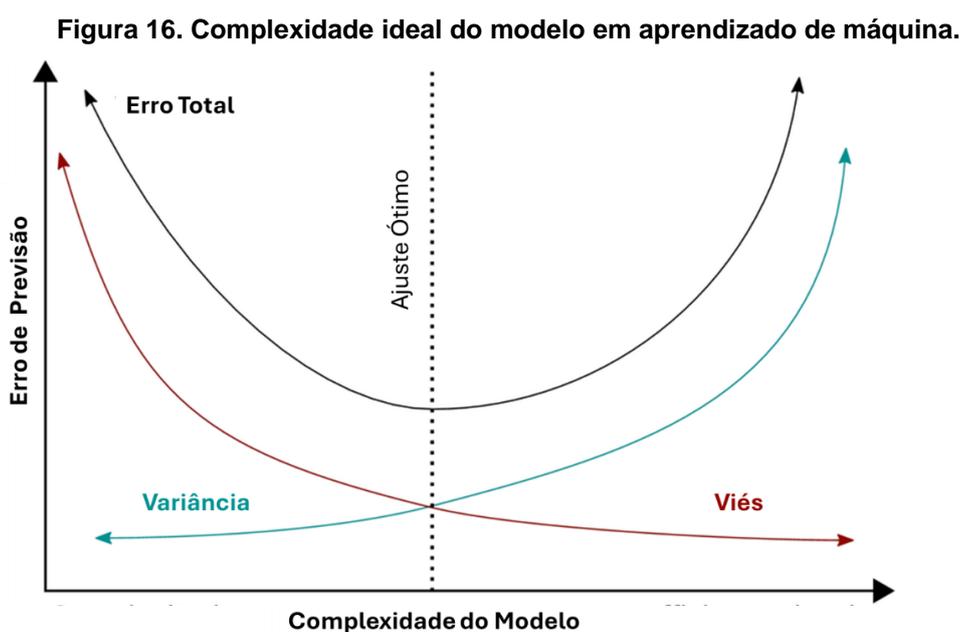
Figura 15. Validação cruzada



Fonte: Autoria própria

O conceito de compensação entre viés e variância (*bias-variance trade-off*) é um dos principais desafios no aprendizado de máquina e está exemplificado na Figura 16. O viés (*bias*), representado na curva vermelha, indica a simplificação excessiva, frequentemente observada em modelos simples, o que pode resultar em subajuste (*underfitting*).

Em contrapartida, a variância (*variance*), ilustrada na curva azul, reflete a sensibilidade do modelo a flutuações nos dados de treinamento, algo comum em modelos mais complexos que podem sofrer de sobreajuste (*overfitting*). A curva preta mostra o erro total como soma dessas duas componentes. Para minimizar esse erro total, é importante ajustar a complexidade do modelo de modo a alcançar o ponto de equilíbrio indicado, utilizando estratégias como regularização e validação cruzada, garantindo melhor generalização e desempenho do modelo (JIANG, 2022).



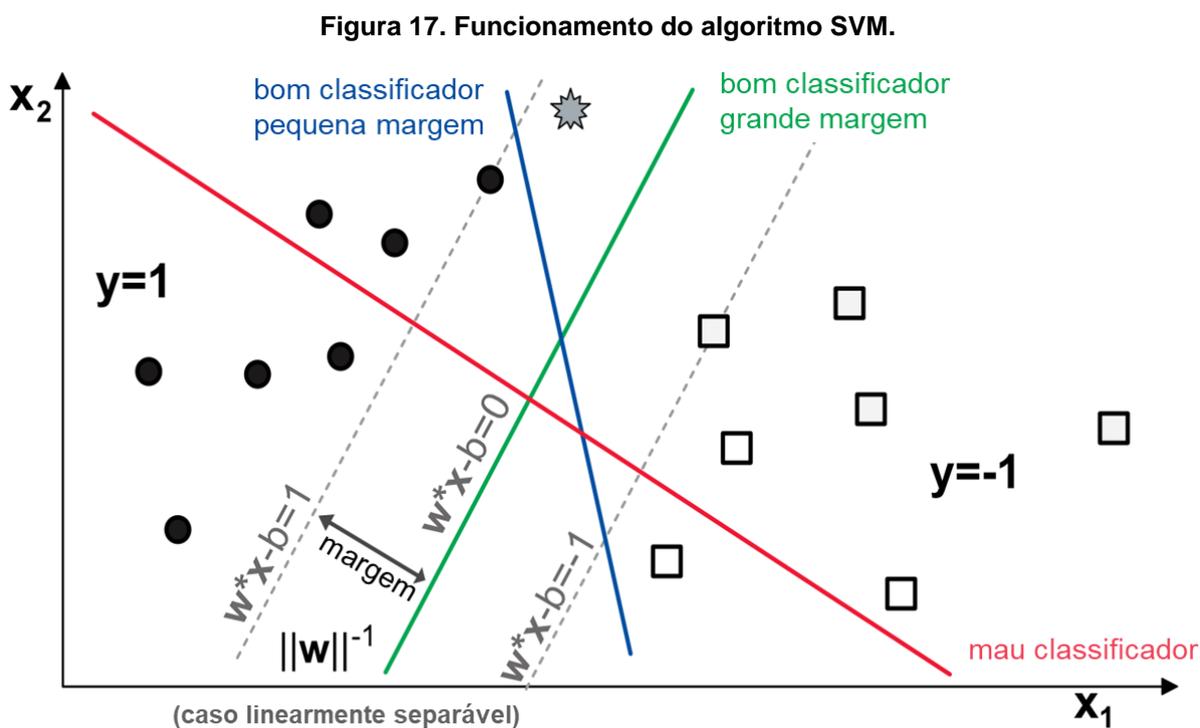
Fonte: Adaptado de Jiang (2022)

O aprendizado de máquina tem impulsionado a inspeção automatizada por ultrassom, aumentando a precisão e eficiência da detecção de falhas estruturais. Mukhti, Gucunski e Kee (2024) desenvolveram um sistema de classificação assistida por IA para avaliar danos em concreto induzidos por corrosão do aço. O desenvolvimento de algoritmos cada vez mais eficientes representa um passo relevante para a automação da inspeção industrial, reduzindo custos e aumentando a segurança da gestão de integridade de equipamentos.

2.3.2 Support Vector Machine (SVM)

A técnica de *Support Vector Machine* (SVM) destaca-se como uma poderosa ferramenta no campo do aprendizado de máquina supervisionado, amplamente utilizada em problemas de classificação e regressão. Baseada na teoria de aprendizado estatístico, a SVM busca encontrar um hiperplano que maximize a margem entre as classes em um espaço de características, minimizando sobreajuste quando comparada a métodos mais simples (ALPAYDIN, 2010).

A Figura 17 ilustra como o SVM identifica o hiperplano de separação em um caso linearmente separável, destacando que um hiperplano com margem maior tende a generalizar melhor, enquanto um com margem menor, embora funcional, é mais suscetível a ruídos. Também é mostrado um exemplo de um classificador, cuja margem inadequada resulta em baixa separação entre as classes (ALI, 2024).



Fonte: Adaptado de Ali (2024)

Essa abordagem é modelada matematicamente pela otimização da margem máxima, conforme Equação 2.24:

$$\min_{\mathbf{w}, b} \frac{1}{2} \|\mathbf{w}\|^2 \quad 2.24$$

sujeito a:

$$y_i(\mathbf{w}^\top \mathbf{x}_i + b) \geq 1, \forall i \quad 2.25$$

onde:

- \mathbf{w} é o vetor normal ao hiperplano,
- b é viés do hiperplano,
- $y_i \in \{-1, 1\}$ são os rótulos das classes,
- \mathbf{x}_i são os vetores de entrada.

Quando os dados não são linearmente separáveis, a técnica de SVM emprega o Truque do *Kernel* (*Kernel Trick*), que permite projetar os dados para um espaço dimensional superior, onde uma separação linear se torna possível. Funções de kernel, como o linear, polinomial e radial (RBF, do inglês *Radial Basis Function*), são amplamente utilizadas para realizar essa projeção (ALPAYDIN, 2010).

O Truque do *Kernel* substitui o produto interno original $(\mathbf{x}_i, \mathbf{x}_j)$ pela função $K(\mathbf{x}_i, \mathbf{x}_j)$ definida conforme Equação 2.26:

$$K(\mathbf{x}_i, \mathbf{x}_j) = \phi(\mathbf{x}_i)^\top \phi(\mathbf{x}_j), \quad 2.26$$

onde $\phi(x)$ é uma transformação implícita dos dados para o espaço de características de maior dimensão.

Kernel linear – Equação 2.27:

$$K(\mathbf{x}_i, \mathbf{x}_j) = \mathbf{x}_i^\top \mathbf{x}_j \quad 2.27$$

Kernel Gaussiano (RBF) – Equação 2.28:

$$K(\mathbf{x}_i, \mathbf{x}_j) = \exp\left(-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|^2}{2s^2}\right) \quad 2.28$$

Kernel Polinomial – Equação 2.29:

$$K(\mathbf{x}_i, \mathbf{x}_j) = (\mathbf{x}_i^T \mathbf{x}_j + 1)^q. \quad 2.29$$

O *Kernel* RBF, em particular, destaca-se por sua capacidade de modelar relações complexas em dados de alta dimensionalidade, sendo amplamente empregado em aplicações como a detecção de defeitos em materiais metálicos e outros sistemas (XIAO, 2020).

No domínio dos ensaios ultrassônicos, a SVM tem sido aplicada com sucesso em problemas como a detecção de defeitos em materiais metálicos e concretos. Por exemplo, foi utilizada para diferenciar inclusões e cavidades, integrando características como amplitude e frequência dos sinais ultrassônicos, alcançando alta precisão em classificações (DE SOUZA, 2023). Liu (2024) demonstrou a eficácia do SVM na detecção e classificação de corrosão em dutos, utilizando processamento de sinais e técnicas de aprendizado de máquina para otimizar a inspeção automatizada.

Estudos adicionais reforçam a eficácia da SVM na análise de sinais ultrassônicos e em outros contextos de classificação complexa. Por exemplo, WANG et al. (2022) combinaram PCA, LDA e SVM para diagnóstico de falhas em reguladores de gás, demonstrando a eficácia do modelo na redução de dimensionalidade e na separação de classes, alcançando resultados robustos em cenários industriais. Da mesma forma, HAYATI et al. (2024) aplicaram a combinação PCA-LDA-SVM. Nesse estudo, a metodologia alcançou até 100% de acurácia, reforçando a capacidade da SVM, em conjunto com técnicas de redução de dimensionalidade, de lidar com dados de alta complexidade.

A aplicação de PCA e LDA aprimora a separação entre padrões espectrais e reduz a complexidade dos dados, tornando o kernel linear do SVM uma escolha vantajosa. Sua simplicidade computacional, boa capacidade de generalização e baixa necessidade de ajuste de hiperparâmetros o tornam uma alternativa viável a modelos mais complexos (MATHWORKS, 2024). Assim, o SVM Linear se destaca pelo equilíbrio entre eficiência, desempenho e interpretação, especialmente após a aplicação de PCA e LDA, que transformam os dados em um espaço de maior separabilidade linear.

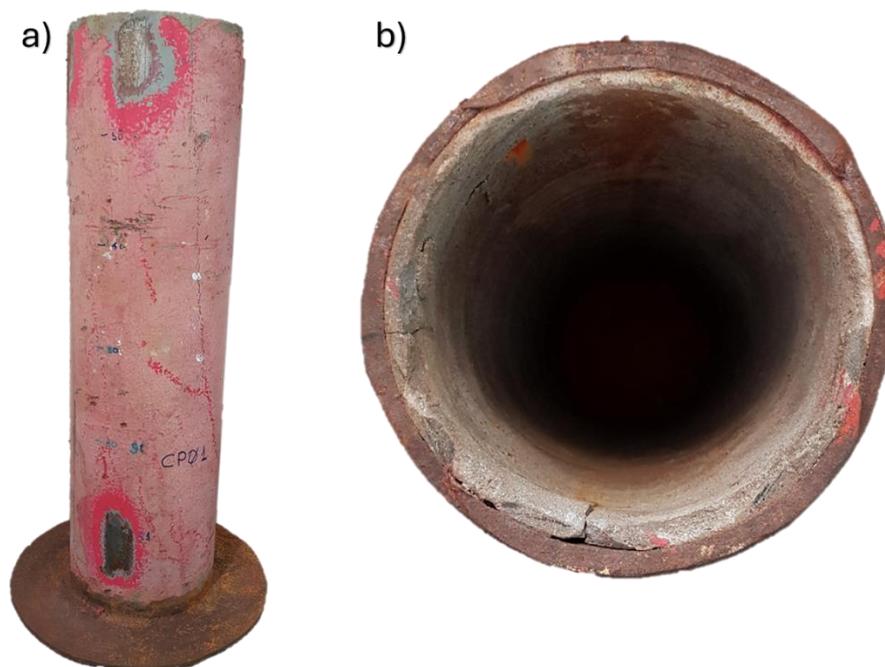
3 MATERIAL E MÉTODOS

3.1 Material

O material utilizado neste estudo consiste em segmentos de tubulação de aço carbono API 5L Gr B, com diâmetro nominal de 150 mm e espessura de parede de 6,02 mm, conforme mostrado na Figura 18. Os tubos foram retirados de um sistema de incêndio, em operação há 21 anos, utilizando água salgada bruta do mar.

A tubulação possui um revestimento externo composto por uma camada de tinta epóxi-fosfato de zinco, com espessura de 100 μm . Sobre esta camada, há um acabamento de tinta de poliuretano acrílico, com espessura de 65 μm , para aumentar a resistência à degradação, causada pela radiação UV e prolongar a durabilidade do revestimento externo. Além disso, a tubulação é revestida internamente com uma argamassa de cimento e areia, com espessura de 5 mm, de acordo com a norma ANSI/AWWA C205-12 (AWWA., 2012). Ambos os revestimentos externo e interno tem como função principal, proteger o tubo de aço carbono contra processos corrosivos.

Figura 18. Vista frontal (a) e posterior (a) do tubo usado para fabricação do corpo de prova.

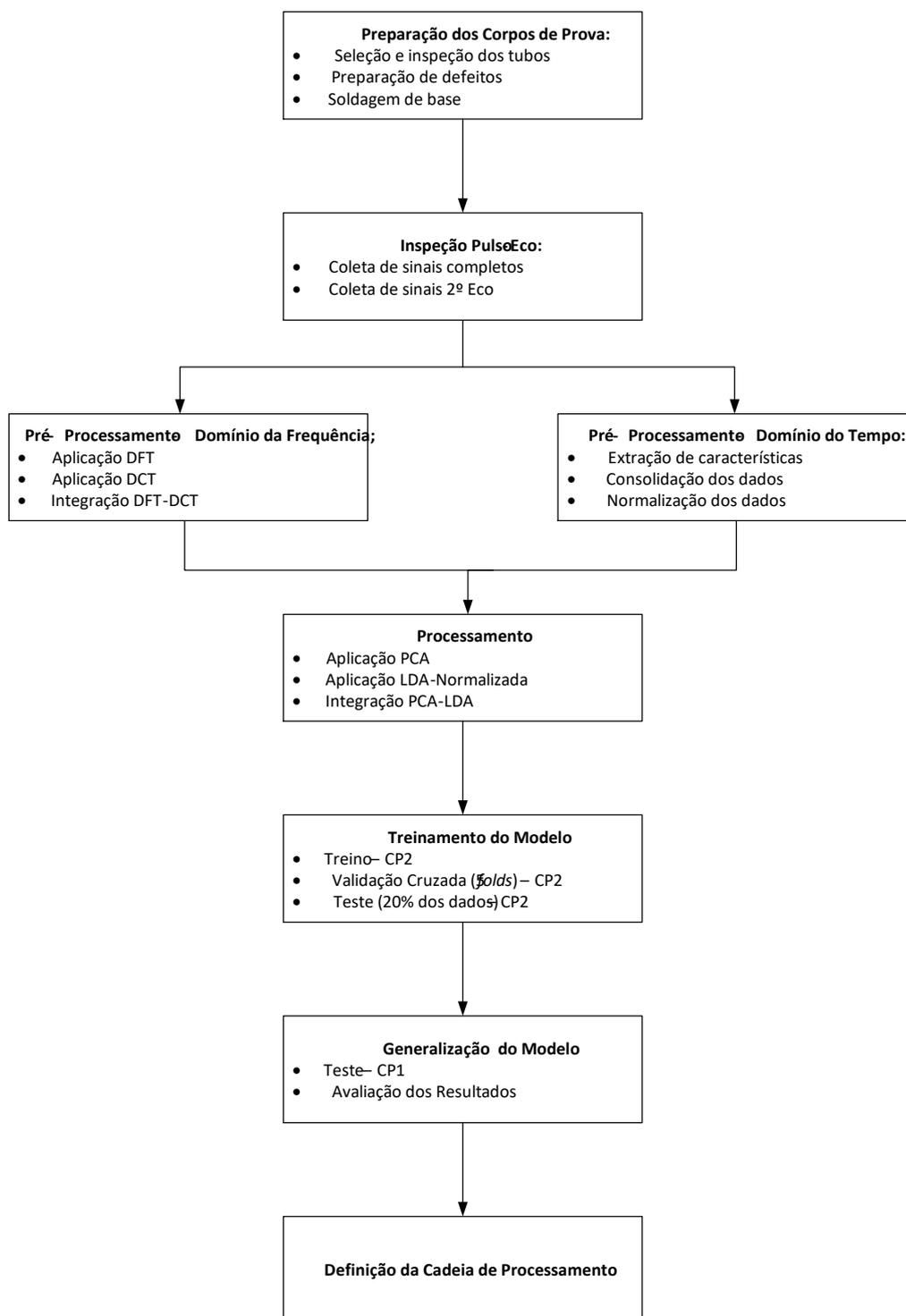


Fonte: Autoria Própria

3.2 Métodos

A metodologia empregada para o desenvolvimento deste trabalho encontra-se resumidamente descrita no fluxograma, exibido na Figura 19.

