

APLICAÇÕES DA INTELIGÊNCIA ARTIFICIAL NA AVALIAÇÃO DA QUALIDADE TÉCNICA DA MAMOGRAFIA: REVISÃO SISTEMÁTICA

APPLICATIONS OF ARTIFICIAL INTELLIGENCE IN THE ASSESSMENT OF MAMMOGRAPHY TECHNICAL QUALITY: A SYSTEMATIC REVIEW

APLICACIONES DE LA INTELIGENCIA ARTIFICIAL EN LA EVALUACIÓN DE LA CALIDAD TÉCNICA DE LA MAMOGRAFÍA: UNA REVISIÓN SISTEMÁTICA

Mathias Cesar de Assis¹
Nauane Gabriela da Silva Lopes²
Letícia Matsuno da Costa³
Ronaldo Martins da Costa⁴

RESUMO: A qualidade técnica da mamografia exerce influência direta na detecção precoce do câncer de mama, sendo a taxa de repetição (TR) um importante indicador da qualidade dos serviços de diagnóstico por imagem. Este estudo teve como objetivo sintetizar as evidências científicas sobre a aplicação de técnicas de Inteligência Artificial (IA), com ênfase em modelos de aprendizado profundo (*Deep Learning* — DL), na avaliação da qualidade técnica da mamografia e em seu potencial para reduzir a TR. Foi realizada uma Revisão Sistemática da Literatura (RSL), conduzida conforme as diretrizes PRISMA 2020, com buscas nas bases ACM Digital Library, IEEE Xplore, PubMed, ScienceDirect, Scopus, SpringerLink e Web of Science. A qualidade metodológica dos estudos foi avaliada segundo critérios adaptados de Kitchenham. Após o processo de seleção, 13 estudos primários compuseram o corpus da revisão. Observou-se predominância de modelos baseados em DL, especialmente Redes Neurais Convolucionais (CNNs) e arquiteturas derivadas da U-Net, aplicados à identificação de marcos anatômicos, classificação da qualidade técnica das imagens e implementação de sistemas de apoio à decisão e *feedback* durante a aquisição mamográfica. Os estudos reportaram acurácia de até 98,5% na identificação do músculo peitoral, redução da proporção de imagens inadequadas de 13,3% para 3,2% e redução da TR para 0,17% no melhor cenário identificado. As evidências indicam que a IA apresenta elevado potencial para apoiar o controle de qualidade da mamografia, promovendo maior padronização das avaliações, redução da variabilidade entre avaliadores e melhoria do fluxo de aquisição das imagens. Esses resultados reforçam o potencial da IA como ferramenta de apoio aos programas de controle de qualidade da mamografia. Entretanto, sua incorporação em larga escala ainda depende de validação clínica multicêntrica, maior interpretabilidade dos modelos e padronização dos protocolos de avaliação.

Palavras-chave: Mamografia. Posicionamento Mamográfico. Taxa de Repetição.

¹Mestrando em Ciência da Computação, Universidade Federal de Goiás - Instituto de Informática.

²Engenharia de Produção, Pontifícia Universidade Católica de Goiás.

³Medicina, Centro Universitário de Goiatuba – UniCerrado.

⁴Orientador: Universidade Federal de Goiás - Instituto de Informática.

