MINISTÉRIO DA EDUCAÇÃO

UNIVERSIDADE FEDERAL DE CIÊNCIAS DA SAÚDE DE PORTO ALEGRE

PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO

PROGRAMA DE PÓS-GRADUAÇÃO EM TECNOLOGIAS DA INFORMAÇÃO E GESTÃO EM SAÚDE

André Gonçalves Jardim

MÉTODOS AUTOMATIZADOS PARA CONTROLE DE QUALIDADE EM EQUIPAMENTOS DE ULTRASSOM *MODO B:* O TESTE DE ESCALA DE CINZA

Porto Alegre

2026

André Gonçalves Jardim

MÉTODOS AUTOMATIZADOS PARA CONTROLE DE QUALIDADE EM EQUIPAMENTOS DE ULTRASSOM *MODO B:* O TESTE DE ESCALA DE CINZA

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Tecnologias da Informação e Gestão em Saúde da Universidade Federal de Ciências da Saúde de Porto Alegre como requisito parcial para a obtenção do título de Mestre.

Orientador: Profa. Dra. Viviane Rodrigues Botelho.

Coorientador: Profa. Dra. Thatiane Alves Pianoschi.

Porto Alegre

2026

Catalogação na Publicação

Jardim, André Gonçalves
MÉTODOS AUTOMATIZADOS PARA CONTROLE DE QUALIDADE EM EQUIPAMENTOS DE ULTRASSOM MODO B : O TESTE DE ESCALA DE CINZA / André Gonçalves Jardim. -- 2026.
114 p. : il., graf., tab. ; 30 cm.

Dissertação (mestrado) -- Universidade Federal de Ciências da Saúde de Porto Alegre, Programa de Pós-Graduação em Tecnologias da Informação e Gestão em Saúde, 2026.

Orientador(a): Viviane Rodrigues Botelho ; coorientador(a): Thatiane Alves Pianoschi.

1. ultrassonografia. 2. controle de qualidade. 3. escala de cinza. 4. processamento de imagem. 5. detecção automatizada. I. Título.

Sistema de Geração de Ficha Catalográfica da UFCSPA com os dados fornecidos pelo(a) autor(a).

André Gonçalves Jardim

MÉTODOS AUTOMATIZADOS PARA CONTROLE DE QUALIDADE EM EQUIPAMENTOS DE ULTRASSOM *MODO B:* O TESTE DE ESCALA DE CINZA

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Tecnologias da Informação e Gestão em Saúde da Universidade Federal de Ciências da Saúde de Porto Alegre como requisito parcial para a obtenção do título de Mestre em 2026.

Orientador: Prof. Profa. Dra. Viviane Rodrigues Botelho

Coorientador: Profa. Dra. Thatiane Alves Pianoschi

Aprovada em: 03 de agosto de 2026.

BANCA EXAMINADORA:

_____

Prof. Dr. Ana Trindade Winck

UFCSPA - Universidade Federal de Ciências da Saúde de Porto Alegre

_____

Prof. Dr. Henrique Trombini

UFCSPA - Universidade Federal de Ciências da Saúde de Porto Alegre

_____

Prof. Dr. Willian Souza Santos

UFS - Universidade Federal de Sergipe

RESUMO

Equipamentos de ultrassom diagnóstico estão sujeitos a exigências regulatórias de controle de qualidade no Brasil, porém o teste de escala de cinza permanece dependente de análise visual, sujeita à subjetividade do avaliador. Este trabalho desenvolve, aprimora e valida metodologias computacionais para a automatização do teste de escala de cinza e do alcance dinâmico exibido em equipamentos de ultrassom Modo B. A metodologia utilizou um banco de 156 imagens DICOM de três phantoms multipropósito (ATS 539, CIRS 040GSE e CIRS 054GS), adquiridas com transdutores convexos e lineares de múltiplos fabricantes e pareadas com máscaras de referência obtidas por segmentação manual. Foram implementadas três estratégias de detecção automática das regiões de interesse — abordagem morfológica bidimensional (AM2D), abordagem diferencial de sinais unidimensionais (ADS1D) e detecção por picos com restrição geométrica (DP) —, avaliadas por varredura combinatória exaustiva de 2.100 combinações de hiperparâmetros, totalizando 327.600 execuções individuais. O desempenho foi quantificado por métricas espaciais, score composto multicritério, testes de Friedman e Wilcoxon pareado e análise de concordância de Bland-Altman aplicada à Grayscale Mapping Function (GMF). A combinação líder (DP com filtro Savitzky-Golay) alcançou IoU médio de 0,7663, área de interseção de 85,06% e score composto de 0,8517. A GMF obtida pelo método automatizado apresentou viés inferior a 1,1 nível de cinza, coeficiente angular estatisticamente equivalente à unidade e coeficiente de determinação superior a 0,959, com ponto ótimo operacional em 75% do raio. Conclui-se que o método automatizado é funcionalmente equivalente à segmentação manual e oferece uma base computacional reprodutível, em código aberto, para a incorporação do teste de escala de cinza ao fluxo de controle de qualidade em ultrassonografia diagnóstica.

Palavras-chave: ultrassonografia; controle de qualidade; escala de cinza; processamento de imagem; detecção automatizada.

Objetivos de Desenvolvimento Sustentável (ODS): ODS 3 – Saúde e bem-estar: garantir uma vida saudável e promover o bem-estar para todos, em todas as idades; ODS 9 – Inovação e infraestrutura: construir infraestruturas resilientes, promover a industrialização inclusiva e sustentável e fomentar a inovação.

ABSTRACT

Diagnostic ultrasound equipment is subject to regulatory quality control requirements in Brazil; however, the grayscale test still relies on visual assessment, which is subject to the evaluator's subjectivity. This work develops, refines, and validates computational methodologies for automating the grayscale and displayed dynamic range test in B-mode ultrasound equipment. The methodology used a database of 156 DICOM images from three multipurpose phantoms (ATS 539, CIRS 040GSE, and CIRS 054GS), acquired with convex and linear transducers from multiple manufacturers and paired with reference masks obtained by manual segmentation. Three strategies for the automatic detection of regions of interest were implemented — a two-dimensional morphological approach (AM2D), a one-dimensional signal differential approach (ADS1D), and peak detection with geometric constraint (DP) —, evaluated through an exhaustive combinatorial sweep of 2,100 hyperparameter combinations, totaling 327,600 individual runs. Performance was quantified using spatial metrics, a multi-criteria composite score, Friedman and paired Wilcoxon tests, and Bland-Altman agreement analysis applied to the Grayscale Mapping Function (GMF). The leading combination (DP with Savitzky-Golay filtering) achieved a mean IoU of 0.7663, an intersection area of 85.06%, and a composite score of 0.8517. The GMF obtained by the automated method showed a bias below 1.1 gray levels, a regression slope statistically equivalent to unity, and a coefficient of determination above 0.959, with the operational optimum at 75% of the target radius. It is concluded that the automated method is functionally equivalent to manual segmentation and provides a reproducible, open-source computational basis for incorporating the grayscale test into the quality control workflow of diagnostic ultrasonography.

Keywords: ultrasonography; quality control; gray scale; image processing; automated peak detection.

Sustainable Development Goals (SDGs): SDG 3 – Good health and well-being: ensure healthy lives and promote well-being for all at all ages; SDG 9 – Industry, innovation

and infrastructure: build resilient infrastructure, promote inclusive and sustainable industrialization and foster innovation.

SUMÁRIO

1 INTRODUÇÃO 11
2 OBJETIVOS 16
3 REFERENCIAL TEÓRICO 17
3.1 FORMAÇÃO DA IMAGEM DE ULTRASSOM 17
3.1.1 O TESTE DE ESCALA DE CINZA E ALCANCE DINÂMICO 19
3.2 SIMULADORES E O CONTROLE DE QUALIDADE 20
3.3 O TESTE DE ESCALA DE CINZA NOS PROTOCOLOS E CONTROLE DE QUALIDADE 29
3.4 O TESTE DE ESCALA DE CINZA NA LEGISLAÇÃO BRASILEIRA 32
3.5 RESUMO COMPARATIVO DE MANUAIS E PROTOCOLOS 33
3.6 PROCESSAMENTO DAS IMAGENS 35
3.6.1 FILTROS NO PROCESSAMENTO DE IMAGENS 36
3.6.1.1 Ruído *speckle* 36
3.6.1.2 Redução de ruídos 37
3.6.1.3 Realce de bordas 38
3.6.1.4 Equalização de histogramas 39
3.6.2 FILTROS NO PROCESSAMENTO DE SINAIS 40
3.6.3 SEGMENTAÇÃO DE IMAGENS 42
3.6.3.1 Transformações Morfológicas 42
3.6.3.2 *Transformada de Hough* 42
4 TRABALHOS RELACIONADOS 43
4.1 IDENTIFICAÇÃO DO PROBLEMA DE PESQUISA 44
4.2 CRITÉRIOS DE INCLUSÃO E EXCLUSÃO E IDENTIFICAÇÃO DOS ESTUDOS PRÉ-SELECIONADOS 45
4.3 ANÁLISE E CATEGORIZAÇÃO DOS ESTUDOS SELECIONADOS 46
4.4 ANÁLISE E INTERPRETAÇÃO DOS ESTUDOS 47
4.4.1 METODOLOGIA DE SASSAROLI ET AL. (2017) 47
4.4.2 METODOLOGIA DE SASSAROLI ET AL. (2019) 50
4.4.3 METODOLOGIA DE TSAPAKI ET AL. (2021) 53
4.4.4 METODOLOGIA DE PAVERD ET AL. (2024) 55
5 METODOLOGIA 56
5.1 ORGANIZAÇÃO DO BANCO DE DADOS 57

5.1.1 DELIMITAÇÃO DAS REGIÕES DE ANÁLISE 59
5.2 ESTRATÉGIAS DE DETECÇÃO AUTOMATIZADA DE *ROIS* 59
5.3 *AM2D* - ABORDAGEM MORFOLÓGICA BIDIMENSIONAL 60
5.4 *ADS1D* - ABORDAGEM DIFERENCIAL DE SINAIS UNIDIMENSIONAIS 63
5.4.1 PROJEÇÃO POR COLUNAS 64
5.4.2 SUAVIZAÇÃO E ANÁLISE ESPECTRAL 66
5.4.3 DERIVADAS E PONTOS DE MUDANÇA 68
5.4.4 ESTIMATIVA DE CENTROS E LIMITAÇÕES 69
5.5 *ADP1D* – ABORDAGEM DE DETECÇÃO POR PICOS COM RESTRIÇÃO GEOMÉTRICA ... 70
5.5.1 REMOÇÃO DE *BASELINE* E TRANSFORMAÇÃO POR VALOR ABSOLUTO 71
5.5.2 DETECÇÃO DE PICOS USANDO RESTRIÇÃO DE POSIÇÃO 72
5.5.2.1 *DP-P* - Detecção direta por picos (picos) 73
5.5.2.2 *DP-E* - Detecção estrutural (estrutural) 73
5.5.2.3 *DP-H* - Detecção híbrida (hibrido) 73
5.5.2.4 *DP-HM* - Detecção híbrida com máximo local (hibridoMedio) 74
5.5.2.5 *DP-HS* - Detecção híbrida sequencial (hibridoSequencial) 74
5.5.3 REFINAMENTO VERTICAL DOS CENTROS 74
5.6 ANÁLISE DOS RESULTADOS 75
5.6.1 *VARREDURA COMBINATÓRIA EXAUSTIVA* 75
5.6.2 MÉTRICAS ESPACIAIS DE VALIDAÇÃO DA DETECÇÃO DE *ROIS* 76
5.6.2.1 Desvios normalizados nos eixos X e Y 76
5.6.2.2 Área de interseção percentual 77
5.6.2.3 Razão de interseção sobre união (*IoU*) 78
5.6.3 CRITÉRIOS PARA IDENTIFICAÇÃO DAS MELHORES COMBINAÇÕES 78
5.6.4 EXTRAÇÃO DE INDICADORES DE ESCALA DE CINZA 78
5.6.5 *RANKING* COMPOSTO MULTICRITÉRIO 79
5.6.6 TESTES ESTATÍSTICOS DE COMPARAÇÃO ENTRE METODOLOGIAS 80
5.6.7 ANÁLISE DE CONCORDÂNCIA PARA A *GMF* 81
6 RESULTADOS 82
6.1 CARACTERIZAÇÃO DO *BANCO DE DADOS* 82
6.2 VARREDURA COMBINATÓRIA EXAUSTIVA METODOLÓGICA 83
6.3 COMPARAÇÃO ESTATÍSTICA ENTRE AS METODOLOGIAS 85
6.3.1 MELHORES COMBINAÇÕES POR FRAÇÃO DO RAIO 87

6.4 Resultados do teste de Escala de Cinza e *GMF* 89
6.4.1 *GMF* com Coordenadas do Método Automatizado 90
6.4.2 Concordância entre Máscara e Método: *Bland-Altman* e Comparação de Inclinações 90
6.5 Perfis de Métricas Individuais das Melhores Combinações 92
6.6 Análise Qualitativa: Exemplos de Detecção 93
7 DISCUSSÃO 96
7.1 Síntese dos Achados Principais 96
7.2 Interpretação do Score Composto e dos Testes Estatísticos 96
7.3 Adequação da Métrica *de variança* e da Extração *roi_edges* 97
7.4 Concordância com a Máscara (*Bland-Altman*) 97
7.5 Implicações Clínicas 98
7.6 Limitações 98
8 CONCLUSÕES 100
9 TRABALHOS FUTUROS 102
10 REFERÊNCIAS 104
APÊNDICE A – PERFIS COMPLETOS DE MÉTRICAS POR ROI DAS CINCO MELHORES COMBINAÇÕES 108

1 Introdução

O ultrassom (US) é uma técnica de diagnóstico por imagem amplamente utilizada na medicina, que emprega ondas mecânicas para gerar imagens (Senra Filho et al., 2014). Esta técnica tem a vantagem de ser menos invasiva e mais econômica do que outras que utilizam radiação ionizante. A imagem é obtida através da emissão e recepção de uma onda sonora por um transdutor, que opera numa faixa de frequência central entre 2 e 15 MHz (Jensen, 2007).

O aparelho de US é conectado a um transdutor, o qual é composto por um ou mais cristais piezoelétricos, produz e detecta a onda de ultrassom. Esses cristais transformam energia elétrica em mecânica para produzir a onda sonora, que é refletida ao encontrar tecidos de diferentes densidades. Os cristais então recebem a onda refletida e convertem sua energia mecânica em pulsos elétricos, que contêm as informações para formar a imagem de ultrassom (MAHESH, 2013).

Os problemas no transdutor são uma das principais causas da baixa qualidade das imagens em ultrassonografias clínicas. Mesmo a falha de um único elemento do transdutor pode afetar o perfil da imagem, resultando em artefatos ou perda de detalhes. Quando um número maior de elementos apresenta defeitos, a resolução e a capacidade de penetração da onda no corpo podem ser comprometidas, resultando em imagens de qualidade inferior e dificultando a interpretação clínica (VAN HORSSEN et al., 2017).

Para assegurar o funcionamento adequado dos equipamentos de ultrassom e garantir a obtenção de imagens diagnósticas de alta qualidade, é essencial avaliar regularmente sua qualidade operacional. Segundo OLIVEIRA et al., 2011, um processo de controle de qualidade rigoroso permite estimar o grau de deterioração de um equipamento, fornece informações sobre os componentes defeituosos, verifica a qualidade dos reparos efetuados e inclusive auxilia no processo de compra de novos equipamentos.

Diversos órgãos internacionais e agências de credenciamento produziram diretrizes de controle de qualidade relacionadas aos padrões mínimos a serem alcançados, conforme citado por GRAZHDANI et al. (2018) as principais referências são: *American Institute of Ultrasound in Medicine* (AIUM), *American College of Radiology* (ACR), *Intersocietal Accreditation Commission* (IAC) e *International Electrotechnical Commission* (IEC).

No Brasil, a Resolução da Diretoria Colegiada (RDC) nº 611, de 9 de março de 2022 (BRASIL, 2022) , estabelece critérios obrigatórios para o Controle de Qualidade (CQ) em sistemas de ultrassom diagnóstico ou intervencionista, reforçando a necessidade de garantir a segurança e a eficácia desses equipamentos. Essa regulamentação, complementada pela Instrução Normativa (IN) nº 96, de 27 de maio de 2021 (BRASIL, 2021), determina que o Supervisor de Proteção Radiológica (SPR) seja responsável por assegurar a implementação e a manutenção desses procedimentos. A introdução dessas normas representa um avanço significativo na padronização dos processos de CQ no país, alinhando-se às práticas internacionais de segurança e qualidade em diagnóstico por imagem.

Nos procedimentos padrões de CQ as imagens geradas a partir da varredura de objetos simuladores, conhecidos como *phantoms*, são avaliadas visualmente por um profissional qualificado. Esses *phantoms* são dispositivos que simulam tecidos humanos, contendo estruturas internas com diferentes densidades e ecogenicidades, como cilindros, filamentos e massas anecóicas (sem eco). O profissional analisa as imagens para detectar a presença desses objetos e contabilizar quantas estruturas são visíveis e distinguíveis em espaços milimétricos. Essa avaliação inclui a verificação da nitidez das bordas, a resolução das estruturas e a capacidade do equipamento de diferenciar contrastes sutis entre os alvos.

No entanto, essas análises são essencialmente visuais e, portanto, subjetivas e operador dependentes. Como as análises devem ser realizadas periodicamente, conforme exigido pela RDC no 611, diferentes operadores podem analisar os resultados de um mesmo equipamento em momentos distintos, dificultando a obtenção de um padrão comparativo de avaliação devido às diferentes percepções dos profissionais. SENRA FILHO et al. (2014) sugerem que o uso de métodos computacionais dedicados a vários testes de CQ pode ser uma alternativa promissora para reduzir o tempo de análise dos testes e garantir a padronização e qualidade dos resultados. Trabalhos como NICK DUDLEY et al., 2010, SENRA FILHO et al., 2014 e VAN HORSSEN et al., 2017 propõem alternativas dessa natureza para CQ de US, demonstrando resultados promissores. Boa parte destas metodologias utilizam *machine learning*, técnicas que se baseiam na construção de algoritmos para predição e tomada de decisões, a partir de um volume massivo de dados. Com base em parâmetros geométricos e estatísticos extraídos das imagens, o algoritmo pode

analisar novas aquisições e realizar os testes de controle de qualidade de forma automatizada.

Diferente de outras técnicas de diagnóstico por imagem, como tomografia computadorizada e ressonância magnética, a falta de padronização no caso do ultrassom não está apenas na avaliação da imagem, mas também na forma como ela é adquirida. A qualidade da imagem depende diretamente de como o operador posiciona e move o transdutor, o que pode introduzir variabilidades significativas nos resultados. Essa dependência da técnica de aquisição ressalta a necessidade de se desenvolver, além de métodos de análise das imagens, um protocolo de aquisição padronizado, que garanta que as imagens sejam obtidas de maneira consistente e reprodutível. Um protocolo bem definido assegura que o método proposto funcione de maneira confiável, desde que as imagens sejam adquiridas seguindo as diretrizes estabelecidas, como o posicionamento correto do transdutor, ajustes de ganho e profundidade, e a utilização de um meio de acoplamento adequado.

Diversos testes são abordados para o aceite de controle de qualidade para os serviços de ultrassonografia no Brasil seguindo a IN nº 96 2021 (BRASIL, 2021), segundo essa instrução normativa os testes requisitados são necessários para garantir a qualidade dos sistemas de ultrassom diagnóstico ou terapêutico.

Dentre os testes abordados na IN nº 96 2021 (BRASIL, 2021), o de escala de cinza e alcance dinâmico exibido que será explorado neste trabalho (abordado como limiar de sensibilidade à baixo contraste na IN), é utilizado para avaliar a capacidade do equipamento de ultrassom em diferenciar variações sutis de contraste e representar fielmente as estruturas anatômicas. Esse teste envolve a análise de alvos com diferentes ecogenicidades presentes em um objeto simulador (*phantom*), onde a qualidade da imagem é avaliada com base na nitidez, formato e variação de brilho dos alvos. No entanto, os desafios associados a esse teste incluem a subjetividade na análise visual, a falta de padronização nos procedimentos de aquisição de imagens e a necessidade de garantir que os resultados sejam consistentes ao longo do tempo e entre diferentes operadores. Além disso, faltam critérios objetivos que definam quando uma imagem é adequada para ser analisada no teste de escala de cinza.

O teste de escala de cinza em ultrassonografia é fundamental para avaliar a qualidade da imagem gerada pelo equipamento, garantindo uma diferenciação adequada entre estruturas de diferentes ecogenicidades. Este teste permite analisar

três características essenciais: (i) o alcance dinâmico exibido, que verifica a capacidade do sistema de distinguir variações sutis na escala de cinza entre alvos com diferentes intensidades de eco, garantindo uma representação fiel da variação dos sinais recebidos; (ii) a nitidez dos alvos, que avalia a clareza e a precisão com que os alvos são representados na imagem, evitando distorções e artefatos que possam comprometer a análise clínica; e (iii) o formato dos alvos, garantindo que as estruturas apareçam com contornos bem definidos e sem deformações que possam interferir na interpretação da imagem.

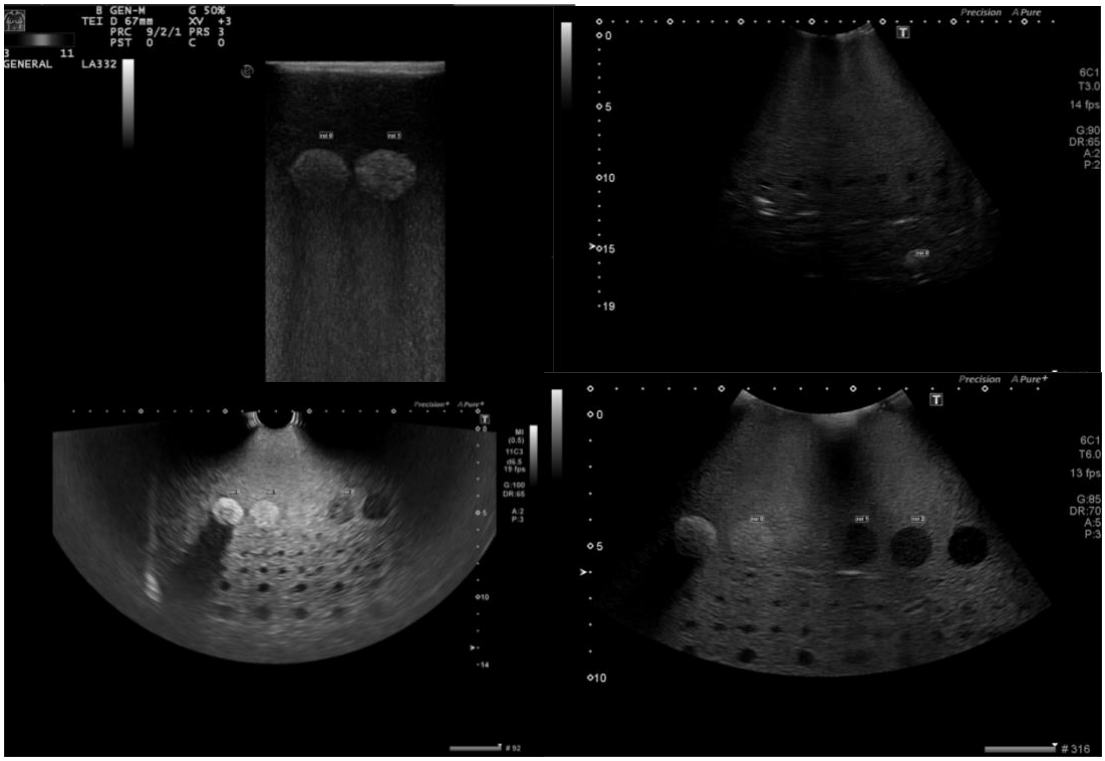
Atualmente, a avaliação desse teste é realizada de forma visual: o profissional adquire imagens do phantom com as configurações padrão do equipamento e observa quantos alvos cilíndricos são distinguidos na imagem, comparando o resultado com os critérios do fabricante do phantom e da legislação vigente. Conforme a IN nº 96/2021 (BRASIL, 2021), o teste de limiar de sensibilidade à baixo contraste (correspondente ao teste de escala de cinza) deve ser realizado periodicamente, porém a normativa não detalha critérios objetivos para o julgamento das imagens, deixando a análise inteiramente a cargo da percepção visual do operador. Essa lacuna reforça a necessidade de métodos automatizados que ofereçam avaliações quantitativas e reprodutíveis.

Portanto, levando em conta a necessidade prática dos setores de radiodiagnóstico e as exigências da legislação atual, este trabalho se propõe, em parceria com Hospital de Clínicas de Porto Alegre (HCPA), a desenvolver e aperfeiçoar metodologias computacionais para a realização dos testes de escala cinza e alcance dinâmico exibido. Nesse contexto, surge a seguinte questão de pesquisa: Como automatizar os testes de escala de cinza e alcance dinâmico exibido em ultrassonografia, utilizando algoritmos de processamento de imagem de modo *a* atender às exigências regulatórias de controle de qualidade.

A principal dificuldade na automatização desse teste reside na grande heterogeneidade das imagens obtidas em campo. Equipamentos de diferentes fabricantes, modelos de transdutores e configurações de aquisição (ganho, profundidade, frequência) produzem imagens com aparências muito distintas entre si. Mesmo para um mesmo phantom, as estruturas cilíndricas de escala de cinza podem apresentar variações significativas de contraste, nitidez e formato, dependendo do equipamento e das condições de aquisição utilizadas. A Figura 1 ilustra essa

variabilidade, apresentando exemplos de imagens de escala de cinza obtidas com diferentes equipamentos e transdutores.

Figura 1: Exemplo de imagens de ultrassom modo b utilizadas no projeto.



Essa diversidade torna impraticável a adoção de abordagens rígidas de reconhecimento de padrões (como técnicas convencionais de visão computacional baseadas em templates fixos) uma vez que não existe um modelo único de imagem que represente todas as condições possíveis de aquisição. Além disso, a presença de ruído speckle, característico de imagens de ultrassom, reduz a nitidez das bordas dos alvos e a uniformidade do fundo, dificultando ainda mais a detecção automatizada das regiões de interesse.

Diante desse cenário, o uso de métodos computacionais para a análise dos testes de controle de qualidade em ultrassom tem o potencial de reduzir significativamente ou até eliminar a subjetividade das análises convencionais. Além disso, pode contribuir para melhorar a qualidade do exame, proporcionando maior segurança no atendimento ao paciente e diagnósticos mais precisos. Esses benefícios estão alinhados com os Objetivos de Desenvolvimento Sustentável (ODS) da ONU, especialmente o ODS 3: Saúde e bem-estar, que visa garantir uma vida saudável e promover o bem-estar para todos em todas as idades, e o ODS 9: Inovação e

infraestrutura, que visa construir infraestruturas resilientes, promover a industrialização inclusiva e sustentável e fomentar a inovação.

O resultado deste trabalho pode ser considerado uma inovação de produto de acordo com o Manual de Oslo, por combinar métodos computacionais para automatizar os testes de controle de qualidade em ultrassom de forma mais eficiente e menos subjetiva do que os métodos convencionais.

Cabe ressaltar que este trabalho está inserido em uma linha de pesquisa mais ampla, intitulada “Métodos computacionais para controle de qualidade para equipamentos de ultrassom diagnóstico e intervencionista”, aprovada pelo Comitê de Ética em Pesquisa (CEP) sob o número CAAE 54743621.8.0000.5345, em conformidade com a Resolução 466/12 do CONEP e com a Lei nº 11.794, de 8 de outubro de 2008.

2 Objetivos

Desenvolver, aprimorar e validar metodologias computacionais para a automatização do teste de escala de cinza e alcance dinâmico exibido em equipamentos de ultrassom em modo-B, visando melhorar a padronização, reduzir subjetividades e contribuir para o atendimento das exigências regulatórias de controle de qualidade.

Para atingir o objetivo geral, os seguintes objetivos específicos são:

- Realizar um levantamento da literatura acerca de testes de escala de cinza e alcance dinâmico exibido;
- Implementar soluções embasadas em processamento de imagem para segmentação automática e identificação de regiões de interesse (*ROIs*) em imagens de ultrassom;
- Desenvolver metodologias embasadas em processamento de imagem para avaliação automática da escala de cinza e alcance dinâmico exibido.
- Avaliar as metodologias propostas com diferentes tipos de transdutores (convexos e lineares) e equipamentos de múltiplos fabricantes, verificando qual delas apresenta melhor comportamento diante dessa variabilidade.

3 Referencial Teórico

3.1 Formação da imagem de ultrassom

A ultrassonografia é uma técnica de imagem médica que utiliza ondas sonoras de alta frequência para obter representações visuais das estruturas internas do corpo. O princípio físico fundamental por trás dessa tecnologia é a propagação das ondas acústicas e suas interações com os tecidos biológicos. As imagens são geradas a partir da reflexão dessas ondas nos limites entre diferentes meios, permitindo a reconstrução das características anatômicas internas (Bushberg et al., 2012; Jensen, 2004).

O transdutor, componente essencial do sistema de ultrassom, contém cristais piezoelétricos que convertem pulsos elétricos em ondas sonoras e, posteriormente, transformam os ecos refletidos em sinais elétricos, que serão processados para a formação da imagem (Mahesh, 2013). Quando a onda sonora atravessa os tecidos, ela sofre reflexão, refração, absorção e espalhamento, fenômenos que influenciam diretamente a qualidade e o contraste da imagem gerada. A reflexão ocorre nos limites entre tecidos com diferentes impedâncias acústicas, sendo essa a base da formação da imagem. Já a absorção e o espalhamento contribuem para a atenuação do sinal, reduzindo a intensidade dos ecos retornados ao transdutor (Rangaraj M. Rangayyan, 2004).

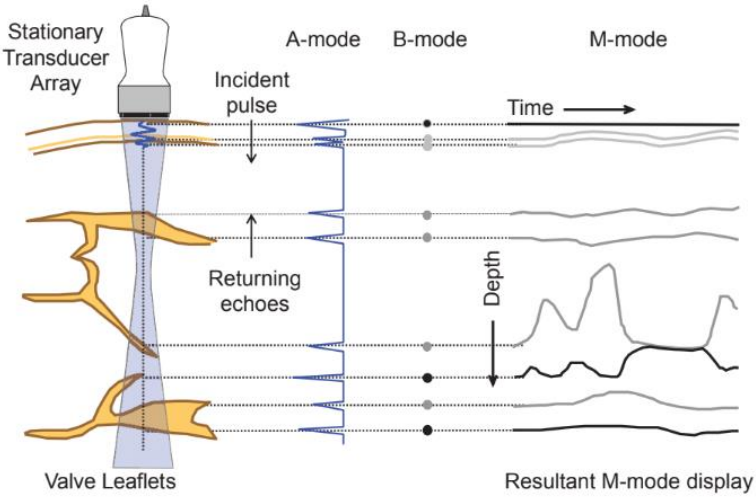
A resolução da imagem ultrassonográfica depende de diversos fatores, incluindo a frequência da onda sonora utilizada. Transdutores de alta frequência (acima de 7 MHz) produzem imagens com maior resolução espacial, porém com menor penetração, sendo indicados para exames superficiais. Por outro lado, frequências mais baixas (entre 2 e 5 MHz) possibilitam maior penetração, mas com perda de detalhes finos (Bushberg et al., 2012). A escolha adequada da frequência é crucial para otimizar a qualidade da imagem em diferentes aplicações clínicas.

Outro fator determinante na qualidade da imagem é a formação do feixe ultrassônico, que pode ser ajustada eletronicamente em sistemas modernos. A focalização dinâmica melhora a resolução lateral ao ajustar o ponto focal do feixe durante a varredura, enquanto técnicas como a compensação de ganho em profundidade (*TGC – Time Gain Compensation*) ajudam a corrigir a atenuação do sinal conforme a profundidade da imagem, garantindo melhor uniformidade de contraste (JENSEN, 2004).

No ultrassom, as imagens são geradas por diferentes modos de aquisição, que variam conforme a forma de emissão e detecção dos ecos. Entre eles, destacam-se o *Modo A*, *Modo B* e *Modo M*. O *Modo A* representa a amplitude do eco como um sinal digital em função do tempo, enquanto o *Modo B* transforma essas informações em pontos modulados por brilho, permitindo a visualização bidimensional de estruturas estáticas. Já para análises dinâmicas, como o monitoramento do fluxo sanguíneo ou a movimentação de órgãos, utilizam-se o *Modo M* e o *Doppler*, que captam variações no tempo e na frequência dos ecos refletidos (Bushberg et al., 2012).

O *Modo M*, especificamente, emprega os dados do *Modo B* para exibir o deslocamento das estruturas em movimento ao longo do tempo. Com os avanços tecnológicos na ultrassonografia, o uso dos Modos A e M tornou-se menos comum, enquanto o *Modo B* e o *Doppler* se consolidaram como as principais técnicas na prática clínica atual. A Figura 2 ilustra as representações dos Modos A, B e M (Mahesh, 2013).

Figura 2: Representação dos Modos A, B e M no imageamento por ultrassom. Fonte: (Bushberg et al., 2012)



O *Modo B* é a técnica fundamental na ultrassonografia diagnóstica, sendo responsável pela formação de imagens bidimensionais a partir da conversão dos ecos de ultrassom em diferentes níveis de brilho. Nesse modo, a intensidade do eco refletido pelos tecidos é traduzida em tons de cinza, onde estruturas mais densas, que refletem mais ondas sonoras, aparecem mais brilhantes, enquanto áreas com menor refletividade são exibidas em tons mais escuros. A qualidade da imagem gerada depende de diversos fatores, como a frequência do transdutor, a resolução espacial e

a atenuação do sinal com a profundidade. Para otimizar esses aspectos, sistemas modernos de ultrassom incorporam técnicas avançadas, incluindo a focalização eletrônica e o processamento digital de sinais, que aumentam a nitidez e o contraste da imagem. Essas inovações tornaram o *Modo B* indispensável na prática clínica, permitindo diagnósticos mais precisos em diversas especialidades médicas (Bushberg et al., 2012; Mahesh, 2013).