Figura 19. Fluxograma de Metodologia para Análise de Sinais Ultrassônicos.



Fonte: Autoria Própria

3.2.1 Preparação dos Corpos de Prova

Para a realização dos experimentos, foram selecionados dois segmentos de tubos, com aproximadamente 600 mm de comprimento cada, provenientes do material descrito anteriormente. Estes segmentos foram cortados transversalmente utilizando técnicas de seccionamento a quente, empregando maçarico e lixadeira, uma vez que estas regiões de extremidades não estariam sob área de interesse do estudo. Os segmentos foram denominados como corpos de prova 01 (CP1) e 02 (CP2).

As extremidades dos corpos de prova foram soldadas a uma chapa de aço carbono ASTM A283 Gr C com 6,35mm de espessura, que desempenhou a dupla função de base para o posicionamento vertical dos tubos e de estanqueidade para armazenamento de água em seu interior, possibilitando a simulação de condições operacionais.

No CP1, após uma inspeção visual, foram selecionadas três regiões de interesse para estudo, buscando correlacionar as características visuais com as respostas obtidas utilizando ensaios ultrassônicos, conforme indicado na Figura 20a. A região S1 é identificada como uma área visualmente íntegra, sem evidências de falhas aparentes no revestimento interno de argamassa. Esta região foi escolhida como referência para análise comparativa. A região S2 consiste numa área onde se observaram fissuras no revestimento interno, além de indícios de separação entre o revestimento de argamassa e a parede interna da tubulação, causada por degradação ao longo da vida útil do material. A região S3 é outra área com presença de fissuras visíveis no revestimento interno, indicando degradação localizada.

O CP2, exibido na Figura 20b, complementa o estudo ao proporcionar um cenário experimental controlado que facilita a análise e a correlação entre os defeitos e os sinais detectados. Neste corpo de prova, também foram definidas três regiões de interesse para estudo. Uma região denominada S1, que foi selecionada como referência por apresentar um revestimento interno visualmente íntegro, sem sinais de degradação ou falhas aparentes. Nas regiões S2 e S3, foram induzidos defeitos no revestimento interno para simular condições de falha que poderiam ocorrer em serviço. Na região S2, o defeito consistiu na remoção total do revestimento, reproduzindo uma condição semelhante ao caso real ilustrado na Figura 4b que

mostra uma tubulação retirada de operação devido à degradação severa do revestimento interno. Na Região S3, o defeito foi projetado para reproduzir uma perda de aderência entre o revestimento de argamassa e a parede interna da tubulação, simulando um cenário de descolamento total associado à degradação do material.

Figura 20. Corpos de prova CP1 e CP2 com indicação das regiões de interesse.



Fonte: Autoria Própria

As características de cada região para ambos os corpos de prova estão organizadas na Tabela 2.

Tabela 2. Descrição das regiões de interesse nos corpos de prova CP1 e CP2.

| REGIÃO | CP1 - DEFEITOS NATURAIS | CP2 - DEFEITOS ARTIFICIAS |
|--------|---|---|
| S1 | Referência, com revestimento íntegro. | Referência, com revestimento íntegro. |
| S2 | Presença de fissuras e indícios de separação do revestimento. | Defeito simulado de perda de aderência do revestimento. |
| S3 | Área com fissuras visíveis, indicando degradação localizada. | Defeito simulado de remoção total do revestimento. |

Fonte: Autoria Própria

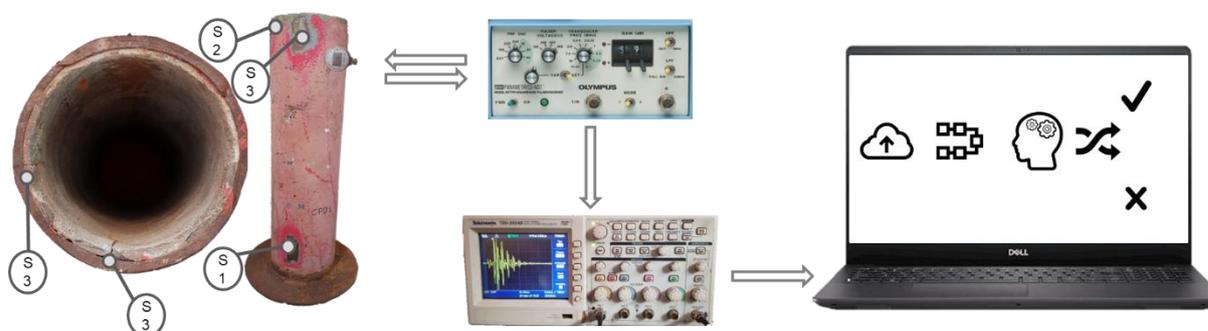
Essa abordagem visa representar cenários reais observados em operações industriais, permitindo avaliar a eficácia da metodologia proposta na identificação dessas áreas comprometidas e na distinção de regiões com integridade preservada.

3.2.2 Inspeção Ultrassônica Pulso-Eco

A técnica de inspeção ultrassônica Pulso-Eco foi utilizada para coletar os sinais nas regiões S1, S2 e S3 de ambos os corpos de prova.

Para a coleta dos sinais ultrassônicos, foram utilizados os seguintes equipamentos exibidos na Figura 21: um gerador de pulso ultrassônico modelo OLYMPUS 5077PR (a), um osciloscópio digital modelo Tektronix TDS 2024B (b), e um cabeçote de ultrassom monocristal modelo GE 026HJ9 (c), operando a uma frequência central de 2,25 MHz, onde buscou-se o equilíbrio entre penetração e resolução. Para garantir a transmissão do sinal ultrassônico entre o cabeçote e a superfície do tubo e minimizar as perdas por impedância, foi utilizado o líquido acoplante SONOTECH Ultragel II. A captura e o arquivamento dos sinais ultrassônicos foram realizados utilizando o aplicativo LabVIEW instalado em um notebook Dell com processador Intel Core i7 e 8 GB de RAM (d).

Figura 21. Configuração experimental e equipamentos utilizados para a coleta dos sinais ultrassônicos.



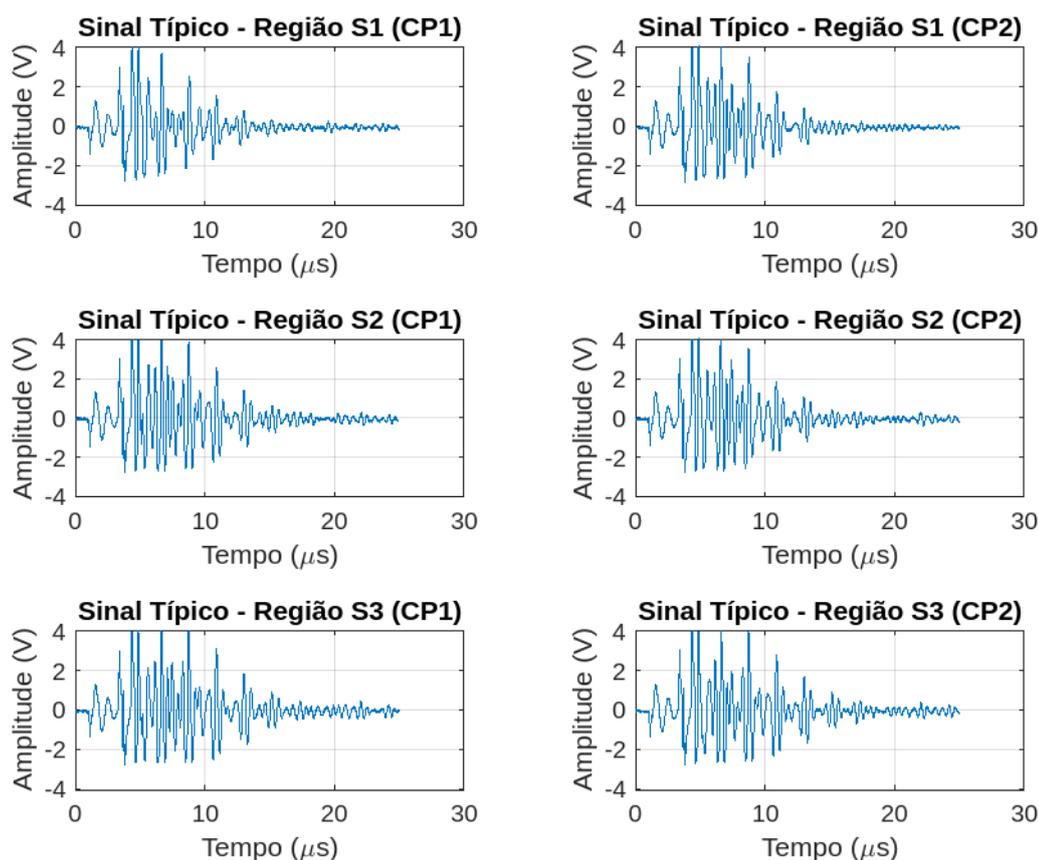
Fonte: Autoria Própria

A coleta dos sinais foi realizada por meio do acoplamento e desacoplamento repetido do cabeçote ultrassônico em uma área previamente delimitada de 1,5 cm x 1,5 cm, garantindo a aquisição dos sinais em diferentes pontos dentro de cada região de interesse dos corpos de prova.

Para análise no domínio do tempo, foram coletados 50 sinais *A-scan* de domínio completo, abrangendo tanto o pulso inicial transmitido quanto os ecos subsequentes refletidos, em cada região de interesse dos corpos de prova CP1 e CP2.

Os sinais registrados apresentaram 2500 pontos amostrados, com intervalo de amostragem de $0,01 \mu\text{s}$, garantindo alta resolução temporal para as análises. A Figura 22 apresenta exemplos típicos dos sinais de onda completa *A-scan* para as três regiões (S1, S2 e S3) dos corpos de prova CP1 e CP2, respectivamente.

Figura 22. Sinais típicos *A-scan* de onda completa nas três regiões de interesse (S1, S2 e S3) dos corpos de prova CP1 e CP2.



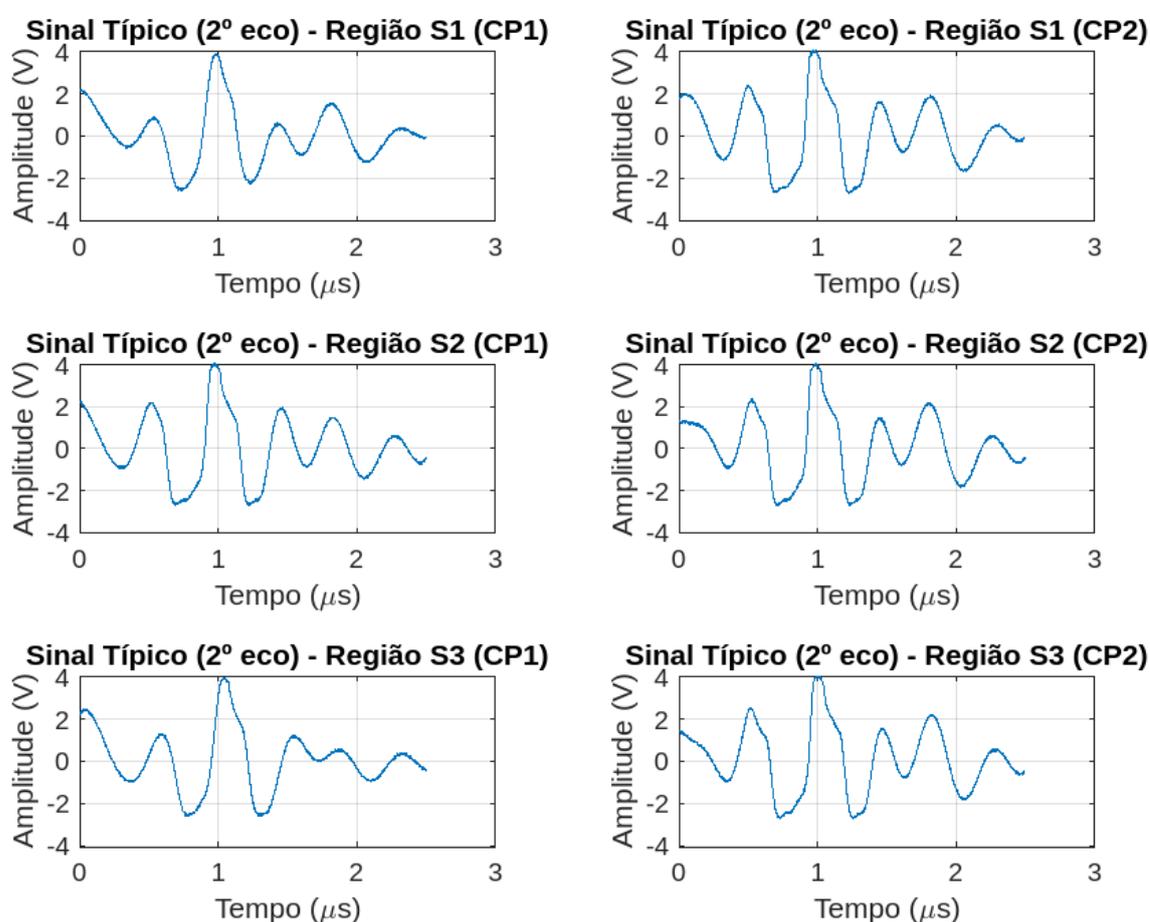
Fonte: Autoria Própria

Adicionalmente, foram coletados 50 sinais referentes ao segundo eco de fundo da parede interna do tubo em cada região de interesse dos corpos de prova CP1 e CP2. Para estes sinais, foi realizada uma janela temporal no osciloscópio, garantindo que apenas a parte correspondente ao segundo eco fosse capturada, resultando em 2500 pontos com intervalo de $0,001 \mu\text{s}$. Isso permitiu uma posterior aplicação de transformadas no domínio da frequência, otimizando a análise espectral sem perda significativa de resolução. Diferentemente do sinal de onda completa, onde a aplicação de transformadas levaria a uma perda de detalhes devido à maior dispersão

de informações, o janelamento do segundo eco possibilitou uma análise mais detalhada e focada na região de interesse.

A Figura 23 exibe os sinais típicos *A-scan* correspondentes ao segundo eco de fundo, também organizados pelas três regiões de interesse nos corpos de prova CP1 e CP2. Os gráficos permitem visualizar as diferenças sutis nas características dos sinais coletados nas diferentes regiões, destacando a variação de amplitude e o comportamento temporal.

Figura 23. Sinais típicos *A-scan* do segundo eco de fundo nas três regiões de interesse (S1, S2 e S3) dos corpos de prova CP1 e CP2.



Fonte: Autoria Própria

Cada corpo de prova, portanto, gerou um total de 150 sinais para onda completa e 150 sinais do segundo eco, distribuídos uniformemente entre as regiões analisadas (S1, S2 e S3).

3.2.3 Processamento dos Sinais

Para o processamento dos sinais ultrassônicos e a aplicação de técnicas de aprendizado de máquina, o MATLAB® foi escolhido como ferramenta principal devido à sua integração nativa entre módulos de processamento de sinais e aprendizado de máquina, além de sua adoção no Programa de Pós-Graduação em Engenharia de Materiais (PPGEM) do IFBA.

Foram utilizadas bibliotecas como o *Signal Processing Toolbox* e o *Statistics and Machine Learning Toolbox*, que permitiram análises eficientes e precisas, além do *Classification Learner*, que simplificou o desenvolvimento dos modelos.

Embora alternativas como Python com SciPy e TensorFlow ofereçam funcionalidades similares, a escolha do MATLAB® foi determinada pela sua disponibilidade institucional.

3.2.3.1 Pré-processamento dos Sinais no Domínio do Tempo

Para a análise dos sinais ultrassônicos Pulso-Eco no domínio do tempo, foram consideradas as informações dos sinais completos coletados nas diferentes regiões de interesse dos corpos de prova CP1 e CP2.