ABSTRACT: The technical quality of mammography directly influences the early detection of breast cancer, with the repeat rate (RR) serving as an important indicator of quality in diagnostic imaging services. This study aimed to synthesize the scientific evidence on the application of Artificial Intelligence (AI) techniques, with an emphasis on deep learning (DL) models, in the assessment of mammography technical quality and their potential to reduce the RR. A Systematic Literature Review (SLR) was conducted following the PRISMA 2020 guidelines, with searches performed in ACM Digital Library, IEEE Xplore, PubMed, ScienceDirect, Scopus, SpringerLink, and Web of Science. The methodological quality of the studies was assessed using criteria adapted from Kitchenham. After the selection process, 13 primary studies were included in the review. DL-based models predominated, particularly Convolutional Neural Networks (CNNs) and U-Net-derived architectures, applied to anatomical landmark identification, technical image quality classification, and the implementation of decision-support and feedback systems during mammographic acquisition. The studies reported accuracy of up to 98.5% for pectoral muscle identification, a reduction in the proportion of inadequate images from 13.3% to 3.2%, and a decrease in the RR to 0.17% in the best-performing scenario identified. The evidence indicates that AI has strong potential to support mammography quality control by improving the standardization of assessments, reducing interobserver variability, and enhancing the image acquisition workflow. These findings reinforce the potential of AI as a valuable support tool for mammography quality assurance programs. However, its widespread adoption still depends on multicenter clinical validation, greater model interpretability, and the standardization of evaluation protocols.

Keywords: Mammography. Mammographic Positioning. Repeat Rate.

RESUMEN: La calidad técnica de la mamografía influye directamente en la detección precoz del cáncer de mama, y la tasa de repetición (TR) constituye un importante indicador de la calidad de los servicios de diagnóstico por imagen. Este estudio tuvo como objetivo sintetizar la evidencia científica sobre la aplicación de técnicas de Inteligencia Artificial (IA), con énfasis en modelos de aprendizaje profundo (*Deep Learning* — DL), en la evaluación de la calidad técnica de la mamografía y su potencial para reducir la TR. Se realizó una Revisión Sistemática de la Literatura (RSL), siguiendo las directrices PRISMA 2020, con búsquedas en las bases ACM Digital Library, IEEE Xplore, PubMed, ScienceDirect, Scopus, SpringerLink y Web of Science. La calidad metodológica de los estudios se evaluó mediante criterios adaptados de Kitchenham. Tras el proceso de selección, 13 estudios primarios conformaron el corpus de la revisión. Se observó un predominio de modelos basados en DL, especialmente Redes Neuronales Convolucionales (CNN) y arquitecturas derivadas de U-Net, aplicados a la identificación de puntos de referencia anatómicos, la clasificación de la calidad técnica de las imágenes y la implementación de sistemas de apoyo a la decisión y retroalimentación durante la adquisición mamográfica. Los estudios reportaron una precisión de hasta el 98,5 % en la identificación del músculo pectoral, una reducción de la proporción de imágenes inadecuadas del 13,3 % al 3,2 % y una disminución de la TR hasta el 0,17 % en el mejor escenario identificado. La evidencia demuestra que la IA presenta un elevado potencial para apoyar el control de calidad de la mamografía, favoreciendo una mayor estandarización de las evaluaciones, la reducción de la variabilidad entre evaluadores y la mejora del flujo de adquisición de imágenes.

Estos resultados refuerzan el potencial de la IA como herramienta de apoyo a los programas de control de calidad de la mamografía. Sin embargo, su incorporación a gran escala aún depende de la validación clínica multicéntrica, una mayor interpretabilidad de los modelos y la estandarización de los protocolos de evaluación

Palabras clave: Mamografía. Posicionamiento mamográfico. Tasa de repetición.

INTRODUÇÃO

O rastreamento mamográfico desempenha papel fundamental na detecção precoce do câncer de mama, pois permite identificar alterações em estágios iniciais da doença e amplia as possibilidades de diagnóstico e tratamento oportunos. Para que esse benefício seja alcançado, entretanto, é essencial que as imagens apresentem qualidade técnica adequada. Falhas de posicionamento, enquadramento ou aquisição podem comprometer a visualização de estruturas anatômicas relevantes, dificultar a interpretação radiológica e levar à repetição do exame, aumentando a exposição da paciente à radiação ionizante (Brahim et al., 2022; Hejduk et al., 2023; INCA, 2024).