Para a gerar essas imagens, o aparelho de ultrassom possui um componente chamado de transdutor, o qual envia e recebe as ondas de ultrassom ao tecido a ser imageado. Os transdutores têm três tipos principais utilizados na aquisição de imagens por ultrassom: linear, convexo e em fases. Os transdutores lineares formam uma imagem retangular através do disparo contínuo de um conjunto de elementos. Os transdutores convexos operam no mesmo princípio, mas possuem uma configuração de superfície convexa do conjunto de elementos, resultando numa maior área da imagem formada. O transdutor de matriz em fases emite um único disparo de todos os elementos e a direção do feixe é determinada pelo atraso elétrico dos sinais desses elementos (Jensen, 2007).

3.1.1 O teste de escala de cinza e alcance dinâmico

O teste de escala de cinza é um procedimento de controle de qualidade que avalia a capacidade do sistema de ultrassom em representar, de forma simultânea e distinguível, ecos com diferentes amplitudes na mesma imagem. Do ponto de vista físico, esse teste verifica se o mapeamento entre a amplitude do eco retornado e o nível de brilho (tom de cinza) exibido na tela é consistente ao longo de toda a faixa dinâmica do equipamento.

O alcance dinâmico exibido corresponde ao intervalo entre o eco de menor amplitude perceptível (nível mínimo de cinza) e o eco de máxima amplitude (brilho máximo). Um sistema com bom desempenho em escala de cinza deve ser capaz de diferenciar alvos com ecogenicidades distintas (tanto hiperecogênicos quanto hipoecogênicos em relação ao fundo) de forma proporcional à diferença de intensidade acústica real. Clinicamente, essa capacidade é essencial para a diferenciação entre tecidos moles com impedâncias acústicas semelhantes e para a detecção de lesões de baixo contraste (Bushberg et al., 2012; International Electrotechnical Commission, 2010).

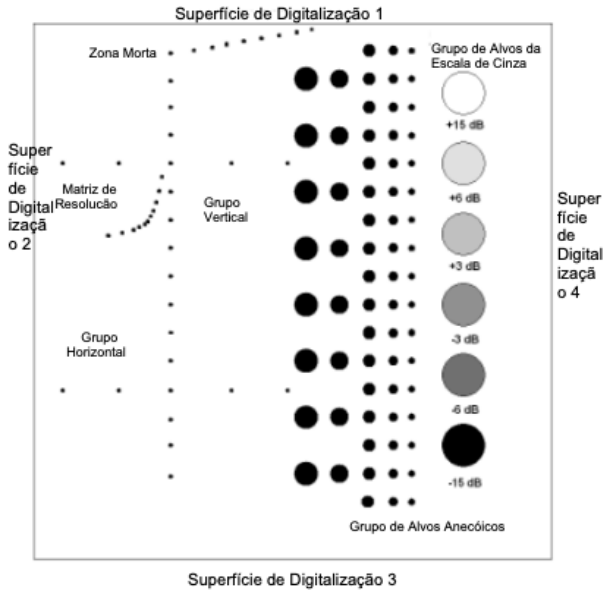
No contexto do controle de qualidade, o teste de escala de cinza serve como indicador da integridade da cadeia de processamento do sinal, desde a conversão do eco em sinal elétrico pelo transdutor, passando pela amplificação e compressão logarítmica, até a exibição final no monitor. Degradações em qualquer etapa dessa cadeia podem resultar em perda de diferenciação de contraste, comprometendo a qualidade diagnóstica da imagem.

3.2 Simuladores e o controle de qualidade

Os testes de controle de qualidade em ultrassom requerem a avaliação de imagens geradas sob condições controladas. Para isso, utilizam-se phantoms, objetos simuladores (OS) feitos de material com características acústicas semelhantes às do tecido humano, contendo estruturas internas com diferentes propriedades acústicas que visam simular a heterogeneidade do corpo humano (Pessôa, 2009).

Há diversos modelos de OS, como o *phantom* Multifuncional Modelo 539 da fabricante *ATS Laboratories Incorporated*, CIRS Company (Figura 3). Este OS foi projetado com uma combinação de alvos de linha de monofilamento e alvos cilíndricos que mimetizam tecidos de tamanhos e contrastes variados. Os alvos de linha de monofilamento possuem um diâmetro de 0,12 mm, para otimizar a imagem exibida em frequências que variam de 2,25 a 7,5 MHz (ATS Laboratories Incorporated, 2016).

Figura 3: Representação da estrutura do *phantom* ATS modelo 539 (ATS Laboratories Incorporated, 2016)



Conforme o manual deste *phantom* (ATS Laboratories Incorporated, 2016), são descritos 8 testes de controle de qualidade, sendo eles os seguintes:

- Zona Morta
- Calibração da Medição Vertical
- Calibração da Medição Horizontal
- Resolução Axial e Lateral
- Zona Focal
- Sensibilidade (Profundidade Máxima de Penetração)
- Resolução Funcional, Definição e Preenchimento
- Escala Cinza e Alcance Dinâmico Exibido

No manual deste *phantom* (ATS Laboratories Incorporated, 2016) o teste de foco desta pesquisa é identificado como "Escala de cinza e alcance dinâmico exibido"**.** Segundo o manual, o teste avalia dois aspectos importantes:

1. Alcance dinâmico exibido: Verifica a capacidade do equipamento de diferenciar a escala de cinza entre alvos com diferentes ecogenicidades, garantindo uma representação fiel da variação dos ecos.
2. Nitidez e forma dos alvos: Analisa se os alvos aparecem de maneira clara, nítida e circular, assegurando a precisão na resolução da imagem e evitando distorções ou artefatos que possam comprometer a qualidade diagnóstica.

O termo alcance dinâmico exibido refere-se à faixa de sinal de eco que vai desde um eco apenas perceptível (menor nível de escala de cinza) até o brilho máximo (máxima amplitude de eco). Clinicamente, essa característica é essencial para a visualização de estruturas com diferentes ecogenicidades, especialmente em exames que envolvem tecidos moles e áreas anecóicas. Para isso, o manual do *phantom* Modelo 539 (ATS Laboratories Incorporated, 2016) descreve as seguintes etapas:

1. Preparação do *Phantom*: O *phantom* é colocado sobre uma superfície plana e limpa, utilizando a superfície de digitalização número 4 do *phantom* (Figura 3) para o exame.
2. Aplicação do Meio de Acoplamento: Uma quantidade adequada de gel de baixa viscosidade ou água é aplicada na superfície de digitalização, tomando cuidado para evitar a formação de bolhas de ar ao utilizar água.

3. Ajuste dos Parâmetros de Digitalização: As configurações do aparelho, como compensação de ganho em profundidade (*TGC*) e saída, são ajustadas para os valores de linha de base recomendados para uma imagem de fígado normal. O ganho deve ser ajustado até que a parte inferior do *phantom* fique completamente preta, e essas configurações são registradas para referência futura.
4. Aquisição da Imagem: O transdutor é posicionado sobre o grupo de alvos em escala cinza até que uma imagem clara seja obtida (Figura 3).
5. Congelamento e Registro da Imagem: Após obter a imagem, ela é congelada e uma cópia impressa é feita para análise.
6. Análise da Imagem: Segundo o manual, os alvos devem aparecer de forma circular, com bordas nítidas e bem definidas, exibindo uma variação no brilho que vai desde níveis de contraste baixos até altos. Além disso, qualquer sombreamento posterior às estruturas deve ser observado e anotado. No entanto, o manual não fornece indicadores quantitativos claros para definir o que seria uma "variação de brilho aceitável" ou o que caracteriza "bordas nítidas". Essa falta de parâmetros objetivos é uma limitação significativa para a automação do método, pois a avaliação desses critérios depende, em grande parte, da interpretação subjetiva do operador. Portanto, para que a automação seja viável, será necessário estabelecer métricas quantitativas que possam ser aplicadas de forma consistente, como a medição do gradiente de intensidade nas bordas ou a análise da distribuição de brilho nos alvos.
7. Registro dos Resultados: Todos os achados da imagem são documentados no registro de garantia de qualidade para futuras comparações.

Cabe salientar que, nos manuais, não existe uma padronização em relação aos nomes dos testes, métodos e objetivos específicos. Dessa forma, há diversos testes em outros manuais que citam a visualização de escala de cinza de objetos de diferentes ecogenicidades com outros nomes.

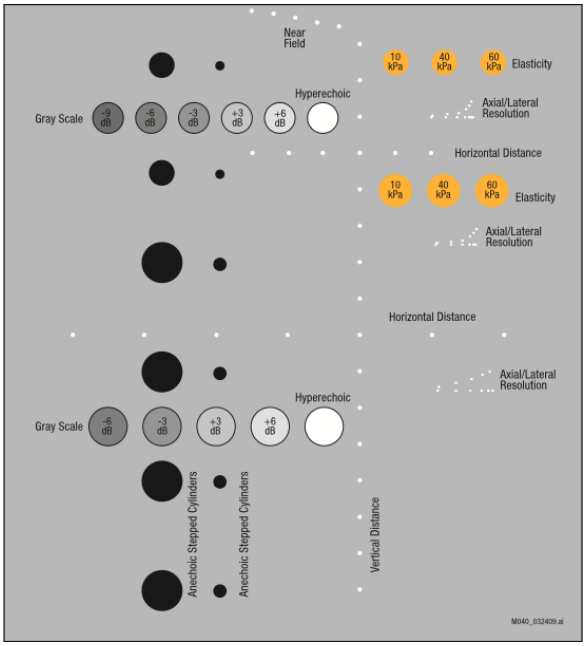
Além do ATS 539, outros simuladores amplamente utilizados em testes de controle de qualidade são o *Multi-Purpose, Multi-Tissue Ultrasound Phantom Model 040GSE* (CIRS) (Figura 4), 054GS General Purpose Ultrasound Phantom (CIRS)

(Figura 5), 403 GS LE Precision Multi-Purpose Phantoms (Figura 6), US-MP (KONEX) (Figura 7) e outros.

O Modelo 040GSE *Multi-Purpose, Multi-Tissue Ultrasound Phantom* (CIRS, 2013a), é projetado para avaliar o desempenho de sistemas de ultrassom diagnóstico, sendo composto pelo material *Zerdine*®, que simula propriedades acústicas de tecidos humanos, como pode ser observado na Figura 4. Ele permite realizar uma ampla gama de testes de qualidade de imagem, incluindo os seguintes descritos em seu manual:

- Uniformidade
- Grupo de Campo Próximo
- Grupo de Distância Vertical
- Grupo de Distância Horizontal
- Grupos de Resolução Axial e Lateral
- Cilindros Escalonados Anecóicos
- Alvos de Escala de Cinza
- Grupos de Alvos para Elasticidade

Figura 4: Representação da estrutura do *phantom* ATS modelo 040GSE (CIRS, 2013a)



Neste manual (CIRS, 2013a), a sensibilidade ao contraste é avaliada através de dois grupos de alvos em escala de cinza (Figura 4). O primeiro grupo, posicionado

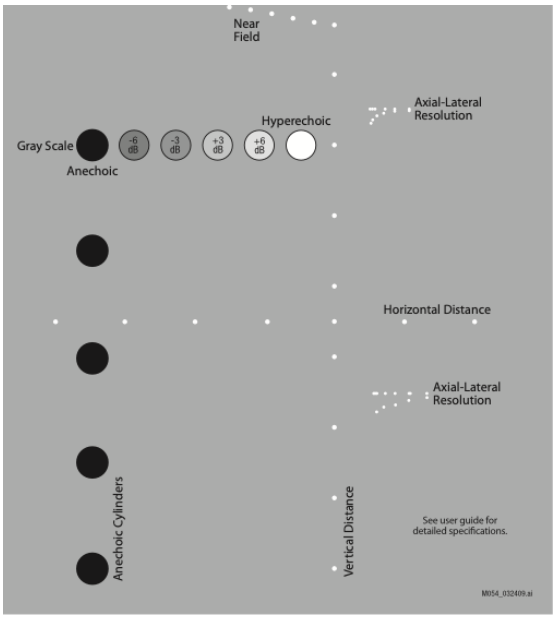
a 3 cm de profundidade, é projetado para sondas de 5 MHz ou superiores e apresenta uma gama de contrastes baixos, variando de -9 *dB* a mais de +15 *dB*. O segundo grupo, localizado a 11,5 cm de profundidade, oferece uma faixa de contraste baixo que vai de -6 *dB* a mais de +15 *dB*. Esses alvos servem, segundo o manual, para avaliar a capacidade do equipamento de distinguir diferentes níveis de contraste em profundidades variadas.

Assim como o Modelo 040GSE (CIRS, 2013a), o Modelo 054GS(CIRS, 2013b) *General Purpose Ultrasound Phantom* é desenvolvido para avaliar o desempenho de sistemas de ultrassom diagnóstico, utilizando o material *Zerdine*®, que simula propriedades acústicas de tecidos humanos, como pode ser observado na Figura 5. Neste modelo, seu manual conta com os seguintes testes descritos:

- Uniformidade
- Profundidade de Penetração
- Perfil do Feixe Ultrassônico / Zona Focal / Largura da Resposta Lateral
- Medição de Distância Vertical
- Medição de Distância Vertical
- Resolução Axial e Lateral
- Resolução no Plano de Elevação
- Detectabilidade de Alvos de Baixo Contraste
- Sensibilidade de Contraste em Escala de Cinza
- Avaliação da Zona Morta

No entanto, o Modelo 054GS se diferencia por sua aplicação mais focada em testes básicos de desempenho em sistemas de ultrassom gerais, com alvos projetados para verificar contraste de imagem, detectabilidade de alvos de baixo contraste (Low-Contrast Target Detectability) e parâmetros como resolução elevacional (Elevational Resolution). Além disso, ele contém cilindros com diferentes densidades e contrastes ecográficos, facilitando análises comparativas da resposta dinâmica do equipamento frente a variações nos parâmetros de ganho e frequência (CIRS, 2013b).

Figura 5: Representação da estrutura do *phantom* ATS modelo 054GS (CIRS, 2013b)



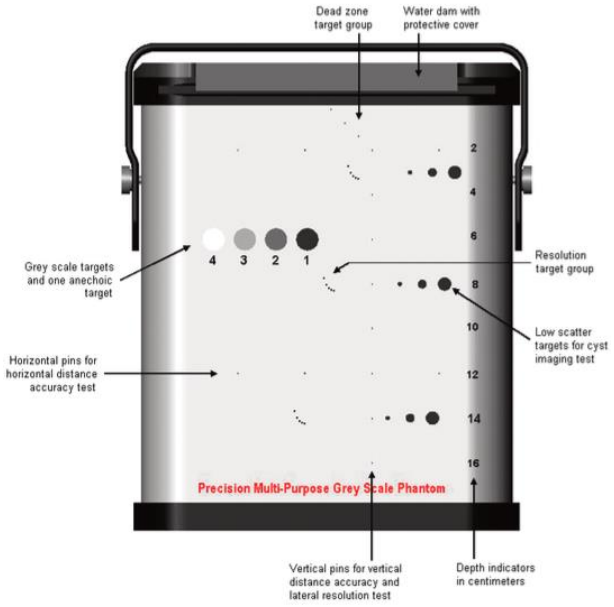
Neste manual (CIRS, 2013b), o teste de sensibilidade ao contraste é realizado utilizando seis cilindros, com variação de ecogenicidade (Figura 5) que vai de anecoico a hiperecoico em relação ao material de fundo. Esses cilindros, com 8 mm de diâmetro e posicionados a uma profundidade de 4 cm, são dispostos no *phantom* para avaliar a faixa dinâmica de um sistema de ultrassom. O teste, segundo o manual, tem como função observar a capacidade do sistema em distinguir alvos de diferentes níveis de escala de cinza, essencial para verificar a precisão da imagem gerada pelo equipamento. Além disso, a precisão na representação do tamanho e formato da massa pela máquina também pode ser determinada através desse teste.

O 403 GS LE *Precision Multi-Purpose Phantom* (Gammex, 2003) foi desenvolvido para avaliar a qualidade de imagem de sistemas de ultrassom de alta resolução, fornecendo alvos cuidadosamente distribuídos para medir parâmetros como resolução, profundidade de penetração e precisão dos cálipers eletrônicos. Seu material de fundo imita tecidos humanos, garantindo que os testes reflitam de maneira mais fiel o desempenho clínico do equipamento. Este *phantom* contém alvos de escala de cinza**,** posicionados a uma profundidade de 6 cm, com contrastes de -6 *dB*, +6 *dB* e alta dispersão em relação ao material de fundo, permitindo a avaliação da capacidade do equipamento em diferenciar intensidades de eco, como pode ser observado na Figura 6.

O manual do 403 GS LE descreve uma série de testes de controle de qualidade que podem ser realizados para monitorar a qualidade da imagem gerada pelo sistema de ultrassom. Esses testes incluem:

- Zona Morta
- Precisão de Medida Vertical e Horizontal
- Resolução Axial e Lateral
- Profundidade de Penetração
- Alvos Císticos
- Alvos em Escala de Cinza
- Linearidade do Sistema
- Imagem Harmônica
- Dessecamento do *phanton*

Figura 6: Representação da estrutura do *phantom* modelo 403 GS LE (Gammex, 2003)



No manual deste *phantom* (Gammex, 2003) o teste de escala de cinza é referido como *Grey Scale Targets*. Esse teste avalia a capacidade do sistema de ultrassom em diferenciar variações sutis de contraste, utilizando alvos de escala de cinza posicionados a uma profundidade de 6 cm (Figura 6). Esses alvos apresentam contrastes de -6 *dB*, +6 *dB* e alta dispersão em relação ao material de fundo, permitindo a avaliação da capacidade do equipamento em representar fielmente diferentes níveis de brilho na imagem.

O manual descreve duas abordagens para a análise dos alvos de escala de cinza: uma quantitativa e outra qualitativa**.**

1. Medição Quantitativa:
    a. Nessa abordagem, o sistema de ultrassom é utilizado para medir a intensidade média dos *pixels* dentro de uma região de interesse (*ROI*) definida sobre cada alvo de escala de cinza. O objetivo é quantificar a relação entre os níveis de contraste nominal dos alvos (em *dB*) e os valores de brilho exibidos na imagem.
    b. O manual recomenda que o operador ajuste o ganho do sistema até que o ruído no alvo anecóico seja minimizado, garantindo que a medição seja realizada em condições consistentes. Em seguida, a intensidade média dos *pixels* é registrada para cada alvo e comparada com valores de referência.
    c. Essa abordagem permite uma avaliação objetiva da linearidade do sistema e da capacidade de diferenciar contrastes, mas depende da disponibilidade de ferramentas de análise de *ROI* no equipamento de ultrassom.
2. Medição Qualitativa:
    a. Na abordagem qualitativa, o operador compara visualmente o brilho dos alvos de escala de cinza com uma barra de cinza exibida no monitor do ultrassom. Cada alvo é associado a um nível de brilho correspondente na barra de cinza, e o operador registra o nível que melhor corresponde ao brilho observado.
    b. Essa abordagem é mais subjetiva, pois depende da percepção visual do operador, mas é útil para uma avaliação rápida da consistência da representação de brilho na imagem.

Além do teste de *Grey Scale Targets*, o manual do 403 GS LE (Gammex, 2003) descreve outro teste alinhado com os objetivos deste trabalho: o *System Linearity* (Linearidade do Sistema). Esse teste avalia a consistência do mapeamento de ecos em níveis de brilho na tela, verificando se o sistema mantém uma relação linear entre a amplitude dos ecos e o brilho exibido. A linearidade do sistema é importante para garantir que variações sutis de contraste sejam representadas de forma fiel na

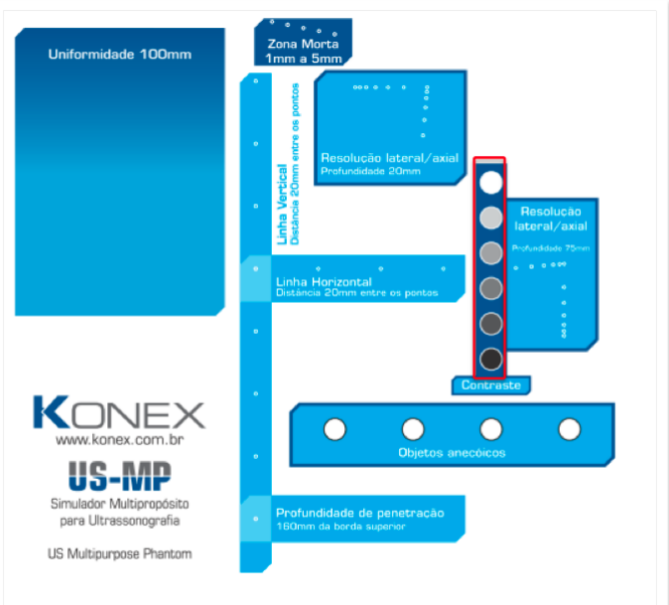
imagem, o que está diretamente relacionado aos objetivos dos testes de escala de cinza descritos nos outros manuais.

O *phantom* Multipropósito de Ultrassonografia US-MP (KONEX, *[S.d.]*) é um simulador projetado para permitir a avaliação de diversos parâmetros. Seu design incorpora alvos como linhas de monofilamento, estruturas cilíndricas e objetos anecóicos distribuídos em profundidades variadas, simulando tecidos humanos para testes realistas. O material de base possuí uma atenuação de 0,5 *dB*/cm a 1 MHz, garantindo compatibilidade com padrões clínicos, como pode ser observado na Figura 7.

Os testes indicados no manual deste *phantom* são os seguintes:

- Zona Morta
- Uniformidade
- Profundidade de Penetração
- Zona Focal
- Exatidão Vertical/Horizontal
- Resolução Axial e Lateral
- Visibilidade de Objetos Anecóicos
- Sensibilidade de Baixo Contraste

Figura 7: Representação da estrutura do *phantom* modelo US-MP (KONEX, *[S.d.]*)



O teste de *System Linearity* é realizado analisando a curva de mapeamento de brilho (S-curve), que relaciona a amplitude dos ecos aos níveis de brilho exibidos no

monitor. Qualquer alteração na forma dessa curva pode indicar problemas no processamento de imagem, como distorções no contraste ou perda de detalhes em regiões de baixa ou alta ecogenicidade. Esse teste complementa a avaliação dos alvos de escala de cinza, fornecendo uma visão mais abrangente da capacidade do sistema em representar variações de contraste de forma consistente.

No manual do modelo US-MP (KONEX, *[S.d.]*) o teste de sensibilidade de baixo contraste tem como objetivo descrito no manual, avaliar a capacidade do sistema de ultrassom em distinguir objetos com diferentes densidades, o que está diretamente relacionado ao desempenho do processamento da escala de cinza e à faixa dinâmica (Figura 7). Os alvos utilizados para essa avaliação são estruturas cilíndricas ecogênicas, com diâmetro de 5,5 mm, distribuídas em diferentes profundidades que variam de 50 mm a 110 mm.

Segundo o manual, para a execução do teste, o transdutor deve ser posicionado sobre o grupo de alvos em escala de cinza até obter uma imagem clara. Os alvos devem se apresentar em formato circular, com diferentes níveis de brilho, possibilitando a avaliação do alcance dinâmico do equipamento. Além disso, é recomendado registrar a presença ou ausência de artefatos como sombreamento posterior, que podem afetar a percepção do contraste.

O teste serve como linha de base para o acompanhamento do desempenho do sistema ao longo do tempo e permite a identificação de possíveis falhas na capacidade do equipamento em diferenciar variações sutis de contraste, aspecto fundamental para a qualidade da imagem diagnóstica.

3.3 O teste de escala de cinza nos Protocolos e controle de qualidade

Os protocolos de controle de qualidade são diretrizes abrangentes para a garantia de qualidade em sistemas de imagem por ultrassom, abordando, por exemplo, procedimentos de teste, calibração de equipamentos e manutenção preventiva. Entre os principais documentos que orientam essas práticas estão:

- ***Quality Assurance of Ultrassond Imaging* Systems** (Nick Dudley et al., 2010)

O livro *Quality Assurance of Ultrasound Imaging Systems*, publicado pelo *Institute of Physics and Engineering in Medicine* (IPEM) como parte de seus relatórios técnicos, apresenta diretrizes abrangentes para a garantia de qualidade em sistemas

de imagem por ultrassom. O IPEM, uma organização britânica dedicada ao desenvolvimento de padrões e diretrizes na física médica e engenharia biomédica, aborda no livro procedimentos de teste, calibração de equipamentos e manutenção preventiva, contribuindo para a padronização e o controle de qualidade desses sistemas.

Esse livro, que pode ser tratado como um protocolo, aborda o teste *Grey-Scale Performance*, que, segundo o protocolo, avalia a capacidade do sistema em distinguir variações de contraste. O teste utiliza um *phantom* que deve conter alvos cilíndricos ou prismáticos com diferentes níveis de ecogenicidade, todos posicionados na mesma profundidade. A visibilidade de cada alvo é registrada em uma escala de três pontos: não visível, visível, mas indistinto e claramente visível. Segundo os autores, embora esse método forneça uma avaliação subjetiva, a repetição do teste em diferentes condições ajuda a estabelecer um critério de tolerância adequado. É importante notar que este relatório foi retirado do site do IPEM e não foi republicado, permanecendo apenas a mensagem de que se encontra em revisão.

- Protocolo Español de control de calidad en radiodiagnóstico (Sociedad Española de Física Médica, 2012)

Elaborado pela Sociedade Espanhola de Física Médica e pela Sociedade Espanhola de Proteção Radiológica, este protocolo estabelece critérios para o controle de qualidade de equipamentos de radiodiagnóstico, incluindo sistemas de ultrassom. O Protocolo Espanhol aborda a avaliação da qualidade da imagem por meio de testes diversos, incluindo o teste EC010 - *Umbral de sensibilidad a bajo contraste*. Este teste, segundo os autores, tem como objetivo verificar a capacidade do equipamento em detectar objetos com baixo contraste em relação ao fundo.

Para realizar o teste, deve-se utilizar um *phantom* contendo objetos com diferentes níveis de ecogenicidade. O transdutor deve ser posicionado e alinhado de *modo a* imagear esses objetos, de modo que sejam visualizados com a máxima nitidez. O número de estruturas visíveis deve permanecer consistente ao longo das avaliações sucessivas, servindo como referência para monitorar possíveis variações na sensibilidade do sistema ao longo do tempo.

O teste deve ser realizado anualmente, na avaliação inicial do equipamento ou sempre que houver modificações significativas no sistema. A execução é de responsabilidade de um especialista em física médica ou profissional capacitado.

Se houver discrepâncias nos resultados em relação aos valores de referência, pode ser um indicativo de degradação do equipamento ou necessidade de ajustes técnicos, sendo recomendada uma investigação mais aprofundada.

- Protocolo INTERNATIONAL STANDARD IEC 61391-2:2010(International Electrotechnical Commission, 2010)

Este protocolo apresenta métodos padronizados para avaliar o desempenho de equipamentos de ultrassonografia de modo-B, abrangendo parâmetros relacionados à sensibilidade do sistema e à capacidade de diferenciação de níveis de intensidade de sinal. Dentre os testes descritos, destaca-se a avaliação da faixa dinâmica local (local *dynamic range*), que se refere ao intervalo de variação da amplitude do eco que o equipamento é capaz de representar na imagem, sem que haja saturação ou perda de informação. Esse parâmetro está diretamente relacionado à capacidade do sistema em exibir variações de intensidade na escala de cinza, característica essencial para a avaliação da qualidade da imagem ultrassonográfica.

Neste protocolo é descrito uma abordagem, que utiliza *phantoms* com inclusões para estimar a capacidade do sistema em diferenciar variações de intensidade na imagem.

Nesse método, o teste é realizado utilizando um *phantom* que contém inclusões esféricas ou cilíndricas de material acústico, cujos coeficientes de densidade são conhecidos e variam em relação ao material de fundo. O protocolo sugere a utilização de inclusões que apresentam contrastes típicos de +15 *dB*, +6 *dB*, +3 *dB*, -3 *dB* e -6 *dB* em relação ao material de fundo. A avaliação consiste em adquirir imagens do *phantom* com as configurações do equipamento ajustadas para uma curva de compressão logarítmica linear, de modo que a relação entre a amplitude do eco e o nível de cinza seja uniforme ao longo da faixa dinâmica.

Após a aquisição das imagens, são definidas regiões de interesse (*ROIs*) sobre as inclusões para calcular o valor médio dos *pixels* na região. Para evitar interferências causadas por sombreamento acústico, recomenda-se que apenas a metade superior

das inclusões seja utilizada na análise. Os valores médios obtidos para cada inclusão são então plotados em função do contraste de densidade, criando uma curva que representa a resposta do sistema para diferentes níveis de sinal.

A partir dessa curva, é possível estimar a faixa dinâmica local por meio de uma regressão linear, que fornece a relação entre os níveis de contraste e os valores de *pixel*. A faixa dinâmica é então definida como o intervalo de contraste que resulta em variações na escala de cinza desde o nível mínimo detectável até o nível máximo exibido pelo equipamento. Este proocolonão define explicitamente critérios rígidos de aprovação ou reprovação para o teste de local *dynamic range*, pois, seu objetivo é fornecer uma metodologia padronizada para a avaliação comparativa entre equipamentos, além de servir como ferramenta de monitoramento da qualidade ao longo do tempo.

3.4 O teste de escala de cinza na legislação brasileira

No Brasil, a Resolução RDC nº 611, de 9 de março de 2022, da Agência Nacional de Vigilância Sanitária (ANVISA) (BRASIL, 2022), estabelece os requisitos sanitários para a organização e o funcionamento de serviços de radiologia diagnóstica ou intervencionista. Embora a RDC 611 abranja uma variedade de modalidades de imagem, incluindo ultrassonografia, ela não especifica detalhadamente os procedimentos ou parâmetros para testes de escala de cinza em equipamentos de ultrassom. A resolução enfatiza a necessidade de um Programa de Garantia da Qualidade (PGQ), que deve incluir avaliações rotineiras dos parâmetros técnicos e de desempenho dos equipamentos, conhecidas como testes de constância.

Para complementar a RDC 611, a ANVISA publicou a Instrução Normativa (IN) nº 96 (BRASIL, 2021), que dispõe sobre os requisitos sanitários para a garantia da qualidade e da segurança em sistemas de ultrassom diagnóstico ou intervencionista. Embora a IN nº 96 forneça orientações mais específicas sobre os procedimentos de controle de qualidade, incluindo a avaliação da resolução e do contraste das imagens, ela não detalha os parâmetros específicos ou a nomenclatura utilizada para os testes de escala de cinza, nela trazida com o nome Limiar de sensibilidade a baixo contraste, somente trazendo como periodicidade do teste uma aceitação anual e quando houver reparos e como tolerância da normativa para esse teste a informação “Segundo especificações do fabricante do equipamento e do simulador”

Dessa forma, apesar das normativas nacionais estabelecerem a obrigatoriedade de programas de garantia da qualidade e testes regulares de desempenho para equipamentos de ultrassom, há uma lacuna em relação à padronização dos procedimentos e parâmetros específicos para a avaliação da escala de cinza. Isso ressalta a necessidade de desenvolvimento de diretrizes mais detalhadas ou a adoção de padrões internacionais para assegurar a consistência e a objetividade na avaliação da qualidade das imagens ultrassonográficas.

Embora existam divergências quanto à metodologia de aplicação e à nomenclatura dos testes, todos compartilham o objetivo de possibilitar a visualização simultânea de ecos com diferentes intensidades na mesma imagem, o que contribui para uma diferenciação mais precisa entre os diversos tecidos e estruturas.

3.5 Resumo comparativo de manuais e protocolos

A Tabela 1 resume as informações elencadas nas seções 3.3, 3.4 e 3.5. O levantamento apresentado nestas seções deixa evidente que, embora existam divergências quanto à metodologia de aplicação e à nomenclatura dos testes, todos compartilham o objetivo de possibilitar a visualização simultânea de ecos com diferentes intensidades na mesma imagem, o que contribui para uma diferenciação mais precisa entre os diversos tecidos e estruturas.

Tabela 1: Resumo dos manuais de *phantos* e protocolos de controle de qualidade a respeito do teste de escala de cinza e alcance dinâmico e testes similares.

Documento Base	Nome do teste no documento	Descrição da região no simulador	Descrição da análise
***Phantom* ATS 539**	Escala de cinza e alcance dinâmico exibido	Grupo de alvos de escala de cinza com 6 cilindros variando de +15db, +6db, +3db, -3db, -6db e -15db.	Avalia a nitidez, formato circular e variação de brilho dos alvos, sem critérios quantitativos definidos.