A análise foi orientada pela hipótese de que regiões com defeitos, devido à predominância da interface aço-água, apresentam maior reflexão e menor atenuação das ondas ultrassônicas, em comparação às regiões íntegras, caracterizadas pela interface aço-concreto.

Conforme discutido na seção 2.1.2.2, a propagação das ondas ultrassônicas é influenciada pela condição do revestimento. A ausência de material ou a presença de fissuras intensifica a reflexão do sinal, pois a diferença de impedância acústica na interface aço-água é significativamente maior do que na interface aço-concreto.

Essa hipótese foi avaliada com base nos coeficientes de reflexão, calculados pela Equação 2.9, considerando valores de impedância da Tabela 1.

Os cálculos apresentados abaixo mostram que a interface aço-concreto reflete 44,4% da energia ($R_{aço-concreto} = 0,44$), enquanto a interface aço-água reflete 88% ($R_{aço-água} = 0,88$). Assim, nas regiões íntegras há maior transmissão de energia ao

concreto, enquanto nas regiões com defeitos predomina a reflexão, favorecendo a propagação do sinal ultrassônico coletado no experimento.

Região Íntegra (aço-concreto)

| | |
|---|---|
| Impedância do aço: | Impedância do concreto: |
| $Z_{\text{aço}} = 46 \times 10^6 \text{ kg/m}^2 \cdot \text{s}$ | $Z_{\text{concreto}} = 9,2 \times 10^6 \text{ kg/m}^2 \cdot \text{s}$ |

$$R = \left(\frac{Z_{\text{aço}} - Z_{\text{concreto}}}{Z_{\text{aço}} + Z_{\text{concreto}}} \right)^2 = \left(\frac{46 - 9,2}{46 + 9,2} \right)^2 = \left(\frac{36,8}{55,2} \right)^2 = 0,444$$

Região com Defeito (aço-água)

| | |
|---|--|
| Impedância do aço: | Impedância da água: |
| $Z_{\text{aço}} = 46 \times 10^6 \text{ kg/m}^2 \cdot \text{s}$ | $Z_{\text{água}} = 1,48 \times 10^6 \text{ kg/m}^2 \cdot \text{s}$ |

$$R = \left(\frac{Z_{\text{aço}} - Z_{\text{água}}}{Z_{\text{aço}} + Z_{\text{água}}} \right)^2 = \left(\frac{46 - 1,48}{46 + 1,48} \right)^2 = \left(\frac{44,52}{47,48} \right)^2 = 0,88$$

A análise dos sinais coletados foi realizada com base nos parâmetros extraídos, buscando correlacionar as condições das interfaces com o comportamento dos sinais ultrassônicos. Os resultados foram representados por meio de *boxplots*, destacando medianas, quartis, valores atípicos e variações entre regiões íntegras e com defeitos.

Além da avaliação estatística, cada parâmetro foi analisado individualmente no classificador SVM Linear, permitindo verificar sua eficácia na distinção das condições do revestimento. Após o cálculo dos parâmetros estatísticos — coeficiente de atenuação, energia do sinal, valor médio, RMS, desvio padrão e curtose, os dados foram organizados em uma matriz para cada corpo de prova. Em seguida, os valores foram normalizados, garantindo a uniformização das escalas e evitando que variáveis com magnitudes elevadas tivessem influência desproporcional nas análises. Essa matriz (denominada TIME, do tipo *double*, 6×150) foi utilizada como base para o treinamento e teste do classificador SVM com kernel linear.

3.2.3.2 Pré-processamento dos Sinas no Domínio da Frequência

Com o objetivo de complementar a análise no domínio do tempo, os sinais ultrassônicos foram transformados para o domínio da frequência utilizando a DFT, a DCT e uma matriz combinada DFT-DCT, com base nos dados do segundo eco de fundo.

A DFT foi empregada para analisar a distribuição de energia ao longo do espectro, permitindo a extração de parâmetros como frequência dominante, amplitude máxima, largura de banda e área total do espectro.

A DCT, por sua vez, foi utilizada para explorar sua capacidade de concentrar a maior parte da energia do sinal em componentes de baixa frequência, oferecendo uma perspectiva complementar à DFT.

Para otimizar o processamento e reduzir a dimensionalidade, foram selecionados os 50 primeiros coeficientes mais representativos dos espectros gerados pelas transformações individuais.

A combinação das transformadas DFT e DCT tem se mostrado uma abordagem eficiente para a extração e preservação de informações relevantes em sinais digitais. Estudos como os de Hamidi et al. (2018) destacam sua capacidade de aprimorar a robustez contra perturbações externas, facilitando a identificação de padrões mesmo em sinais ruidosos. Para garantir que ambas as transformações contribuíssem de forma equilibrada para a análise, as matrizes resultantes foram previamente normalizadas individualmente, evitando que diferenças na escala dos coeficientes influenciassem desproporcionalmente a classificação. Conforme Raghu et al. (2017), essa fusão amplia a discriminação das características extraídas, tornando-se uma estratégia valiosa para o reconhecimento de padrões e aprendizado de máquina.

Os dados transformados, após a aplicação das diferentes cadeias de processamento exploradas neste trabalho, serão utilizados como entrada para o classificador SVM Linear. A análise abrangerá tanto a eficácia das transformadas DFT e DCT aplicadas isoladamente quanto a concatenação das matrizes DFT-DCT, permitindo avaliar o desempenho de cada abordagem na separação das classes de defeitos no revestimento interno da tubulação.

3.2.3.3 Análise dos Componentes Principais (PCA)

A PCA foi utilizada para reorganizar as matrizes de dados, facilitando sua análise e aprimorando a separação entre classes. Sua aplicação permitiu tanto a avaliação direta dos dados no classificador SVM Linear quanto a preparação para a aplicação da LDA, otimizando a representatividade das características extraídas. O objetivo dessa abordagem foi explorar as informações mais relevantes de cada representação, otimizando a distinção entre diferentes condições estruturais dos corpos de prova. Após as etapas de extração de características e normalização, as matrizes consolidadas foram preparadas para a PCA. Cada linha representava uma amostra (sinal) e cada coluna, uma característica. A PCA foi então aplicada utilizando a função *'pca'* do MATLAB, que decompôs os dados em componentes principais ordenados pela variância explicada. As projeções das amostras no espaço de componentes principais, geradas pela PCA, foram utilizadas integralmente para a análise subsequente.

3.2.3.4 Análise Discriminante Linear (LDA)

A LDA foi aplicada como ferramenta para projetar os dados em um espaço discriminante que maximizasse a separação entre as três classes definidas, correspondentes a diferentes condições de integridade dos revestimentos dos corpos de prova. A técnica foi aplicada em dois cenários distintos.

No primeiro cenário, a LDA foi aplicada diretamente às matrizes de características extraídas, incluindo dados transformados por DFT, DCT, a matriz combinada DFT-DCT e as características extraídas no domínio do tempo. No entanto, os resultados iniciais revelaram discrepâncias significativas nas escalas dos discriminantes lineares entre os conjuntos CP1 e CP2, o que levou à atribuição incorreta de todas as amostras a uma única classe, comprometendo a generalização do modelo. Para mitigar esse problema, foi implementada a normalização dos discriminantes lineares após a aplicação da LDA, garantindo compatibilidade das projeções entre os dois conjuntos.

No segundo cenário, a PCA foi aplicada previamente à LDA, com o objetivo de reorganizar os dados em componentes principais e reduzir possíveis discrepâncias na escala dos discriminantes.

3.2.4 Aprendizado de Máquina

Neste trabalho, diferentes cadeias de processamento foram aplicadas aos sinais ultrassônicos coletados nos CP1 e CP2, com o objetivo de avaliar a eficácia do SVM Linear na classificação das condições de integridade dos revestimentos internos.

As cadeias de processamento foram organizadas conforme as etapas de redução de dimensionalidade empregadas: PCA, LDA ou PCA-LDA. As combinações analisadas são apresentadas na Tabela 3:

Tabela 3. Cadeias de processamento avaliadas em combinação com o algoritmo SVM Linear

| Domínio da Frequência | Domínio do Tempo |
|---|--|
| <ul style="list-style-type: none"> ○ DFT-PCA ○ DFT-LDA ○ DFT-PCA-LDA ○ DCT-PCA ○ DCT-LDA ○ DCT-PCA-LDA ○ DFT-DCT-PCA ○ DFT-DCT-LDA ○ DFT-DCT-PCA-LDA | <ul style="list-style-type: none"> ○ TIME-PCA ○ TIME-LDA ○ TIME-PCA-LDA |

O objetivo principal foi identificar a cadeia de processamento mais eficiente para distinguir sinais provenientes de regiões com defeitos de sinais oriundos de áreas íntegras. O treinamento, validação e teste foram realizados utilizando o recurso *Classification Learner* do MATLAB®, que oferece uma interface gráfica intuitiva para experimentação e análise dos dados. O corpo de prova CP2 foi utilizado para as etapas de treinamento e validação, sendo que 20% de suas amostras foram reservadas como um conjunto de teste independente. Para garantir uma avaliação mais equilibrada, foi utilizada a estratégia de validação cruzada com 5 divisões (*folds*),

na qual os dados são divididos em cinco partes iguais. Em cada ciclo, quatro partes foram usadas para treinamento e a parte restante para validação, repetindo-se o processo cinco vezes.

Os classificadores foram treinados e testados cinco vezes, gerando médias e desvios padrão das métricas de desempenho, incluindo as acurácias de validação e teste, para avaliar a consistência dos resultados.

Após o treinamento no CP2, as cadeias de processamento foram aplicadas aos dados do corpo de prova CP1 para avaliar a capacidade de generalização do modelo. Esse teste utilizou as mesmas configurações empregadas no CP2, analisando o desempenho do classificador em um conjunto de dados não visto previamente. Os resultados de cada cadeia de processamento foram consolidados em tabelas comparativas, destacando as médias e desvios padrão das acurácias de validação e teste. Adicionalmente, foram geradas matrizes de confusão e gráficos de dispersão, ilustrando o desempenho das combinações PCA, LDA e PCA-LDA na classificação das condições de integridade do revestimento. As métricas principais utilizadas na análise foram as apresentadas nas equações 3.1, 3.2 e 3.3 (SELVI e KAMALAKANNAN, 2024):

$$\text{Acurácia} = \frac{TP + TN}{TP + TN + FP + FN} \quad (3.1)$$

$$\text{Sensibilidade}_{s_1} = \frac{TP}{TP + FN} \quad (3.2)$$

$$\text{Especificidade}_{s_1} = \frac{TN}{TN + FP} \quad (3.3)$$

onde:

- *TP (True Positives)*: Verdadeiros Positivos – previsões corretas da classe positiva.
- *TN (True Negatives)*: Verdadeiros Negativos – previsões corretas da classe negativa.
- *FP (False Positives)*: Falsos Positivos – previsões incorretas da classe positiva (quando a classe verdadeira é negativa).
- *FN (False Negatives)*: Falsos Negativos – previsões incorretas da classe negativa (quando a classe verdadeira é positiva).

Essas métricas foram calculadas com base nas matrizes de confusão geradas para cada cadeia de processamento, permitindo uma avaliação detalhada do desempenho do classificador. A especificidade da classe S1 (Região Íntegra) é fundamental para evitar falsos positivos, onde áreas com defeitos podem ser classificadas como íntegras, comprometendo a segurança operacional. Por outro lado, a sensibilidade da classe S1 é especialmente relevante em aplicações industriais, pois erros ao classificar áreas íntegras como defeituosas (falsos negativos) podem resultar em gastos desnecessários com inspeções e reparos preventivos. Essa análise possibilitou a identificação da cadeia de processamento mais adequada, equilibrando acurácia de validação e capacidade de generalização.

4 RESULTADOS E DISCUSSÃO

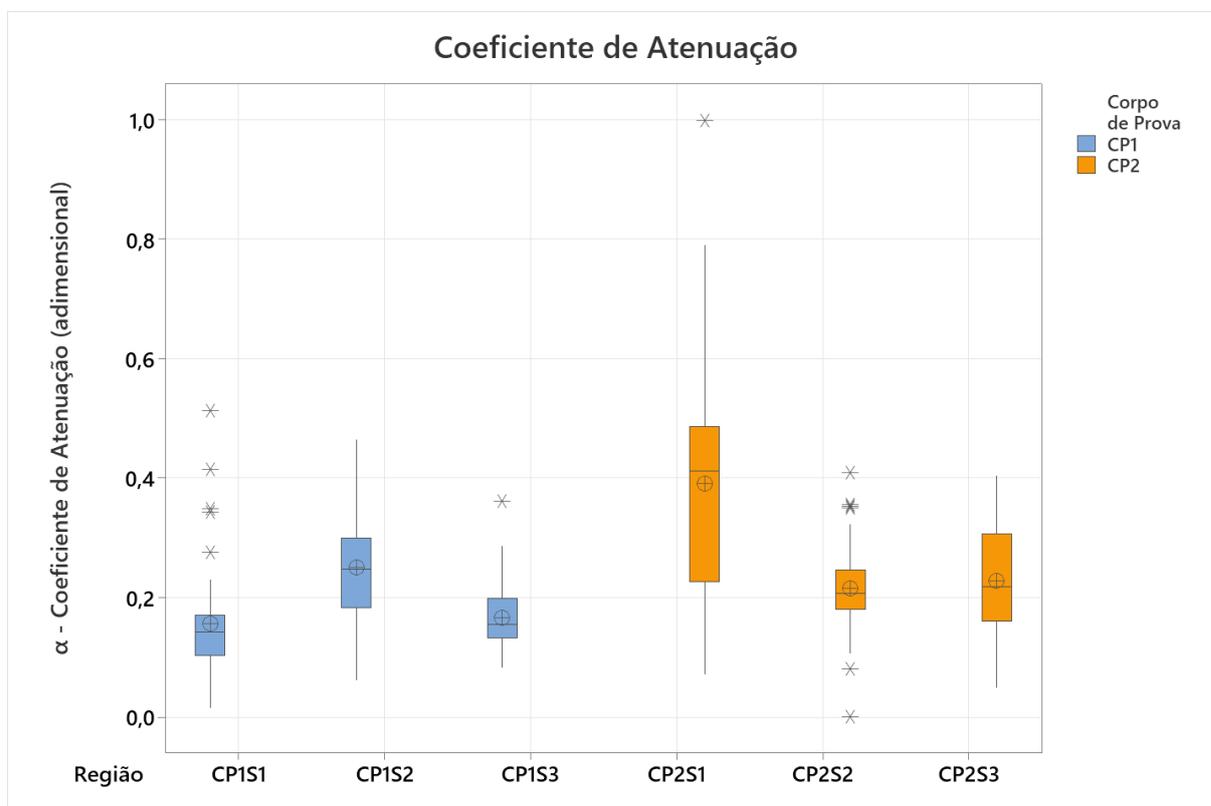
4.1 Avaliação das Características dos Sinais no Domínio do Tempo

Na etapa descrita, foram analisados parâmetros como coeficiente de atenuação, energia do sinal, média, RMS, desvio padrão e curtose, fornecendo informações sobre a distribuição de energia e a variabilidade dos sinais nas diferentes regiões dos corpos de prova.

4.1.1 Coeficiente de Atenuação

A Figura 24 apresenta a distribuição dos coeficientes de atenuação normalizados para as três regiões analisadas em ambos os corpos de prova.

Figura 24 - Gráfico *boxplot* - Coeficiente de atenuação das diferentes regiões dos CP1 e CP2.