Entre os fatores que influenciam a qualidade técnica da mamografia, o posicionamento adequado da mama nas incidências craniocaudal (CC) e mediolateral oblíqua (MLO) ocupa papel de destaque. Sua avaliação considera critérios anatômicos e geométricos, como a visualização do músculo peitoral, o posicionamento do mamilo e a inclusão adequada do tecido mamário. Sistemas como o PGMI (*Perfect, Good, Moderate, Inadequate*) são utilizados para organizar essa avaliação. Contudo, por dependerem predominantemente da análise visual de especialistas, essas avaliações estão sujeitas à variabilidade entre observadores, o que pode dificultar sua padronização e tornar o controle da qualidade mais complexo, especialmente em serviços com grande volume de exames (Waade et al., 2021; Santner et al., 2025).

Nesse cenário, a IA, especialmente por meio de técnicas de aprendizado profundo, tem ampliado as possibilidades de automatização da avaliação da qualidade técnica da mamografia. Diferentes abordagens vêm sendo aplicadas à classificação da qualidade das imagens, segmentação de estruturas anatômicas, identificação de marcadores relacionados ao posicionamento e desenvolvimento de sistemas capazes de fornecer apoio e *feedback* durante a aquisição mamográfica. Os resultados relatados na literatura indicam avanços tanto no desempenho técnico dessas soluções quanto em sua aproximação com o ambiente clínico (Gupta et al., 2020; Eby et al., 2023; Watanabe et al., 2023; Gennaro et al., 2024; Tanyel et al., 2024; Kawasaki et al., 2025; Sexauer et al., 2025).

Apesar dos resultados promissores, as evidências permanecem distribuídas entre estudos com diferentes objetivos, arquiteturas de IA, critérios de qualidade, bases de dados e métricas de desempenho. Além disso, enquanto algumas pesquisas se concentram em critérios anatômicos específicos, outras avaliam a qualidade global do posicionamento ou investigam o impacto da IA no fluxo de aquisição e na taxa de repetição dos exames. Essa diversidade dificulta a comparação direta dos resultados e a compreensão do atual nível de maturidade dessas tecnologias para aplicação no controle da qualidade mamográfica.

Diante desse contexto, torna-se relevante reunir e analisar sistematicamente as evidências disponíveis, permitindo compreender não apenas o desempenho alcançado pelos modelos, mas também as estratégias empregadas, os critérios técnicos avaliados, os impactos observados na prática e as limitações ainda existentes. Essa síntese pode contribuir para identificar tendências, lacunas de pesquisa e desafios que precisam ser superados para uma adoção mais consistente dessas tecnologias na prática clínica.

Assim, este estudo tem como objetivo sintetizar as evidências científicas sobre a aplicação de técnicas de IA, com ênfase em modelos de aprendizado profundo, na avaliação da qualidade técnica da mamografia e em seu potencial para reduzir a taxa de repetição dos exames. Para isso, foi conduzida uma RSL, buscando identificar as principais abordagens de IA empregadas, os critérios de qualidade avaliados, as métricas e resultados reportados, bem como os desafios, limitações e oportunidades relacionados à sua aplicação no contexto clínico.

2 FUNDAMENTAÇÃO TEÓRICA

A avaliação da qualidade técnica da mamografia envolve aspectos relacionados à aquisição da imagem, ao posicionamento da mama e à adequada representação das estruturas anatômicas de interesse. A compreensão desses critérios é essencial para analisar como técnicas de IA vêm sendo incorporadas aos processos de controle da qualidade. Nesta seção, são apresentados os principais conceitos relacionados ao posicionamento mamográfico, aos marcos anatômicos utilizados em sua avaliação e às aplicações do aprendizado profundo nesse contexto.

2.1. MAMOGRAFIA E QUALIDADE DO POSICIONAMENTO

A mamografia é um dos principais exames de imagem utilizados no rastreamento e na detecção precoce do câncer de mama. Sua efetividade depende, entre outros fatores, da

qualidade técnica das imagens adquiridas, sendo o posicionamento adequado da mama fundamental para garantir a representação do tecido mamário e das estruturas anatômicas necessárias à interpretação radiológica (INCA, 2024; Santner et al., 2025; Margolies et al., 2025).