***Phantom* CIRS 040GSE**	*Gray scale targets*	Dois grupos de alvos cilíndricos posicionados a 3 cm e 11,5 cm de profundidade, em que cada grupo contém 6 cilindros variando os contrastes de -9 *dB* a +15 *dB*.	Verificar a capacidade de distinguir contrastes em diferentes profundidades.
***Phantom* CIRS 054GS**	*Grayscale contrast sensitivity*	Seis cilindros (8 mm de diâmetro) a 4 cm de profundidade, variando de anecoico a hiperecoico.	Avaliar a faixa dinâmica e a representação de tamanho/formato dos alvos.
***Phantom* Gammex 403 GS LE**	*Grey Scale Targets*	Alvos posicionados a 6 cm de profundidade, com 4 cilindros variando o contraste de -6 *dB*, +6 *dB* e alta dispersão.	Análise quantitativa (média de *pixels*) ou qualitativa (comparação visual com barra de cinza).
***Phantom* Konex US-MP**	Sensibilidade de baixo contraste	Cilindros ecogênicos (5,5 mm de diâmetro) distribuídos entre 50 mm e 110 mm de profundidade.	Observação da forma circular, variação de brilho e presença de artefatos como sombreamento.
***Quality Assurance* of Ultrassond**	*Grey-Scale Performance*	Menciona que deve conter alvos cilíndricos ou prismáticos com	Análise visual em 3 classificações distintas Não visível

Imaging Systems		diferentes níveis de ecogenicidade	Visível, mas indistinto Claramente visível
Protocolo Español de control de calidad en radiodiagnóstico	*Umbral de sensibilidad a bajo contraste*	O protocolo sugere *phantoms* com diferentes níveis de contraste para avaliar a capacidade do sistema em distinguir variações de intensidade.	O protocolo recomenda a verificação da capacidade do sistema em distinguir variações sutis de intensidade
Protocolo international standard iec 61391-2:2010	*local dynamic range*	O protocolo sugere a utilização de inclusões que apresentam contrastes típicos de +15 *dB*, +6 *dB*, +3 *dB*, -3 *dB* e -6 *dB* em relação ao material de fundo.	Regressão linear entre contrastes nominais e níveis de cinza para estimar a faixa dinâmica.
Resolução RDC nº 611/2022	Limiar de sensibilidade ao baixo contraste	-	Segundo especificações do fabricante do equipamento e do simulador

3.6 Processamento das imagens

Conforme mencionado, o desafio deste estudo consiste na falta de padronização e heterogeneidade das imagens. Assim, para lidar com esse problema, a metodologia desenvolvida neste trabalho é centrada em técnicas de processamento de imagens. Nesta seção, será apresentada a fundamentação das principais técnicas de processamento utilizadas ao longo do trabalho. A descrição do uso de cada uma delas será a bordado na seção de metodologia.

Segundo (Gonzalez, 2018), o processamento de imagem refere-se a um conjunto de técnicas aplicadas a imagens digitais para melhorar sua qualidade, extrair informações relevantes ou facilitar sua interpretação. Esse processo pode envolver operações como realce de contraste, filtragem, segmentação, restauração e transformação de formatos, sendo amplamente utilizado em áreas como medicina, vigilância, sensoriamento remoto e reconhecimento de padrões. No contexto deste estudo, o processamento de imagens desempenha um papel fundamental na análise de imagens de ultrassom, especialmente em aplicações voltadas à automação de testes de qualidade. Sua relevância está na capacidade de transformar dados subjetivos em métricas objetivas e reprodutíveis, permitindo uma avaliação mais precisa dos parâmetros de qualidade.

Para lidar com essas questões, diversas técnicas de processamento podem ser empregadas. Neste trabalho o enfoque são técnicas para redução de ruído, realce de bordas, segmentação e extração de características, considerado que, conforme já mencionado nas seções anteriores, a qualidade das imagens está diretamente relacionada a análise da forma e contraste dos alvos cilíndricos imageado.

3.6.1 Filtros no processamento de imagens

Filtros são operações matemáticas aplicadas às imagens para realçar características específicas ou remover informações indesejadas. Eles são essenciais para preparar as imagens para etapas subsequentes de análise (Gonzalez, 2018). No contexto de ultrassom, destacam-se filtros para redução de ruído (como filtro e mediano) e detecção de bordas devido às particularidades das imagens deste trabalho: baixo contraste, presença de ruído *speckle* e necessidade de precisão na delimitação dos alvos em *phantoms.*

3.6.1.1 Ruído *speckle*

O ruído speckle é um padrão granular de interferência inerente a sistemas de imageamento coerente, como o ultrassom, sendo resultado da interação construtiva e destrutiva de ondas acústicas retroespalhadas por estruturas menores que o comprimento de onda do feixe transmitido (Goodman, 1976) . Diferentemente de ruídos aditivos convencionais, o speckle possui natureza multiplicativa, o que significa que sua intensidade varia proporcionalmente ao sinal local, tornando sua remoção

mais complexa. Em imagens de ultrassom Modo B, o speckle manifesta-se como flutuações aleatórias de intensidade que degradam o contraste entre regiões adjacentes, dificultando a delimitação precisa de bordas e a identificação de estruturas de baixo contraste, como os alvos cilíndricos presentes em *phantoms* de controle de qualidade (Wagner et al., 1983). A presença desse ruído impõe desafios significativos ao processamento automático de imagens, exigindo etapas específicas de filtragem para viabilizar a detecção confiável dos alvos de interesse.

3.6.1.2 Redução de ruídos

A presença de ruído pode comprometer a identificação de estruturas e a precisão dos testes. Alguns dos principais filtros para atenuação de ruído incluem os filtros de mediana e gaussiano, bilateral

O filtro da mediana é um operador espacial não-linear que substitui o valor de cada *pixel* pela mediana estatística dos valores contidos em uma vizinhança pré-definida (HARALICK; DINSTEIN; SHANMUGAM, 1973).

Seja $f(x, y)$ a imagem de entrada e $g(x, y)$ a imagem filtrada resultante. O filtro espacial da mediana é formalmente definido como:

$$g(x, y) = mediana\{f(x + s, y + t)\} \quad (1)$$

$$(s, t) \in W, s \in [-a, a], t \in [-b, b] \quad (2)$$

Onde W é a janela da vizinhança.

O filtro Gaussiano realiza a suavização da imagem por meio de uma operação de convolução baseada na função gaussiana bidimensional, definida por (Gonzalez, 2018):

$$G(x, y) = \frac{1}{2\pi\sigma^2} \cdot e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (3)$$

onde x e y representam as coordenadas espaciais da imagem e σ é o desvio padrão da distribuição gaussiana, que controla a intensidade da suavização. Quanto maior o valor de σ, maior será o efeito de desfoque aplicado à imagem.

A suavização gaussiana reduz variações abruptas de intensidade, eliminando ruídos e pequenas imperfeições, mas preservando as transições de maior contraste. No entanto, um efeito colateral desse método é a possível suavização excessiva de detalhes finos, o que pode comprometer a identificação de pequenas estruturas ou bordas sutis em imagens médicas ou científicas (DUDA; HART; STORK, 2001).

Embora o filtro Gaussiano seja eficiente na redução de ruídos, sua principal limitação reside na suavização indiscriminada da imagem, o que pode degradar bordas e detalhes estruturais relevantes. Para superar essa limitação, (Tomasi; Manduchi, *[S.d.]*) propuseram o filtro bilateral, um operador não-linear que suaviza a imagem preservando as bordas por meio de uma combinação ponderada dos valores dos *pixels* vizinhos, levando em conta simultaneamente a proximidade espacial e a similaridade fotométrica entre os *pixels*.

O filtro bilateral pode ser utilizado como uma alternativa para reduzir o ruído *speckle* preservando transições espaciais relevantes. Diferentemente de uma suavização linear simples, o filtro bilateral pondera a contribuição dos *pixels* por dois critérios simultâneos: proximidade espacial e similaridade de intensidade. Assim, *pixels* próximos e com intensidade semelhante contribuem mais para o valor filtrado, enquanto *pixels* com intensidade muito diferente têm contribuição reduzida, realizado pelas equações (4) e (5).

$$I_f(x) = \frac{1}{W(x)} \sum_{y \in \Omega(x)} I(y)\ w_s(\| x - y \|)\, w_r(|\ I(x) - I(y)\ |) \qquad (4)$$

$$W(x) = \sum_{y \in \Omega(x)} w_s\ (\| x - y \|)\, w_r(|\ I(x) - I(y)\ |) \qquad (5)$$

Nas equações (4) e (5), $I_f(x)$ representa a intensidade filtrada no *pixel* x enquanto $I(y)$ corresponde à intensidade dos *pixels* vizinhos pertencentes à vizinhança $\Omega(x)$. Os termos w_s e w_r representam, respectivamente, as funções de ponderação associadas à distância espacial e à diferença de intensidade entre *pixels*. O termo $W(x)$ corresponde ao fator de normalização responsável por manter os pesos adequadamente escalados.

Essa propriedade é particularmente importante em imagens de ultrassom, pois o *speckle* produz oscilações locais que podem gerar falsas bordas, mas a localização dos alvos depende da preservação das transições entre o alvo e o fundo do *phantom.*

3.6.1.3 Realce de bordas

A detecção de bordas consiste na identificação de regiões da imagem onde ocorrem variações abruptas de intensidade, correspondendo frequentemente a contornos de objetos e estruturas relevantes. Entre os métodos mais utilizados destacam-se o filtro de *Canny* e o operador *Laplaciano*.

O filtro de *Canny* é um detector de bordas baseado em múltiplas etapas: suavização gaussiana para redução de ruído, cálculo do gradiente de intensidade, supressão de não-máximos e limiarização por histerese. A magnitude do gradiente é calculada a partir das derivadas parciais da imagem suavizada, definida por (Canny, 1986).

$$M(x,y) = \sqrt{Gx^2 + Gy^2} \quad (6)$$

onde Gx e Gy representam as respostas do gradiente nas direções horizontal e vertical, respectivamente, obtidas pela convolução da imagem com máscaras derivativas.

O operador *Laplaciano* é um filtro diferencial de segunda ordem que detecta bordas por meio do cálculo da segunda derivada da intensidade da imagem. Para uma imagem $f(x,y)$, o *Laplaciano* é definido como (Gonzalez, 2018)

$$\nabla^2 f(x,y) = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} \quad (7)$$

que em sua forma discreta, aplicável a imagens digitais, é aproximado por:

$$\nabla^2 f(x,y) = f(x+1,y) + f(x-1,y) + f(x,y+1) + f(x,y-1) - 4f(x,y) \quad (8)$$

Regiões de borda são identificadas pelos pontos de cruzamento por zero (zero-crossings) da resposta do operador. Por ser sensível a ruídos, o *Laplaciano* é frequentemente precedido de suavização gaussiana, combinação conhecida como *Laplaciano* do Gaussiano (LoG) (Gonzalez, 2018).

3.6.1.4 Equalização de histogramas

A equalização de histograma é uma técnica de processamento de imagens que busca melhorar o contraste por meio da redistribuição dos níveis de intensidade ao longo de toda a faixa dinâmica disponível. Essa abordagem é especialmente relevante em imagens de ultrassom, nas quais certas regiões podem apresentar baixo contraste, dificultando a identificação de estruturas anatômicas e padrões de interesse (Gonzalez, 2018).

Matematicamente, a técnica baseia-se na aplicação de uma função de transformação derivada da função de distribuição cumulativa (FDA) dos níveis de intensidade da imagem. Seja $f(x,y)$ uma imagem com L níveis de intensidade possíveis e n o número total de *pixels*. A probabilidade de ocorrência de cada nível de intensidade r_k é estimada por:

$$p(r_k) = \frac{n_k}{} n, k = 0, 1, \dots, L-1 \quad (9)$$

onde n_k é o número de *pixels* com intensidade r_k. A transformação equalizadora é então definida como:

$$s_k = (L-1) \cdot \Sigma\, p(r_j), j = 0, 1, \dots, k \quad (10)$$

O resultado é uma imagem com distribuição de intensidades aproximadamente uniforme, o que amplia o contraste global e torna a imagem mais adequada para análises quantitativas (Gonzalez, 2018)

Embora eficiente para aprimoramento global do contraste, a equalização de histograma convencional pode amplificar ruídos e produzir resultados inadequados em imagens com variações locais de iluminação ou contraste, como ocorre frequentemente em imagens médicas (Rangaraj M. Rangayyan, 2004). Para contornar essa limitação, (Zuiderveld, 1994) propôs a Equalização de Histograma Adaptativa com Limite de Contraste (*CLAHE*, do inglês *Contrast Limited Adaptive Histogram Equalization*), originalmente desenvolvida para imagens médicas, demonstrando desempenho superior no realce de imagens de baixo contraste.

Ao contrário da equalização global, o *CLAHE* opera sobre sub-regiões locais da imagem, denominadas *tiles*, aplicando a equalização de forma independente em cada bloco. Esse processamento local torna o método eficaz na redução de ruídos e no aprimoramento da qualidade da imagem. Para evitar a superamplificação do contraste, o método impõe um limite máximo (*clip limit*) ao histograma local antes do cálculo da FDA, redistribuindo os valores excedentes uniformemente entre os demais níveis. A transformação aplicada a cada tile é definida por:

$$s_k = s_{\min + \left(\frac{(s_{\max} - s_{\min})}{}n\right)} \cdot \Sigma\, p_{clip(r_j)}, j = 0, 1, \dots, k \quad (11)$$

onde p_{clip} representa o histograma local após a aplicação do limite de contraste, $s_{\min} s_{\max}$ lores mínimo e máximo da faixa de saída, e n o número total de *pixels* no tile. A interpolação bilinear entre tiles adjacentes garante a continuidade espacial da imagem resultante (Zuiderveld, 1994).

3.6.2 Filtros no Processamento de Sinais

No processamento de sinais unidimensionais, a suavização é uma etapa fundamental para reduzir ruído e variações de alta frequência sem comprometer as características macroscópicas do sinal, como posição de máximos, mínimos e pontos

de inflexão. Dois filtros amplamente utilizados para essa finalidade são o filtro Savitzky-Golay e o filtro Butterworth passa-baixa.

O filtro Savitzky-Golay opera no domínio espacial por meio do ajuste local de um polinômio de grau definido a uma janela deslizante de amostras consecutivas. Os coeficientes resultantes desse ajuste são aplicados como uma convolução sobre o sinal original, produzindo uma versão suavizada que preserva a forma e a posição de picos e inflexões com maior fidelidade do que uma média móvel simples, conforme a Equação 19 (Rangaraj M. Rangayyan, 2004).

$$S_f(x_i) = \sum_{k=-m}^{m} c_k \; S(x_{i+k}) \qquad (19)$$

O filtro Butterworth passa-baixa, por sua vez, atua no domínio da frequência. Ele atenua progressivamente as componentes de frequência acima de uma frequência de corte definida, com uma transição suave entre a banda de passagem e a banda de rejeição. A ordem do filtro controla a taxa de atenuação: ordens mais altas produzem um corte mais abrupto. Essa característica torna o filtro Butterworth adequado para eliminar componentes associadas a ruído sem introduzir oscilações (ringing) no sinal filtrado (Gonzalez, 2018).

$$X(k) = \sum_{\{n=0\}}^{\{N-1\}} x(n) e^{\left\{-\frac{j2\pi kn}{N}\right\}} \qquad (20)$$

A seleção da frequência de corte pode ser guiada pela análise espectral do sinal, realizada por meio da Transformada Rápida de Fourier (FFT). A FFT decompõe um sinal discreto no domínio espacial em suas componentes de frequência, conforme a Equação 20, permitindo identificar quais faixas concentram a energia associada à estrutura macroscópica do sinal e quais correspondem a variações de alta frequência atribuíveis a ruído. A partir da distribuição de energia acumulada no espectro, é possível definir uma frequência de corte que preserve as componentes estruturais e suprima as componentes ruidosas (Rangaraj M. Rangayyan, 2004).

3.6.3 Segmentação de imagens

A segmentação de imagens consiste no processo de particionamento da imagem em regiões semanticamente significativas, permitindo isolar regiões de interesse (*ROIs*) e diferenciar estruturas presentes no plano de fundo. Em imagens de ultrassom, esse processo apresenta desafios particulares devido à presença de ruído *speckle* (um ruído granular característico de sistemas de imageamento coerente, como ultrassom, radar e laser, originado pela interferência construtiva e destrutiva das ondas sonoras refletidas pelas microestruturas dos tecidos), bem como pela baixa relação sinal-ruído e pelas variações na intensidade do eco (Gonzalez, 2018).

3.6.3.1 Transformações Morfológicas

As operações morfológicas são técnicas de processamento baseadas na teoria dos conjuntos, aplicadas sobre a geometria das estruturas presentes na imagem com o auxílio de um elemento estruturante B. As duas operações fundamentais são a erosão e a dilatação (Gonzalez, 2018).

A erosão de uma imagem $f(x, y)$pelo elemento estruturante B é definida como:

$$f \ominus B = \{ z \mid (B)z \subseteq f \} \quad (12)$$

onde $(B)z$ representa o elemento estruturante transladado para a posição z. A erosão reduz regiões segmentadas e elimina pequenos elementos indesejados. A dilatação, operação dual à erosão, é definida como:

$$f \oplus B = \{ z \mid (\hat{B})z \cap f \neq \emptyset \} \quad (13)$$

onde $\hat{B}$ é o elemento estruturante refletido. A dilatação expande as regiões segmentadas, preenchendo lacunas e conectando componentes próximos. A combinação sequencial dessas operações dá origem a operações compostas de maior utilidade prática: a abertura morfológica $(f \ominus B) \oplus B$, que remove pequenos objetos preservando a forma geral das estruturas, e o fechamento morfológico $(f \oplus B) \ominus B$, que preenche buracos internos e suaviza contornos (Gonzalez, 2018).

3.6.3.2 *Transformada de Hough*

A *Transformada de Hough* foi originalmente proposta por Duda; Hart, 1972como um procedimento computacionalmente eficiente para a detecção de retas em imagens. Duda; Hart, 1972 aprimoraram o método ao adotar a parametrização por ângulo e raio (ρ, θ) em lugar da representação por coeficiente angular e intercepto,

tornando a transformada aplicável a retas de qualquer orientação, incluindo linhas verticais. Essa parametrização é definida por:

$$\rho = x \cdot \cos(\theta) + y \cdot \sin(\theta) \quad (14)$$

onde ρ é a distância perpendicular da reta à origem e θ é o ângulo de inclinação da normal à reta, com $\theta \in [0°, 180°)$. Cada ponto de borda (x, y) da imagem contribui com uma curva senoidal no espaço de parâmetros (ρ, θ), denominado espaço de acumulação ou espaço de *Hough*. O princípio fundamental é que cada ponto detectado vota no conjunto de parâmetros correspondente a todas as retas possíveis que passam por aquele ponto. Quando múltiplos pontos estão alinhados, os parâmetros daquela reta acumulam um número elevado de votos, indicando suporte significativo para sua detecção.

A detecção de estruturas circulares, também relevante em imagens de ultrassom, é realizada pela *Transformada de Hough* Circular, parametrizada por:

$$(x - a)^2 + (y - b)^2 = r^2 \quad (15)$$

onde (a, b) é o centro e r o raio do círculo. Nesse caso, o espaço de acumulação é tridimensional, e cada ponto de borda gera um cone no espaço (a, b, r) (Duda; Hart, 1972).

4 Trabalhos Relacionados

Com o objetivo de se avaliar o estado da arte acerca de métodos computacionais aplicados ao teste de escala de cinza e alcance dinâmico exibido (e testes com outras nomenclaturas, mas objetivos similares) uma revisão da literatura foi realizada, seguindo uma metodologia estruturada em seis etapas conforme proposta por (Tavares De Souza; Dias Da Silva; De Carvalho, 2010)

1. **Identificação do problema de pesquisa:** definição das questões norteadoras, seleção de palavras-chave, formulação da equação de busca e escolha das bases de dados.
2. **Estabelecimento dos critérios de inclusão e exclusão:** definição dos critérios e das bases de busca.
3. **Identificação dos estudos pré-selecionados:** triagem inicial e seleção final dos estudos relevantes.
4. **Análise e categorização dos estudos selecionados:** organização dos artigos conforme suas temáticas e metodologias.

5. **Análise e interpretação dos estudos:** avaliação crítica dos achados.

As subseções a seguir detalham os desenvolvimentos realizados em relação a cada uma das etapas da metodologia.

4.1 IDENTIFICAÇÃO DO PROBLEMA DE PESQUISA

Com base na metodologia descrita, as questões norteadoras para a pesquisa foram formuladas da seguinte maneira:

- Quais metodologias estão sendo utilizadas para realizar o teste de escala de cinza e alcance dinâmico exibido (e similares) no controle de qualidade de equipamentos de ultrassom no *modo B*?
- Existem métodos computacionais capazes de analisar os testes de forma automatizada?
- Como são obtidas as imagens para a realização dos testes?
- De que forma os resultados dos testes são classificados e interpretados?

Baseado nas perguntas norteadores, foram desenvolvidas as seguintes equações de palavras-chave para busca nas bases de dados científicas: *“ultrasound AND B-mode AND phantom AND "Image quality" AND ("quality assurance" OR "quality control")”.* Estes termos deveriam estar contidos no título, palavras-chave, resumo ou introdução.

Essas palavras-chave foram selecionadas para filtrar artigos que apresentavam métodos computacionais e não computacionais, de modo que a seleção dos computacionais foi realizada após a aplicação dos critérios de inclusão e exclusão. Isso foi realizado devido à dificuldade de se encontrar termos restritivos e genéricos que englobassem apenas métodos computacionais na busca. Os termos *ultrasound AND B-mode* foram incluídos para garantir que os estudos estivessem focados em imagens de ultrassom no modo bidimensional (*modo B*), que é amplamente utilizado na avaliação clínica. Além disso, os termos *quality assurance OR quality OR quality control* foram aplicados para garantir que os estudos envolvessem controle de qualidade e avaliações de desempenho dos equipamentos de ultrassom.

Levando em conta que existem métodos com nomenclaturas distintas e objetivos similares, termos como *grayscale dynamic range econtrast resolution* não

foram incluídos na busca, apesar de estarem diretamente relacionados à análise de contraste em tons de cinza, aspecto fundamental no teste estudado, devido à tendência de esses termos restringirem excessivamente a busca. Muitos artigos relevantes, que aplicam testes de controle de qualidade em ultrassom, podem não mencionar esses termos explicitamente no título, palavras-chave, resumo ou introdução, embora abordem tais aspectos no conteúdo do estudo.

A equação de busca foi aplicada nas bases de dados PubMed, IEEE Xplore e ScienceDirect.

4.2 CRITÉRIOS DE INCLUSÃO E EXCLUSÃO E IDENTIFICAÇÃO DOS ESTUDOS PRÉ-SELECIONADOS

Considerando os aspectos descritos na seção 4.1, foram adicionados aos critérios de exclusão as condições específicas para filtrar estudos que não abordam análises que envolvam contraste em tons de cinza ou resolução de contraste.

Assim, foram estabelecidos os critérios de inclusão e exclusão, conforme detalhado na Tabela 2. Primeiramente foi feito uma leitura dos *Abstracts* dos artigos resultantes da *string* de busca, com a finalidade de classificar os artigos de acordo com os critérios de exclusão e Inclusão.

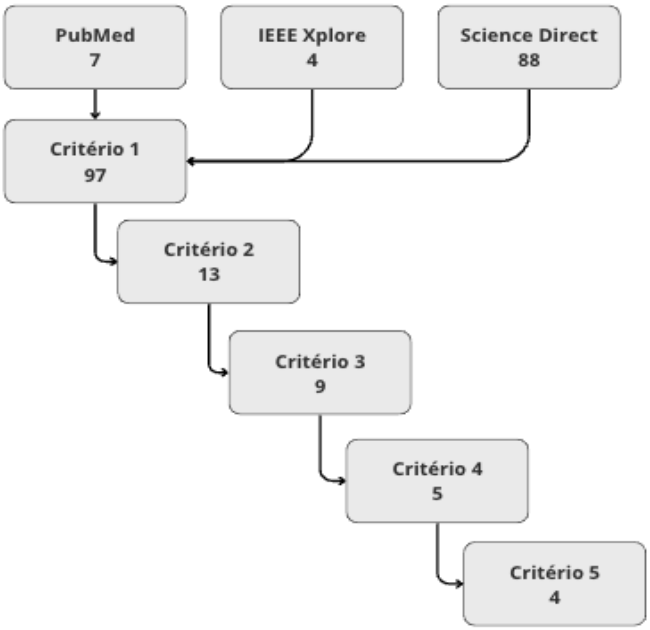
Tabela 2: Tabela apresentando os critérios de inclusão e exclusão da revisão realizada

Critérios de Inclusão	Critérios de Exclusão
Apenas artigos publicados em periódicos indexados	Critério 1 Artigos em duplicatas
Publicados de 2015 até 2026	Critério 2 Artigo que não aborda métodos para controle de qualidade em equipamentos de ultrassom diagnóstico ou intervencionista.
Linguagem de português, inglês ou espanhol	Critério 3 Este artigo trata-se de um estudo focado na prototipação de *Phantom*.
Artigo Original	Critério 4 Artigos que não trata sobre ultrassom de *modo B*

--	Critério 5 Este artigo não aborda testes de CQ envolvendo a escala de cinza e/ou visualização de objetos de baixa frequência e/ou contraste

Ao final da análise dos artigos retornados pela equação de buscas e na aplicação dos critérios de inclusão e exclusão, restaram-se 5 (Figura 8)

Figura 8: Fluxograma da aplicação dos critérios na busca de artigos.



4.3 ANÁLISE E CATEGORIZAÇÃO DOS ESTUDOS SELECIONADOS

Os estudos selecionados na Figura 8 foram categorizados de acordo com as 3 perguntas abaixo:

1) Questão 1: Modelo de simulador utilizado para o teste alvo;

2) Questão 2: A abordagem do método que auxilia no controle de qualidade de imagem é manual, semiautomática ou automática?

3) Questão 3: Utiliza parâmetros operacionais como dados de entrada, além da imagem?

A Tabela 3 apresenta um resumo dos artigos e suas respectivas categorizações de acordo com as 3 perguntas apresentadas anteriormente.

Tabela 3: Análise e categorização dos artigos selecionados.

Publicação	Questão 1	Questão 2	Questão 3
PAVERD et al., (2024)	040GSE	Manual com suporte de *software*	Sim.
TSAPAKI et al. (2021)	Gammex 403GS LE,	Manual	Sim
SASSAROLI et al. (2019)	040GSE	semiautomática	Sim
SASSAROLI et al. (2017)	Não especificado	semiautomática	Não

4.4 ANÁLISE E INTERPRETAÇÃO DOS ESTUDOS

4.4.1 METODOLOGIA DE SASSAROLI ET AL. (2017)

Sassaroli et al., (2017) desenvolveu um estudo abrangente sobre a tecnologia de ultrassom de mama e a avaliação do desempenho de equipamentos de ultrassom em *modo B* com objetivo de desenvolver e propor metodologias mais eficazes para o CQ de sistemas de ultrassom, especialmente no contexto de imagens de mama em *modo B*. O autor enfatiza a necessidade de substituir os métodos de CQ atuais, que muitas vezes são subjetivos, por abordagens mais objetivas.

Entre os testes de CQ discutidos, além do teste de baixo contraste, que mede a capacidade do sistema de ultrassom de diferenciar sutis variações na ecogenicidade entre estruturas próximas, o autor também menciona outros testes essenciais, como o de uniformidade, que avalia a consistência das imagens obtidas, e o teste de sensibilidade, que quantifica a profundidade máxima de visualização dos sinais de eco em relação ao ruído eletrônico. Além disso, são abordados o teste de resolução espacial, que examina a capacidade do sistema em distinguir entre alvos próximos, e o teste de precisão na medição de distância, que verifica se as medidas feitas pelo sistema estão corretas em comparação com padrões conhecidos.

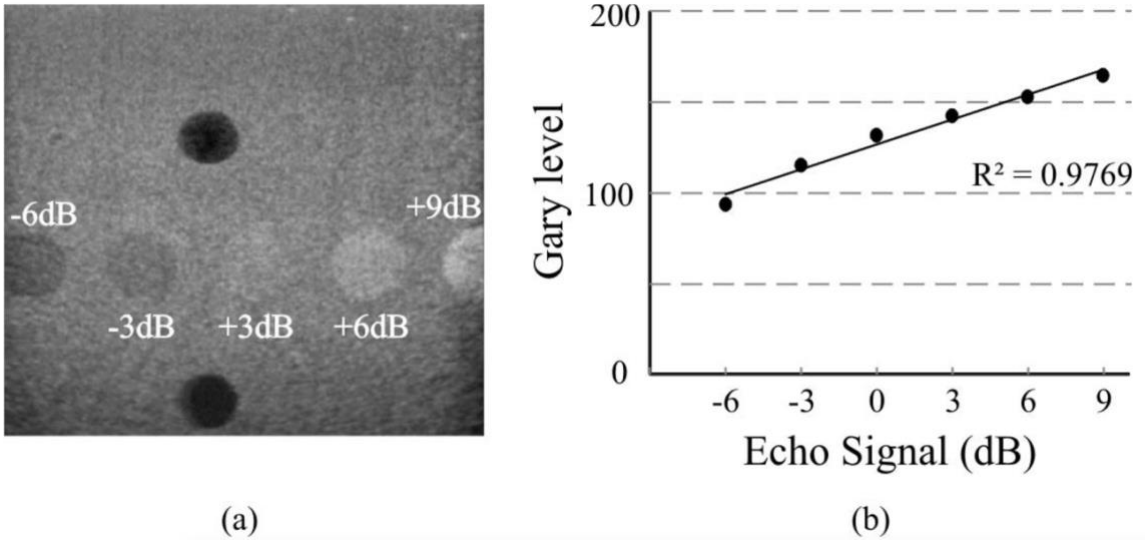
O método de teste de baixo contraste proposto no artigo mede a capacidade do sistema de ultrassom de diferenciar variações sutis na ecogenicidade entre estruturas próximas. Para realizar o teste, foram utilizados *phantoms* com alvos de contraste embutidos, consistindo em cilindros com diferentes níveis de ecogenicidade, variando de +9 *dB* a -9 *dB* [Figura 9 (a)]. A metodologia possui uma etapa manual

relacionada a definição de *ROIs* sobre a área de cada cilindro no *phantom*. Nesta etapa, são realizados ajustes manuais para garantir que a *ROI* esteja corretamente posicionada sobre o alvo de contraste, a fim de evitar interferências, como artefatos de sombreamento.

Uma vez definida a *ROI*, a análise é computacional. O sistema mediu níveis de cinza dentro da *ROI* de forma automatizada, fornecendo uma média de intensidade de *pixels* para cada nível de contraste do cilindro, sendo uma média para cada *ROI*. Esses dados de intensidade de cinza foram analisados por meio de um *software* de processamento de imagem, desenvolvido pelos autores, que utilizou regressão linear para relacionar os valores de contraste nominal dos cilindros (em *dB*) com os níveis de cinza medidos. A inclinação da regressão linear (Figura 9 (b)) é o parâmetro que indica a capacidade do sistema de detectar pequenas diferenças de contraste. Um sistema com maior inclinação na regressão linear demonstra maior sensibilidade ao contraste.

Além da análise da inclinação da regressão linear, o *software* também calculou a faixa dinâmica local do sistema de ultrassom, um parâmetro que indica a capacidade do equipamento em detectar e diferenciar níveis de contraste dentro de um intervalo específico. Essa faixa representa o intervalo de variação de contraste no qual o sistema consegue distinguir diferenças de intensidade de forma confiável, sem que as variações de brilho na imagem sejam comprimidas ou distorcidas pelo processamento do equipamento. Para determinar essa faixa, os autores analisaram a região linear da curva de mapeamento de escala de cinza *(do inglês Grayscale Mapping Function – GMF*), que descreve como os valores reais de contraste dos cilindros no *phantom* são convertidos em níveis de cinza na imagem. Dessa forma, a faixa dinâmica local reflete o limite no qual o sistema mantém a fidelidade das diferenças de contraste, garantindo que estruturas com pequenas variações de ecogenicidade possam ser visualizadas adequadamente sem perda de detalhe ou saturação na escala de cinza.

Figura 9: (a) Alvos de contraste em tons de cinza em um *phantom* generalista. (b) Valores de *pixels* em tons de cinza versus valores nominais de contraste dos alvos (Sassaroli et al., 2017).



A abordagem proposta pelos autores é mais objetiva e padronizada, comparando aos manuais dos *phantoms*, visando superar as limitações dos métodos subjetivos utilizados atualmente.

Embora o estudo demonstre a eficácia do método na quantificação da resposta do sistema ao contraste, não são apresentados valores de acurácia ou métricas quantitativas de desempenho, como o coeficiente de determinação (R^2) da regressão linear ou erros médios nas medições. A ausência desses indicadores limita a avaliação objetiva da precisão e confiabilidade do modelo proposto.

Apesar da eficácia do método de teste de baixo contraste proposto no artigo, algumas lacunas ainda persistem, afetando a eficiência e precisão dos resultados. Em primeiro lugar, o processo de controle de qualidade não é totalmente automatizado, uma vez que a definição das *ROIs* deve ser realizada manualmente pelo operador. Essa dependência humana pode introduzir variabilidade nos resultados devido a diferenças na execução entre operadores, além de aumentar o tempo necessário para a realização dos testes.

Outra limitação importante é a ausência de um critério de aprovação automatizado. Embora os dados sejam analisados por meio de métodos computacionais, o sistema não gera automaticamente uma avaliação final sobre a conformidade do equipamento no controle de qualidade. A decisão ainda requer interpretação humana, o que pode resultar em inconsistências no julgamento dos resultados. A implementação de critérios automáticos de aprovação ou reprovação poderia eliminar essa subjetividade e tornar o processo mais confiável.

Outra limitação observada é a restrição na validação do método para diferentes tipos de transdutores. Embora o estudo tenha demonstrado a eficácia do procedimento para transdutores lineares, não há informações sobre sua aplicabilidade a transdutores curvilíneos ou matriciais, amplamente utilizados em diversas aplicações clínicas. Como esses transdutores possuem características acústicas e de formação de imagem diferentes, é fundamental que a metodologia seja validada para um conjunto mais amplo de equipamentos, garantindo maior generalização e aplicabilidade dos resultados obtidos. Essas lacunas indicam que, embora o método represente um avanço na padronização e objetividade dos testes de controle de qualidade, ele ainda precisa ser aprimorado para atingir níveis mais elevados de automação, flexibilidade e precisão. Melhorias nesse sentido tornariam o método mais robusto e adaptado às diversas exigências da prática clínica atual.

4.4.2 METODOLOGIA DE SASSAROLI et al. (2019)

Sassaroli et al., (2019) apresenta um método computacional para a avaliação da qualidade de imagem em modos avançados de ultrassonografia *B-mode*. Diferentemente do *modo B* convencional, que forma imagens bidimensionais a partir da amplitude dos ecos refletidos pelos tecidos, os modos avançados do *B-mode* incorporam técnicas adicionais para melhorar a qualidade da imagem, reduzindo artefatos e aprimorando a visualização de estruturas anatômicas. Esses modos incluem a imagem harmônica (HI), que explora frequências não lineares para aumentar a resolução e o contraste; a imagem composta espacial (SCI), que combina múltiplos ângulos de varredura para reduzir *speckle* e melhorar a definição de bordas; a redução adaptativa de *speckle* (SR), que aplica filtros para reduzir o ruído sem comprometer os detalhes estruturais; e a correção de aberração de tecido (TAC), que ajusta o tempo de propagação do som para compensar variações na velocidade do ultrassom em diferentes tecidos.