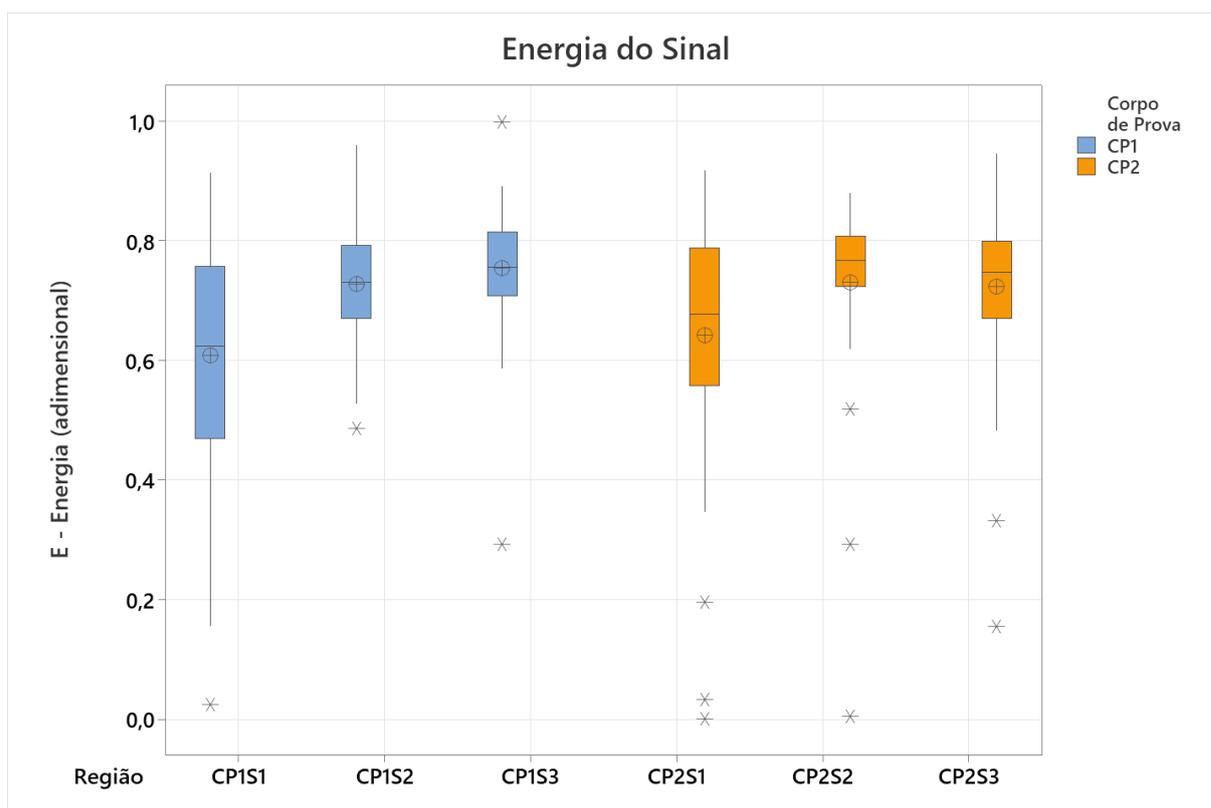
Os coeficientes de atenuação apresentaram comportamentos distintos nos corpos de prova. No CP1, as regiões com defeitos mostraram maior atenuação do que

as íntegras, contrariando a hipótese inicial. Em contraste, no CP2, a hipótese foi confirmada, com regiões com defeitos exibindo menor atenuação, atribuída à predominância da interface aço-água, que favorece a reflexão e reduz a absorção do sinal. No entanto, a sobreposição dos dados entre regiões íntegras e defeituosas dificultou a distinção clara, destacando as limitações dos coeficientes de atenuação como indicadores isolados de integridade do revestimento.

4.1.2 Energia do Sinal

A Figura 25 apresenta os resultados, avaliando a hipótese de que regiões íntegras, devido à maior atenuação, apresentam menor energia integrada.

Figura 25. Gráfico *boxplot* – Energia do sinal das diferentes regiões dos CP1 e CP2.



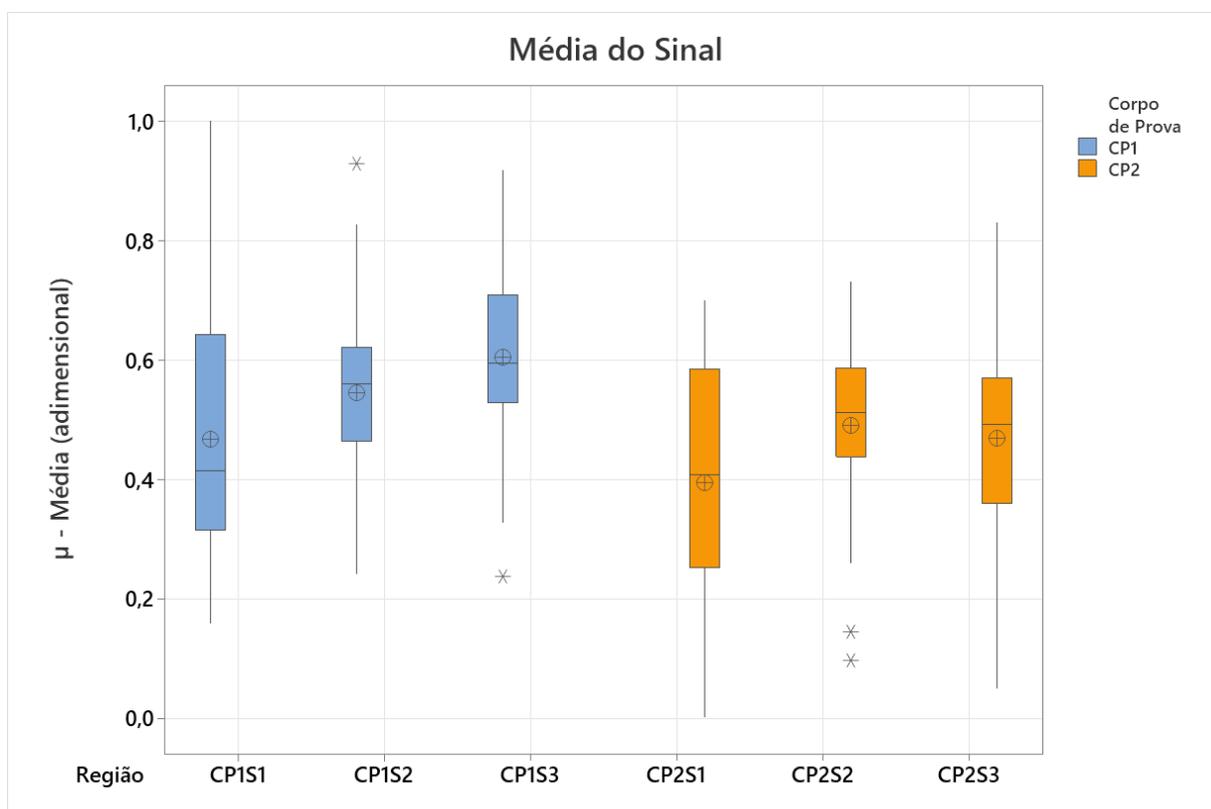
A análise da energia refletida confirmou que regiões com defeitos apresentam maior reflexão, alinhando-se à hipótese inicial de que a interface aço-água favorece esse comportamento, devido à maior diferença de impedância acústica. Contudo, a sobreposição dos valores entre as condições dificultou a distinção precisa, especialmente em CP2, onde a variabilidade foi mais expressiva. Esses resultados

indicam que, embora a energia refletida seja um parâmetro relevante, sua utilização como métrica isolada é insuficiente.

4.1.3 Média do Sinal

A Figura 26 apresenta a distribuição das médias normalizadas do sinal para as três regiões avaliadas em ambos os corpos de prova (CP1 e CP2).

Figura 26. Gráfico *boxplot* – Média do sinal das diferentes regiões dos CP1 e CP2.

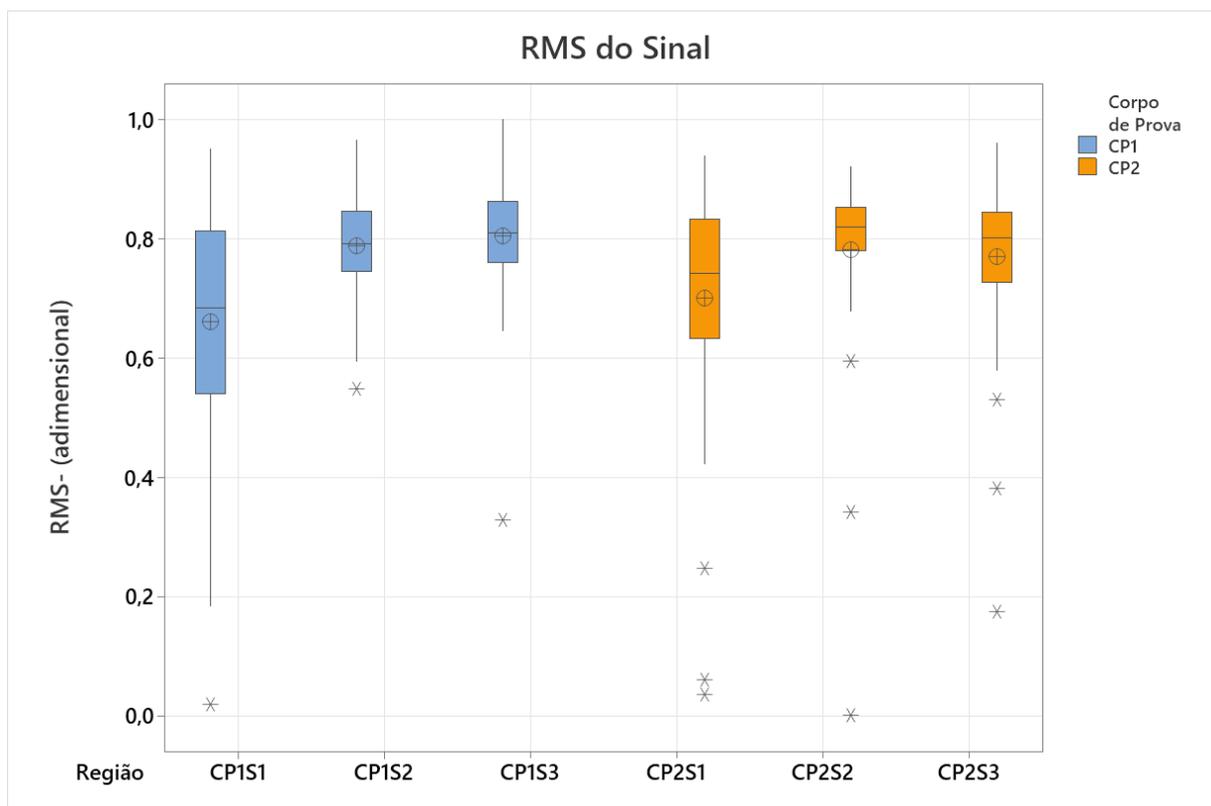


A análise da média normalizada do sinal mostrou que as regiões com defeitos apresentaram valores médios superiores às íntegras, indicando menor atenuação ultrassônica e uma separação mais clara entre as regiões no CP1. Contudo, em CP2, as diferenças entre as condições foram menos expressivas, e a maior dispersão nos dados resultou em significativa sobreposição de valores entre regiões íntegras e com defeitos. Essa limitação evidenciou que, embora útil para identificar tendências gerais, a média normalizada do sinal não é suficientemente para diferenciar de forma confiável as condições quando utilizada como métrica isolada.

4.1.4 RMS (*Root Mean Square*) do Sinal

A Figura 27 apresenta a distribuição dos valores RMS dos sinais ultrassônicos para as três regiões avaliadas em ambos os corpos de prova (CP1 e CP2).

Figura 27. Gráfico *boxplot* – RMS do sinal das diferentes regiões dos CP1 e CP2.

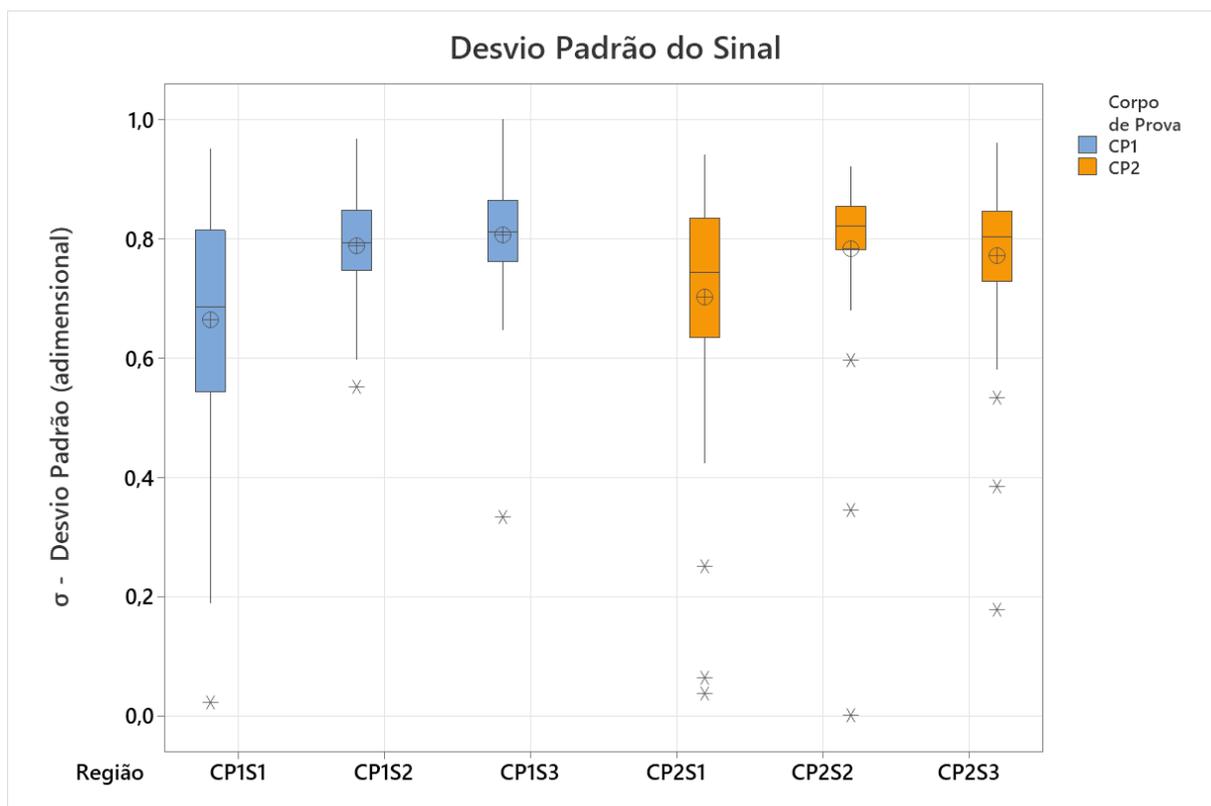


A análise dos valores RMS destacou diferenças sutis entre regiões íntegras e com defeitos, especialmente no CP1, onde as regiões com defeitos apresentaram valores médios de RMS mais elevados e menor dispersão. Isso sugere que as interfaces defeituosas, como a interface aço-água, refletem mais consistentemente o sinal ultrassônico. Em contrapartida, as regiões íntegras exibiram valores de RMS mais baixos, com maior variabilidade, refletindo a absorção irregular do sinal pelo concreto. No CP2, as diferenças foram menos evidentes, e a sobreposição significativa dos intervalos comprometeu a distinção clara entre as condições. Assim, o RMS demonstrou utilidade na identificação de tendências gerais, mas sua aplicação como métrica isolada para distinção de regiões íntegras e com defeitos é limitada.

4.1.5 Desvio Padrão do Sinal

A Figura 28 apresenta a distribuição do desvio padrão dos sinais ultrassônicos para as três regiões analisadas em ambos os corpos de prova (CP1 e CP2).

Figura 28. Gráfico *boxplot* – Desvio padrão do sinal das diferentes regiões dos CP1 e CP2.



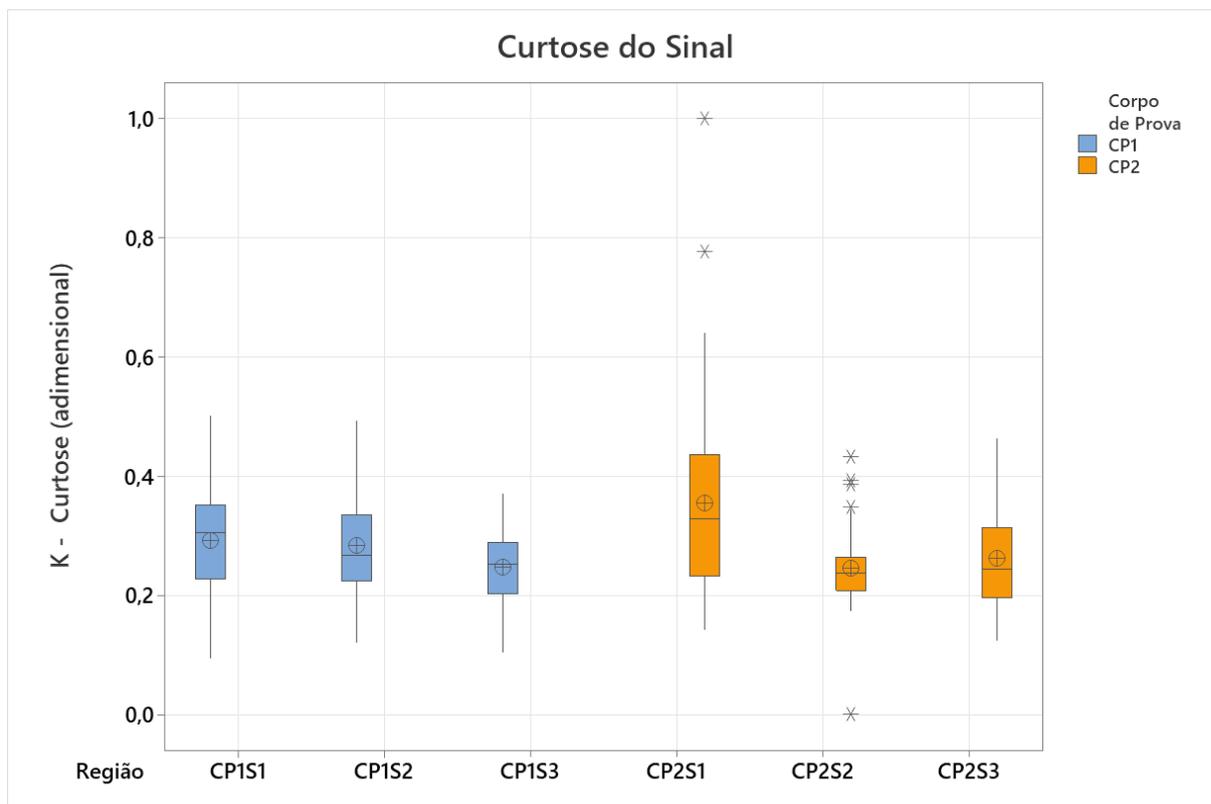
A análise do desvio padrão dos sinais ultrassônicos mostrou que, em ambos os corpos de prova, as regiões com defeitos apresentaram valores mais altos e menor dispersão, indicando menor atenuação e maior consistência na reflexão do sinal.