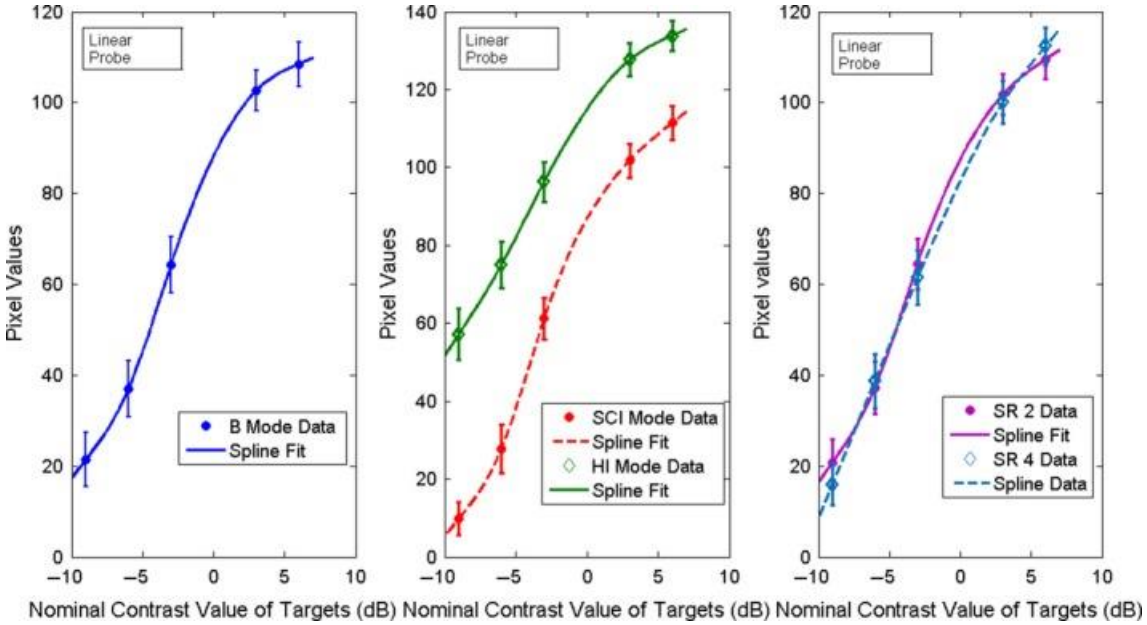
O objetivo do estudo é propor e avaliar métodos computacionais para a CQ desses modos avançados, uma vez que ainda não existem protocolos padronizados para sua validação. Os testes são focados na avaliação de parâmetros como função de mapeamento em escala de cinza, contraste de imagem, razão de contraste-ruído (CNR) e resolução espacial de alto contraste (HCSR). A pesquisa busca, assim, contribuir para o desenvolvimento de diretrizes objetivas e reprodutíveis para a

calibração e manutenção de sistemas de ultrassom que utilizam esses modos avançados, garantindo imagens de melhor qualidade para a prática clínica.

O artigo descreve métodos para diversos testes de qualidade de imagem, incluindo função de mapeamento em escala de cinza, contraste de imagem, razão de contraste-ruído (CNR) e resolução espacial de alto contraste (HCSR). Para a aquisição de imagens, foram utilizados sistemas de ultrassom, como o GE LOGIQ E9 e o Philips iU22, em combinação com *phantoms* comerciais (Modelo 040GSE, da CIRS, com material *Zerdine* que simula tecidos moles). O *phantom* possui alvos de contraste e filamentos de nylon, empregados para avaliar o desempenho dos modos avançados nas diferentes combinações de scanner e sonda.

No teste de contraste, a metodologia foi baseada na análise de alvos cilíndricos de contraste conhecido, dispostos em profundidades específicas no *phantom* 040GSE. O contraste foi calculado a partir dos valores de cinza dos *pixels* dentro de uma região de interesse (*ROI*) elíptica em torno de cada alvo definida manualmente, com os valores de contraste plotados em função do contraste nominal dos alvos (**Figura 10**) para avaliar a curva de mapeamento em escala de cinza (*GMF*).

Figura 10: Ajuste de curva de ajuste de *spline* da média ponderada e desvio padrão dos valores de *pixel* para os alvos de contraste com contraste nominal: −9,−6,−3,3,3,6 *dB* para os modos B, SCI, HI, SR 2, SR 4, para a sonda linear.



A aprovação ou desaprovação no teste de contraste é baseada na capacidade de o sistema diferenciar alvos com contrastes próximos, observando a presença de saturação nas regiões de contraste baixo ou alto. Em termos práticos, as imagens que apresentam uma curva *GMF* com uma região linear ampla e uma inclinação acentuada são consideradas satisfatórias, pois demonstram capacidade de resolver alvos de baixo contraste em uma ampla gama de intensidades.

Na conclusão, os autores destacam que os parâmetros *GMF*, CNR e HCSR demonstraram sensibilidade às variações nos modos avançados de ultrassonografia, tornando-se indicadores viáveis para a avaliação objetiva da qualidade da imagem. O estudo também revelou que a função de mapeamento em escala de cinza (*GMF*) varia entre os modos, com a imagem composta espacial (SCI) apresentando maior inclinação na região linear, sugerindo maior capacidade de diferenciação de contrastes sutis.

Para a avaliação, foram adquiridas entre 5 e 10 imagens por modo, utilizando diferentes combinações de transdutores e ajustes operacionais. A acurácia da metodologia foi analisada por meio da comparação entre os valores medidos e os contrastes nominais dos alvos, sendo observada uma variação máxima de ±1 *dB* nos valores de contraste. Além disso, a dispersão dos resultados foi avaliada pelo desvio padrão das medições, indicando que o CNR apresentou maior variabilidade para alvos de contraste positivo (+3 *dB*), enquanto os alvos hiperecóicos apresentaram maior estabilidade entre os testes.

Apesar da eficácia da metodologia, os autores identificaram limitações, principalmente relacionadas à dependência da seleção manual das *ROIs* e à necessidade de ajustes específicos nos parâmetros operacionais para garantir a reprodutibilidade dos testes. A falta de um processo totalmente automatizado também foi apontada como um desafio, uma vez que a padronização da aquisição de imagens ainda depende de ajustes manuais do operador, o que pode introduzir variabilidade interobservador. Como solução, sugere-se a implementação de um sistema de análise automatizado para permitir a avaliação de um maior volume de imagens, melhorando a robustez da metodologia e facilitando sua aplicação em protocolos clínicos de controle de qualidade.

4.4.3 METODOLOGIA DE TSAPAKI et al. (2021)

Tsapaki et al., (2021) sugere o desenvolvimento e implementação de um protocolo de controle de qualidade para equipamentos de ultrassom em *modo B*, com o objetivo de garantir a precisão e confiabilidade desses dispositivos na prática clínica. O artigo detalha os principais testes de controle de qualidade para avaliar o desempenho desses equipamentos, contribuindo para um uso seguro e eficiente na medicina diagnóstica por imagem. Os testes abordados são: uniformidade (consistência da iluminação da imagem); penetração máxima (profundidade máxima de visualização clara); sensibilidade (detecção de pequenos sinais ecogênicos); resolução axial e lateral (distinção de objetos próximos em diferentes planos); registro de distância (precisão das medições); largura do feixe (foco do feixe de ultrassom); e contraste de objetos anecóicos (diferenciação de áreas anecóicas, como cistos, do tecido circundante).

O objetivo principal do estudo é estabelecer um protocolo padronizado para o controle de qualidade de equipamentos de ultrassom em *modo B*, com base em testes de desempenho que possam ser facilmente aplicados por profissionais de saúde.

Para a aquisição das imagens, o autor utiliza um *phantom* de ultrassom de modelo Gammex 403GS LE, amplamente empregado em testes de controle de qualidade de ultrassom, devido à sua capacidade de simular estruturas anatômicas e fornecer condições de teste consistentes e replicáveis. Esse *phantom* permite a avaliação de vários parâmetros, incluindo resolução, contraste e sensibilidade dos equipamentos.

A metodologia do teste de escala de cinza neste artigo, consistiu na avaliação da capacidade do equipamento de exibir variações progressivas de tons de cinza, garantindo que ele consiga diferenciar estruturas de diferentes ecogenicidades sem perda de detalhes. Para isso, imagens do *phantom* foram capturadas sob diferentes configurações de parâmetros do ultrassom, como ganho, compressão dinâmica e potência de transmissão. O teste foi realizado observando se a transição entre os níveis de cinza ocorre de maneira contínua e progressiva, sem saltos abruptos ou saturação da imagem. Além disso, foi analisada a capacidade do equipamento de manter a distinção entre regiões de alto e baixo contraste dentro da escala de cinza disponível.

A aprovação ou desaprovação no teste de escala de cinza foi determinada pela clareza com que os diferentes tons são representados na imagem. Para que o equipamento seja aprovado, deveria haver uma transição suave entre os níveis de cinza, sem degradação da qualidade da imagem por excesso de ruído ou artefatos. Caso a diferenciação entre tons de cinza tenha sido insuficiente ou a transição ocorra de maneira abrupta, o equipamento foi considerado inapto para uso clínico, necessitando ajustes nos parâmetros operacionais ou manutenção.

Na conclusão, o autor ressalta a importância da realização periódica de testes de controle de qualidade para garantir a precisão diagnóstica dos exames de ultrassom. O teste de escala de cinza é destacado como essencial para avaliar a capacidade do equipamento de representar adequadamente as variações de intensidade acústica, fator crítico para a correta interpretação das imagens médicas. O estudo conclui que a implementação de protocolos padronizados, como o descrito, pode melhorar significativamente a confiabilidade e segurança dos exames de ultrassom.

Apesar da robustez da metodologia proposta, o autor reconhece algumas lacunas no processo. O controle de qualidade descrito não é totalmente automatizado, exigindo a intervenção e avaliação subjetiva de um operador humano para interpretar os resultados dos testes. Esse fator pode introduzir variações nos resultados, dependendo da experiência do profissional. Além disso, o artigo não aborda soluções automatizadas que poderiam otimizar o processo de aquisição e análise das imagens de controle de qualidade, o que poderia reduzir o tempo necessário para a realização dos testes e aumentar a precisão na detecção de falhas nos equipamentos.

Essa limitação aponta para a necessidade de futuras pesquisas focadas na automação dos testes de controle de qualidade em ultrassom, incorporando tecnologias mais avançadas, como aprendizado de máquina e inteligência artificial, que poderiam realizar uma análise mais precisa e objetiva das imagens adquiridas durante os testes. A automatização dos processos pode melhorar a consistência dos resultados e permitir uma detecção mais rápida e eficiente de problemas nos equipamentos, contribuindo para um atendimento de saúde mais seguro e de alta qualidade.

4.4.4 METODOLOGIA DE PAVERD et al. (2024)

PAVERD et al. (2024) desenvolveu um estudo com o objetivo de avaliar a qualidade de imagem em exames de ultrassonografia utilizando *pads* de contato de gel e líquidos, comparando-os com imagens obtidas sem o uso de *pads*. O foco da pesquisa é verificar o impacto dos *pads* em métricas de imagem, como resolução lateral, intensidade de alvos próximos ao transdutor e a relação sinal-ruído de contraste (CNR). As imagens foram adquiridas em um sistema de ultrassom Canon *Aplio* i800 utilizando um *phantom* de referência padronizado, o "*Multi-Purpose, Multi-Tissue Ultrasound Phantom*, modelo 040GSE" da Computerized Imaging Reference Systems, Inc. (CIRS). Além deste, foram usados *phantoms* customizados de agarose para análise de foco elevacional.

Os testes de controle de qualidade desenvolvidos no artigo incluem a avaliação da resolução lateral, intensidade de alvos na região próxima ao transdutor (*nearfield*), e a relação de contraste-sinal (CNR) de alvos hiperecogênicos e hipoecogênicos. O estudo também avalia a largura de feixe elevacional e sua influência na visualização de vasos e cistos.

A aquisição das imagens foi realizada utilizando o *phantom* CIRS e pads comerciais, como o "Aquaflex™ Pad", um pad de gel, e o "aiSon™ Focus Pad", um pad líquido. Esses pads foram colocados em diferentes alturas acima do *phantom*, e o transdutor foi mantido em posições controladas por um sistema de trilhos para garantir estabilidade e repetibilidade nos experimentos.

Um dos principais testes descritos é o de contraste, onde a relação sinal-ruído (CNR) é avaliada para alvos ecogênicos e hipoecogênicos dentro do *phantom*. A CNR foi calculada com base na diferença entre a intensidade média do sinal dentro e fora das regiões de interesse (alvos circulares), normalizada pela variação do ruído (desvio-padrão) em ambas as regiões. Essa métrica foi utilizada para quantificar a diferença de contraste entre alvos hipo- e hiperecogênicos, tanto para o "Aquaflex™ Pad" quanto para o "aiSon™ Focus Pad".

A aprovação ou desaprovação no teste de contraste foi determinada pela amplitude da CNR em diferentes profundidades, com o maior intervalo de CNR representando um melhor contraste. Os resultados mostram que o "Aquaflex™ Pad" apresentou o maior intervalo de CNR nas imagens de alta posição, seguido pelo "aiSon™ Focus Pad" em posição média.

Na conclusão, os autores destacam que o uso de pads pode influenciar a qualidade da imagem em termos de intensidade e contraste, mas as diferenças em resolução lateral e alcance de contraste são mínimas. A principal influência é a alteração na intensidade do sinal devido à mudança de altura do transdutor, e não necessariamente à presença do pad em si. Para estruturas superficiais, o "Aquaflex™ Pad" oferece melhor qualidade interna de imagem para alvos pequenos, como vasos, localizados até 25 mm de profundidade. No entanto, o estudo ressalta que a adição de pads pode alterar o perfil de foco elevacional, o que pode afetar a visualização de estruturas menores.

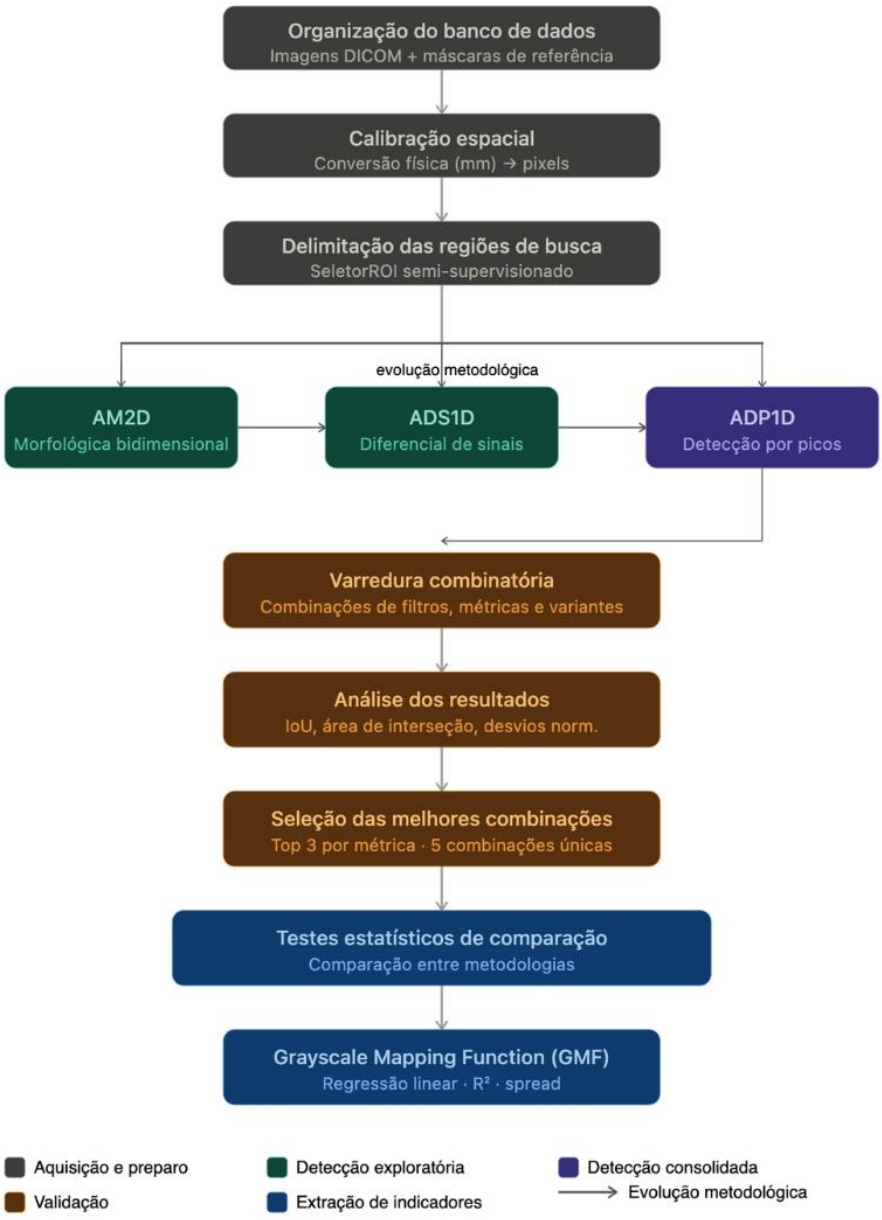
Como lacunas, os autores apontam que o processo de aquisição de imagens não é totalmente automatizado, exigindo ajustes manuais de ganho para compensar variações na intensidade do sinal próximo à superfície.

5 Metodologia

Este capítulo descreve a metodologia empregada para o desenvolvimento e a validação computacional de métodos automatizados para o teste de escala de cinza em imagens de ultrassom *Modo B*. O procedimento foi estruturado em etapas como pode ser visualizado no seguinte fluxograma (Figura 11).

Uma vez que a metodologia foi desenvolvida com foco em imagens de *phantom* de controle de qualidade, os alvos presentes no objeto simulador funcionam como alvos de contraste e tamanho conhecidos, e o objetivo computacional é localizar automaticamente os centros dessas estruturas para permitir a extração padronizada de medidas de intensidade em escala de cinza.

Figura 11: Fluxograma macro do desenvolvimento metodológico cobrindo as três famílias metodológicas. Fonte: Elaborado pelo autor (2026).



5.1 Organização do banco de dados

As imagens utilizadas neste trabalho foram obtidas pela equipe do grupo de pesquisa em instituições parceiras do projeto, a partir de equipamentos de ultrassonografia Modo B de três fabricantes distintos, identificados pelos metadados DICOM. Os parâmetros de aquisição variaram em termos de frequência, profundidade, ganho, TGC e foco. A Tabela 4 apresenta os principais equipamentos e transdutores empregados.

Tabela 4: Equipamentos *de ultrassonografia utilizados nas aquisições das imagens.*

Fabricante	Modelo do Equipamento	Transdutor Principal
Canon Medical Systems	Canon CUS-AA000	PVT-375BT
Toshiba Medical Systems	Toshiba TUS-A300	PVT-375BT / PLT-805AT
Philips Medical Systems	Philips Affiniti 70G	L12_4

O objeto de ensaio variou conforme o conjunto de imagens utilizado. Foram empregados três modelos de *phantoms*, cujas características relevantes para o teste de escala de cinza estão sumarizadas na Tabela 5.

Tabela 5: Phantoms *utilizados no ensaio de escala de cinza.*

Phantom	Fabricante	Alvos de Escala de Cinza	Profundidades dos Alvos
ATS Modelo 539	ATS Laboratories Incorporated	6 alvos circulares: −15, −6, −3, +3, +6, +15 dB em relação ao fundo ecogênico	Organizados lateralmente (superfície nº 4)
Model 040GSE	CIRS	2 grupos: a 3 cm (−9 a +15 dB) e a 11,5 cm (−6 a +15 dB)	3 cm e 11,5 cm
Model 054GS	CIRS	5 alvos circulares (−6 a +6 dB)	4 cm

As imagens foram armazenadas em formato DICOM, preservando a matriz de pixels e os metadados de aquisição. Quando disponível, cada imagem foi associada a uma máscara de referência (*ground truth*), essa marcação consistiu na delimitação espacial dos alvos de escala de cinza diretamente sobre as imagens, produzindo máscaras binárias que indicam a localização esperada de cada ROI.

As máscaras utilizadas como referência neste estudo foram geradas manualmente a partir da identificação visual dos alvos presentes nos phantoms por membros do grupo de pesquisa. A anotação manual foi adotada por representar a melhor estimativa disponível da posição e dimensão dos alvos, constituindo o padrão de referência (*ground truth*) para avaliação dos métodos automáticos. Dessa forma, todas as métricas geométricas apresentadas neste trabalho foram calculadas em relação a essas máscaras de referência.

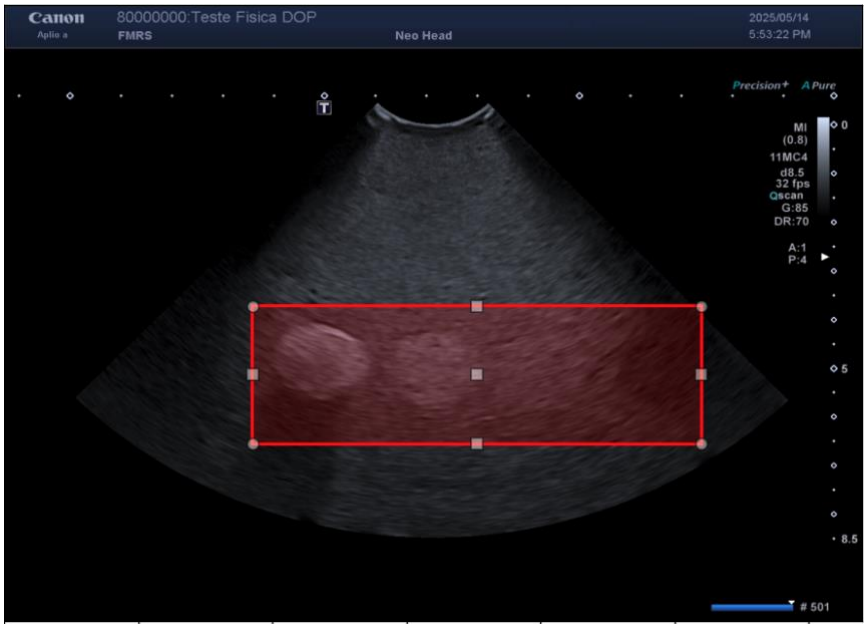
A localização automática das *ROIs* depende da conversão entre o domínio físico, expresso em milímetros, e o domínio das imagens, expresso em *pixels*. Essa

conversão foi realizada a partir dos metadados *DICOM* de calibração espacial. Quando disponível, foi utilizado o atributo *SequenceOfUltrasoundRegions*, em especial o valor de *PhysicalDeltaX*, convertido para milímetros por *pixel* conforme a unidade registrada no cabeçalho. Quando esse atributo não estava disponível, foi utilizado o campo *PixelSpacing* como alternativa (DICOM Library, *[S.d.]*).

5.1.1 Delimitação das regiões de análise

Embora o objetivo do método seja automatizar a localização dos centros dos alvos, adotou-se uma etapa inicial manual para delimitar regiões amplas de busca. Essa seleção restringe o campo de busca do algoritmo subsequente, eliminando interferências de artefatos não pertinentes ao teste de escala de cinza. As coordenadas selecionadas são armazenadas para reutilização e, quando disponível uma máscara de referência, aplicadas também a esta, mantendo a correspondência espacial entre o método automático e a referência de validação. A Figura 12 ilustra a ferramenta.

Figura 12: Exemplo de delimitação das regiões de análise. Fonte: Elaborado pelo autor (2026).



5.2 Estratégias de Detecção Automatizada de *ROIs*

O principal desafio no desenvolvimento desta metodologia reside na ausência de padronização das imagens. Embora os phantoms possuam dimensões conhecidas e controladas, fatores como o transdutor empregado, o equipamento de aquisição e a técnica do operador introduzem variabilidade considerável entre as imagens.

Por esse motivo, foram investigadas três famílias metodológicas, denominadas neste trabalho como:

- *AM2D* - Abordagem Morfológica Bidimensional.
- *ADS1D* - Abordagem Diferencial de Sinais Unidimensionais.
- *ADP1D* – Abordagem de Detecção por Picos com Restrição Geométrica.

A *AM2D* e a *ADS1D* foram estratégias exploratórias: não conduziram à solução final, mas contribuíram para delimitar as escolhas metodológicas adotadas. A primeira avaliou se os alvos poderiam ser localizados por filtros espaciais, contornos e modelos circulares diretamente na imagem bidimensional. A segunda avaliou se a redução da imagem a perfis unidimensionais permitiria detectar bordas por variações abruptas do sinal. A *ADP1D* foi definida como a metodologia consolidada do projeto, pois combina informação de intensidade extraída das *ROIs* com restrições geométricas conhecidas do *phantom*.

O raio nominal dos alvos r_{px} e o espaçamento entre eles e_{px} foram utilizados como parâmetros metodológicos, dados oriundos dos manuais dos respectivos *phantoms*. A partir desses valores, foi definida a distância centro-a-centro esperada, conforme descrito na Equação (16):

$$D_{px} = 2\, r_{px} + e_{px} \qquad (16)$$

em que D_{px} representa a distância esperada entre os centros de dois alvos consecutivos, r_{px} corresponde ao raio nominal dos alvos e e_{px} representa o espaçamento entre eles, todos descritos em *pixels*.

5.3 *AM2D* - Abordagem Morfológica Bidimensional

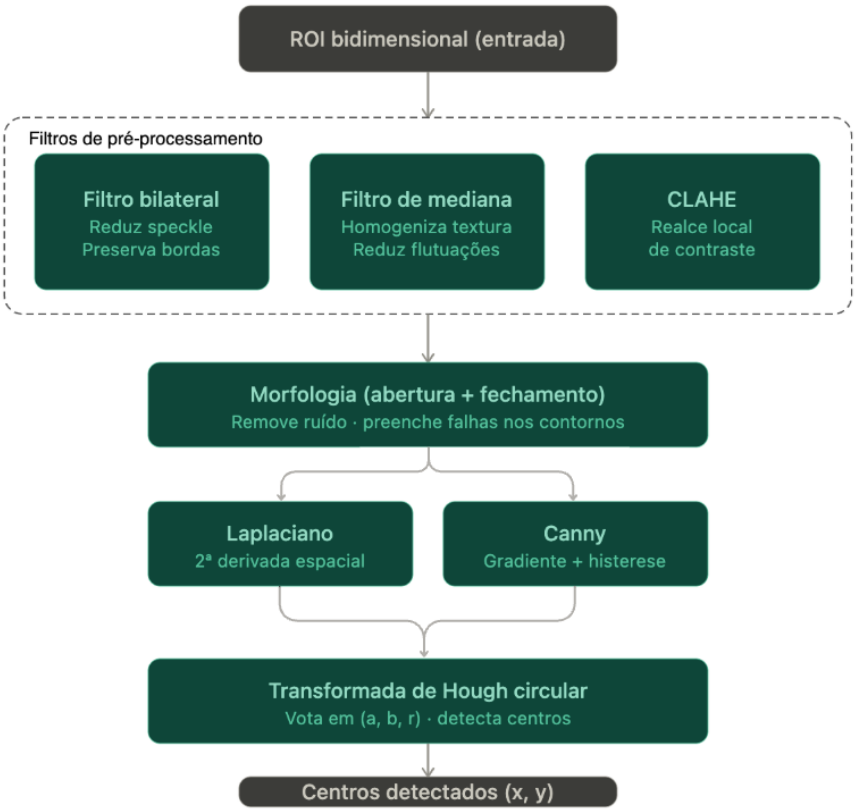
Como primeira abordagem experimental temos a “Abordagem Morfológica Bidimensional” (*AM2D*) tratou cada região de interesse (*ROI)* como uma imagem bidimensional e buscou localizar os alvos por meio de técnicas clássicas de processamento de imagens (detalhadas a seguir). A hipótese inicial foi que, após uma etapa adequada de filtragem e realce, os alvos poderiam ser representados como estruturas aproximadamente circulares, detectáveis por contornos ou por ajuste de modelos geométricos.

A AM2D foi organizada em quatro etapas principais: pré-processamento, realce de estruturas, extração de bordas e ajuste geométrico. A definição da pipeline foi conduzida de forma empírica e iterativa, avaliando individualmente e em combinação

diferentes filtros aplicáveis a imagens de ultrassom, tendo como critério a capacidade de cada configuração de favorecer a detecção de contornos circulares pela Transformada de Hough.

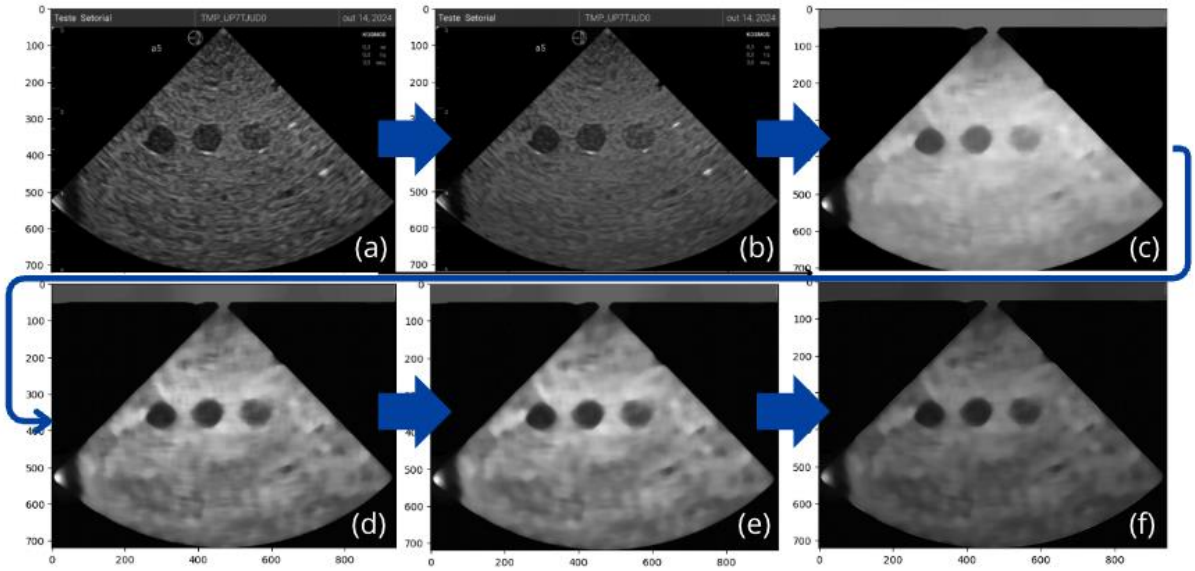
O filtro bilateral foi posicionado como primeira etapa por reduzir o ruído speckle preservando as bordas das estruturas, ponderando a contribuição de cada pixel vizinho tanto pela distância espacial quanto pela similaridade de intensidade. Em seguida, aplicou-se um filtro de mediana para suavizar a textura interna das ROIs sem introduzir gradientes artificiais nas bordas. O CLAHE foi inserido após as etapas de suavização para amplificar o contraste local entre as ROIs e o tecido adjacente, compensando a variação espacial de brilho típica das imagens de ultrassom, ao contrário da equalização global, ele atua por blocos e limita o realce para evitar que o speckle gere bordas artificiais. Operações morfológicas com elementos estruturantes elípticos foram empregadas para adequar a pipeline à geometria circular dos alvos: a abertura removeu estruturas espúrias, o fechamento reconectou bordas fragmentadas e o gradiente morfológico foi avaliado como alternativa de realce de contornos.

Figura 13: Fluxograma detalhado do desenvolvimento da Abordagem Morfológica Bidimensional.



Para a extração de bordas, dois detectores foram avaliados. O Laplaciano, sempre precedido de suavização dada sua sensibilidade ao ruído, destacou regiões de mudança abrupta de intensidade em qualquer direção, sendo útil para o realce de contornos circulares. O detector de Canny, por combinar suavização, cálculo de gradiente, supressão de não máximos e histerese, produziu bordas mais finas e seletivas, sendo empregado para obter candidatos de contorno para o ajuste geométrico via RANSAC. A definição dos limiares do Canny mostrou-se crítica: valores baixos multiplicam falsas detecções por speckle, enquanto valores altos podem suprimir bordas reais pouco contrastadas.

Figura 14: Fluxo de filtros aplicados em uma imagem amostral. (a) Imagem Original; (b) Filtro bilateral; (c) Fonte: Elaborado pelo autor (2026).



Detecção circular por *Transformada de Hough*

Após a filtragem e a detecção de bordas, aplicou-se a Transformada de Hough para círculos. Nesse método, cada ponto de borda vota em um espaço de parâmetros (centro, raio), e os máximos desse espaço indicam os círculos candidatos.