Em contrapartida, as regiões íntegras exibiram desvios padrão mais baixos, com maior variabilidade, devido à maior absorção e dispersão do sinal na interface aço-concreto. Apesar dessas tendências, a presença de *outliers* e a sobreposição dos valores entre as condições dificultaram a distinção clara entre regiões íntegras e defeituosas. Assim, o desvio padrão, embora tenha apresentado alguma capacidade de identificar características gerais dos sinais, não se mostrou suficientemente consistente como métrica isolada para distinção confiável das condições estruturais.

4.1.6 Curtose do Sinal

A Figura 29 apresenta a distribuição dos valores de curtose dos sinais ultrassônicos para as três regiões avaliadas em ambos os corpos de prova (CP1 e CP2).

Figura 29. Gráfico *boxplot* – Curtose do sinal das diferentes regiões dos CP1 e CP2.



A análise da curtose mostrou-se ineficaz para distinguir de forma clara as regiões íntegras das com falhas nos revestimentos em ambos os corpos de prova. Apesar de as regiões íntegras apresentarem médias de curtose mais altas e maior dispersão, indicando uma maior concentração de picos e variabilidade associada à atenuação do sinal, a significativa sobreposição dos valores entre as condições comprometeu a confiabilidade da distinção.

Essa sobreposição, amplificada pela variabilidade intraclasse, evidenciou as limitações da curtose como métrica para a identificação precisa das condições estruturais do revestimento.

4.1.7 Análise Consolidada e Limitações dos Parâmetros Ultrassônicos

A Tabela 4 apresenta uma visão consolidada dos resultados obtidos para os diferentes parâmetros analisados nos corpos de prova CP1 e CP2.

Tabela 4. Resultados consolidados dos parâmetros ultrassônicos

| Parâmetro | CP1 Tendência Observada | CP2 Tendência Observada | Eficácia na Distinção |
|--------------------------|--|---|--|
| Coeficiente de Atenuação | Defeitos com maior atenuação do que as regiões íntegras, contrariando a hipótese inicial | Defeitos com menor atenuação, alinhando-se à hipótese inicial | Hipótese não confirmada. Inconclusiva. |
| Energia Refletida | Defeitos com maior reflexão, alinhando-se à hipótese inicial | Defeitos com maior reflexão, mas com sobreposição entre condições | Hipótese confirmada. Mas baixa eficácia, limitada pela sobreposição dos valores |
| Média do Sinal | Defeitos com valores médios superiores às íntegras e boa separação | Diferenças menos evidentes, com maior dispersão nos dados | Hipótese confirmada. Mas baixa eficácia, limitada pela sobreposição dos valores |
| RMS do Sinal | Defeitos com valores médios mais elevados e menor dispersão | Diferenças sutis com sobreposição entre condições | Hipótese confirmada. Eficácia moderada no CP1, baixa no CP2 |
| Desvio Padrão | Defeitos com maior desvio padrão e menor dispersão | Diferenças menos claras, limitadas por outliers | Hipótese confirmada. Eficácia baixa no CP2 |
| Curtose | Regiões íntegras com médias mais altas e maior dispersão, mas sem clara distinção | Sem diferenças marcantes entre condições | Inconclusiva |

A análise dos parâmetros avaliados demonstrou limitações significativas para diferenciar de forma clara e confiável as regiões íntegras das com defeitos. Embora alguns parâmetros, como a energia refletida e o RMS, tenham apresentado tendências alinhadas à hipótese inicial, a sobreposição dos valores entre as condições e a variabilidade intraclasse comprometeram a eficácia das métricas isoladamente. Assim, os resultados indicam que nenhum dos parâmetros analisados é adequado como ferramenta única para distinção entre regiões íntegras e defeituosas.

4.1.8 Desempenho do SVM Linear com as Características Isoladas

Os resultados apresentados nesta seção avaliam a capacidade do classificador SVM Linear de identificar áreas íntegras e com defeitos no CP2, utilizando parâmetros extraídos no domínio do tempo. A validação cruzada foi aplicada e 20% dos dados foram reservados para teste. As métricas consideradas foram a acurácia geral, a sensibilidade e a especificidade, ambas avaliadas em relação à S1 (área íntegra).

Conforme mostrado na Tabela 5, nenhum parâmetro isolado conseguiu classificar com precisão as áreas analisadas, mesmo com o aprendizado de máquina.

Tabela 5. Desempenho SVM Linear com as Característica no Domínio do Tempo no CP2.

| Parâmetro | Acurácia Validação (%) | Acurácia Teste (%) | Sensibilidade Região S1 (%) | Especificidade Região S1 (%) |
|------------------|---------------------------------------|-----------------------------------|--|---|
| Atenuação | 52,5 | 56,7 | 70,0 | 100,0 |
| Curtose | 48,3 | 46,7 | 60,0 | 90,0 |
| Desvio Padrão | 40,0 | 30,0 | 0,0 | 90,0 |
| RMS | 38,3 | 43,3 | 40,0 | 85,0 |
| Média | 37,5 | 40,0 | 30,0 | 90,0 |
| Energia | 37,5 | 33,3 | 30,0 | 75,0 |

O melhor desempenho foi obtido com o parâmetro Atenuação, com acurácia de 56,7% no teste, mas sua sensibilidade limitada (70,0%) resultou em falsos negativos significativos. Outros parâmetros, como Curtose, Desvio Padrão, RMS, Média e

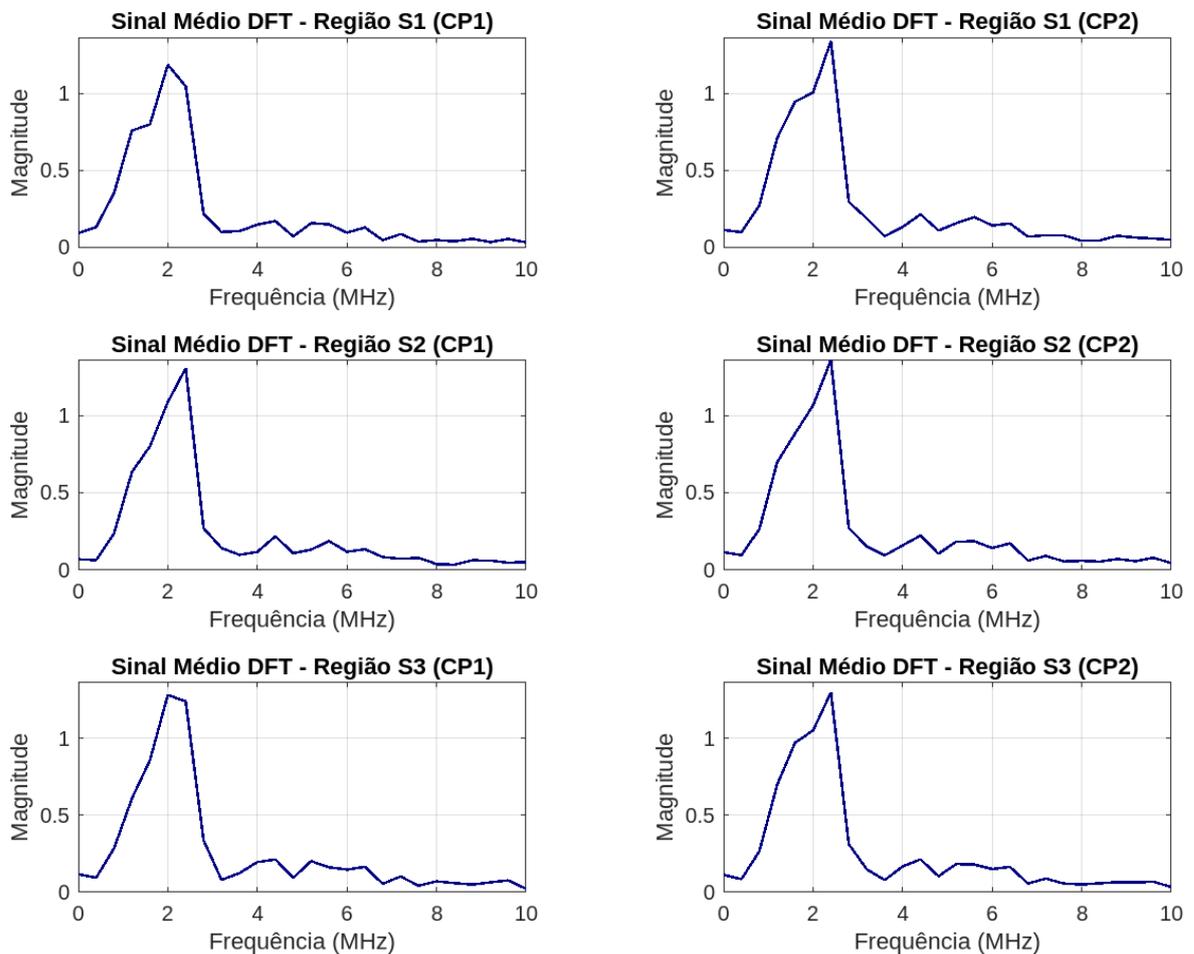
Energia, apresentaram acurácias e sensibilidades ainda mais baixas, confirmando que a análise isolada de características não é suficiente para atender aos objetivos do trabalho.

Os resultados destacam as limitações dessa abordagem e a necessidade de estratégias mais consistentes para melhorar a separação entre as classes.

4.2 Análise Espectral e Extração de Parâmetros no Domínio da Frequência

A análise dos espectros de frequência teve como objetivo caracterizar o comportamento das regiões íntegras e com defeitos dos corpos de prova, utilizando a DFT. A Figura 30 apresenta os espectros médios de frequência das regiões analisadas nos corpos de prova CP1 e CP2. A análise dos parâmetros revelou diferenças visualmente sutis entre as regiões íntegras e aquelas com defeitos, indicando limitações na identificação apenas pela avaliação visual dos espectros.

Figura 30. Espectros Médios de Frequência nas Regiões de CP1 e CP2.



A partir dos espectros, foram extraídos parâmetros como amplitude máxima, largura de banda, área sob a curva e frequência dominante, conforme apresentado na Tabela 6. Esses indicadores foram avaliados para identificar variações que permitissem diferenciar as condições dos revestimentos e, assim, auxiliar na detecção de defeitos. A análise dos espectros de frequência dos corpos de prova CP1 e CP2 demonstrou que, ao utilizar os parâmetros amplitude máxima, largura de banda, área sob a curva e frequência dominante, as diferenças observadas entre regiões íntegras e com defeitos não foram suficientemente para uma distinção clara e consistente.

Tabela 6. Parâmetros Extraídos dos Espectros de Frequência para CP1 e CP2.

| Cadeia de Processamento | Amplitude Máxima (u.a.) | Largura de Banda (MHz) | Área Sob a Curva (u.a.) | Frequência Dominante (MHz) |
|--------------------------------|--------------------------------|-------------------------------|--------------------------------|-----------------------------------|
| CP1S1 | 1,2 | 0,4 | 2,6 | 2,0 |
| CP1S2 | 1,3 | 0,4 | 2,7 | 2,4 |
| CP1S3 | 1,3 | 0,4 | 2,9 | 2,0 |
| CP2S1 | 1,3 | 0,4 | 2,8 | 2,4 |
| CP2S2 | 1,4 | 0,4 | 2,9 | 2,4 |
| CP2S3 | 1,3 | 0,8 | 2,9 | 2,4 |

A análise buscou identificar diferenças entre as áreas íntegras e defeituosas dos revestimentos. A área sob a curva foi o único parâmetro que apresentou uma leve distinção entre S1 e as demais regiões (S2 e S3) em ambos os corpos de prova (CP1 e CP2), embora com variações sutis. Os demais parâmetros, como amplitude máxima, frequência dominante e largura de banda, apresentaram valores muito próximos entre as regiões analisadas, reforçando a dificuldade em discriminar condições diferentes dos revestimentos apenas com base nesses indicadores.

Esses resultados ressaltam a necessidade de integrar os dados espectrais a técnicas de aprendizado de máquina para melhorar a classificação e identificação de falhas nos revestimentos internos. Para isso, etapas como a redução de dimensionalidade utilizando PCA e a LDA serão aplicadas nas transformadas espectrais, permitindo destacar as características mais relevantes dos dados e facilitar a separação entre as classes.

4.3 SVM Linear com Diferentes Métodos de Processamento no CP2

Os resultados apresentados na Tabela 7 destacam o desempenho do classificador SVM Linear utilizando diferentes cadeias de processamento para os sinais coletados no CP2, em ordem decrescente de acurácia de teste. Os métodos analisados incluem transformadas DCT e DFT, a consolidação das características dos sinais no domínio do tempo (TIME) e combinações dessas transformadas (DFT-DCT), associadas às técnicas PCA e LDA. A avaliação considerou a média e o desvio padrão das acurácias de validação e teste, com 20% dos dados reservados para a etapa de teste.

Tabela 7. Desempenho SVM Linear com diferentes métodos de processamento dos sinais CP2.

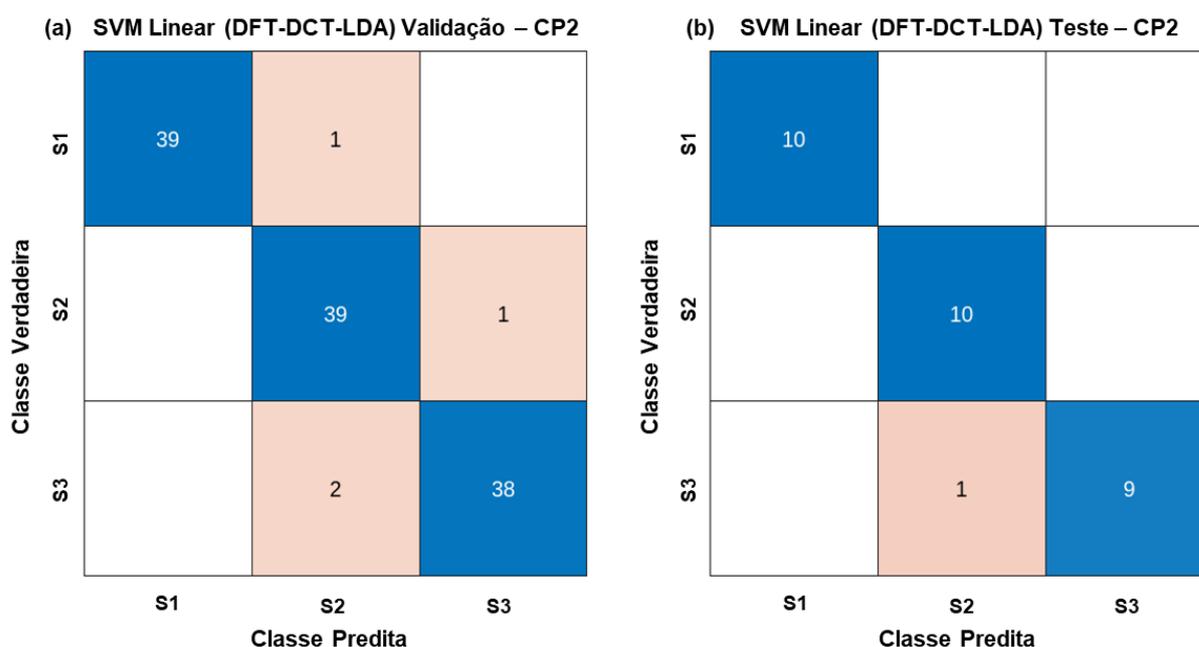
| # | Cadeia de Processamento | Média Acurácia % (Validação) | Desvio Padrão Acurácia % (Validação) | Média Acurácia % (Teste) | Desvio Padrão Acurácia % (Teste) |
|----|-------------------------|------------------------------|--------------------------------------|--------------------------|----------------------------------|
| 1 | DFT-DCT-LDA | 97,2 | 1,0 | 96,0 | 2,5 |
| 2 | DFT-DCT-PCA-LDA | 97,3 | 1,0 | 95,3 | 1,6 |
| 3 | DCT-PCA-LDA | 90,7 | 0,6 | 92,0 | 3,4 |
| 4 | DFT-PCA-LDA | 85,7 | 1,0 | 91,3 | 4,5 |
| 5 | DCT-LDA | 90,5 | 2,7 | 88,0 | 3,4 |
| 6 | DFT-LDA | 88,2 | 1,1 | 87,3 | 4,4 |
| 7 | DCT-PCA | 67,2 | 5,6 | 66,7 | 7,6 |
| 8 | DFT-PCA | 65,6 | 3,9 | 64,0 | 9,3 |
| 9 | TIME-PCA-LDA | 58,7 | 1,7 | 60,7 | 8,3 |
| 10 | TIME-LDA | 57,2 | 2,7 | 60,0 | 5,2 |
| 11 | TIME-PCA | 60,4 | 4,4 | 56,7 | 7,6 |
| 12 | DFT-DCT-PCA | 53,3 | 3,7 | 49,3 | 8,5 |

Os resultados evidenciam que a combinação das transformadas DFT e DCT com a LDA potencializa a separação entre as classes, otimizando a classificação de regiões íntegras e defeituosas.