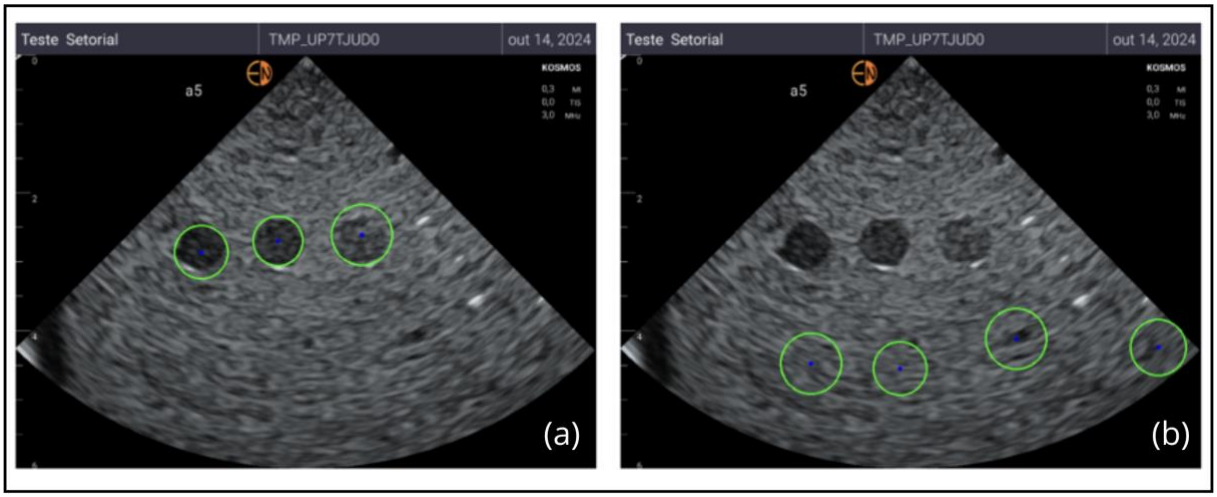
O método foi adequado como primeira abordagem porque o raio físico conhecido dos alvos permite restringir a faixa de busca. Além disso, a distância esperada entre centros serve como critério para eliminar detecções duplicadas de uma mesma estrutura.

A principal limitação é que o método depende fortemente dos parâmetros de pré-processamento. Os valores configurados para o filtro bilateral, filtro de mediana e

CLAHE influenciam diretamente a qualidade das bordas extraídas e, consequentemente, o desempenho da Transformada de Hough circular. Dada a grande variabilidade entre as imagens, não foi viável definir um critério sistemático para esses parâmetros, tornando o método dependente de ajuste manual, o que contraria o objetivo de automatização e padronização deste trabalho.

A Figura 15 ilustra essa sensibilidade: a mesma imagem de ultrassom, processada com kernel 3×3 e 5×5 no filtro de mediana, produz, respectivamente, detecção adequada e inadequada dos círculos.

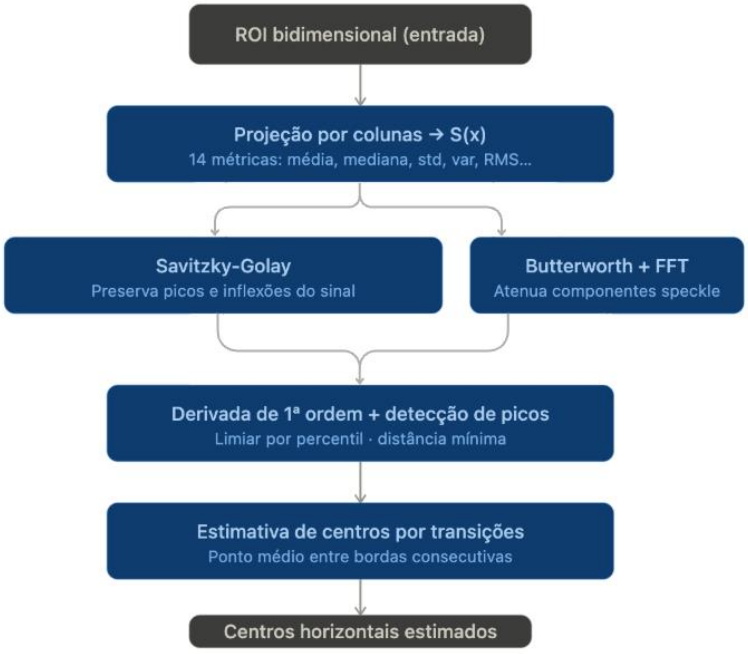
Figura 15: Sensibilidade da Transformada *de Hough* Circular à configuração dos parâmetros de filtragem antecedentes: (a) obtido com *kernel* de tamanho 3×3; (b) obtido com *kernel* de tamanho 5×5,. Fonte: Elaborado pelo autor (2026).



5.4 *ADS1D* - Abordagem Diferencial de Sinais Unidimensionais

A Abordagem Diferencial de Sinais Unidimensionais (ADS1D), ao invés de analisar a ROI como uma imagem bidimensional, resume cada imagem em um perfil unidimensional, calculando um indicador de intensidade a partir de cada uma das colunas da matriz de pixels. Dessa forma, o problema de localizar os alvos passa a ser tratado como a identificação de variações em uma curva. Como os alvos estão dispostos lado a lado no eixo horizontal, cada um produz uma variação característica nesse perfil. A ADS1D busca, então, detectar transições bruscas associadas às bordas laterais dos alvos e estimar seus centros a partir dessas transições. A Figura 16 sumariza o procedimento que será detalhado a seguir.

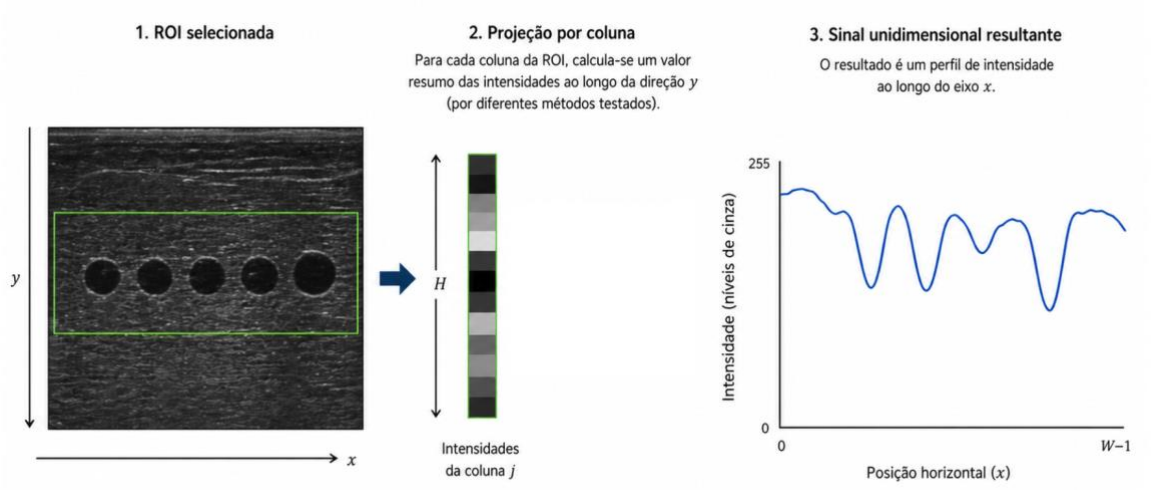
Figura 16: Fluxograma detalhado do desenvolvimento da Abordagem Diferencial de Sinais Unidimensionais. Fonte: Elaborado pelo autor (2026).



5.4.1 Projeção por colunas

A primeira etapa da *ADS1D* consiste em transformar a *ROI* retangular extraída da imagem de ultrassom em um sinal unidimensional, por meio da determinação de indicadores ao longo das colunas. Esse procedimento está ilustrado na Figura 17.

Figura 17: Representação esquemática da projeção por colunas: (esquerda) *ROI* selecionada na imagem original; (centro) agregação das intensidades por coluna; (direita) perfil unidimensional resultante, com depressões correspondentes aos alvos de menor intensidade. Fonte: Elaborado pelo autor (2026).



Foram implementados e avaliados os seguintes indicadores por coluna: Média aritmética; Valor máximo; Valor mínimo; Desvio padrão; Variância; Mediana; Valor Quadrático Médio do Perfil de Intensidade (RMS); Potência média do sinal; Energia do sinal; Moda; Assimetria; Curtose.

Além deste, foi determinada a média ponderada por pesos gaussianos centrados verticalmente na ROI. Os pesos são calculados por uma função gaussiana da forma $w(i) = \exp\left[-\frac{1}{2} \cdot \left(\frac{i - \frac{H}{2}}{\sigma}\right)^2\right]$, onde i é o índice vertical de cada linha e σ é o desvio padrão que controla a largura da curva de pesos. Dessa forma, pixels mais próximos ao centro vertical recebem peso máximo (≈ 1), enquanto os das extremidades superior e inferior recebem pesos progressivamente menores, reduzindo a contribuição de regiões periféricas que não pertencem à faixa central dos alvos. Após o cálculo, os pesos são normalizados pela sua soma, de modo que o resultado corresponda a uma média ponderada genuína.

O resultado é um vetor $S_i(x)$ de comprimento igual à largura da *ROI* (W), cujos elementos representam a intensidade projetada de cada coluna. Esse perfil unidimensional converte a variação bidimensional da ROI em uma curva ao longo do eixo horizontal x, viabilizando a aplicação de técnicas de análise de sinais para detectar os alvos. A Tabela 6 resume a notação dos sinais unidimensionais empregada na ADS1D e na ADP1D.

A Figura 18 apresenta um exemplo de imagem de ultrassom adquirida com transdutor convexo, com a região de interesse recortada contendo os alvos circulares de diferentes ecogenicidades, mostrando as curvas extraídas pelos diferentes métodos de cálculo.

Figura 18: Exemplificação de imagem de ultrassom com um recorte da região de interesse onde se encontram as circunferências de diferentes ecogenicidades do *phantom* imageado e o gráfico contendo as extrações de métricas da imagem recortada. Fonte: Elaborado pelo autor (2026).

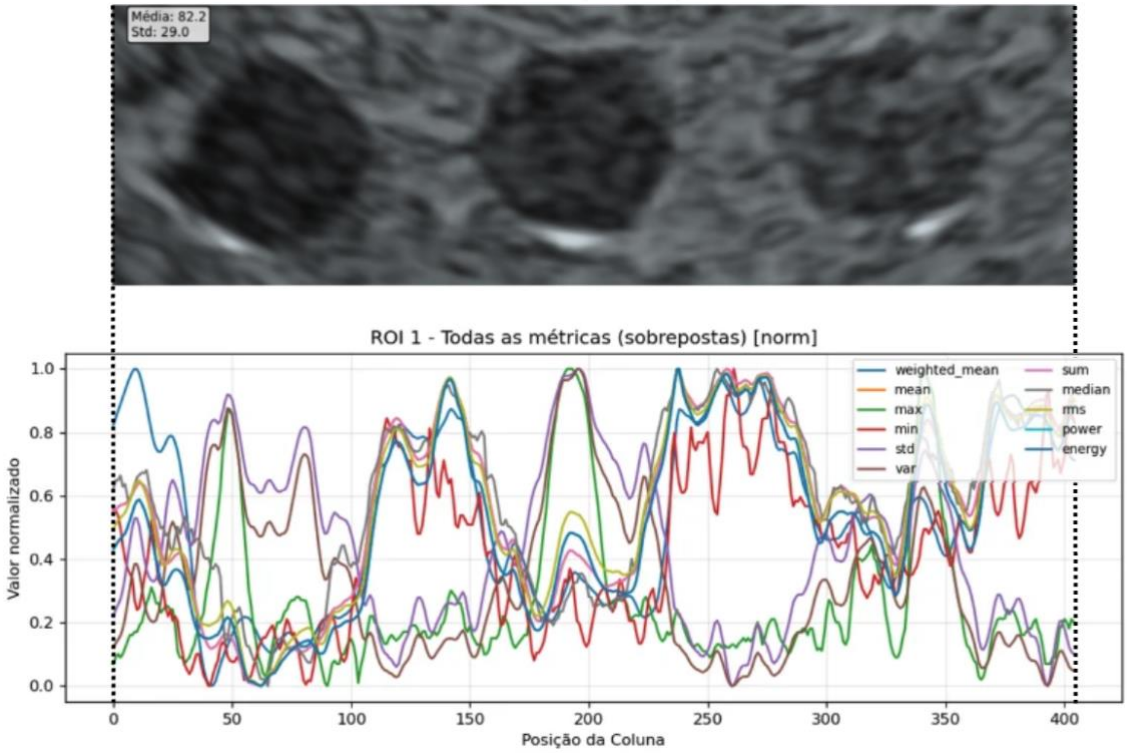


Tabela 6: Notação dos sinais unidimensionais utilizados na ADS1D e na ADP1D.

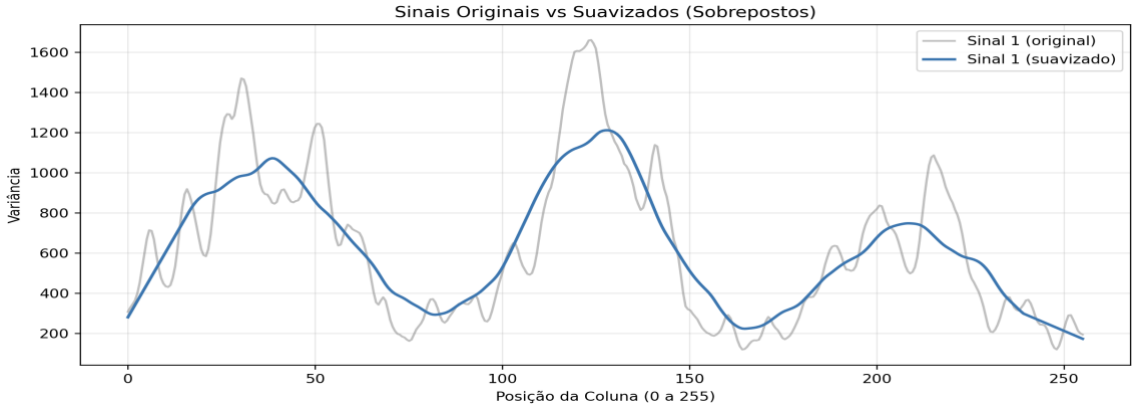
Símbolo	Descrição
$S_i(x)$	Perfil unidimensional bruto; i identifica o método de projeção (ex.: média, variância, máximo, RMS…)
$Sfilt_{i,j}(x)$	Sinal filtrado; j identifica o filtro (SG = Savitzky-Golay; BW = Butterworth)
$Sfilt'_{i,j}(x)$	Derivada numérica do sinal filtrado (diferenças finitas centrais)
$Sbase_{i,j}(x)$	Sinal após remoção de baseline b e transformação em valor absoluto; b estimado pela média do sinal ou pelo valor de fundo na borda da ROI

5.4.2 Suavização e análise espectral

O sinal unidimensional foi submetido a suavização por meio de dois tipos distintos de filtros.

O filtro *Savitzky-Golay* foi aplicado ao perfil projetado para suavizar oscilações de alta frequência preservando a posição das transições associadas às bordas dos alvos. A Figura 19 ilustra o efeito desse filtro sobre a imagem da Figura 18, com o processo de projeção de coluna utilizando o valor de variação.

Figura 19: Gráfico as curvas de sinal 1 e 2 extraídas do processo de projeção por coluna utilizando a métrica de variância, no qual o sinal 2 passou pelo filtro *Savitzky-Golay*. Fonte: Elaborado pelo autor (2026).



Também foi avaliado o filtro Butterworth passa-baixa. A frequência de corte foi definida a partir da energia acumulada do espectro obtido pela FFT separando as componentes associadas à geometria dos alvos daquelas atribuíveis ao *speckle* (Figura 20).

Figura 20: Espectro e frequências do sinal extraído com a Transformada Rápida de *Fourier*, com uma linha de *cutoff* na qual limita as frequências que vão compor o sinal a partir dessa filtragem. Fonte: Elaborado pelo autor (2026).

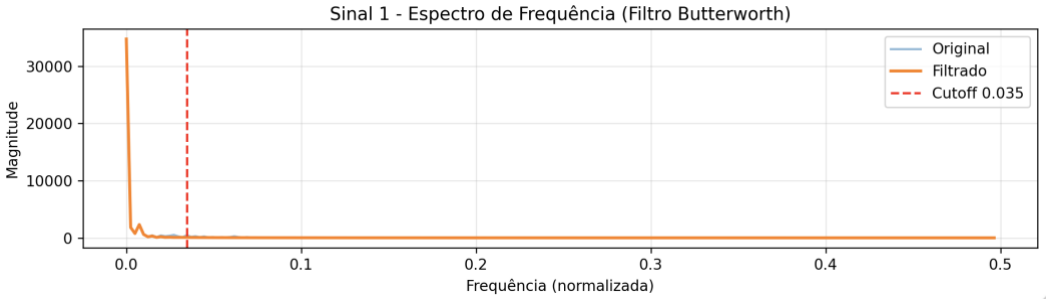


Figura 21: Gráfico as curvas de sinal 1 e 2 extraídas do processo de projeção por coluna utilizando a métrica de variância, no qual o sinal 2 passou pelo filtro *Butterworth*. Fonte: Elaborado pelo autor (2026).

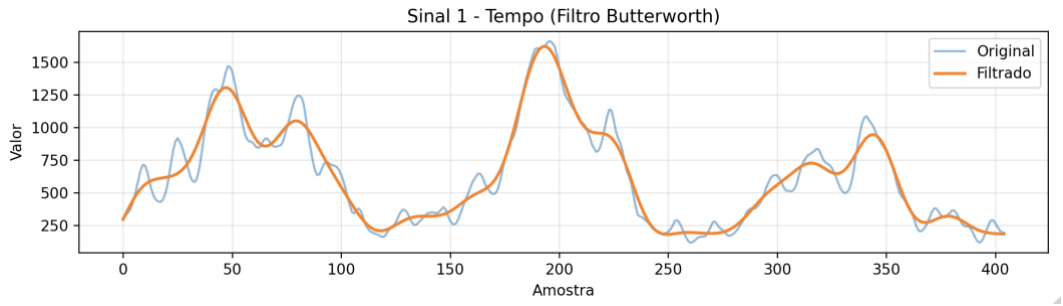
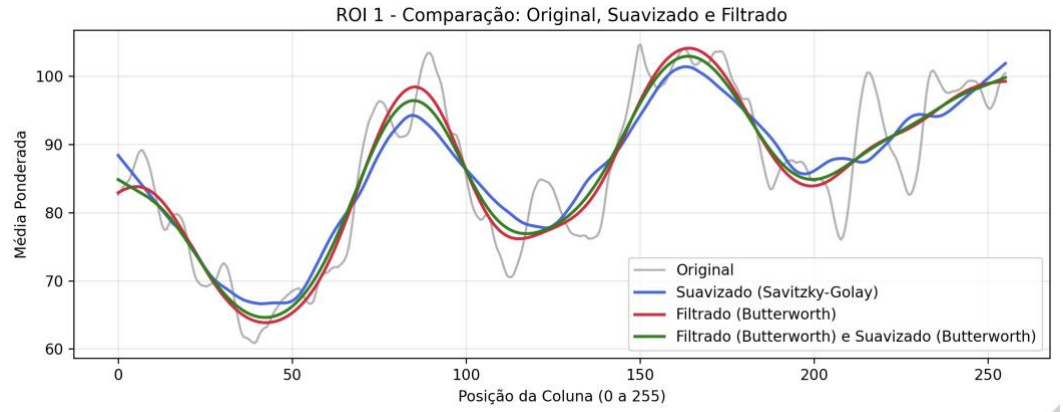


Figura 22: Gráfico comparativo das métricas de suavização *Savitzky-Golay* e *Butterworth* em relação ao sinal original. Fonte: Elaborado pelo autor (2026).

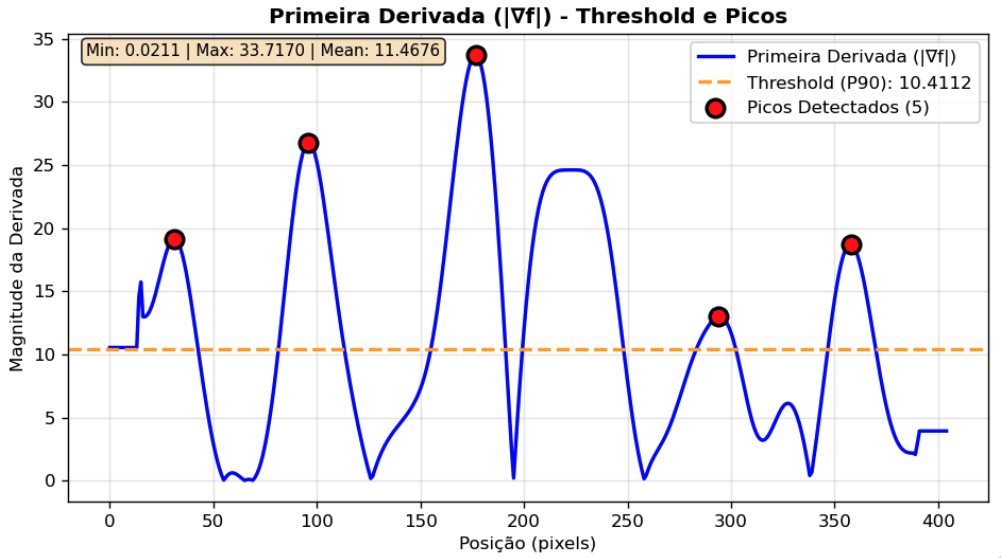


Essas opções foram tratadas como possíveis etapas de processamento dentro das combinações de filtros avaliadas, com o objetivo metodológico de tornar mais comparáveis respostas de alvos com diferentes ecogenicidades.

5.4.3 Derivadas e pontos de mudança

Após o condicionamento do sinal, foi calculada a derivada numérica do sinal por meio de diferenças finitas centrais nos pontos internos do sinal e diferenças unilaterais nas extremidades, preservando o mesmo comprimento do sinal original. A derivada resultante estima a inclinação local da curva e destaca regiões onde a intensidade muda rapidamente; os picos dessa derivada foram identificados por meio de um algoritmo de detecção de picos locais (Figura 23). Na hipótese da *ADS1D*, as bordas laterais dos alvos deveriam produzir picos na magnitude dessa derivada, pois representam transições entre fundo e alvo.

Figura 23: Primeira derivada do perfil de intensidade unidimensional, destacando as regiões de transição associadas às bordas laterais dos alvos e pontos candidatos detectados a partir dos picos de maior magnitude na derivada. Fonte: Elaborado pelo autor (2026).

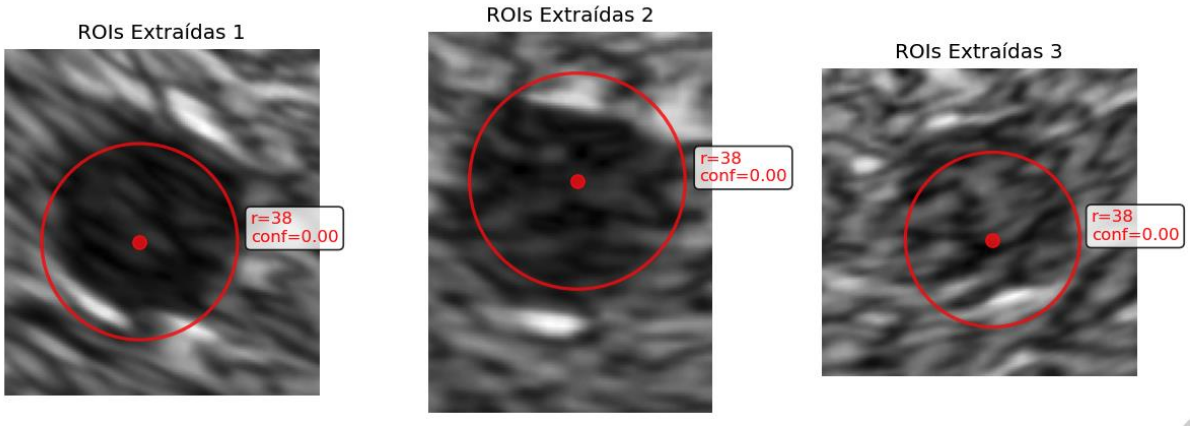


Os pontos candidatos foram obtidos pela detecção de picos na magnitude da derivada do sinal de intensidade (Figura 23). Ao se tomar o valor absoluto da derivada, tanto as bordas de entrada quanto as de saída de cada alvo produzem picos. Para reduzir a detecção de variações irrelevantes, foram aplicados dois critérios de seleção: uma distância mínima entre picos consecutivos, definida a partir do raio esperado das ROIs, que impede múltiplas detecções dentro de uma mesma estrutura; e a retenção apenas dos N picos de maior magnitude, sendo N o número esperado de alvos na imagem.

5.4.4 Estimativa de centros e limitações

Com os pontos de mudança detectados, a *ADS1D* buscou estimar limites laterais dos alvos. Quando duas transições consecutivas são interpretadas como bordas de uma mesma estrutura, considera-se que entre elas pode haver uma circunferência inscrita. Dessa forma, essas regiões foram recortadas na imagem e, a partir dos fragmentos resultantes, aplicou-se o algoritmo de *transformada de Hough* para a detecção de circunferências.

Figura 24: Estimativa dos centros horizontais dos alvos pelo ponto médio entre pares de transições consecutivas detectadas no perfil unidimensional pela abordagem *ADS1D*. Fonte: Elaborado pelo autor (2026).

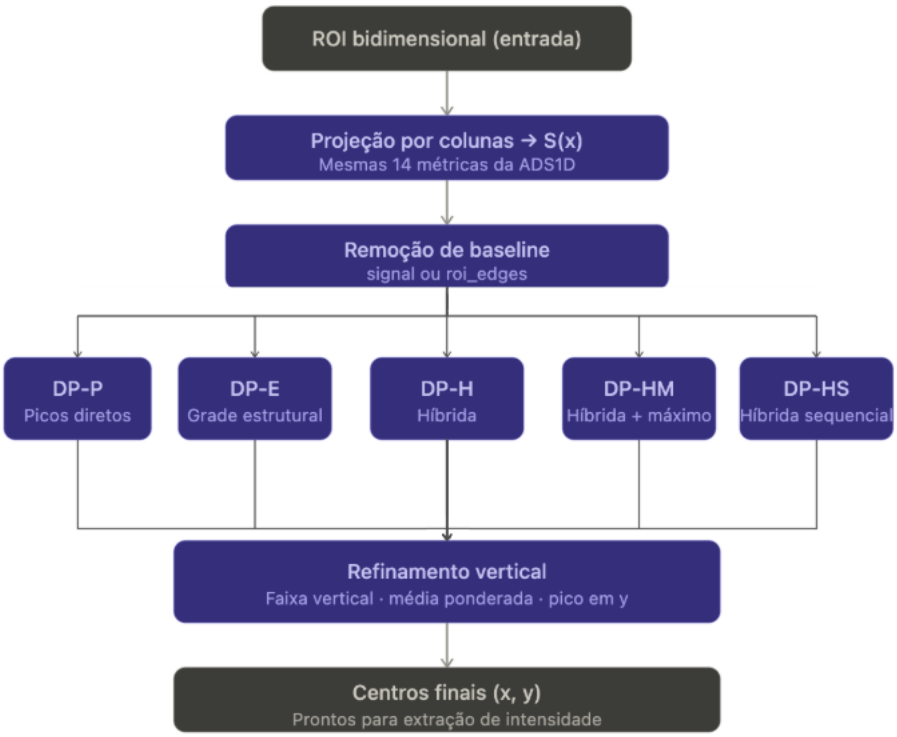


Entretanto, em imagens de ultrassom, a correspondência entre pontos candidatos obtidos e as bordas reais nem sempre é direta. O *speckle* pode produzir picos de derivada dentro do próprio alvo ou no fundo do *phantom*. Além disso, a borda visual pode ser gradual, fazendo com que a derivada apresente múltiplos máximos próximos em vez de uma transição única. A *ADS1D* também depende da escolha do nível de suavização, do percentil usado como limiar e do número esperado de pontos de mudança. Por essas razões, a *ADS1D* foi mantida como abordagem exploratória e serviu como base conceitual para a formulação posterior da A*DP1D*.

5.5 *ADP1D* – Abordagem de Detecção por Picos com Restrição Geométrica

A Abordagem de Detecção por Picos com Restrição Geométrica (*ADP1D*) foi definida como a metodologia consolidada do trabalho. A principal diferença em relação à *ADS1D* é que ela procura localizar os centros das circunferências (alvos) por meio de extremos no perfil de intensidade da *ROI*. Essa hipótese foi adotada porque o centro do alvo tende a ser uma referência espacial mais estável do que sua borda aparente, que pode ser degradada por exemplo, pelo ruído *speckle* ou sombra acústica. Desse modo, as etapas de Projeção por Colunas (5.4.1) e Suavização e Análise Espectral (5.4.2) foram utilizadas tal qual na abordagem *ADS1D*.

Figura 25: Fluxograma detalhado do desenvolvimento da Abordagem de Detecção por Picos com Restrição Geométrica. Fonte: Elaborado pelo autor (2026).



5.5.1 Remoção de *baseline* e transformação por valor absoluto

O centro de cada circunferência corresponde a um extremo no perfil de intensidade, pois concentra o maior número de pixels com intensidade distinta do fundo. Entretanto, ele pode ocorrer em um pico ou vale, dependendo se o alvo é mais claro ou mais escuro que o fundo, condição que não pode ser determinada a priori e que, em algumas imagens, ocorre simultaneamente para alvos distintos. Para contornar essa ambiguidade e garantir que todos os centros sejam representados como picos, aplicou-se a remoção de baseline seguida da transformação em valor absoluto, conforme descreve a Equação 17.

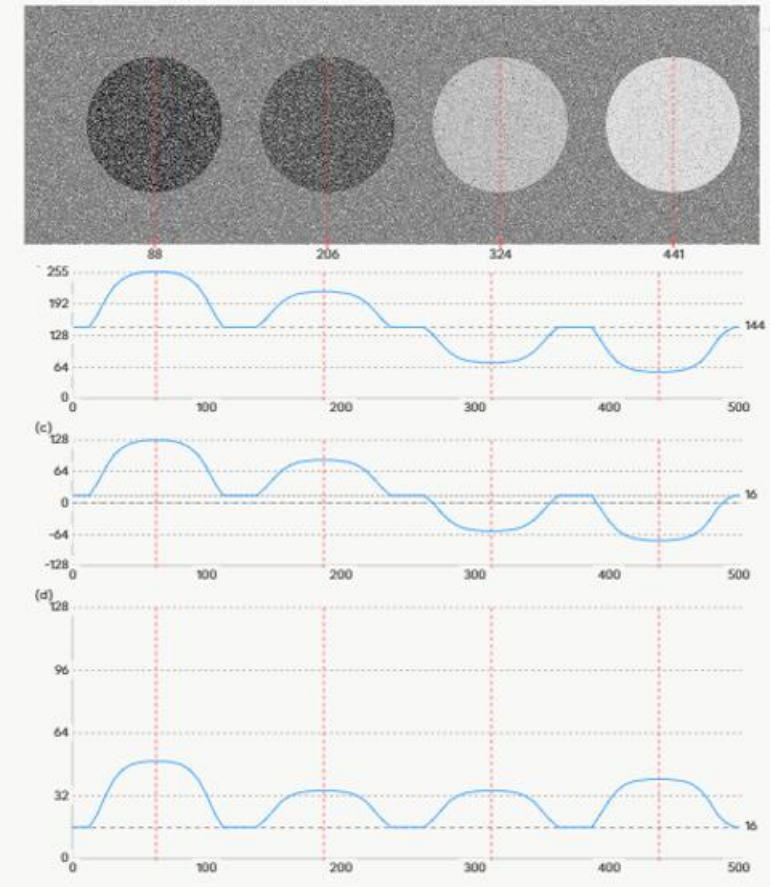
$$Sbase(x) = |S(x) - b| \quad (17)$$

onde S(x) é o sinal original, b é o valor de baseline estimado e $Sbase\ (x)$ é o sinal com baseline removido.

Foram avaliadas duas estratégias de estimativa do baseline: pelo valor médio do próprio sinal projetado na projeção de colunas (tratada a seguir como abordagem

’signal’) e pelo valor de fundo extraído das bordas da do retângulo extraído na seção 5.1.1 tratada a seguir como abordagem ‘roi_edges’). Essas abordagens podem ser vistas na Figura 26.

Figura 26: Demonstração das etapas para remoção da baseline e transformação por valor absoluto. Fonte: Elaborado pelo autor (2026).



.

5.5.2 Detecção de picos usando restrição de posição

A etapa central da *ADP1D* é a detecção de máximos no sinal processado. Para isso, é utilizada metodologia embasada na detecção de picos com restrição de distância mínima entre candidatos, proveniente do espaçamento entre os alvos no *phantom*. A restrição impede que pequenas oscilações dentro de um mesmo alvo sejam interpretadas como múltiplos centros. Na implementação, a distância centro-a-centro esperada (D_{px}) é usada como referência para restringir a busca, admitindo tolerância para variações de imagem e pequenos desvios de posicionamento.

A *ADP1D* utiliza o pico de maior amplitude como candidato de maior confiança, denominado pico principal. A partir dele, o algoritmo pode gerar posições esperadas

à esquerda e à direita, separadas por D_{px}, e verificar se há evidência local no sinal para ajustar essas posições. As variantes dessa lógica são descritas a seguir.

5.5.2.1 *DP-P* - Detecção direta por picos (`picos`)

Na DP-P, os centros são detectados diretamente como picos no sinal processado, com restrição de separação mínima igual a D_{px}. Não há âncora geométrica: cada pico é aceito ou rejeitado pela sua proeminência local. O método é adequado quando os alvos produzem extremos evidentes, mas pode falhar quando o speckle gera picos com amplitude comparável à dos alvos ou quando algum alvo apresenta baixo contraste.

5.5.2.2 *DP-E* - Detecção estrutural (`estrutural`)

Nessa abordagem, primeiro é identificado o pico principal, ou seja, o extremo de maior amplitude no sinal processado. A partir dele, são geradas posições esperadas para os demais alvos, avançando para a esquerda e para a direita com passo igual à distância centro-a-centro D_{px}. O objetivo dessa variante é priorizar a regularidade geométrica do *phantom*. Mesmo que um alvo específico apresente baixo contraste ou produza um pico pouco evidente, sua posição pode ser estimada a partir da disposição física esperada. A *DP-E*, portanto, reduz a dependência da visibilidade individual de cada alvo. Sua limitação é que erros na posição do pico principal ou desvios reais em relação ao espaçamento ideal podem se propagar para todas as posições geradas.

5.5.2.3 *DP-H* - Detecção híbrida (`hibrido`)

A variante *DP-H* corresponde ao método hibrido e combina a regularidade geométrica da *DP-E* com refinamento local baseado no sinal. O algoritmo identifica o pico principal, gera posições estruturais esperadas a partir de D_{px} e, para cada posição esperada, define uma janela de busca local. Dentro dessa janela, o método procura o pico mais proeminente do sinal processado.