Os resultados indicam que a combinação das transformadas DFT e DCT supera o uso isolado de cada uma. Esse achado corrobora estudos prévios, como Raghu et al. (2017) e Hamidi et al. (2018), que demonstraram que a fusão dessas transformadas melhora a robustez do modelo e a extração de padrões espectrais. No presente estudo, essa abordagem possibilitou um melhor aproveitamento das informações presentes nos sinais ultrassônicos, otimizando a separação entre regiões íntegras e defeituosas.

A aplicação da LDA mostrou-se eficiente na redução de variabilidade intra-classe e no aumento da discriminação interclasse, refletindo diretamente no desempenho do modelo. A Figura 31 apresenta o desempenho consistente da cadeia DFT-DCT-LDA, com elevada precisão tanto na etapa de validação (Figura 31a), quanto no teste (Figura 31b).

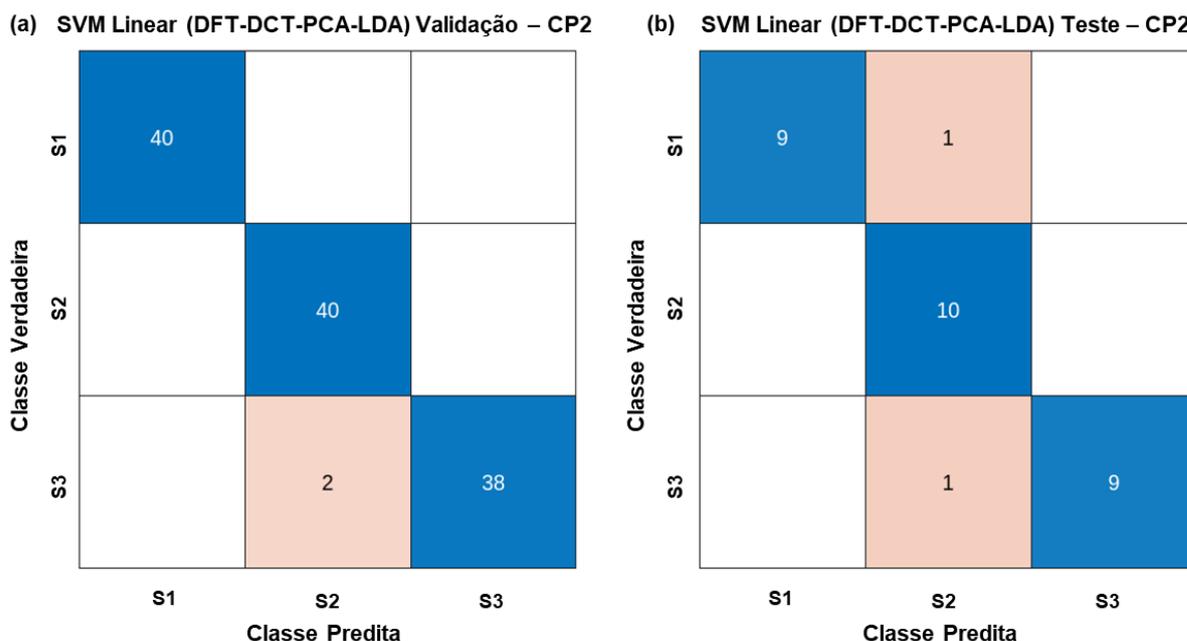
Figura 31. Matrizes de confusão do SVM Linear com a cadeia DFT-DCT-LDA nas etapas de (a) validação e (b) teste com dados do CP2.



A segunda melhor cadeia, DFT-DCT-PCA-LDA, também apresentou resultados expressivos, alcançando 95,3% de acurácia no teste. A aplicação do PCA antes da LDA auxiliou no alinhamento das escalas e na organização dos dados em componentes principais, sem realizar uma seleção específica dos componentes mais relevantes. Esse procedimento facilitou a posterior aplicação da LDA, otimizando a separação entre as classes.

As matrizes de confusão dessa cadeia, mostradas na Figura 32, indicam um desempenho consistente, com apenas dois erros na validação (Figura 32a) e dois no teste (Figura 32b), evidenciando sua eficiência e capacidade de generalização em diferentes conjuntos de dados.

Figura 32. Matrizes de confusão do SVM Linear com a cadeias DFT-DCT-PCA-LDA nas etapas de (a) validação e (b) teste com dados do CP2.



As cadeias de processamento DCT-PCA-LDA e DFT-PCA-LDA demonstraram um desempenho sólido, alcançando acurácias de teste de 92,0% e 91,3%, respectivamente. As cadeias DCT-LDA e DFT-LDA também apresentaram desempenhos consistentes, com acurácias de 88,0% e 87,3%, respectivamente. Apesar de não incluírem a etapa de PCA, a LDA isolada demonstrou boa capacidade de proporcionar a separação entre as classes. Embora apresentem um desempenho ligeiramente inferior às cadeias que aplicam PCA antes da LDA, essas abordagens ainda são viáveis. No entanto, a necessidade de normalização pós-LDA torna essencial um ajuste adicional para garantir uniformização dos dados e a separação eficaz entre as classes.

Em contraste, cadeias que usaram apenas os dados extraídos no domínio do tempo ou o PCA apresentaram desempenhos inferiores na classificação, evidenciando as limitações dessa técnica quando aplicada isoladamente no contexto

deste estudo. Conforme discutido por Jolliffe (2002), o PCA prioriza a maximização da variância total dos dados, sem considerar diretamente a separação entre as classes.

As figuras abaixo apresentam a distinção dos resultados pelos métodos PCA e LDA. O PCA (Figura 33) prioriza a maximização da variância total dos dados, independentemente da separação entre as classes, enquanto o LDA (Figura 34) foca na maximização de separação entre grupos, validando conceitos discutidos por Jombart, Devillard e Balloux (2010).

Figura 33. Gráfico de dispersão (*Scatter Plot*) DFT-PCA.

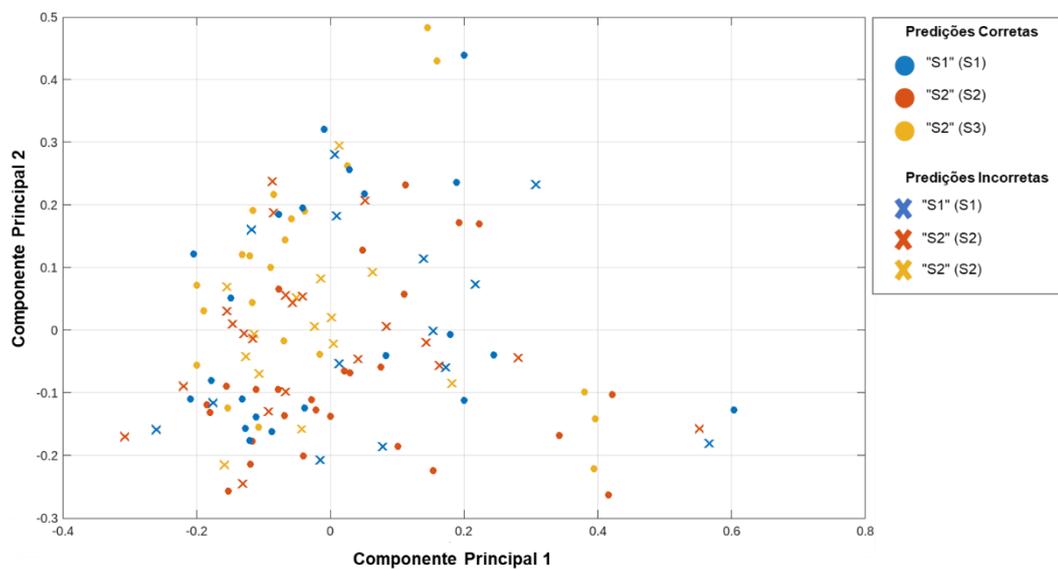
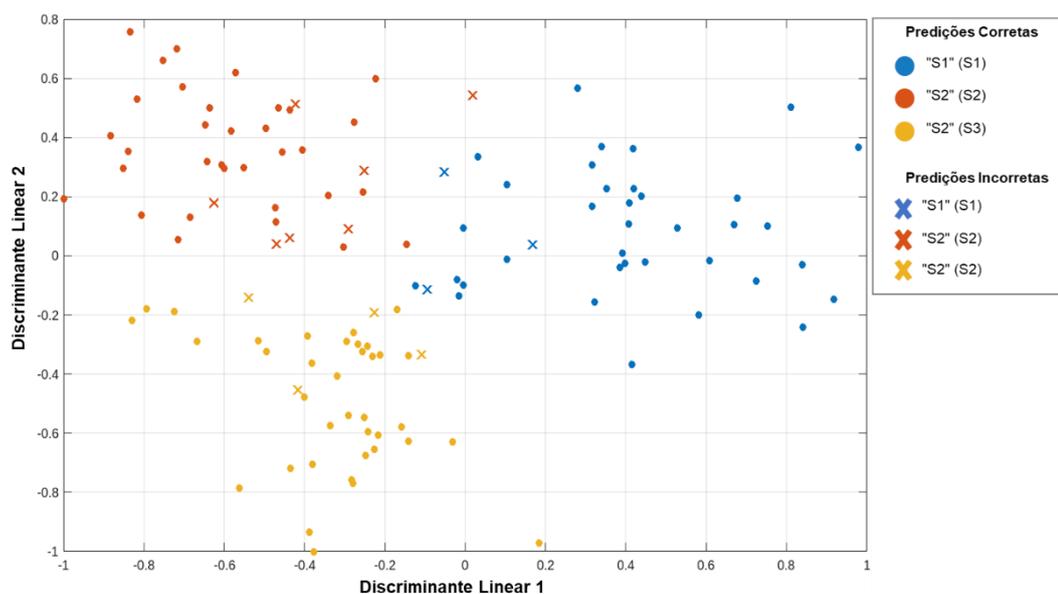


Figura 34. Gráfico de dispersão (*Scatter Plot*) DFT-LDA.



4.4 Generalização do Modelo com Dados do CP1

Nesta seção, são apresentados os resultados do teste das cadeias de processamento no CP1, avaliando a capacidade do modelo de generalizar para novos dados. O teste utilizou o modelo previamente treinado e validado no CP2, aplicando as mesmas cadeias de processamento e configurações. O desempenho do SVM Linear com os dados do CP1 está detalhado na Tabela 8. As métricas consideradas para essa análise incluem:

- **Acurácia Geral do Modelo:** A proporção total de classificações corretas realizadas pelo modelo.
- **Sensibilidade S1:** A capacidade do modelo de identificar corretamente as áreas íntegras (Classe 1) no segundo corpo de prova.
- **Especificidade S1:** A capacidade do modelo de identificar corretamente as áreas defeituosas como não pertencentes à Classe 1.

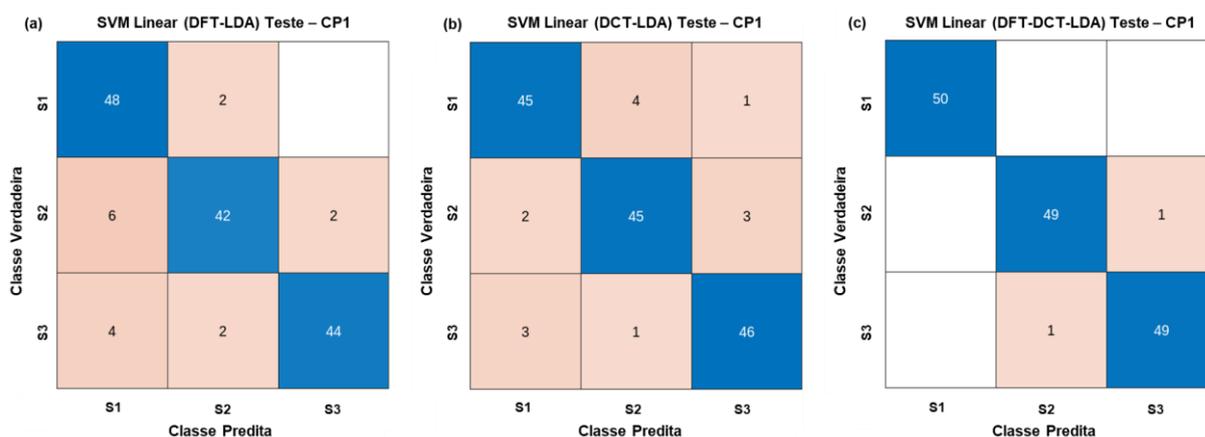
Tabela 8. Desempenho SVM Linear com Dados do CP1.

| # | Cadeia de Processamento | Acurácia (%) | Sensibilidade Classe 1 (%) | Especificidade Classe 1 (%) |
|----|-------------------------|--------------|----------------------------|-----------------------------|
| 1 | DFT-DCT-LDA | 98,7 | 100,0 | 100,0 |
| 2 | DFT-DCT-PCA-LDA | 98,0 | 98,0 | 100,0 |
| 3 | DCT-LDA | 90,7 | 90,0 | 95,0 |
| 4 | DFT-PCA-LDA | 90,7 | 94,0 | 94,0 |
| 5 | DCT-PCA-LDA | 90,0 | 84,0 | 98,0 |
| 6 | DFT-LDA | 89,3 | 96,0 | 90,0 |
| 7 | TIME-PCA-LDA | 63,3 | 74,0 | 80,0 |
| 8 | TIME-LDA | 54,7 | 94,0 | 43,0 |
| 9 | DCT-PCA | 45,3 | 48,0 | 66,0 |
| 10 | TIME-PCA | 37,3 | 40,0 | 64,0 |
| 11 | DFT-PCA | 31,3 | 30,0 | 70,0 |
| 12 | DFT-DCT-PCA | 28,7 | 28,0 | 60,0 |

Os métodos baseados em transformações no domínio da frequência continuaram demonstrando alta eficácia na classificação com os dados do CP1, consolidando-se como abordagens consistentes e confiáveis. A cadeia DFT-DCT-LDA destacou-se como a de melhor desempenho, alcançando uma acurácia de 98,7%, com sensibilidade e especificidade de 100,0%. Esses resultados evidenciam a capacidade da abordagem em identificar com precisão regiões íntegras e defeituosas, mesmo em um conjunto de dados não utilizado durante o treinamento, conforme mostrado na Figura 35c.

Além disso, as cadeias DFT-LDA e DCT-LDA também apresentaram desempenhos consistentes, com acurácias de 89,3% e 90,7%, respectivamente. Embora não tenham atingido o mesmo nível de eficácia da DFT-DCT-LDA, ambas confirmam a importância da LDA na maximização da separabilidade entre as classes. As Figura 35a e Figura 35b apresentam as matrizes de confusão para as cadeias DFT-LDA e DCT-LDA, respectivamente, demonstrando uma classificação eficiente com poucos erros, especialmente nas categorias de regiões íntegras.

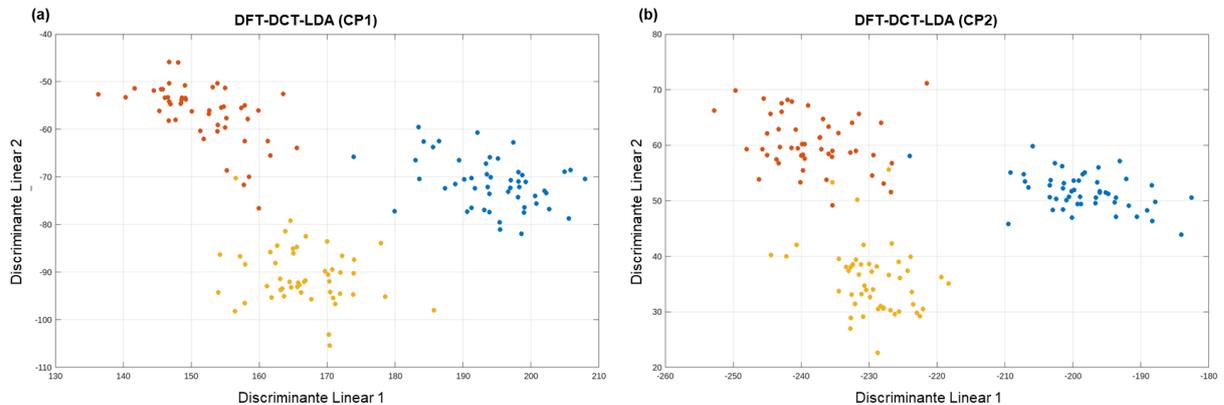
Figura 35. Matrizes de confusão do SVM Linear com as cadeias (a) DFT-LDA, (b) DCT-LDA e (c) DFT-DCT-LDA no teste com dados do CP1.