Quando um pico é encontrado dentro da janela, a posição estrutural é substituída pela posição real desse pico. Quando nenhum pico é identificado, a posição estrutural é mantida como *fallback*. Assim, a *DP-H* permite corrigir pequenos deslocamentos locais quando há evidência no sinal, mas preserva uma estimativa geométrica quando a evidência local é insuficiente.

5.5.2.4 *DP-HM* - Detecção híbrida com máximo local (`hibridoMedio`)

A variante *DP-HM* corresponde ao método hibridoMedio. Ela mantém a lógica híbrida de gerar posições esperadas e procurar evidência local, mas altera o comportamento quando a detecção formal de picos falha dentro da janela. Em vez de retornar diretamente a posição estrutural, o algoritmo seleciona o máximo local do sinal naquela janela.

Essa alteração foi incluída porque, em alguns casos, o alvo não satisfaz os critérios de proeminência ou separação da rotina de detecção, embora ainda exista um máximo local associado a ele. Assim, a DP-HM aproveita a informação do sinal mesmo quando o pico não é destacado o suficiente para ser detectado automaticamente.

5.5.2.5 *DP-HS* - Detecção híbrida sequencial (`hibridoSequencial`)

A variante *DP-HS* corresponde ao método hibridoSequencial. Assim como nas demais variantes híbridas, o ponto de partida é o pico principal. A diferença é que a busca por centros vizinhos é realizada de forma iterativa. Cada centro detectado passa a ser a nova âncora para estimar a próxima posição esperada, em vez de todas as posições serem calculadas diretamente a partir do pico principal original.

Esse procedimento foi adotado para lidar com deslocamentos acumulados. Se os alvos estiverem levemente deslocados em relação à geometria ideal, uma grade fixa baseada apenas no pico principal pode se afastar progressivamente dos centros reais. Na *DP-HS*, o refinamento sequencial permite que a próxima busca incorpore o deslocamento observado na detecção anterior. O método realiza essa busca tanto para a direita quanto para a esquerda e combina, ao final, os pontos detectados em ordem espacial.

5.5.3 Refinamento vertical dos centros

As etapas anteriores da ADP1D fornecem apenas a coordenada horizontal (eixo x) de cada alvo. Para determinar a coordenada vertical (eixo y), foi implementado um refinamento baseado em sub-imagens individuais por alvo.

A partir da região de análise delimitada pelo usuário e dos N centros horizontais detectados na etapa anterior, a região é dividida em N sub-imagens pelo eixo x. Para cada centro detectado c_i, o recorte horizontal é definido pelo intervalo com a distância

esperada entre centros consecutivos. Dessa forma, cada sub-imagem contém exatamente um alvo em seu interior.

Em cada sub-imagem, aplica-se a mesma sequência de processamento do sinal unidimensional descrita nas etapas anteriores agora operando na direção vertical. Como cada sub-imagem contém um único alvo, espera-se um único pico no sinal resultante. A posição desse pico fornece a estimativa da coordenada y do centro do alvo. Quando a detecção de picos não retorna candidatos, utiliza-se a posição do máximo global do sinal processado como alternativa.

5.6 Análise dos resultados

A análise dos resultados da detecção automática através do método ADP1D foi realizada por comparação entre os centros detectados pelo método e os centros extraídos das máscaras de referência.

5.6.1 *Varredura Combinatória Exaustiva*

Para avaliar sistematicamente as alternativas desenvolvidas, foi implementada uma varredura metodológica exaustiva. Esse procedimento combina as diferentes métricas de extração de sinal, variantes de detecção, métodos de *baseline* e sequências de filtros apresentadas nas seções anteriores. O objetivo é identificar quais combinações de métodos produzem detecções espacialmente mais consistentes em relação às máscaras.

As dimensões avaliadas na busca foram:

Tabela 8: Resumo do espaço de busca das hipóteses metodológicas.

Dimensão	**Opções Avaliadas**	**Quantidade**
Métricas de sinal por coluna (5.4.1)	Média ponderada, média, máximo, mínimo, desvio padrão, variância, soma, mediana, RMS, potência, energia, moda, assimetria e curtose	14
Variantes de detecção (5.5.2)	DP-P, DP-E, DP-H, DP-HM, DP-HS	5

Dimensão	Opções Avaliadas	Quantidade
Métodos de baseline (5.5.1)	signal, roi_edges	2
Combinações de filtros (5.4.2)	Nenhum, butterworth, savgol, regularizar e combinações sequenciais	15
Total de combinações		2.100

5.6.2 Métricas espaciais de validação da detecção de *ROIs*

Para cada par centro detectado–centro de referência, foram calculados treze métricas organizadas em três grupos: desvios normalizados por eixo e distância, área de interseção e razão de interseção sobre união (*IoU*).

5.6.2.1 Desvios normalizados nos eixos X e Y

Os desvios entre o centro detectado pelo método e o centro de referência da máscara foram calculados separadamente para os eixos X e Y, conforme as Equações (18) e (19):

$$D_x = \frac{x_{detectado} - x_{\text{mascara}}}{r_{px}} \qquad (18)$$

$$D_y = \frac{y_{detectado} - y_{\text{mascara}}}{r_{px}} \qquad (19)$$

onde $x_{detectado}$ e $y_{detectado}$ representam as coordenadas do centro estimado pelo método, $x_{mascara}$ e $y_{mascara}$ correspondem às coordenadas do centro de referência obtido a partir da máscara, e r_{px} é o raio nominal do alvo em pixels.

A normalização pelo raio permite comparar os desvios entre imagens com diferentes resoluções ou diferentes tamanhos de alvo, tornando os resultados adimensionais e interpretáveis em relação à dimensão característica da estrutura de interesse. Para ambos os indicadores, valores absolutos menores indicam maior proximidade entre o centro detectado e o centro de referência, caracterizando melhor desempenho do método.

5.6.2.2 Área de interseção percentual

Foi calculada a área de interseção geométrica entre dois círculos: o círculo da máscara, com centro em $(x_{\text{mascara}}, y_{\text{mascara}})$ e raio r_{px} (100% do raio nominal) descrito no manual do *phantom*, e o círculo do método detectado, com centro em $(x_{detectado}, y_{detectado})$ e raio $r_{detectado} = p \cdot r_{px}$, em que p é um fator percentual que ajusta o tamanho da ROI detectada ou a ser extraída futuramente. A área da região de interseção foi determinada a partir da distância euclidiana d entre os centros e dos raios dos círculos ($r_{detectado}$ e r_{px}), sendo estruturada em três cenários geométricos possíveis:

1. Inclusão Total: Caso a distância d seja menor ou igual à diferença absoluta entre os raios ($d \leq |\, r_{detectado} - r_{px}|\,$), um círculo encontra-se totalmente contido no outro. Desse modo, a área de interseção é a própria área do menor círculo, calculada por $\pi * \min\big(\, r_{detectado}, r_{px}\big)^2$.
2. Exclusão Total: Caso a distância d seja maior ou igual à soma dos raios ($d \geq |\, r_{detectado} - r_{px}|\,$), os círculos não compartilham nenhuma região comum, resultando em uma área de interseção igual a 0.
3. Interseção Parcial: Quando os círculos se cruzam parcialmente, a área da "lente" geométrica formada é calculada por meio da soma de dois segmentos circulares, subtraindo-se a área excedente do quadrilátero central.

Para este último caso, determinam-se primeiramente os ângulos de abertura (α e β) correspondentes aos setores circulares de cada circunferência que invadem o raio oposto, definidos pelas funções de arco cosseno limitadas numericamente ao intervalo [-1, 1] para estabilidade do algoritmo:

$$\alpha = \arccos\left(\frac{(d2 + r_{detectado2} - r_{px2})}{(2 \cdot d \cdot r_{detectado})}\right) \quad (20)$$

$$\beta = \arccos\left(\frac{(d2 + r_{px2} - r_{detectado2})}{(2 \cdot d \cdot r_{px})}\right) \quad (21)$$

A área final da interseção A é obtida pela combinação desses setores deduzidos da área dos triângulos centrais, conforme a equação:

$$A = {r_{detectado}}^2 \cdot \alpha + {r_{px}}^2 \cdot \beta - {}^1\!/_2 \sqrt{(-d + r_{detectado} + r_{px})(d + r_{detectado} - r_{px})(d - r_{detectado} + r_{px})(d + r_{detectado} + r_{px})} \quad (22)$$

O cálculo foi realizado para cinco percentuais de raio do círculo do método p: 100%, 75%, 50%, 25% e 10%. O uso de raios menores captura um critério de validação progressivamente mais exigente: com raio reduzido a 10%, somente detecções com erro de posição muito pequeno em relação ao raio nominal obtêm valores elevados de interseção.

5.6.2.3 Razão de interseção sobre união (*IoU*)

A razão de interseção sobre união foi calculada com a seguinte equação:

$$IoU = \frac{A_{\cap}}{(A_{mascara} + A_{detectada} - A_{\cap})} \quad (23)$$

onde $A_{mascara} = \pi \cdot R^2$ é a área do círculo de referência da máscara, $A_{detectada} = \pi \cdot r^2 = \pi \cdot (p \cdot R)^2$ é a área do círculo do método para o percentual $p, e\ A \cap$ é a área de interseção calculada conforme a equação anterior.

O *IoU* é uma métrica simétrica e mais exigente do que a área de interseção percentual, pois penaliza tanto erros de posição quanto diferenças de área entre os círculos comparados. Os valores variam entre 0 e 1; valores mais altos correspondem a melhor desempenho.

5.6.3 Critérios para identificação das melhores combinações

A identificação das melhores combinações de métodos foi realizada a partir das métricas IoU, área de intersecção e os desvios no eixo x e no eixo y (em todos os casos necessários foram aplicadas as normalizações dos indicadores e utilizado o raio em seu tamanho integral de 100% para os indicadores de IoU e área de intersecção). Para cada métrica o valor médio do indicador foi calculado, levando em conta todas as imagens do banco de dados. Em seguida, foram elencadas as 3 melhores combinações para cada indicador e unificadas em uma única lista, sem duplicatas.

5.6.4 Extração de Indicadores de Escala de Cinza

Após a detecção dos centros, foram extraídas regiões circulares ao redor de cada ponto detectado, considerando diferentes percentuais do raio nominal (100%, 75%, 50%, 25% e 10%). Essa variação de raio permite investigar a sensibilidade das medidas de escala de cinza à inclusão de *pixels* periféricos, onde a borda do alvo pode ser menos definida.

A fração de 100% corresponde à utilização integral da região detectada, enquanto as demais frações permitem avaliar progressivamente o efeito da redução

da área amostrada sobre os indicadores de interseção e sobre a representatividade dos níveis de cinza extraídos. A fração de 10% representa um cenário limite de elevada sobreposição geométrica e baixa representatividade espacial.

Em cada região circular, foram calculadas média e desvio padrão da intensidade dos pixels. Esse procedimento foi aplicado às coordenadas obtidas pela máscara e às coordenadas obtidas pelo método automático, permitindo comparação metodológica.

Além disso, a linearidade da *GMF* foi quantificada por meio do coeficiente de determinação (R^2) da regressão, permitindo avaliar a previsibilidade do mapeamento de tons de cinza do equipamento. Complementarmente, a amplitude da resposta dinâmica foi representada pelo *spread*, definido como a diferença absoluta entre os níveis médios de cinza dos alvos de maior e menor contraste da sequência, fornecendo uma medida da extensão da escala de cinza efetivamente utilizada.

5.6.5 *Ranking* Composto Multicritério

Para comparar as combinações metodológicas, os indicadores de desempenho foram consolidados em um *score* composto ponderado. Para cada indicador k, os valores médios por combinação foram normalizados para o intervalo $[0,1]$:

$$\bar{v}_{ik}^{\text{norm}} = \begin{cases} \dfrac{\bar{v}_{ik} - \min\limits_j \bar{v}_{jk}}{\max\limits_j \bar{v}_{jk} - \min\limits_j \bar{v}_{jk}} & \text{se } k \text{ é benefício (maximizar)} \\ \dfrac{\max\limits_j \bar{v}_{jk} - \bar{v}_{ik}}{\max\limits_j \bar{v}_{jk} - \min\limits_j \bar{v}_{jk}} & \text{se } k \text{ é custo (minimizar)} \end{cases} \quad (24)$$

O *score* composto Si de cada combinação i é a média ponderada dos indicadores normalizados (Equação 25):

$$S_i = \frac{\sum_{k=1}^{K} w_k \cdot \bar{v}_{ik}^{\text{norm}}}{\sum_{k=1}^{K} w_k} \quad (25)$$

Os pesos w_k foram definidos como uniformes ($w_k = 1{,}0$ para todos os indicadores), pois não foram identificadas evidências na literatura que justificassem pesos diferenciados entre os indicadores geométricos avaliados. Dessa forma, buscou-se uma avaliação equilibrada entre precisão posicional e sobreposição espacial, selecionando os seguintes indicadores: IoU, interseção percentual, desvio

no eixo x e desvio no eixo y e para os dois primeiros indicadores utilizando o raio descrito no manual do phantom de forma integral.

Cabe distinguir duas escalas de apresentação do mesmo score, ambas reportadas na Tabela 10. O Score A (normalização min–max) primeiro calcula a média de cada combinação e então normaliza esses cinco valores médios para o intervalo [0, 1]; trata-se de uma escala puramente relativa, que estica qualquer diferença entre as combinações para todo o intervalo, independentemente de sua magnitude absoluta. O Score B (média por imagem) normaliza os indicadores sobre todas as observações imagem × combinação (tendo, portanto, como régua a variabilidade real entre imagens) e só então calcula a média por combinação. Por essa razão o Score A varia de 0,25 a 0,75 entre as cinco combinações, enquanto o Score B permanece na faixa de 0,833 a 0,852: a diferença real entre a 1ª e a 5ª combinação é de apenas 0,019, pequena frente à variabilidade intra-método (aproximadamente 0,13).

5.6.6 Testes Estatísticos de Comparação entre Metodologias

A unidade experimental adotada foi a imagem, utilizando-se a média das ROIs de cada imagem em cada combinação metodológica. Essa escolha é motivada pela não-independência das ROIs extraídas de uma mesma imagem: elas compartilham fundo ecogênico, condições de aquisição e padrão de ruído *speckle*. Tratar cada ROI como observação independente inflaria artificialmente o tamanho amostral.

Para avaliar a significância das diferenças entre combinações, optou-se por testes não paramétricos. Essa opção é fundamentada em duas características do contexto: (i) os indicadores de desempenho não apresentam distribuição normal garantida (métricas como IoU exibem comportamento assimétrico, com valores concentrados nos extremos em raios pequenos) e (ii) o tamanho amostral por combinação é limitado. Inicialmente, utilizou-se o teste de *Friedman* (Friedman, 1937), alternativa não paramétrica à ANOVA de medidas repetidas, apropriada para comparar múltiplas condições avaliadas sobre as mesmas unidades experimentais. O teste baseia-se na ordenação dos resultados em rankings para cada imagem e verifica se há diferenças sistemáticas entre as combinações analisadas. Foram consideradas as seguintes hipóteses:

- $\boldsymbol{H_0}$: As medianas do indicador são iguais para todas as combinações metodológicas;

- $\boldsymbol{H_1}$: Pelo menos uma combinação apresenta mediana diferente;
- Nível de significância: $\boldsymbol{\alpha = 0,05.}$

Se p < $\boldsymbol{\alpha}$, rejeitamos $\boldsymbol{H_0}$ e concluímos que existe pelo menos uma diferença significativa entre as combinações, mas o teste de *Friedman* não identifica quais pares diferem. Para isso, são necessários testes *post-hoc*.

Quando o teste de *Friedman* rejeita H_0, são conduzidos testes de *Wilcoxon* pareado (Wilcoxon, 1945) para cada par de combinações. Em termos simples, esse teste compara duas condições da seguinte forma: para cada imagem, calcula-se a diferença entre os *scores* das duas combinações e verifica-se se essas diferenças tendem sistematicamente para um lado ou se flutuam ao redor de zero. Se a maioria das diferenças é positiva (ou negativa), há evidência de que uma combinação é distinta da outra; caso contrário, as duas são consideradas equivalentes.

As análises estatísticas foram conduzidas sobre o score composto por imagem (Seção 5.6.5), calculado com normalização global min-max e pesos uniformes. Aplicar os testes de Friedman e Wilcoxon separadamente sobre cada um dos quatro indicadores (IoU, interseção percentual, desvio em x, desvio em y) elevaria a taxa de erro tipo I global (inflação por múltiplos testes). Ao condensar o desempenho em um único score composto, mantém-se a integridade do nível de significância adotado (α = 0,05) e preserva-se a interpretabilidade dos resultados.

5.6.7 Análise de Concordância para a *GMF*

A comparação entre os indicadores de escala de cinza extraídos com coordenadas da máscara (*ground truth*) e com coordenadas do método automatizado foi realizada por análise de concordância, e não apenas por correlação.

Para isso, foi adotada a análise de Bland-Altman (Martin Bland; Altman, 1986). Para cada par de valores de *grayscale* $(g_{\text{mascara}}, g_{\text{detectado}})$, escala de cinza média da circunferência da máscara de referência e da circunferência detectada, respectivamente, calcularam-se a diferença ($d = g_{\text{mascara}} - g_{\text{detectado}}$) e a média ($\bar{m} = \frac{(g_{\text{mascara}} + g_{\text{detectado}})}{2}$).

A diferença d avalia a magnitude e a direção do viés sistemático entre os dois métodos, valores próximos de zero indicam boa concordância absoluta, enquanto a

média m fornece a referência de magnitude do fenômeno medido para cada imagem, permitindo verificar se o viés varia em função do nível de sinal (proporcionalidade).

Adicionalmente, foi realizada a comparação das inclinações (slopes) da GMF, conforme proposto por (Altman; Bland, 1983). Para cada imagem, ajustou-se uma regressão linear entre o índice da ROI e o *grayscale* médio, tanto para as coordenadas da máscara quanto para as do método. A diferença relativa entre as inclinações constitui uma medida funcionalmente relevante da concordância, pois indica se ambas as extrações produzem a mesma avaliação da resposta dinâmica do equipamento (Altman; Bland, 1983).

6 Resultados

6.1 Caracterização do *Banco de dados*

O *banco de dados* foi composto por 156 imagens *DICOM* de *phantom* de controle de qualidade, adquiridas em diferentes protocolos e configurações de equipamento, conforme caracterizado na Tabela 7. Cada imagem foi pareada com uma máscara de referência contendo a localização manual dos alvos de escala de cinza, conforme descrito na seção 5.1.

Tabela 7: Caracterização do banco de dados.

Phantom	Nº Imagens	Percentual
ATS 539	110	70,5%
CIRS 040GSE	39	25,0%
CIRS 054GS	7	4,5%
Total	**156**	**100%**

Para cada imagem, foram extraídas, a partir da máscara de referência, as ROIs claramente visíveis ao aplicador do teste. Imagens com apenas 1 *ROI* detectada na máscara foram excluídas da análise de *GMF* (93 imagens com ≥ 2 *ROIs* foram utilizadas nessa etapa).

6.2 Varredura Combinatória Exaustiva Metodológica

A varredura combinatória exaustiva avaliou sistematicamente 2.100 combinações de parâmetros, resultando em 327.600 execuções totais (2.100 combinações × 156 imagens).

A partir das quatro métricas-chave descritas na seção 5.6.3, foram selecionadas as 3 melhores combinações por indicador. Após unificação e remoção de duplicatas, resultaram 5 combinações candidatas, apresentadas na Tabela 9. As combinações são identificadas pela notação DP-X(filtros), onde DP indica a abordagem por Detecção por Picos, X representa a variante de detecção (P = picos simples, HM = híbrido médio, H = híbrido, HS = híbrido sequencial), e os sufixos indicam os filtros aplicados: SG = Savitzky-Golay, BW = Butterworth, R = regularização.

A convergência da varredura combinatória exaustiva revelou um padrão predominante: todas as 5 melhores combinações identificadas utilizam a métrica de variância (*var*) como extração de sinal por coluna e o *baseline* por bordas da *ROI* (*roi_edges*). Essa convergência reduziu o espaço efetivo de decisão às dimensões de variante de detecção e combinação de filtros.

Tabela 9: Cinco melhores combinações metodológicas por score composto

Posição	Combinação (métrica · variante · *baseline* · filtros)	Número de repetições em diferentes indicadores
1º	*DP-P* (SG)	4
1º	*DP-HM* (BW + SG)	4
2º	*DP-HS* (SG + R)	2
2º	*DP-P* (SG + R)	1
2º	*DP-H* (SG + R)	1

A Tabela 10 apresenta os indicadores de desempenho médio das 5 melhores combinações exibindo os resultados do indicador composto.

Tabela 10: Matriz dos indicadores de desempenho médio das cinco melhores combinações sobre n = 156 imagens válidas.

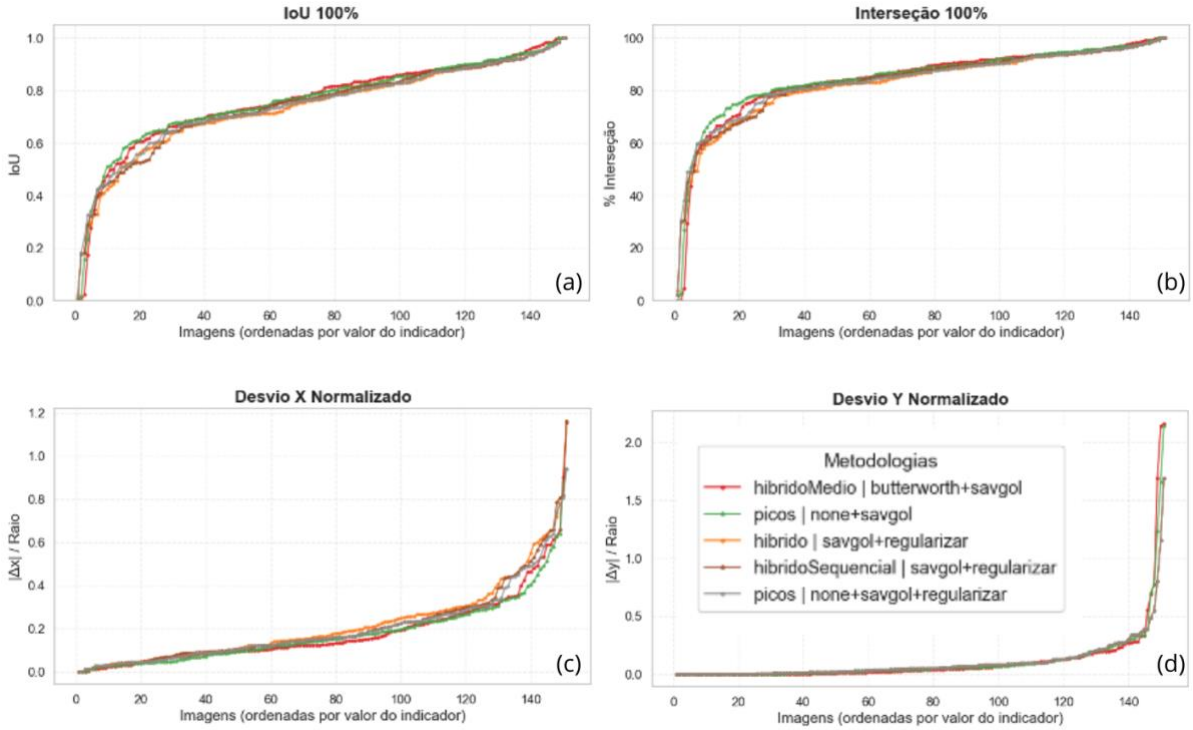
Combinação	*IoU*	Intersec.	Δx norm.	Δy norm.	Score A (min–max)	Score B (média/img)
***DP-P* (SG)**	0.7663	85.0561	0.1706	0.1272	0.7500	0.8517
***DP-HM* (BW + SG)**	0.7620	84.3820	0.1737	0.1263	0.5817	0.8486
***DP-HS* (SG + R)**	0.7512	84.1268	0.1980	0.1099	0.4837	0.8381
***DP-P* (SG + R)**	0.7466	83.8898	0.1986	0.1099	0.3861	0.8408
***DP-H* (SG + R)**	0.7454	83.5103	0.2075	0.1099	0.2500	0.8330

A combinação DP-P com filtro savgol apresentou IoU médio de 0,7663, o que indica que, em média, 76,6% da área de cada círculo de amostragem do método coincide com a área do *ground truth*. A área de interseção percentual indica que 85,1% da região detectada pelo método coincide com a região de referência.

As componentes direcionais revelam que o desvio horizontal ($\Delta x_{norm} = 0{,}171$) é ligeiramente superior ao vertical ($\Delta y_{norm} = 0{,}127$), provavelmente pela maior dificuldade em resolver a posição horizontal exata a partir do perfil de intensidade por colunas.

A leitura das curvas da Figura 27 evidencia dois comportamentos relevantes. Primeiro, em todos os indicadores observa-se uma cauda de aproximadamente vinte imagens em que o desempenho se deteriora acentuadamente; como os valores estão ordenados de forma crescente, essa pequena fração de casos difíceis (tipicamente imagens com alvos pouco visíveis, baixo contraste ou artefatos) puxa as médias dos indicadores para baixo; na maioria das imagens, a detecção é substancialmente melhor do que a média sugere.

Figura 27: Desempenho das cinco melhores combinações metodológicas em quatro indicadores de qualidade de detecção de *ROIs*. Cada *subplot* apresenta um indicador (*IoU*, Interseção, Δxnorm e Δynorm), com os valores agregados por imagem (média das *ROIs*) e ordenados de forma crescente no eixo horizontal. Fonte: Elaborado pelo autor (2026).



6.3 Comparação estatística entre as metodologias

O teste de *Friedman* foi aplicado sobre o *score* composto calculado anteriormente, tomando como unidade experimental a imagem (média das ROIs por imagem para cada combinação), com N=156 imagens e k=5 grupos. Conforme a Tabela 11, o teste retornou χ^2=36,383 e p=2,38×10^{-7}, valor muito inferior ao nível de significância adotado (α=0,05). Rejeita-se, portanto, a hipótese nula de que as cinco combinações têm desempenho equivalente: ao menos uma delas difere sistematicamente das demais quanto ao score composto. Esse resultado justifica a etapa seguinte de comparações par-a-par, necessária para identificar especificamente quais combinações se distinguem.

Tabela 11: Teste de *Friedman* aplicado ao *score* composto por imagem**.** Fonte: Elaborado pelo autor (2026)

Métrica	χ^2 *Friedman*	*p*-valor	Significativo ($\alpha = 0{,}05$)
Score Composto	36,383	$2{,}38 \times 10^{-7}$	Sim

Na etapa seguinte, os testes *post-hoc* de *Wilcoxon* compararam cada par de combinações individualmente, identificando quatro pares com diferença estatisticamente significativa. Todos os quatro envolvem a combinação *DP-HM* (butterworth + savgol) em contraste com cada uma das demais, o que significa que *DP-HM* apresenta desempenho sistematicamente diferente do restante do grupo. Os seis pares restantes que comparam entre si *(DP-P* savgol, *DP-H*, *DP-HS* e *DP-P* savgol+reg) não atingiram significância, indicando que essas quatro combinações são estatisticamente equivalentes entre si e que as diferenças numéricas observadas entre elas não são generalizáveis além da amostra utilizada.

A Tabela 12 apresenta os resultados completos dos testes de *Wilcoxon* pareados para os 10 pares possíveis entre as cinco melhores combinações.

Tabela 12: Comparações par-a-par pelo teste de *Wilcoxon* sobre o *score* composto por imagem (n = 156. Valores de p ajustados; α = 0,05. Fonte: Elaborado pelo autor (2026).

Combinação A	Combinação B	W	p-valor ajustado	Significativo
DP-HM (BW+SG)	*DP-H* (SG+R)	3146	0,000092	Sim
DP-HM (BW+SG)	*DP-P* (SG+R)	3246	0,000200	Sim
DP-HM (BW+SG)	*DP-HS* (SG+R)	3566	0,003511	Sim
DP-HM (BW+SG)	*DP-P* (SG)	3565	0,006754	Sim
DP-P (SG)	*DP-H* (SG+R)	3475	0,5501	Não
DP-P (SG)	*DP-P* (SG+R)	3336	0,5501	Não
DP-H (SG+R)	*DP-HS* (SG+R)	246	0,6867	Não
DP-H (SG+R)	*DP-P* (SG+R)	325	0,7597	Não
DP-P (SG)	*DP-HS* (SG+R)	3729	0,7597	Não
DP-HS (SG+R)	*DP-P* (SG+R)	274	0,9075	Não

A leitura da Tabela 12 revela que DP-HM (BW+SG) é a única combinação com perfil estatisticamente distinto: todos os seus quatro pares resultaram em rejeição de H_0, indicando que essa combinação se distingue de forma consistente das demais em frequência de vitórias por imagem. Em síntese, dos 10 pares possíveis, 4 atingiram significância (todos envolvendo DP-HM) e 6 não, configurando um cluster de quatro combinações estatisticamente equivalentes (DP-P, DP-H, DP-HS e DP-P com

regularização) e uma combinação isolada (DP-HM) com desempenho significativamente diferenciado.

6.3.1 Melhores Combinações por fração do Raio

Conforme descrito na Seção 5.6.2, a área de interseção percentual foi calculada para cinco frações do raio do círculo detectado (100%, 75%, 50%, 25% e 10%). Essa análise por frações foi motivada por dois aspectos complementares.

Primeiramente, frações menores do raio tendem a elevar os valores de área de interseção percentual. Como a métrica considera a proporção da área do círculo detectado contida na máscara de referência, pequenas discrepâncias na posição do centro podem fazer com que parte da circunferência detectada ultrapasse os limites da máscara quando o raio completo é utilizado. Ao reduzir o raio, diminui-se a probabilidade de que a região analisada extrapole a máscara.

Em segundo lugar, a região interna ao círculo detectado é utilizada para a extração dos indicadores de escala de cinza e da Função de Mapeamento de Cinza (GMF). Dessa forma, a avaliação em diferentes frações do raio permite analisar se a região efetivamente utilizada para a extração dessas características permanece contida na área correspondente ao alvo de referência, reduzindo a influência de bordas ou de estruturas adjacentes.

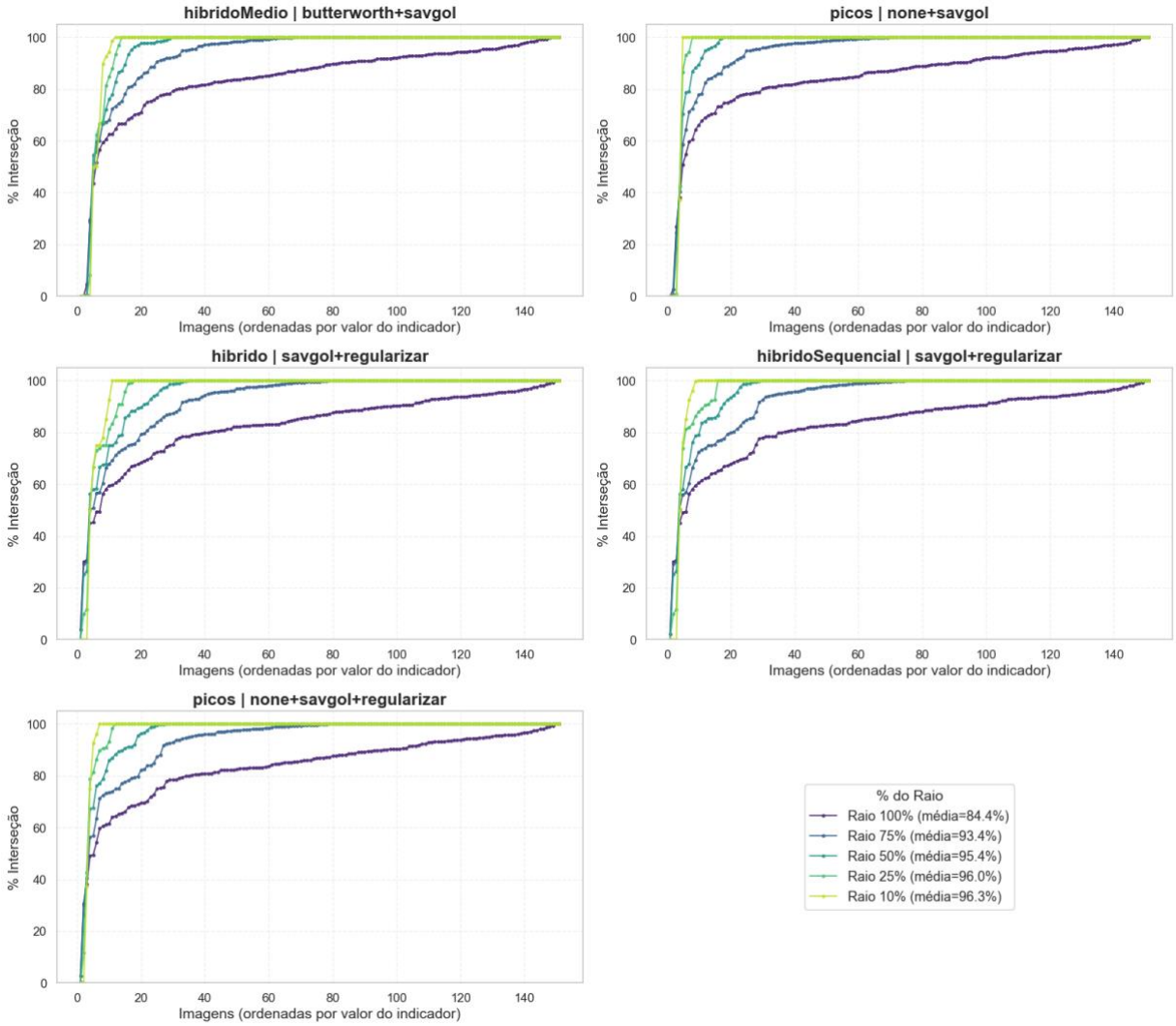
Entretanto, a redução progressiva do raio também implica uma diminuição da representatividade da região amostrada. Embora frações menores possam produzir valores mais elevados de área de interseção, a região analisada passa a corresponder a uma parcela cada vez menor do alvo real. Em casos extremos, uma região muito pequena pode apresentar interseção de 100% com a máscara e, ainda assim, não ser representativa das características radiométricas do alvo como um todo. A escolha da fração do raio deve, portanto, equilibrar esses dois aspectos. A Figura 28 exibe a porcentagem de intersecção média de cada imagem em função do raio de avaliação, para cada combinação de abordagem, e a Tabela 13 apresenta as melhores combinações por percentual de raio, com a respectiva área de intersecção.

Tabela 13: Sensibilidade dos indicadores de qualidade ao percentual do raio do círculo de amostragem do método, por raio.

Raio %	Combinação	Intersecção

100	*DP-P* (SG)	85.056147
75	*DP-P* (SG)	94.654046
50	*DP-P* (SG + R)	96.800085
25	*DP-P* (SG + R)	97.935352
10	*DP-P* (SG + R)	98.250791

Figura 28: Sensibilidade da interseção percentual à redução do raio de análise para cada uma das combinações metodológicas selecionadas. Fonte: Elaborado pelo autor (2026).



A Tabela 14 sumariza os resultados da Figura 28, exibindo a dominância do par (*var*, *roi_edges*). A tabela consolida, para cada raio, a combinação líder em dois critérios complementares: área de interseção percentual e *IoU*. Em ambos os

critérios, a combinação vencedora alterna apenas os filtros aplicados, enquanto a variante de detecção é sempre a DP-P (picos).

Tabela 14: Melhor combinação metodológica por raio segundo o critério de área de interseção percentual (média ± desvio padrão sobre n imagens). Fonte: Elaborado pelo autor (2026).

Raio (%)	Combinação	Área média	Área std	IoU média	IoU std
100	*DP-P* (SG)	85.06	16.50	0.7663	0.1887
75	*DP-P* (SG)	94.65	16.62	0.5260	0.1012
50	*DP-P* (SG+R)	96.80	13.01	0.2412	0.0343
25	*DP-P* (SG+R)	97.94	12.33	0.0612	0.0078
10	*DP-P* (SG+R)	98.25	12.56	0.0098	0.0013

Os valores de IoU em raios pequenos (10%–25%) apresentam magnitudes na ordem de 10^{-2}, ao passo que a área de interseção percentual correspondente ultrapassa 98% para as mesmas combinações. Essa divergência decorre de *uma propriedade geométrica da IoU* detalhada na Seção 7.3.

6.4 Resultados do teste de Escala de Cinza e *GMF*

A GMF foi calculada conforme a descrição da Seção 5.6.4. O objetivo desta etapa não foi avaliar o desempenho deste indicador como ferramenta para controle de qualidade, mas verificar se a interpretação obtida a partir das ROIs demarcadas manualmente coincide com aquela produzida pelas ROIs geradas automaticamente pelo método proposto. Assim, para fins comparativos, a Tabela 15 apresenta os indicadores de linearidade (R^2) e amplitude (spread, conforme definido na Seção 5.6.4) da GMF calculada com as coordenadas da máscara de referência, para as 93 imagens com 2 ou mais *ROIs*.

Tabela 15: Indicadores de linearidade (R^2) e amplitude (*spread*, em níveis de cinza) da *GMF* calculada com as coordenadas da máscara de referência, por percentual de raio (n = 93 imagens com ≥ 2 *ROIs*).

Raio (%)	R^2 (máscara)	Spread (máscara)
100	0.8514	38.73
75	0.8772	43.73
50	0.8688	45.19

25	0.8562	45.02
10	0.8538	46.15

Os valores medianos de R² situam-se entre 0,85 e 0,88, o que indica que a relação entre o índice do alvo e o *grayscale* médio é predominantemente linear para a maioria das imagens; entre os raios avaliados, o raio de 75% apresenta o maior R² (0,877). O spread, por sua vez, aumenta ligeiramente à medida que o raio diminui, o que reflete a menor contribuição do fundo ecogênico nas regiões mais centrais dos alvos.

6.4.1 *GMF* com Coordenadas do Método Automatizado

A Tabela 16 apresenta os mesmos indicadores utilizando as coordenadas detectadas pelo método automatizado (melhor combinação para cada percentual de raio).

Tabela 16: Indicadores de linearidade (R^2) e amplitude (*spread*) da *GMF* calculada com as coordenadas detectadas pelo método automatizado (melhor combinação por raio), por percentual de raio.

Raio (%)	Combinação	R^2 (método)	Spread (método)
100	*DP-P* (SG)	0.8573	38.47
75	*DP-P* (SG)	0.8770	43.78
50	*DP-P* (SG+R)	0.8652	45.30
25	*DP-P* (SG+R)	0.8549	44.90
10	*DP-P* (SG+R)	0.8498	46.31

6.4.2 Concordância entre Máscara e Método: *Bland-Altman* e Comparação de Inclinações

A concordância entre os valores de *grayscale* extraídos com coordenadas da máscara e com coordenadas do método automatizado foi avaliada por meio de duas análises complementares descritas na Seção 5.6.7: a análise de *Bland-Altman*, que quantifica viés e limites de concordância (LoA) das diferenças método − máscara, e a comparação das inclinações (*slopes*) da regressão linear por imagem entre o índice da *ROI* e o *grayscale* médio. As duas análises foram conduzidas para os cinco percentuais de raio (10%, 25%, 50%, 75% e 100%) utilizando-se a combinação líder

(*DP-P*+Savgol) e n = 281 *ROIs* pareadas em cada raio, visto que só foram utilizadas imagens que tinham ao menos 2 ROIs inscritas.

A Tabela 17 apresenta os parâmetros de Bland-Altman por raio. O viés médio (bias) é inferior a 1 nível de cinza na maioria dos raios, variando de +0,38 (raio 100%) a +1,08 (raio 25%), com sinal sistematicamente positivo (o que indica uma leve tendência do método em superestimar o grayscale em relação à máscara). O desvio padrão das diferenças cresce de forma contínua à medida que o raio diminui, de 4,83 (raio 100%) a 9,53 (raio 10%), refletindo a maior sensibilidade da amostragem em pequenos raios ao deslocamento espacial dos centros detectados. A proporção de pares dentro dos limites de concordância permanece elevada em todos os raios, entre 95,0% (raio 10%) e 97,2% (raio 25%), valores compatíveis com a expectativa teórica de aproximadamente 95% sob normalidade das diferenças.

Tabela 17: Análise de *Bland-Altman* da diferença (método − máscara) do *grayscale* médio por *ROI*, por percentual de raio (n = 281 *ROIs*). LoA = limites de concordância de 95%. Fonte: Elaborado pelo autor (2026).

Raio (%)	Bias (níveis de cinza)	SD	LoA inferior	LoA superior	% dentro dos LoA
100	0.3774	4.8329	-9.0950	9.8498	96.44
75	0.4638	4.8457	-9.0339	9.9614	96.80
50	0.8860	5.4288	-9.7545	11.5265	96.80
25	1.0750	6.6993	-12.0556	14.2056	97.15
10	0.9113	9.5308	-17.7690	19.5917	95.02

A Tabela 18 apresenta os resultados da regressão linear simples entre o *grayscale* extraído pela máscara (X) e pelo método (Y), por raio. Os slopes situam-se entre 0,997 e 1,001 em todos os percentuais de raio avaliados, com |slope − 1| inferior a 0,003, o que indica concordância quase perfeita entre os valores obtidos pelas duas fontes de coordenadas. Os valores de R^2 são superiores a 0,95 em todos os raios, atingindo 0,989 no raio de 75%; a redução progressiva de R^2 nos raios menores (0,959 no raio de 10%) é coerente com o aumento do desvio padrão observado na análise de Bland-Altman, confirmando que a dispersão cresce à medida que a região amostrada se afasta das bordas do alvo.

Tabela 18: Comparação das inclinações da *GMF* (regressão linear máscara X vs. método Y) por percentual de raio. Fonte: Elaborado pelo autor (2026).

Raio (%)	*Slope*	Intercepto	R^2	$\lvert slope - 1 \rvert$
100	0.9973	0.5452	0.9873	0.0027
75	1.0004	0.4396	0.9886	0.0004
50	1.0005	0.8571	0.9864	0.0005
25	0.9999	1.0819	0.9794	0.0001
10	0.9995	0.9390	0.9592	0.0005

A Figura A.6 (Apêndice A) complementa a análise de concordância com a dispersão direta dos valores de *grayscale* obtidos por máscara e por método, em cada raio, oferecendo visualização ponto-a-ponto da concordância e dos pares com maior afastamento.

A Figura A.7 (Apêndice A) apresenta a distribuição conjunta de R^2 e *spread* por raio para as combinações avaliadas, sintetizando a relação entre linearidade da *GMF* e amplitude da escala de cinza utilizada.

6.5 Perfis de Métricas Individuais das Melhores Combinações

Os perfis completos de métricas por ROI (3 × 5 grids) para cada uma das cinco melhores combinações são apresentados no Apêndice A (Figuras A.1 a A.5). Esses gráficos mostram, para cada combinação, as curvas ordenadas de todos os 13 indicadores calculados, permitindo identificar o comportamento típico e a proporção de *ROIs* com desempenho abaixo do aceitável.

A leitura individual das cinco figuras revela um padrão estrutural comum: para cada combinação, os quatro indicadores agregados ao *score* composto (*IoU*, área de interseção, *desvio no eixo x normalizado* e *desvio no eixo y normalizado*) apresentam curvas ordenadas com transição suave, sem quebras abruptas, indicando que o desempenho do método varia de forma contínua ao longo do *dataset* e não é dominado por um pequeno subconjunto de imagens excepcionais. Os indicadores de área de interseção em raios menores (50%, 25% e 10%) exibem o platô esperado próximo de 100% para uma proporção elevada das *ROIs*, e degradam apenas nas extremidades direita das curvas, correspondentes às *ROIs* com maior erro posicional. As curvas de *IoU* exibem o decaimento geométrico discutido anteriormente, com valores próximos de zero em raios pequenos mesmo para *ROIs* com alta área de

interseção, reforçando que a inspeção visual desses perfis deve sempre considerar conjuntamente os dois grupos de indicadores.

A Figura A.1 (DP-P + Savgol, combinação líder) exibe os patamares mais elevados nos indicadores de área de interseção, com cerca de 80% das *ROIs* apresentando área de intersecção percentual utilizando 100% do raio acima de 90% e queda concentrada nas últimas 30 *ROIs* ordenadas. A Figura A.2 (DP-HS + Savgol + Regularizar) é praticamente sobreponível à anterior nas regiões de alta qualidade, separando-se apenas no terço final, onde o procedimento de regularização adiciona suavização extra que penaliza *ROIs* com geometria já bem alinhada. A Figura A.3 (DP-HM + *Butterworth* + Savgol), apesar de ter sido a combinação separada estatisticamente pelos testes *post-hoc*, mostra uma curva ligeiramente mais baixa em *IoU* e *desvio no eixo x normalizado*, com inflexão mais pronunciada (comportamento coerente com o uso do método hibridoMedio), que substitui posições estruturais por máximos locais não-restritos quando a detecção de picos falha. A Figura A.4 (DP-P + Savgol + Regularizar) reproduz o padrão geral da líder com pequena perda de estabilidade nas *ROIs* finais, enquanto a Figura A.5 (DP-H + Savgol + Regularizar) apresenta o perfil mais conservador, com indicadores levemente abaixo das demais combinações na mediana, mas variabilidade reduzida.

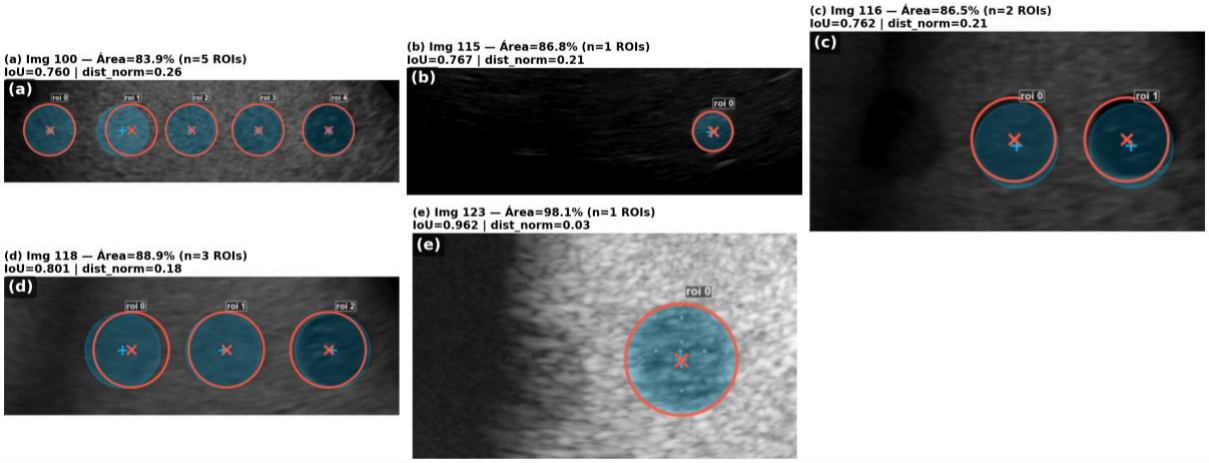
A síntese desses cinco perfis confirma que o *cluster* top-5 da varredura combinatória exaustiva apresenta desempenho equivalente nas *ROIs* facilmente detectáveis e que as diferenças apontadas pelos testes estatísticos se concentram nas *ROIs* difíceis, onde a interação entre filtragem e variante de detecção determina se o método mantém a estimativa estrutural ou se desloca para o máximo local mais próximo. Essa observação é coerente com a recomendação operacional de adotar *DP-P* + Savgol como combinação padrão, reservando a *DP-HM* (butterworth + savgol) como alternativa quando o objetivo for maximizar a robustez em imagens de baixa qualidade.

6.6 Análise Qualitativa: Exemplos de Detecção

A interpretação quantitativa apresentada nas seções anteriores é complementada pela inspeção visual direta de casos individuais, selecionados para ilustrar o comportamento típico do método em situações de sucesso e de desvio posicional. As Figuras 29 e 30 apresentam, respectivamente, casos representativos

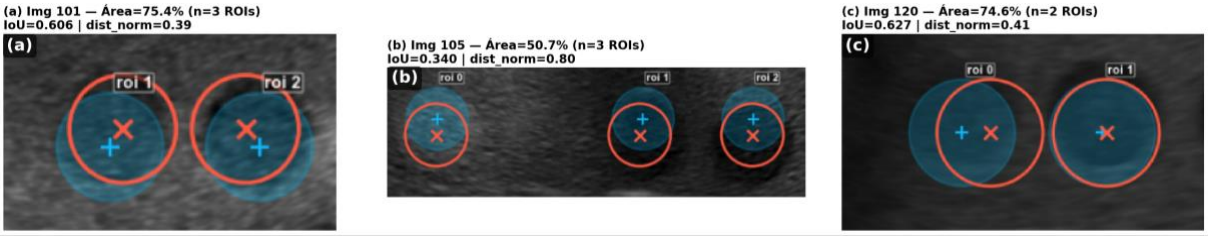
de detecção bem-sucedida e casos com desempenho inferior, ambos obtidos com a combinação líder *DP-P* + Savgol. Em cada imagem, o círculo azul com marcador + representa a região de referência da máscara manual (ground truth), o círculo vermelho com marcador × representa a detecção automática, e o sombreamento azul indica a área de sobreposição efetiva.

Figura 29: Exemplos de detecção bem-sucedida (*DP-P* + Savgol) em cinco imagens do *dataset*. Azul com + = máscara manual (GT); vermelho com × = detecção automática. A sobreposição visual dos círculos é pronunciada em todos os casos. Cada subfigura é identificada por letra: (a) Img 100 - 5 *ROIs*; (b) Img 115; (c) Img 116; (d) Img 118; (e) Img 123. Fonte: Elaborado pelo autor (2026).



A Figura 29 reúne cinco imagens com características imagiológicas distintas, a saber: número variável de alvos (1 a 5), raios que variam de 10 a 53 px, e contextos de aquisição que incluem tecidos com speckle denso (Figuras 29b, 29c, 29d) e alvos de grande diâmetro com contraste moderado (Figura 29e, raio~37 px, área de interseção de 98,1%). Em todos os casos, os círculos vermelho e azul são visualmente sobrepostos, e as métricas confirmam qualidade de detecção acima do limiar operacional: área média entre 83% e 98% e *IoU* entre 0,76 e 0,96. A Figura 29a, que possui 5 *ROIs* na mesma imagem, é particularmente ilustrativa: as *ROIs* centrais (2, 3 e 4) são praticamente perfeitas, enquanto a *ROI* 0, localizada na extremidade da linha de alvos, apresenta desvio ligeiramente maior (área=54%), evidenciando que a confiabilidade do método é mais alta nos alvos com melhor contraste relativo ao fundo local.

Figura 30: Exemplos de detecção com desvio posicional ou condições imagiológicas adversas (*DP-P* + Savgol). Azul com + = máscara (GT); vermelho com x = detecção automática. Subfiguras: (a) Img 101; (b) Img 105; (c) Img 120. Fonte: Elaborado pelo autor (2026).



A Figura 30 ilustra três padrões distintos de desvio posicional, identificáveis na inspeção visual mesmo quando parte das métricas agregadas ainda se encontra em faixa aceitável:

Fig. 30a - Img 101 (área média=75,4%): Ambas as *ROIs* exibem o centro detectado (×, vermelho) deslocado para cima em relação ao centróide da máscara (+, azul), com adicionalmente um leve desvio horizontal para a direita na primeira *ROI* e para a esquerda na segunda. O deslocamento vertical para cima é um padrão recorrente em imagens que apresentam regiões de menor intensidade (mais escuras) na borda superior dos alvos: nesses casos, o sinal de variância por colunas produz um pico deslocado superiormente ao centro geométrico real da alvo, resultando no viés observado.

Fig. 30b - Img 105 (área média=50,7%): Este é o caso de maior desvio do conjunto. O problema principal foi a não identificação de uma das regiões entre as *ROIs* 0 e 1 - o método falhou em separar corretamente os alvos adjacentes, comprometendo a localização dos centróides. Além disso, toda a detecção sofreu um deslocamento sistemático para **baixo**, associado à presença de uma região mais escura na parte inferior da imagem, que desloca o pico do sinal de variância inferiormente ao centro real dos alvos. A combinação desses dois fatores explica a menor área de sobreposição (50,7%) entre todos os casos apresentados.

Fig. 30c - Img 120 (área média=74,6%): As duas *ROIs* exibem comportamentos completamente distintos na mesma aquisição. A *ROI* 0 (à esquerda) apresenta o × deslocado expressivamente para a **direita** do +, com sobreposição apenas parcial entre os dois círculos. Em contraste, a *ROI* 1 (à direita) mostra × e + praticamente coincidentes, com sobreposição quase perfeita. O motivo pelo qual o método falhou especificamente para a *ROI* 0 não é claro; trata-se de uma falha pontual do *pipeline*

cuja causa não foi determinada, evidenciando que o método pode apresentar erros isolados mesmo em imagens de boa qualidade geral.

7 Discussão

7.1 Síntese dos Achados Principais

A análise das três famílias metodológicas (AM2D, ADS1D e DP) por *varredura combinatória exaustiva* com 2.100 combinações mostrou que a Detecção por Picos com Restrição Geométrica é a estratégia mais eficaz para localizar alvos de phantom em imagens de ultrassom Modo B. As cinco melhores combinações utilizam a métrica de variância por colunas e o baseline por bordas da ROI (roi_edges), restringindo a decisão prática às escolhas de variante de detecção e cadeia de filtros. A combinação líder (DP-P + Savgol) obteve IoU médio de 0,7663, área de interseção percentual utilizando 100% do raio de 85,06% e desvios normalizados médios de 0,171 no eixo x e 0,127 no eixo y sobre 339 ROIs em 156 imagens.

Embora os desvios médios de localização observados correspondam a aproximadamente 0,17 do raio do alvo no eixo x e 0,13 no eixo y, os resultados obtidos para a Função de Mapeamento de Cinza demonstram que essa discrepância geométrica não produz impacto proporcional sobre os indicadores radiométricos extraídos. Isso sugere que a região central dos alvos apresenta relativa homogeneidade de intensidade, permitindo que pequenas variações na posição da região amostrada preservem adequadamente a caracterização em escala de cinza necessária para a aplicação proposta.

O teste de Friedman rejeitou a equivalência global (χ^2=36,38; $p<10^{-6}$). O post-hoc de Wilcoxon encontrou 4 dos 10 pares significativos (todos envolvendo DP-HM com butterworth + savgol), confirmando que as quatro demais combinações são estatisticamente equivalentes.

A concordância via GMF (Bland-Altman e regressão mask vs método) mostrou viés máximo de 1,08 nível de cinza, inclinações equivalentes à unidade ($|slope-1| \leq 0,0027$) e R^2 entre 0,959 e 0,989. Esses resultados confirmam que o método produz escala de cinza equivalente ao ground truth manual.

7.2 Interpretação do Score Composto e dos Testes Estatísticos

Embora o score composto normalizado varie de 0,75 (1º) a 0,25 (5º), essa amplitude é artificial: a normalização sobre apenas cinco pontos infla as diferenças.

As médias reais por imagem são muito próximas: DP-P (SG) = 0,8517; DP-HM (BW + SG) = 0,8486; DP-P (SG + R) = 0,8408; DP-HS (SG + R) = 0,8381; DP-H (SG + R) = 0,8330. A diferença entre o 1º e o 5º colocados é de apenas 0,019 (1,9%), enquanto a variabilidade intra-método (~0,13) é cerca de sete vezes maior.

A análise das diferenças pareadas por imagem mostra que DP-P (SG), 1º na média, vence apenas 46–48% das imagens contra os métodos 3º a 5º: sua média superior decorre de margens maiores quando vence, não de maior consistência. Já DP-HM (BW + SG) vence 60–66% das imagens contra esses mesmos métodos, o que explica por que o teste de Wilcoxon (baseado em ranks) detecta significância apenas nos pares que o envolvem.

7.3 Adequação da Métrica *de variança* e da Extração *roi_edges*

A dominância do par (var, roi_edges) nas 5 melhores combinações tem explicação técnica. A variância por coluna funciona como descritor de heterogeneidade local: no fundo ecogênico, o speckle produz dispersão estacionária; sobre os alvos, transições de contraste geram dispersão diferente. Ao contrário da média ou mediana, a variância responde igualmente a alvos hiper e hipoecogênicos (sinal sempre positivo), sem necessidade de tratamento polar para contrastes opostos, o que é relevante para phantoms com alvos em ambas as direções (−15 a +15 dB). O baseline roi_edges estima o nível de fundo a partir das bordas superior e inferior do retângulo selecionado, evitando contaminação pelos alvos. Comparado ao baseline por média do sinal (signal), produz referência mais estável em imagens com alvos de amplitudes assimétricas.

7.4 Concordância com a Máscara (*Bland-Altman*)

A análise de Bland-Altman das diferenças (método − máscara) em 281 ROIs pareadas por raio mostra concordância forte. O viés é positivo e inferior a 1,1 nível de cinza em todos os raios, uma leve tendência do método a estimar valores um pouco mais altos, atribuível ao deslocamento residual dos centros em direção a regiões mais brilhantes do interior dos alvos. O desvio-padrão das diferenças cresce de 4,83 (raio 100%) para 9,53 (raio 10%), refletindo maior sensibilidade a erros posicionais em regiões menores, e a proporção de pares dentro dos limites de concordância de 95% (LoA) permanece entre 95,0% e 97,2%. A regressão linear confirma relação 1:1 entre

método e máscara, sem distorção de escala: inclinações praticamente unitárias (|slope−1| ≤ 0,0027) e R^2 entre 0,959 e 0,989.

As Tabelas 15 a 18 apontam o raio de 75% como o melhor ponto de operação. Ele reúne os melhores valores em todos os critérios: inclinação próxima de 1 (1,0004), maior R^2 de regressão (0,9886), viés baixo (+0,46), limites de concordância estreitos (−9,0 a +10,0 níveis de cinza), 96,8% dos pares dentro desses limites e maior R^2 da GMF (0,877). Os demais raios mostram por que essa fração é a mais adequada: raios menores (10–25%) amostram uma região pequena e sensível a erros de posição, o que alarga os limites de concordância (±13 a ±19 níveis de cinza); o raio de 100% inclui mais pixels fora do alvo quando há pequenos erros na localização do centro. O raio de 75% é, assim, o que melhor combina precisão de posição e fidelidade ao conteúdo do alvo, sendo a escolha mais adequada para extrair os indicadores de escala de cinza.

7.5 Implicações Clínicas

A automatização da detecção das ROIs e da extração dos indicadores de escala de cinza elimina a variabilidade da seleção manual operador-dependente, que constitui a principal limitação à reprodutibilidade dos testes de constância em ultrassom. Ao remover essa dependência, o método confere objetividade e rastreabilidade aos resultados, condições essenciais para que o teste de escala de cinza cumpra seu papel na garantia da qualidade diagnóstica.

Além do ganho de reprodutibilidade, a abordagem tem implicações diretas na gestão de equipamentos e na manutenção preventiva. Como o processamento por imagem leva poucos segundos em hardware convencional, é viável executá-lo em lote sobre coletas periódicas e acompanhar a evolução dos indicadores ao longo do tempo. Esse monitoramento permite detectar precocemente a degradação de desempenho de um transdutor ou equipamento (antes que ela comprometa o diagnóstico) e subsidiar decisões de calibração, reparo ou substituição com base em evidência objetiva, e não apenas em inspeção visual pontual.

7.6 Limitações

O presente trabalho apresenta limitações que devem ser consideradas na interpretação dos resultados. Em primeiro lugar, o tamanho amostral efetivo é de 156

imagens válidas para a comparação espacial das combinações e de 93 imagens (com 2 ou mais *ROIs*) para a análise da *GMF*. Embora o conjunto englobe equipamentos de diferentes fabricantes e gerações, há concentração nas plataformas Canon/Toshiba, o que pode introduzir viés de plataforma nas estatísticas agregadas, em especial quanto às características de filtragem proprietária e ao perfil de *speckle* de cada sistema. Além disso, as imagens que apresentaram piores classificações no *ranking* composto foram frequentemente aquelas com alvos de baixo contraste em relação ao fundo ou com artefatos de reverberação na região inferior da imagem, onde o ruído acústico se acumula. Essa degradação em profundidade é inerente à física da formação da imagem ultrassonográfica, mas não foi modelada explicitamente pelo *pipeline*, que trata todos os alvos com a mesma cadeia de filtros independentemente de sua posição axial.

Em segundo lugar, a análise de *Bland-Altman* e a comparação de inclinações da *GMF* foram conduzidas exclusivamente para a combinação líder *DP-P*+Savgol. A generalização desses resultados de concordância para as demais combinações do *cluster* top-5 é plausível em função da equivalência estatística observada no *score* composto, mas não foi verificada explicitamente. Adicionalmente, as imagens da coleção não foram adquiridas sob um protocolo de aquisição de rigorosa padronização, os parâmetros de frequência, profundidade, ganho, *TGC* e foco variaram entre equipamentos e sessões. Essa heterogeneidade, que é presente na rotina clínica, trás um aspecto mais variado ao teste, entretanto, impede a separação do efeito do método de detecção do efeito das condições de aquisição sobre os indicadores de desempenho. Replicar a análise de concordância para as cinco combinações forneceria evidência adicional sobre a robustez funcional do *cluster*.

Em terceiro lugar, a *IoU* em raios pequenos é geometricamente penalizada conforme discutido na Seção 7.3. Embora a literatura de detecção de objetos forneça o arcabouço para essa interpretação, a presente avaliação não inclui métricas alternativas (como Dice ponderado por raio ou *IoU* normalizada) que poderiam mitigar a limitação. A escolha de manter *IoU* clássica foi deliberada para preservar a comparabilidade com a literatura, mas restringe a expressividade do indicador em regimes de raio reduzido.

Em quarto lugar, a ordenação dos testes estatísticos seguiu uma lógica estruturada, a saber, teste global (*Friedman*) seguido de comparações par-a-par

(*Wilcoxon*), mas não se discutiu em que medida a escolha de um *score* composto único, em vez de múltiplos indicadores avaliados separadamente, pode ter mascarado diferenças significativas em dimensões individuais (e.g., *IoU* versus desvio posicional). Trabalhos futuros podem contemplar análises complementares por indicador para verificar se a equivalência estatística observada no *score* composto se mantém em cada dimensão isoladamente.

Por fim, o presente trabalho não realiza validação prospectiva em equipamentos não incluídos no conjunto de treinamento metodológico. Embora o *pipeline* seja, por construção, independente de aprendizado supervisionado, a validação externa em equipamentos novos, em particular em plataformas de fabricantes não representados na coleta atual, é necessária antes da adoção do método como referência para teste de constância em rotina clínica.

8 Conclusões

Este trabalho propôs, implementou e validou um *pipeline* computacional reprodutível para a automatização do teste de escala de cinza em imagens de ultrassom *Modo B*, aplicado a *phantoms* multipropósito de controle de qualidade. O *pipeline* cobre todas as etapas operacionais do ensaio, desde a leitura dos arquivos *DICOM*, calibração espacial e seleção semi-supervisionada das regiões de busca até a detecção automática dos centros dos alvos, a extração dos indicadores de *grayscale* e a estimativa da *Grayscale Mapping Function*, e foi avaliado sistematicamente por meio de varredura combinatória exaustiva abrangendo 2.100 combinações de hiperparâmetros e 327.600 execuções individuais.

Os objetivos específicos centrais enunciados na introdução foram atendidos no escopo do presente trabalho. Foi desenvolvida e comparada uma família de três estratégias metodológicas (*AM2D*, *ADS1D* e *DP*), com a escolha justificada da Detecção por Picos com Restrição Geométrica como *pipeline* consolidado. Foi formulado um conjunto de quatro indicadores espaciais (desvios normalizados nos eixos, área de interseção percentual e *IoU* em cinco percentuais de raio) e um *score* composto multicritério para o ranqueamento das combinações. Foram conduzidas análises estatísticas robustas (*Friedman*, *Wilcoxon* pareado) capazes de distinguir efeitos sistemáticos de variabilidade amostral, e foi implementada a análise de

concordância *Bland-Altman* + regressão linear entre as extrações com coordenadas da máscara e do método para validar a equivalência funcional do *pipeline*.

A combinação líder identificada (designada como *DP-P*+Savgol: *var* | picos | *roi_edges* | ['savgol']) obteve *IoU* médio de 0,7663 utilizando 100% do raio, área de interseção percentual em 100% do raio de 85,06%, desvios normalizados médios de 0,171 (eixo x) e *0,127 (eixo y) e score* composto médio por imagem (Score B) de 0,8517, o maior entre as cinco combinações avaliadas. As cinco melhores combinações compartilham, de forma robusta, a métrica de variância por coluna e o *baseline* calculado pelas bordas da *ROI*, o que demonstra a importância dessas duas escolhas estruturais. O teste de *Friedman* rejeitou a equivalência global das cinco combinações ($\chi^2 = 36{,}38$; $p < 10^{-6}$), e o *post-hoc Wilcoxon* identificou quatro pares significativos, todos envolvendo a combinação *DP-HM* (*var*+hibridoMedio+butterworth+savgol). Os seis pares restantes não apresentaram diferenças estatisticamente significativas, indicando que as quatro melhores combinações são estatisticamente equivalentes em desempenho médio. A escolha da DP-P (SG) como pipeline consolidado apoia-se em sua maior média agregada; a DP-HM (BW + SG), por vencer a maioria dos confrontos diretos por imagem (60–66%), permanece como alternativa recomendada quando o objetivo for maximizar a robustez caso a caso.

A análise de concordância pela *GMF* confirmou que o método automatizado é funcionalmente equivalente à extração com coordenadas da máscara. Em todos os percentuais de raio avaliados (10%, 25%, 50%, 75% e 100%), o viés sistemático permaneceu inferior a 1,1 nível de cinza, o coeficiente angular da regressão linear foi estatisticamente equivalente à unidade ($|\ slope - 1\ |$ máximo de 0,0027), e o coeficiente de determinação R^2 superou 0,959, atingindo 0,9886 no raio de 75%. A proporção de pares dentro dos limites de concordância de 95% manteve-se entre 95,02% e 97,15%, em conformidade com a expectativa teórica de aproximadamente 95% sob normalidade. O raio de 75% foi identificado como o ponto ótimo operacional, conciliando estreita concordância com a máscara e robustez ao erro posicional residual.

As contribuições principais deste trabalho são de três naturezas. Do ponto de vista metodológico, a formulação da Detecção por Picos com Restrição Geométrica, em suas cinco variantes (*DP-P*, *DP-E*, *DP-H*, *DP-HM* e *DP-HS*), oferece um método

unificado para a localização de alvos em *phantoms* de ultrassom, com graus crescentes de incorporação de informação geométrica do objeto físico. Segundo, uma contribuição de validação experimental: a avaliação sistemática de 2.100 combinações sobre um *dataset* multi-equipamento, respaldada por testes estatísticos (*Friedman* e *Wilcoxon* pareado), permitiu identificar a configuração mais robusta com fundamento quantitativo. Terceiro, contribuição de infraestrutura: o *pipeline* foi implementado em código aberto (*Python*, com pydicom, *OpenCV*, *scikit-image*, *SciPy*, ruptures e *matplotlib*).

Em síntese, a extração automática das regiões de interesse reproduz, com elevado grau de concordância, os indicadores de escala de cinza obtidos pela segmentação manual, com precisão posicional compatível com a leitura por máscara. O método proposto cumpre, assim, o requisito funcional de extrair esses indicadores de forma reprodutível e oferece uma base computacional sólida para incorporar de modo automatizado o teste de constância em escala de cinza ao fluxo de controle de qualidade em ultrassonografia diagnóstica, com potencial para reduzir a subjetividade das avaliações visuais tradicionais e contribuir para a padronização desses procedimentos.