Importante destacar que os resultados iniciais da aplicação da LDA sem o PCA demonstraram desempenho insatisfatório no teste com os dados do CP1, com todas as amostras sendo atribuídas a uma única classe, comprometendo a generalização do modelo. Para investigar as razões desse comportamento, foram avaliados os gráficos de dispersão dos discriminantes lineares (LDAs) 1 e 2 para os dados do CP1 (Figura 36a) e CP2 (Figura 36b). Apesar de visualmente separados em ambos os

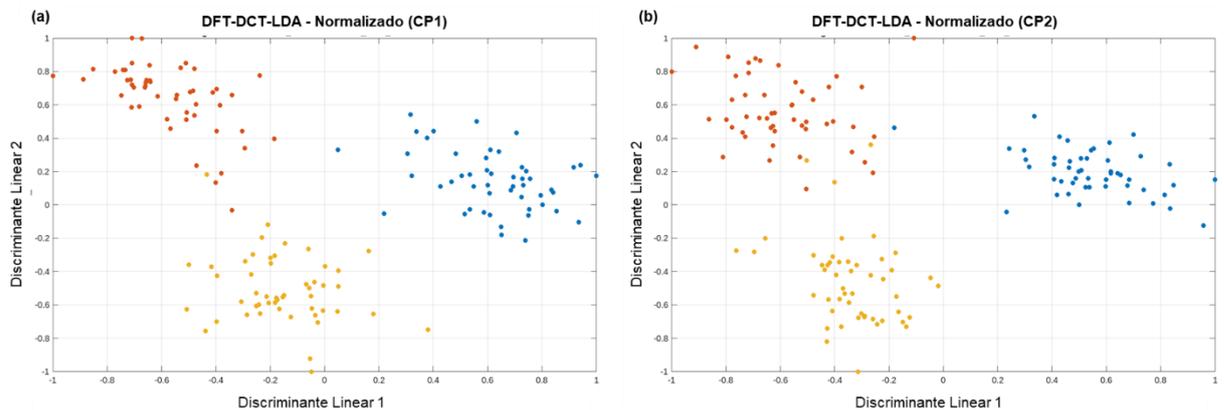
casos, os valores das escalas dos discriminantes apresentaram discrepâncias entre os dois conjuntos, sugerindo uma incompatibilidade entre os dados projetados.

Figura 36. Gráficos de dispersão dos discriminantes lineares gerados pelas cadeias (a) DFT-DCT-LDA sem normalização para o (a) CP1 e (b) CP2.



Para contornar esse problema, foi implementada a normalização dos discriminantes lineares após a aplicação da LDA, visando manter a compatibilidade das escalas entre os conjuntos CP1 (Figura 37a) e CP2 (Figura 37b).

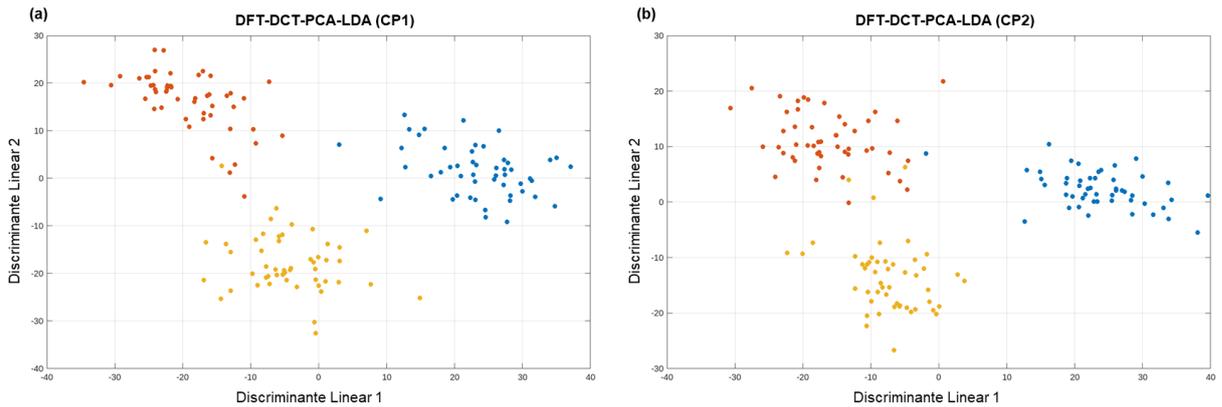
Figura 37. Gráficos de dispersão dos discriminantes lineares gerados pelas cadeias DFT-DCT-LDA com normalização dos dados para o (a) CP1 e (b) CP2.



A normalização ajustou os valores dos discriminantes para um intervalo uniforme. Os resultados pós-normalização indicaram uma melhoria significativa no desempenho, conforme demonstrado na Figura 35 e Tabela 8. Esse ajuste reforça a importância da normalização como uma etapa complementar para o alinhamento das escalas em cenários onde o PCA não é utilizado antes do LDA.

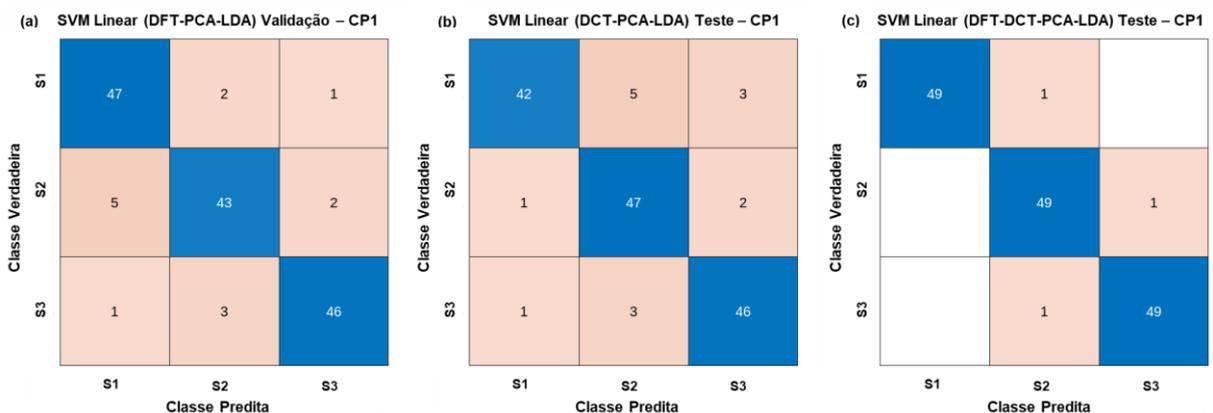
A Figura 38 apresenta os discriminantes lineares para a combinação DFT-DCT-PCA-LDA nos dados do CP1 (Figura 38a) e CP2 (Figura 38b). Nestes gráficos, nota-se que o PCA prévio ajustou as escalas dos dados, garantindo a compatibilidade entre os conjuntos sem a necessidade de uma etapa adicional de normalização.

Figura 38. Gráficos de dispersão dos discriminantes lineares gerados pelas cadeias DFT-DCT-PCA-LDA para (a) CP1 e (b) CP2.



A Figura 39 apresenta os resultados obtidos nas matrizes de confusão para as cadeias de processamento (a) DFT-PCA-LDA, (b) DCT-PCA-LDA e (c) DFT-DCT-PCA-LDA, confirmando o desempenho consistente desses métodos no teste com os dados do CP1. A cadeia DFT-DCT-PCA-LDA apresentou o melhor desempenho, com uma acurácia de 98,0%, evidenciando sua capacidade de generalização ao combinar duas transformadas espectrais. A cadeia DFT-PCA-LDA também demonstrou bons resultados, alcançando 90,7% de acurácia. Por sua vez, a cadeia DCT-PCA-LDA obteve 90,0% de acurácia, com desempenho ligeiramente inferior.

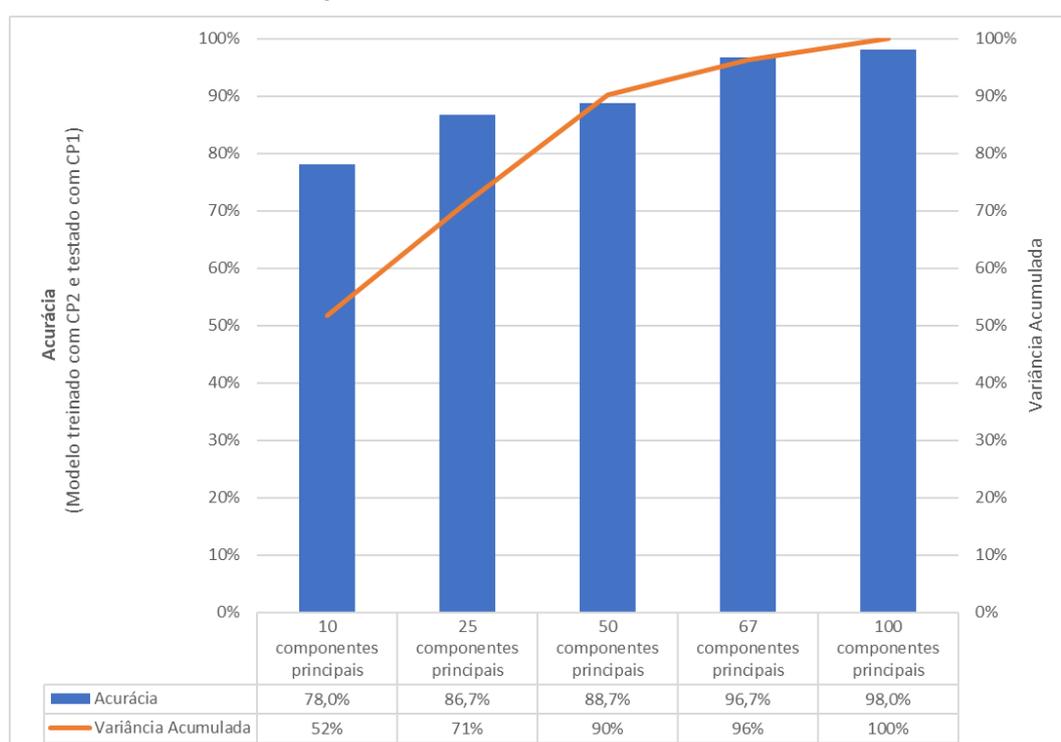
Figura 39. Matrizes de confusão do SVM Linear com as cadeias (a) DFT-PCA-LDA, (b) DCT-PCA-LDA e (c) DFT-DCT-PCA-LDA no teste com dados do CP1.



A PCA foi utilizada neste trabalho como uma etapa intermediária de pré-processamento antes da aplicação da Análise Discriminante Linear (LDA), conforme sugerido no trabalho de Jombart, Devillard e Balloux (2010). Ainda assim, avaliou-se o impacto da seleção de subconjuntos de componentes principais sobre o desempenho do classificador.

A Figura 40 apresenta o desempenho do classificador SVM linear, treinado com os dados do corpo de prova CP2 e testado com CP1, em função do número de componentes principais utilizados antes da aplicação da LDA. Para cada configuração, também é apresentada a variância acumulada explicada pelos componentes selecionados.

Figura 40. Efeito da redução de dimensionalidade via PCA sobre a acurácia do modelo SVM Linear para a cadeia DFT-DCT-PCA-LDA.



Observa-se que, embora a acurácia máxima (98,0%) tenha sido alcançada com a utilização dos 100 componentes principais, resultados muito próximos (96,7%) foram obtidos com apenas 67 componentes, que explicam 96% da variância dos dados. Esse comportamento evidencia que o uso de todos os componentes principais não é condição necessária para o bom desempenho da cadeia PCA-LDA.

Os testes realizados no CP1 demonstraram que as cadeias de processamento baseadas em transformadas espectrais apresentaram desempenho superior na classificação das condições de integridade dos revestimentos. As abordagens DFT-DCT-LDA e DFT-DCT-PCA-LDA destacaram-se como as mais eficazes, evidenciando sua capacidade de generalização para um conjunto de dados não utilizado no treinamento.

A combinação das transformadas DFT e DCT já foi explorada em diferentes aplicações, demonstrando sua eficácia na extração de características relevantes e na melhoria da separabilidade entre classes. Estudos prévios apontam que essa abordagem permite capturar informações complementares do sinal: enquanto a DFT preserva a distribuição global das frequências, a DCT melhora a compactação da informação e a robustez contra ruídos e variações locais, conforme apresentado por Raghu et al. (2017) e Hamidi et al. (2018). Esses achados foram igualmente observados neste estudo, onde a integração DFT-DCT possibilitou uma representação mais discriminativa dos dados.

Os testes com LDA sem normalização revelaram inconsistências devido a discrepâncias nas escalas dos discriminantes lineares entre CP1 e CP2, comprometendo a generalização do modelo. A implementação da normalização pós-LDA se mostrou uma solução eficiente, garantindo compatibilidade entre os conjuntos de dados e restaurando a capacidade do modelo de classificação.

Por outro lado, a integração do PCA antes da LDA demonstrou ser uma estratégia ainda mais consistente para o alinhamento das escalas, eliminando a necessidade de normalização adicional e proporcionando resultados consistentes. Esses achados corroboram a importância da estratégia de Análise Discriminante de Componentes Principais (DAPC) proposta por Jombart, Devillard e Balloux (2010), que combina PCA e LDA para maximizar a separabilidade entre grupos e garantir estabilidade nas projeções discriminantes.

Portanto, a análise corrobora que as combinações DFT-DCT-LDA e DFT-DCT-PCA-LDA, dentre as cadeias de processamento exploradas neste estudo, foram as estratégias mais adequadas para uma metodologia prática para detecção de defeitos em revestimentos internos de argamassa de tubulações industriais, garantindo alta acurácia, consistência entre diferentes conjuntos de dados e menor necessidade de ajustes adicionais.

5 CONCLUSÕES

Com o objetivo de investigar a aplicação de ensaios ultrassônicos e aprendizado de máquina na inspeção de revestimentos internos feitos de argamassa em tubulações de aço carbono, com foco na identificação e classificação de defeitos que possam comprometer a integridade estrutural e a segurança operacional, foi possível concluir que:

A análise espectral nos domínios do tempo e da frequência não foi capaz de distinguir com clareza as regiões íntegras e defeituosas. Parâmetros como largura de banda, amplitude máxima, frequência dominante e área sob a curva apresentaram valores próximos entre as regiões, dificultando a separação das classes. No domínio do tempo, os parâmetros avaliados demonstraram limitações significativas para diferenciar de forma clara e confiável as condições do revestimento. Embora, a energia refletida e o RMS, tenham exibido tendências alinhadas à hipótese inicial, a sobreposição dos valores e a variabilidade intraclasse comprometeram a distinção clara entre as diferentes regiões.

A avaliação do SVM Linear mostrou que a análise no domínio do tempo teve desempenho insatisfatório, com TIME-PCA-LDA (63,3%), TIME-LDA (54,7%) e TIME-PCA (37,3%) apresentando baixa acurácia. Embora a TIME-LDA tenha alcançado 94,0% de sensibilidade, sua especificidade reduzida (43,0%) comprometeu a confiabilidade da classificação, reforçando as limitações dessa abordagem para a distinção entre revestimentos íntegros e defeituosos.

Por outro lado, as cadeias baseadas no domínio da frequência demonstraram resultados significativamente superiores. A combinação DFT-DCT-LDA foi a mais eficiente, atingindo 98,7% de acurácia com os dados de CP1, com 100% de sensibilidade e especificidade para a região íntegra (S1), destacando-se como a abordagem mais promissora. Da mesma forma, a cadeia DFT-DCT-PCA-LDA apresentou desempenho robusto, com 98,0% de acurácia, mantendo 98,0% de sensibilidade e 100% de especificidade. Os resultados evidenciam que a combinação das transformadas DFT e DCT potencializa a separação entre as classes, otimizando a classificação de regiões íntegras e defeituosas.

Os testes com LDA sem normalização revelaram inconsistências devido a discrepâncias nas escalas entre CP1 e CP2, comprometendo a generalização do

modelo. A normalização pós-LDA mostrou-se uma solução eficiente, garantindo compatibilidade entre os conjuntos de dados. No entanto, a integração da PCA antes da LDA provou ser a estratégia mais eficiente, eliminando a necessidade de normalização adicional e assegurando maior confiabilidade na classificação.