9 Trabalhos Futuros

A continuidade deste trabalho contempla sete eixos de desenvolvimento prioritários, que aprofundam, generalizam e operacionalizam as contribuições aqui apresentadas.

O primeiro eixo é a investigação do papel do teste de uniformidade de campo como etapa potencialmente precursora ao teste de escala de cinza. Diferentemente da escala de cinza, que avalia a resposta dinâmica do equipamento a contrastes de atenuação variáveis, o teste de uniformidade caracteriza a estabilidade da resposta do equipamento ao longo da extensão lateral e em profundidade. Uma pergunta em aberto é se eventuais não uniformidades do campo comprometem sistematicamente os indicadores extraídos pelo *pipeline* de escala de cinza; caso isso se confirme, a execução do teste de uniformidade deveria preceder a análise de escala de cinza como etapa obrigatória de validação. Trabalhos futuros deverão examinar essa relação de dependência, avaliando em que medida a presença de artefatos de não uniformidade afeta os resultados do método aqui proposto.

O segundo eixo diz respeito à definição de um protocolo padronizado de aquisição para a metodologia de detecção de centros. Uma decisão metodológica deliberada deste trabalho foi a de incorporar imagens obtidas com variados parâmetros de aquisição e configurações de equipamento, o que, embora torne o desenvolvimento mais complexo (pois exige que o método opere sobre uma maior diversidade de condições de imagem), foi uma escolha que demonstra a robustez do *pipeline* frente a cenários heterogêneos, aproximando-o das condições reais de uso clínico. Esse caminho mais desafiador foi adotado precisamente para evidenciar que o método não depende de condições idealizadas para funcionar. No entanto, trabalhos futuros poderiam investigar os benefícios de uma coleta com protocolo mais rígido e controlado: a padronização de parâmetros como ganho, profundidade, frequência e modelo de *phantom* poderia reduzir a variabilidade espúria nas métricas de detecção e permitir uma avaliação mais precisa dos limites do método em condições otimizadas.

O terceiro eixo é a exploração de arquiteturas híbridas que combinem a metodologia de detecção de centros aqui proposta com representações extraídas por redes convolucionais. A abordagem consiste em utilizar as coordenadas de centros detectadas pelo *pipeline* como entradas explícitas, junto com mapas de características obtidos nas camadas de convolução de uma rede treinada sobre as mesmas imagens, para alimentar um modelo que realize a detecção final de forma mais precisa. Essa fusão entre informação geométrica de alto nível, provida pela metodologia de detecção de pontos, e informação local de baixo nível, extraída automaticamente por convoluções, poderia ampliar a acurácia posicional e a robustez do método em imagens com maior variabilidade de contraste ou artefatos de aquisição.

10 Referências

ATS LABORATORIES INCORPORATED, CIRS Company. **Fantoma Multifuncional Modelo 539**. Norfolk, Virginia: *[S.n.]*.

BRASIL. **INSTRUÇÃO NORMATIVA - IN Nº 96, DE 27 DE MAIO DE 2021 - INSTRUÇÃO NORMATIVA - IN Nº 96, DE 27 DE MAIO DE 2021 - DOU - Imprensa Nacional**. Diário Oficial da União, , 2021. Disponível em: <https://www.in.gov.br/en/web/dou/-/instrucao-normativa-in-n-96-de-27-de-maio-de-2021-323003884>. Acesso em: 27 jan. 2025

BRASIL. **RESOLUÇÃO RDC Nº 611, DE 9 DE Março DE 2022 - RESOLUÇÃO RDC Nº 611, DE 9 DE Março DE 2022 - DOU - Imprensa Nacional**. Diário Oficial da União, , 2022. Disponível em: <https://www.in.gov.br/en/web/dou/-/resolucao-rdc-n-611-de-9-de-marco-de-2022-386107075>. Acesso em: 27 jan. 2025

BUSHBERG, Jerold T. *et al.* **The Essential Physics of Medical Imaging**. 3. ed. *[S.l.]*: Wolters Kluwer | Lippincott Williams & Wilkins., 2012.

CANNY, John. A Computational Approach to Edge Detection. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. PAMI-8, n. 6, p. 679–698, 1986.

CIRS. **Multi-Purpose, Multi-Tissue Ultrasound Phantom Model 040GSE**. *[S.l.: S.n.]*.

CIRS. **General Purpose Ultrasound Phantom Model 054GS**. Norfolk, Virginia: *[S.n.]*.

DICOM LIBRARY. **DICOM Tags - DICOM Library**. Disponível em: <https://www.dicomlibrary.com/dicom/dicom-tags/>. Acesso em: 27 maio. 2026.

DUDA, Richard O.; HART, Peter E. USE OF THE HOUGH TRASFORMTION TO DETECT LINES AND CURVES IN PICTURES. **Comm. ACM**, v. 15, p. 11–15, 1972.

DUDA, Richard O.; HART, Peter E.; STORK, David G. Chapter 9 - Algorithm-Independent Machine Learning. **Pattern Classification**, p. 680, 2001.

FRIEDMAN, Milton. The Use of Ranks to Avoid the Assumption of Normality Implicit in the Analysis of Variance. **Journal of the American Statistical Association**, v. 32, n. 200, p. 675, dez. 1937.

GAMMEX, Inc. **403 GS LE Precision Multi-Purpose Phantoms**. Middleton: *[S.n.]*.

GONZALEZ, Rafael C. Support Package for Digital Image Processing. 2018.

GOODMAN, J. W. Some fundamental properties of speckle*. 1976.

GRAZHDANI, H. *et al.* Quality assurance of ultrasound systems: current status and review of literature. **Journal of Ultrasound**, v. 21, n. 3, p. 173, 1 set. 2018.

HARALICK, Robert M.; DINSTEIN, Its'hak; SHANMUGAM, K. Textural Features for Image Classification. **IEEE Transactions on Systems, Man and Cybernetics**, v. SMC-3, n. 6, p. 610–621, 1973.

INTERNATIONAL ELECTROTECHNICAL COMMISSION. **Protocolo INTERNATIONAL STANDARD IEC 61391-2:2010**. *[S.l.: S.n.]*.

JENSEN, Jørgen Arendt. **Estimation of blood velocities using ultrasound : a signal processing approach**. *[S.l.]*: Cambridge University Press, 2004.

JENSEN, Jørgen Arendt. **Medical ultrasound imaging**. **Progress in Biophysics and Molecular Biology**, jan. 2007.

KONEX. **MANUAL DO USUÁRIO FANTOMA MULTIPROPÓSITO DE ULTRASSONOGRAFIA MODELO US-MP**. *[S.l.: S.n.]*.

MAHESH, Mahadevappa. The Essential Physics of Medical Imaging, Third Edition. **Medical physics**, v. 40, n. 7, jul. 2013.

NICK DUDLEY, Lincolnshire *et al.* **IPEM Report 102 - Quality Assurance of Ultrasound Imaging Systems**. *[S.l.: S.n.]*.

OLIVEIRA, Lutero Marques *et al.* Avaliação da qualidade de imagens de equipamentos de ultrassom modo B. **Research on Biomedical Engineering**, v. 26, n. 1, p. 11–24, 16 jun. 2011.

PAVERD, Catherine *et al.* Assessment of ultrasound image quality in a reference phantom using gel and liquid standoff pads. **WFUMB Ultrasound Open**, v. 2, n. 2, p. 100051, 1 dez. 2024.

PESSÔA, Anderson Guedes. Avaliação dos equipamentos de ultrassonografia da região metropolitana do Recife. 2009.

RANGARAJ M. RANGAYYAN. Biomedical Image Analysis. 30 dez. 2004.

SASSAROLI, Elisabetta *et al.* Breast ultrasound technology and performance evaluation of ultrasound equipment: B-mode. **IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control**, v. 64, n. 1, p. 192–205, 1 jan. 2017.

SASSAROLI, Elisabetta *et al.* Image quality evaluation of ultrasound imaging systems: advanced B-modes. **Journal of Applied Clinical Medical Physics**, v. 20, n. 3, p. 115–124, 1 mar. 2019.

SENRA FILHO, Antonio Carlos Silva da *et al.* A computational tool as support in B-mode ultrasound diagnostic quality control. **Revista Brasileira de Engenharia Biomedica**, v. 30, n. 4, p. 402–405, 2014.

SOCIEDAD ESPAÑOLA DE FÍSICA MÉDICA. **Protocolo Español de control de calidad en radiodiagnóstico**. Madrid: *[S.n.]*.

TAVARES DE SOUZA, Marcela; DIAS DA SILVA, Michelly; DE CARVALHO, Rachel. Revisão integrativa: o que é e como fazer Integrative review: what is it? How to do it? v. 8, n. 1, p. 102–108, 2010.

TOMASI, C.; MANDUCHI, R. Bilateral Filtering for Gray and Color Images. *[S.d.]*.

TSAPAKI, Virginia *et al.* Development and implementation of a quality control protocol for B-mode ultrasound equipment. **Journal of Ultrasound**, v. 25, n. 2, p. 155–165, 8 abr. 2021.

VAN HORSSEN, Pepijn *et al.* Automated quality control of ultrasound based on in-air reverberation patterns. **Ultrasound: Journal of the British Medical Ultrasound Society**, v. 25, n. 4, p. 229, 1 nov. 2017.

WAGNER, Robert F. *et al.* Statistics of Speckle in Ultrasound B-Scans. **IEEE Transactions on Sonics and Ultrasonics**, v. 30, n. 3, p. 156–163, 1983.

WILCOXON, Frank. Individual Comparisons by Ranking Methods. **Biometrics Bulletin**, v. 1, n. 6, p. 80, dez. 1945.

ZUIDERVELD, Karel. Contrast Limited Adaptive Histogram Equalization. **Graphics Gems**, p. 474–485, 1 jan. 1994.

Apêndice A – Perfis Completos de Métricas por ROI das Cinco Melhores Combinações

Este apêndice apresenta os perfis completos de métricas por ROI (3 × 5 grids) para cada uma das cinco melhores combinações identificadas na Seção 6.5, referenciados no texto principal como Figuras A.1 a A.5.

Figura A.1: Perfil completo de métricas por ROI - DP-P (savgol). Fonte: Elaborado pelo autor (2026).

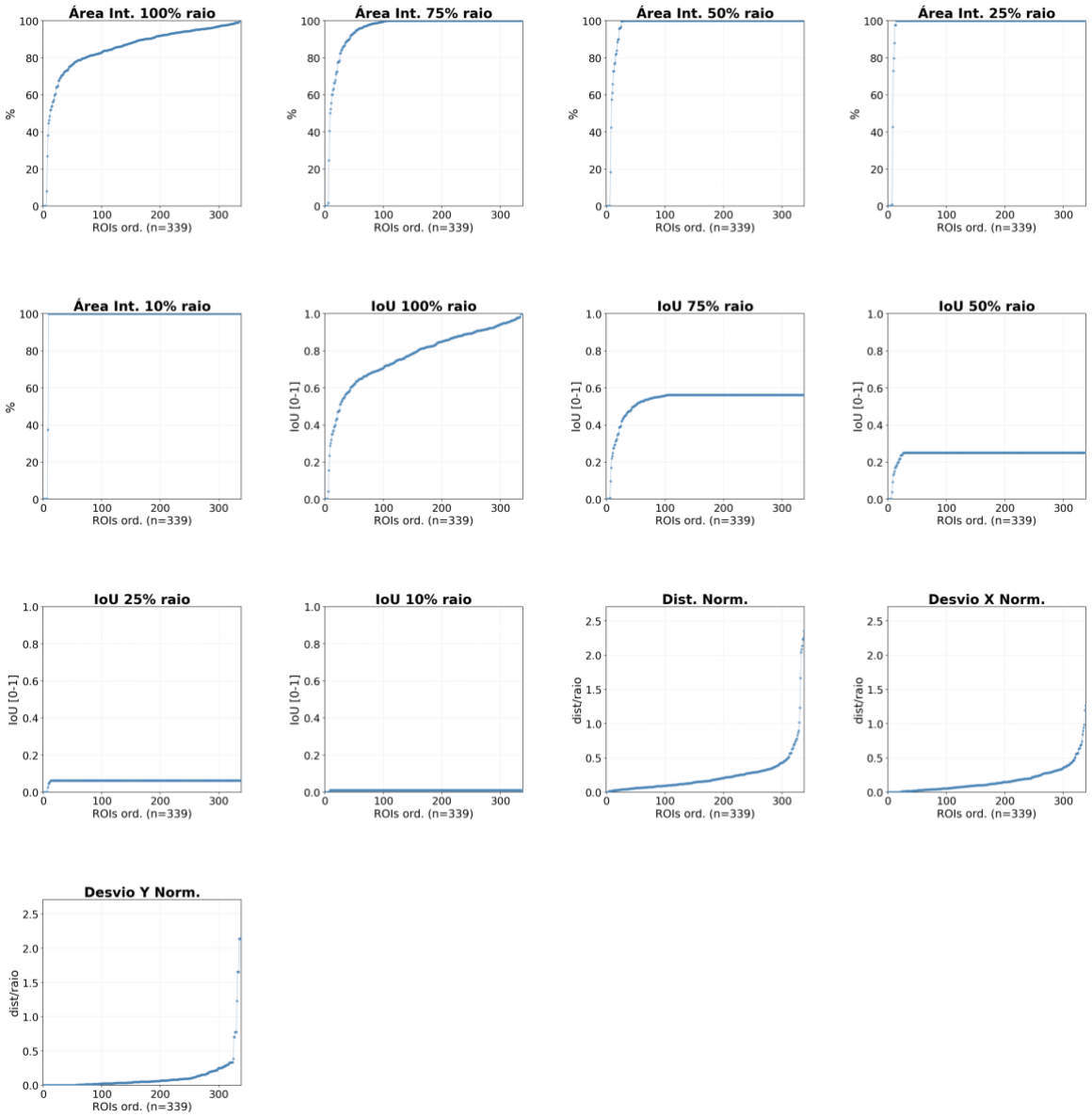


Figura A.2: Perfil completo de métricas por ROI - DP-HS (savgol + regularizar). Fonte: Elaborado pelo autor (2026).

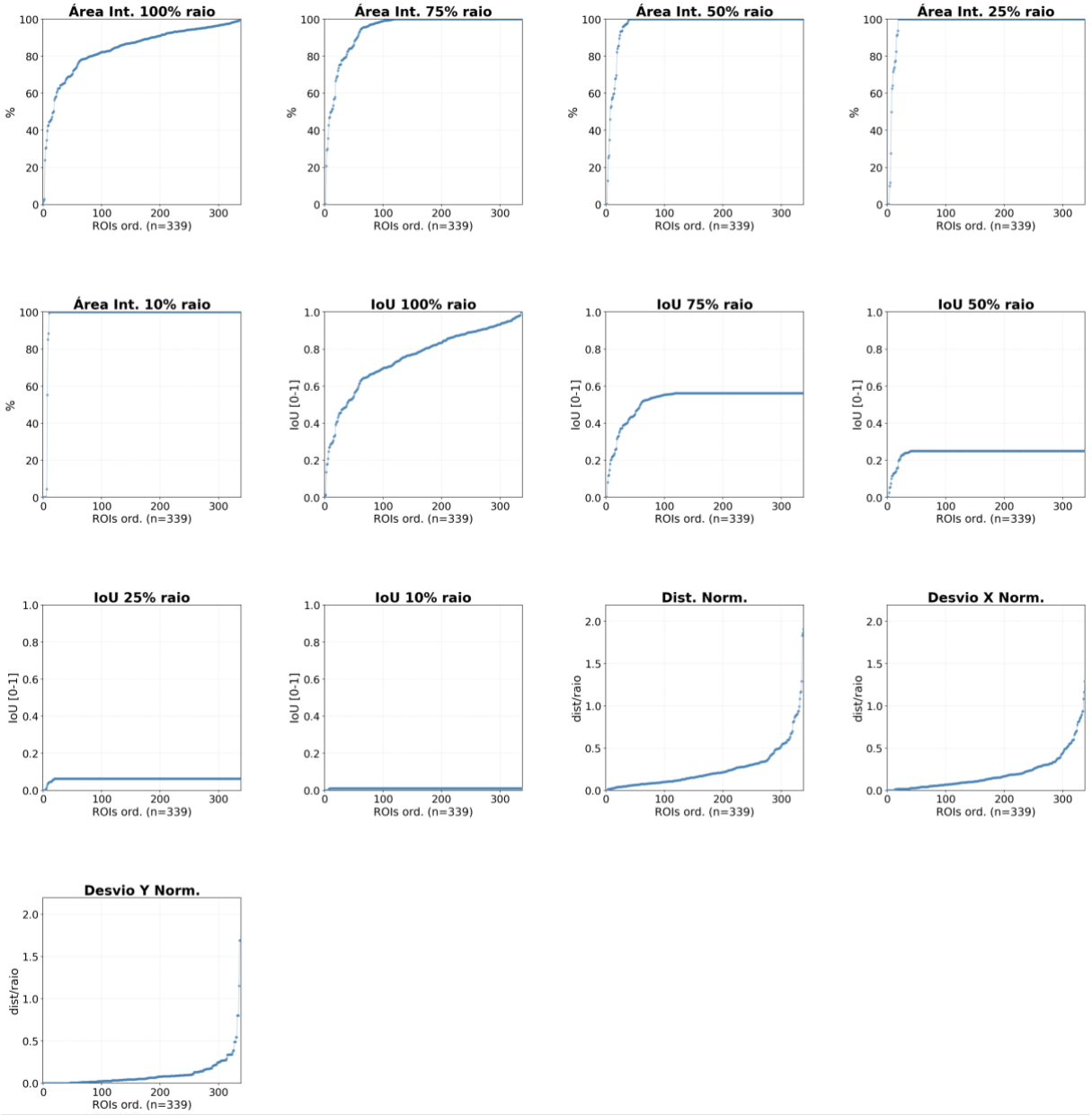


Figura A.3: Perfil completo de métricas por ROI - DP-HM (butterworth + savgol). Fonte: Elaborado pelo autor (2026).

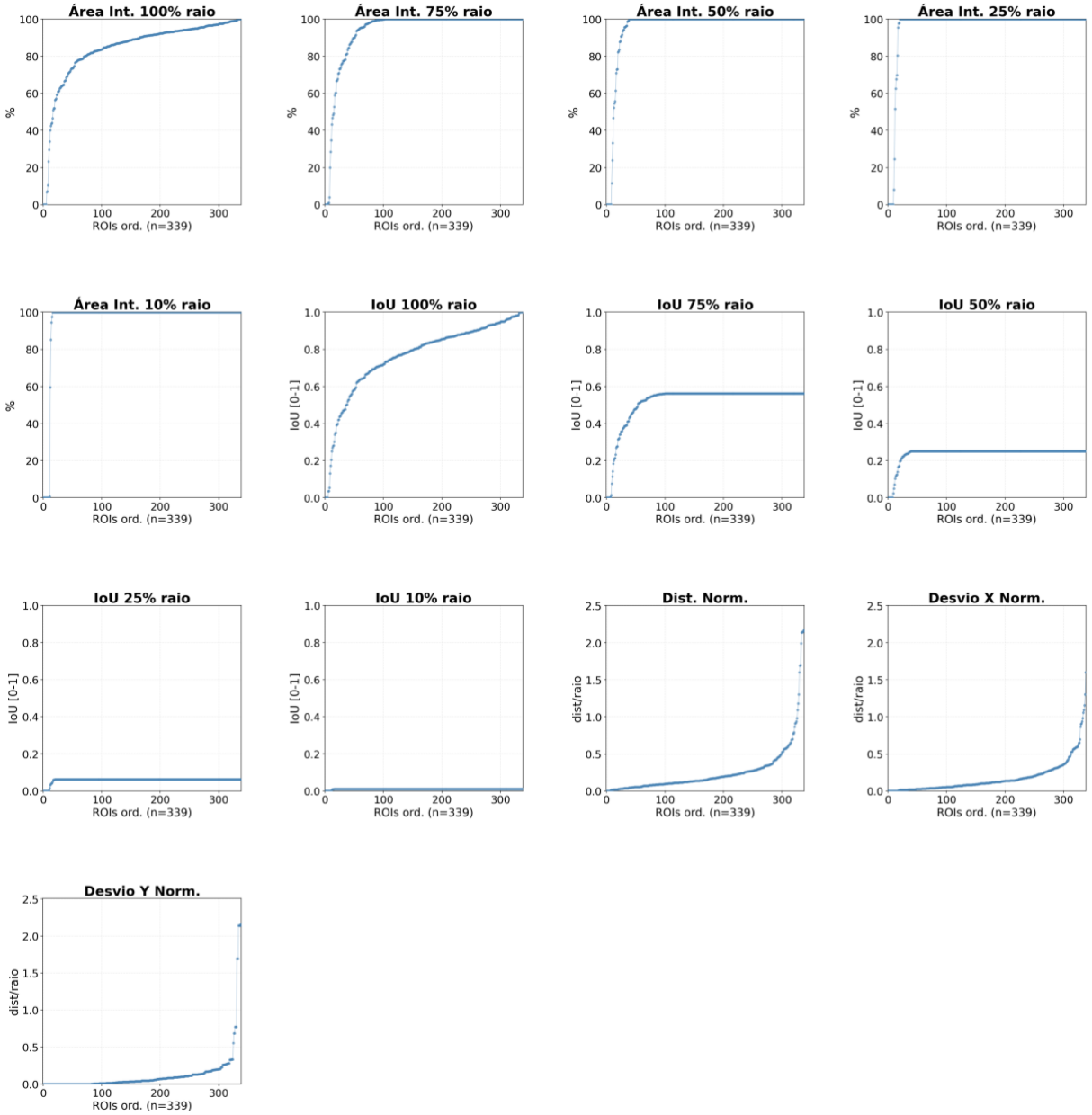


Figura A.4: Perfil completo de métricas por ROI - DP-P (savgol + regularizar). Fonte: Elaborado pelo autor (2026).

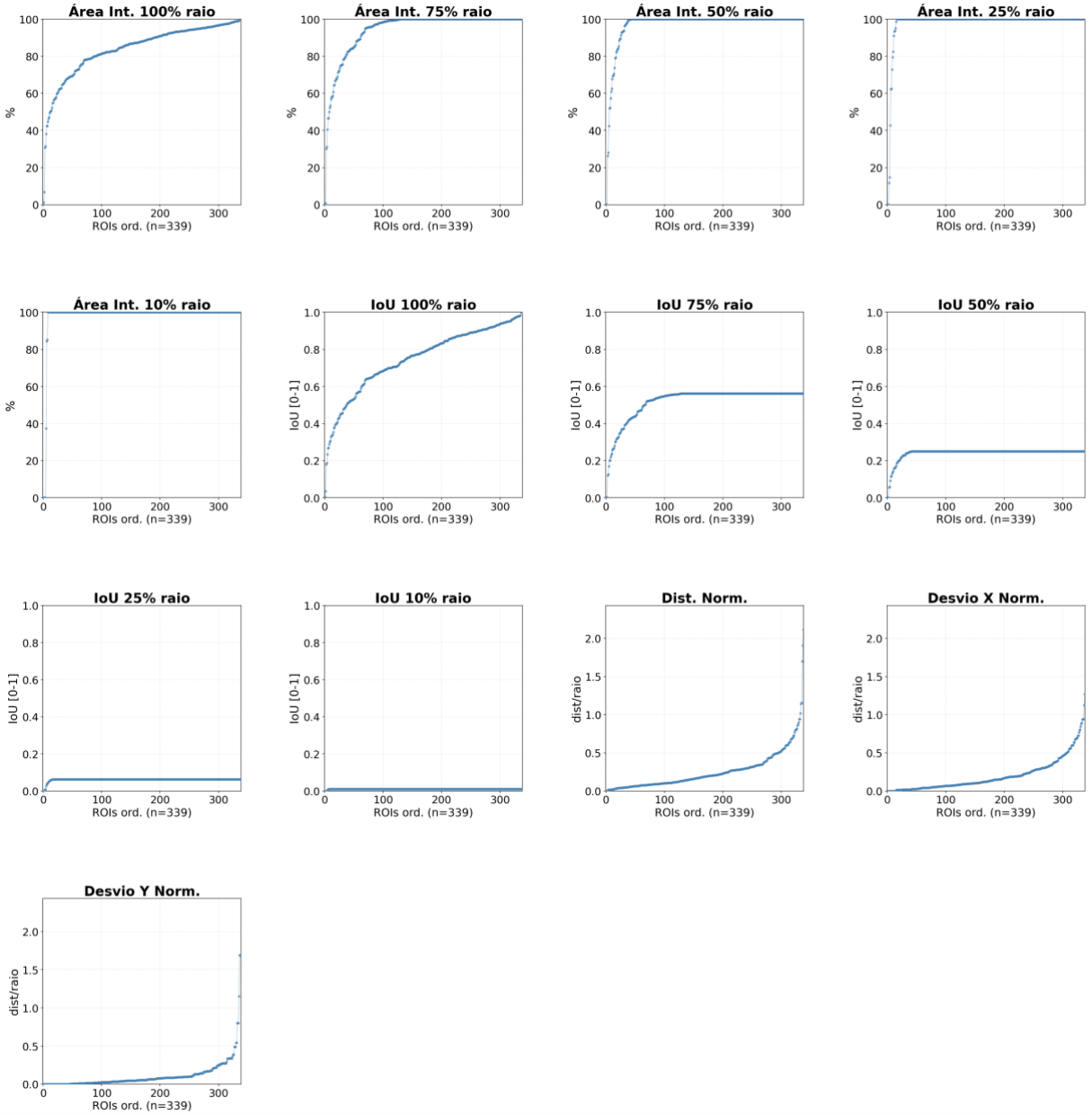


Figura A.5: Perfil completo de métricas por ROI - DP-H (savgol + regularizar). Fonte: Elaborado pelo autor (2026).

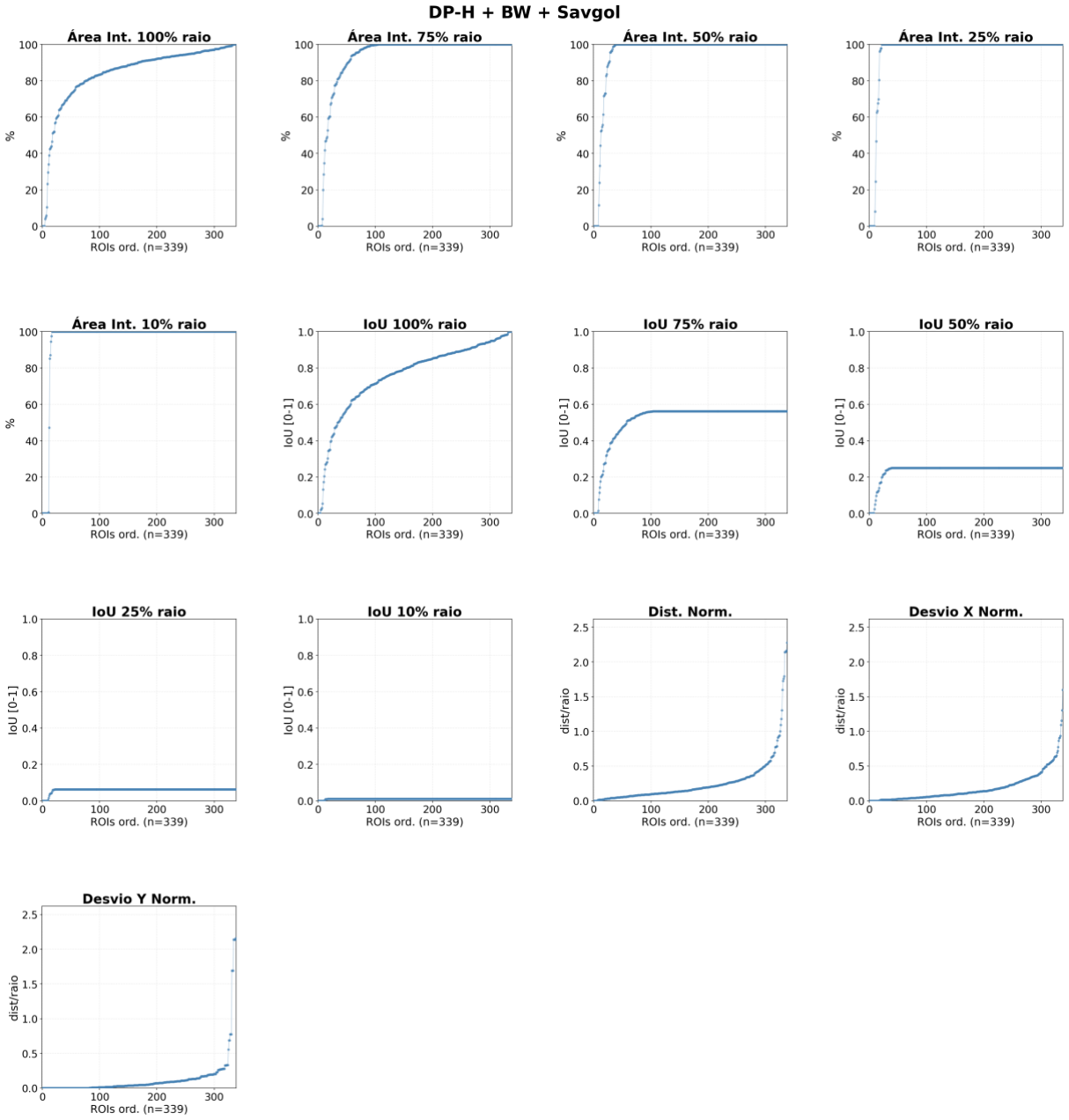


Figura A.6: Diagrama de dispersão entre os valores de grayscale extraídos pela máscara (eixo X) e pelo método automático (eixo Y) para cada percentual de raio avaliado (100%, 75%, 50%, 25% e 10%). Cada ponto representa uma ROI individual. A linha vermelha tracejada indica a concordância perfeita (máscara = método). Fonte: Elaborado pelo autor (2026).

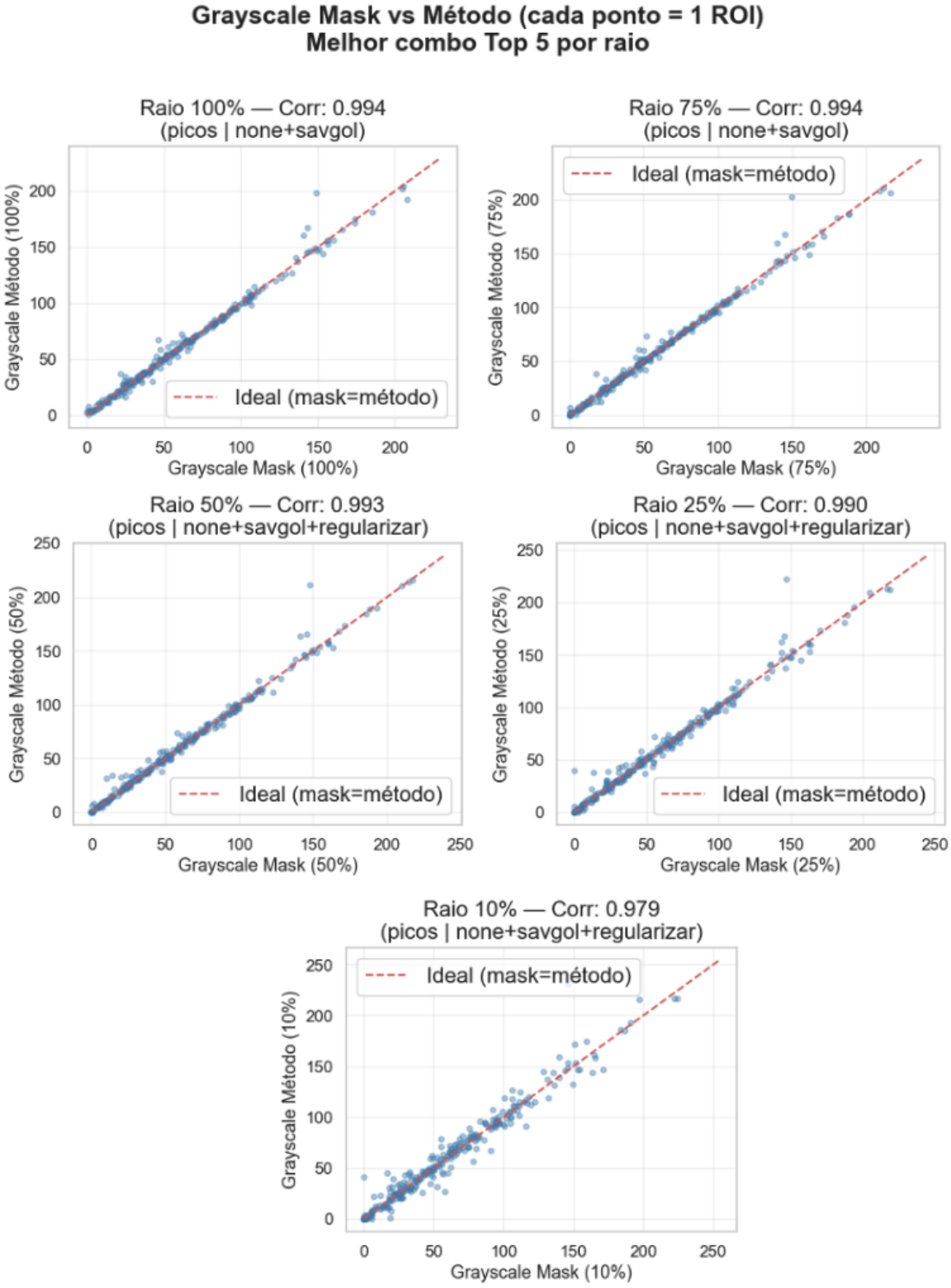


Figura A.7: Distribuição conjunta de R^2 e spread por raio de amostragem, evidenciando o trade-off entre linearidade e amplitude da GMF ao longo dos percentuais de raio considerados. Fonte: Elaborado pelo autor (2026).