Os resultados obtidos reforçam o potencial da combinação de transformadas espectrais, redução de dimensionalidade e aprendizado de máquina para a inspeção não destrutiva na detecção de defeitos em revestimentos de argamassa em tubulações industriais. O estudo indica uma metodologia consistente para futuras pesquisas voltadas à aplicação em cenários industriais ainda mais complexos e para a integração com sistemas de monitoramento preditivo.

6 PERSPECTIVAS PARA TRABALHOS FUTUROS

Os achados desta pesquisa fornecem uma base para futuras investigações, contribuindo para o avanço das técnicas de inspeção não destrutiva e aprendizado de máquina na indústria. Espera-se que a metodologia desenvolvida possa ser ampliada e aplicada em diferentes contextos, aprimorando a segurança e eficiência dos processos de manutenção industrial.

Diante disso, como forma de contribuição para trabalhos futuros, sugere-se:

Aplicar o método em tubulações industriais em operação, considerando ruídos de campo e variações ambientais.

Desenvolver soluções embarcadas para inspeções contínuas e automatizadas, permitindo um monitoramento em tempo real da integridade dos revestimentos internos. Estratégias baseadas em aprendizado de máquina embarcado podem ser exploradas para reduzir a dependência de infraestrutura computacional externa e otimizar o processamento dos sinais diretamente no campo.

REFERÊNCIAS BIBLIOGRÁFICAS

ALI, T. E. A. Smart agriculture: utilizing machine learning and deep learning for drought stress identification in crops. **Scientific Reports**, v. 14, n. 1, p. 1-16, 2024.

ALPAYDIN, E. **Introduction to Machine Learning**. 2ª. ed. Cambridge: MIT Press, 2010.

API. **API 570: Piping Inspection Code – In-service Inspection, Rating, Repair, and Alteration of Piping Systems**. American Petroleum Institute. Washington. 2016.

AWWA. **Cement-mortar protective lining and coating for steel water pipe—4 in. (100 mm) and larger—Shop applied. ANSI/AWWA C205-12**. American Water Works Association. Denver. 2012.

BAHIA. **Decreto nº 14.024, de 6 de junho de 2012**. Salvador. 2012. Aprova o Regulamento da Lei nº 10.431, de 20 de dezembro de 2006, e da Lei nº 11.612, de 8 de outubro de 2009.

BONDS, R. W. Cement-mortar linings for ductile iron pipe. **Ductile Iron Pipe Research Association**, Birmingham, 2005.

BRASIL. **Resolução nº 810, de 16 de março de 2020**. Agência Nacional do Petróleo, Gás Natural e Biocombustíveis. [S.l.], p. 53. 2020. Resolução nº 810, de 16 de março de 2020. Institui a gestão de segurança operacional de terminais para movimentação e armazenamento de petróleo, derivados, gás natural e biocombustíveis nos termos do Regulamento Técnico de Terminais para Movimentação e Ar.

BRODA, M.; HAJDUK, V.; LEVICKÝ, D. Universal statistical steganalytic method. **Journal of Electrical Engineering**, v. 68, n. 2, p. 117-124, 2017.

BUNGEY, J. H.; GRANTHAM, M. G. **Testing of concrete in structures**. 4ª. ed. Oxon: Taylor & Francis e-Library, 2006.

CARELLI, J. M. **Análise do comportamento de ondas ultrassônicas em elementos fissurados de concreto e argamassa**. Universidade Federal de Santa Catarina. Florianópolis. 2014.

CARVALHO JR, E. A. **Uso de ultrassom para análise das propriedades de uma haste cobreada de aterramento**, São José dos Pinhais, 15, n. 39, 2024. 3996-4011.

CHEEKE, J. D. N. **Fundamentals and applications of ultrasonic waves**. 2ª. ed. New York: CRC press, 2012.

CRUZ, F. C. E. A. Efficient feature selection for neural network based detection of flaws in steel welded joints using ultrasound testing. **Ultrasonics**, 73, 2017.

DE SOUZA, L. P. B. E. A. **Machine learning-based cement integrity evaluation with a through-tubing logging experimental setup**, 227, 2023. 211882.

DELGADILLO, H. H. E. A. **Ultrasonic inline inspection of a cement-based drinking water pipeline**, v. 210, p. 110413, 2020.

DELGADILLO, H. H. E. A. Ultrasonic inline inspection of a cement-based drinking water pipeline. **Engineering Structures**, 2020. 110413.

EMERY, J. **An Investigation of Concrete Porosity and Material Interface Interactions Using High Frequency Ultrasonic Attenuation**. University of Maine. Orono, p. 73. 2024.

FREITAS, V. L. D. A. E. A. **Avaliação da potencialidade das técnicas de ultrassom na caracterização de aços AISI e aço inoxidável duplex**. Universidade Federal de Campina Grande. Campina Grande, p. 81. 2009.

GRAF, R.; ZELDOVICH, M.; FRIEDRICH, S. Comparing linear discriminant analysis and supervised learning algorithms for binary classification—A method comparison study. **Biometrical Journal**, v. 66, n. 1, p. 2200098, 2024.

GUNTURI, S. K.; SARKAR, D. **Ensemble machine learning models for the detection of energy theft**, v. 192, p. 106904, 2021.

HAMIDI, M. E. A. Hybrid blind robust image watermarking technique based on DFT-DCT and Arnold transform. **Multimedia Tools and Applications**, v. 77, p. 27181-27214, 2018.

HAYATI, R. E. A. Combination of PCA with LDA and SVM classifiers: A model for determining the geographical origin of coconut in the coastal plantation, Aceh Province, Indonesia. **Case Studies in Chemical and Environmental Engineering**, 2024. 100552.

HESPELER, S. C. E. A. **Deep Learning–Based Time-Series Classification for Robotic Inspection of Pipe Condition Using Non-Contact Ultrasonic Testing**, v. 7, n. 1, 2024.

HO, T. K. The random subspace method for constructing decision forests. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, 20, 1998. 832-844.

HOFMANN, M. D. A. **Atenuação da energia do sinal ultrassônico na detecção de danos por fissuração no concreto**. Universidade Federal de Santa Catarina. Santa Catarina, p. 140. 2015.

HUANG, J. E. A. Detection and quantification of artificial delaminations in CFRP composites using ultrasonic thermography. **Infrared Physics & Technology**, 130, 2023. 104579.

JIANG, H. **Machine Learning Fundamentals: A Concise Introduction**. Cambridge: Cambridge University Press, 2022.

JOHNSON, P. **Fundamentals of Machine Learning: An Introduction to Neural Networks**. Framingham: HiTeX Press, 2024.

JOLLIFFE, I. T. **Principal component analysis for special types of data**. New York: Springer New York, 2002.

JOMBART, T.; DEVILLARD, S.; BALLOUX, F. Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. **BMC genetics**, v. 11, p. 1-15, 2010.

KALHARA, P. G. E. A. **TreeSpirit**: Illegal logging detection and alerting system using audio identification over an IoT network. International Conference on Software, Knowledge, Information Management and Applications. [S.I.]: IEEE. 2017. p. 1-7.

KHAYAM, S. A. The discrete cosine transform (DCT): theory and application. **Michigan State University**, 114, 2003. 31.

KRISHNAMURTHY, R.; MILLIGAN, R.; SLUDER, S. **Integrity Assessment of Distribution Pipelines**. Oak Ridge National Laboratory, Buildings and Transportation Science Division. Oak Ridge. 2023.

LARI, S.; HAN, J.-W.; KWON, H. J. Evaluation of Internal Flaw Geometric Characteristics Using Artificial Neural Networks, 2024.

LATHI, B. P. **Sinais e sistemas lineares-2**. Tradução de Gustavo Guimarães Parma. 2ª. ed. Porto Alegre: Bookman, 2008.

LI, S. E. A. Linear discriminant analysis with generalized kernel constraint for robust image classification. **Pattern Recognition**, v. 136, p. 109196, 2023.

LIANG, M. E. A. **Interpretable Ensemble-Machine-Learning models for predicting creep behavior of concrete**, v. 125, p. 104295, 2022.

LIU, L. **Machine Learning-Driven Corrosion Detection and Classification in Pipelines**. University of Wales Trinity Saint David. [S.I.]. 2024.

MALHOTRA, V. M.; CARINO, N. J. **Handbook on nondestructive testing of concrete**. 2ª. ed. [S.I.]: CRC press, 2003.

MALLIPEDDI, R.; LEE, M. **Ensemble based face recognition using discriminant PCA features**. Proceedings of IEEE World Congress on Computational Intelligence. [S.I.], p. 1-8. 2012.

MATHWORKS. **DCT (Discrete Cosine Transform)**, 2024. Disponível em: <<https://www.mathworks.com/help/signal/ref/dct.html>>. Acesso em: jun abr. 2024.

MATHWORKS. **mean (MATLAB) - Função para calcular a média**, 2024. Disponível em: <<https://www.mathworks.com/help/matlab/ref/double.mean.html>>. Acesso em: 3 Jun 2024.

MATHWORKS. **var (MATLAB) - Função para calcular a variância**, 2024. Disponível em: <<https://www.mathworks.com/help/matlab/ref/double.var.html>>. Acesso em: 5 Jun 2024.

MATHWORKS. **std (MATLAB) - Função para calcular o desvio padrão**, 2024. Disponível em: <<https://www.mathworks.com/help/matlab/ref/double.std.html>>. Acesso em: 2 Jul 2024.

MATHWORKS. **rms (MATLAB) - Função para calcular o valor RMS**, 2024. Disponível em: <<https://www.mathworks.com/help/signal/ref/rms.html>>. Acesso em: 8 Jul 2024.

MATHWORKS. **kurtosis (MATLAB) - Função para calcular a curtose**, 2024. Disponível em: <<https://www.mathworks.com/help/stats/kurtosis.html>>. Acesso em: 3 Set 2024.

MATHWORKS. **Choose a Classifier**, 2024. Disponível em: <<https://www.mathworks.com/help/stats/choose-a-classifier.html>>. Acesso em: 13 dez 2024.

MOSHREFI, A.; NABKI, F. **Advanced industrial fault detection: A comparative analysis of ultrasonic signal processing and ensemble machine learning techniques**, v. 14, n. 15, p. 6397, 2024.

MUKHTI, J. A.; GUCUNSKI, N.; KEE, S.-H. AI-assisted ultrasonic wave analysis for automated classification of steel corrosion-induced concrete damage. **Automation in Construction**, 2024. 105704.

NUMXL. **Principal Component Analysis (PCA) 101**, 2024. Disponível em: <<https://numxl.com/blogs/principal-component-analysis-pca-101>>. Acesso em: 31 dez 2024.

OLYMPUS, I. 39DL PLUS Thickness Gage, 2025. Disponível em: <<https://www.olympus-ims.com/pt/39dl-plus/>>. Acesso em: 7 jan 2025.

OMARZAI, F. Principal Component Analysis (PCA) In Depth. Medium, 2025. Disponível em: <<https://medium.com/@fraidoonomarzai99/principal-component-analysis-pca-in-depth-93c871f25dfa>>. Acesso em: 08 jan 2025.

ONUAHA, C. E. A. Enhancement of pipeline integrity assessment of buried unpiggable pipelines with non-contact magnetic gradient tomography method (MTM-G). **Journal of Materials and Manufacturing**, 2024. 25-37.

ORUKLU, E.; SANIIE, J. **High speed design and performance evaluation of frequency-diverse orthogonal transforms for ultrasonic imaging applications**. IEEE Symposium on Ultrasonics. [S.l.]: IEEE. 2023. p. 258-261.

PENG, H. E. A. Diffusion-based conditional wind power forecasting via channel attention.. **IET Renewable Power Generation**, 2024. 306-320.

PETROBRAS. **N-2782: Técnicas Aplicáveis à Análise de Riscos Industriais**. [S.l.]. 2015.

PETROBRAS. **N-76: Materiais de tubulação para instalações de refino e transporte**. PETROBRAS. [S.l.], p. 151. 2020.

POLIKAR, R. **Ensemble based systems in decision making**, v. 6, n. 3^a, p. 21-45, 2006.

PYLE, R. J.; HUGHES, R. R.; WILCOX, P. D. **Interpretable and explainable machine learning for ultrasonic defect sizing**, v. 70, n. 4, p. 277-290.

RAGHU, A.; GUNDLAPALLI, M.; MANIKANTAN, K. **DFT-DCT Combination Based Novel Feature Extraction Method for Enhanced Iris Recognition**. Proceedings of the International Conference on Signal, Networks, Computing, and Systems. [S.l.]: [s.n.]. 2017. p. 3-12.

RAJ, B.; AL, E. **Practical non-destructive testing**. 3^a. ed. New Delhi: Narosa Publishing House, 2007.

RAVEENDRAN, T. E. A. **Inspection of Concrete/Steel/Concrete pipes by ultrasound methods affected by corrosion**, Zurich, p. 9, 2022.

RAVEENDRAN, T. E. A. **Inspection of Concrete/Steel/Concrete pipes by ultrasound methods affected by corrosion**. International Conference on Non-destructive Evaluation of Concrete in Nuclear Applications. Espoo,: [s.n.]. 2023. p. 244-251.

ROSE, J. L. **Ultrasonic Guided Waves in Solid Media**. [S.l.]: Cambridge University Press, 2014.

SAH, J. P.; AKHTER, M. T. **Integrity assessment of non-piggable pipeline through direct assessment**. ASME India Oil and Gas Pipeline Conference. [S.l.]: American Society of Mechanical Engineers. 2013. p. V001T02A006.

SANTIN, J. L. **Ultra-som - técnica e aplicação**. [S.l.]: Artes Gráficas e Editora Unificado, 2003.

SCHMERR, L. W. **Fundamentals of ultrasonic nondestructive evaluation**. New York: Plenum Press, 1998.

SELMERS. **Selmers**, 2025. Disponível em: <<https://www.selmers.com/product-type/mechatronics/page/3/>>. Acesso em: 3 jan 2025.

SELVI, S. A.; KAMALAKANNAN, T. **The Effect of Emotions and Parkinson's Disease Prediction Using Classification Learner Models with EEG Dataset Generation**. International Conference on Digital Transformation in Business: Navigating the New Frontier. [S.l.]: Atlantis Press. 2024. p. 566-577.

SILVA, G. F.; FURTADO, L. O plano por trás do dano e o dano por trás do plano: o licenciamento ambiental do Terminal Norte Capixaba e seus impactos socioambientais em São Mateus – ES. **Revista Brasileira de Gestão Ambiental e Sustentabilidade**, 2021.

SILVA, L. C. E. A. Embedded decision support system for ultrasound nondestructive evaluation based on extreme learning machines. **Computers & Electrical Engineering**, 2021. 106891.

SMAGULOVA, D.; SAMAITIS, V.; JASIUNIENE, E. **Machine learning based approach for automatic defect detection and classification in adhesive joints**, v. 148, p. 103221, 2024.

SMITH, S. W. **The Scientist and Engineer's Guide to Digital Signal Processing**. 2ª. ed. San Diego: California Technical Pub, 1999.

SONG, Y. E. A. Corrosion process of ductile iron with cement mortar linings as coatings in reclaimed water. **International Journal of Electrochemical Science**, v. 11, p. 7031-7047, 2016. ISSN 8.

SOUZA, F. C. D.; PINTO, R. C. D. A. **Ultrasonic investigation on the effectiveness of crack repair in concrete**, v. 13, p. e13508, 2020.

STEFENON, S. F. E. A. **Analysis of the ultrasonic signal in polymeric contaminated insulators through ensemble learning methods**, v. 10, p. 33980-33991, 2022.

SUN, H.; RAMUHALLI, P.; JACOB, R. E. **Machine learning for ultrasonic nondestructive examination of welding defects: A systematic review**, v. 127, p. 106854, 2023.

TIAN, Z. E. A. A machine learning-assisted nondestructive testing method based on time-domain wave signals. **International Journal of Rock Mechanics and Mining Sciences**, 2024. 105731.

WANG, S.; LAMBORN, L.; CHEN. Weixing. Stress corrosion crack initiation and propagation before proceeding to Stage 2 for hydrostatically tested pipeline steels. **Journal of Materials Science**, 2022. 15967-15989.

WANG, W. E. A. **Fault diagnosis of gas regulator based on pca-lda-svm**. Chinese Automation Congress (CAC). [S.I.]: IEEE. 2019. p. 3594-3598.

WU, Y. E. A. **Two-step detection of concrete internal condition using array ultrasound and deep learning**, v. 139, p. 102945, 2023.

XIAO, H. E. A. **Defects identification using the improved ultrasonic measurement model and support vector machines**, v. 111, p. 102223, 2020.

XU, Q. E. A. **Compressive sensing of ultrasonic array data with full matrix capture in nozzle welds inspection**, 134, 2023. 107085.

YUAN, X.; LI, W.; DENG, M. Detection of defects in girth welded joints of tubular structure by circumferential feature-guided waves. **NDT & E International**, 2024. 103283.

YUAN, X.; LI, W.; DENG, M. Quantitative assessment of corrosion-induced wall thinning in L-shaped bends using ultrasonic feature guided waves. **Thin-Walled Structures**, 196, 2024. 111493.