As incidências CC e MLO são utilizadas rotineiramente na aquisição mamográfica e apresentam critérios próprios para avaliação do posicionamento. Esses critérios consideram a inclusão adequada do tecido mamário e a representação de estruturas anatômicas específicas. Na prática clínica, sua análise é realizada predominantemente por profissionais especializados e pode ser organizada por sistemas de classificação, como o PGMI, além de critérios estabelecidos por programas e diretrizes de controle da qualidade. Embora a avaliação humana seja essencial, estudos demonstram que alguns desses critérios podem apresentar variabilidade entre avaliadores, especialmente aqueles que dependem de maior interpretação visual. Essa característica tem motivado o desenvolvimento de métodos automatizados capazes de fornecer medidas mais objetivas e reproduzíveis, apoiando profissionais na avaliação da qualidade técnica das imagens mamográficas (Gupta et al., 2020; Hejduk et al., 2023; Tanyel et al., 2024; Sexauer et al., 2025; Santner et al., 2025).

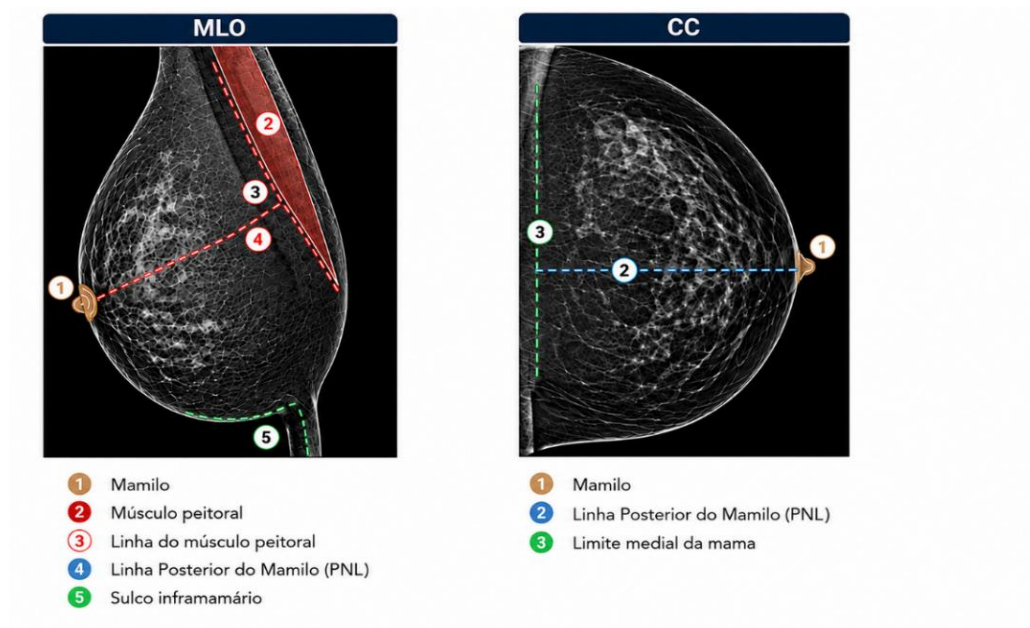
2.2. MARCOS ANATÔMICOS MAMOGRÁFICOS

5

A avaliação do posicionamento mamográfico baseia-se na visualização e na relação espacial entre estruturas anatômicas que permitem verificar se a mama foi adequadamente representada durante a aquisição. Entre os principais referenciais descritos na literatura estão o mamilo, o músculo peitoral, a linha posterior do mamilo (*Posterior Nipple Line* — PNL) e o sulco inframamário, cuja representação está relacionada à inclusão do tecido mamário e à qualidade técnica da imagem, conforme ilustrado na Figura 1 (Santner et al., 2025).

Na incidência MLO, a representação do músculo peitoral constitui um importante critério de posicionamento, contribuindo para avaliar a inclusão do tecido mamário posterior. Outros elementos, como a posição do mamilo e a visualização do sulco inframamário, também são considerados na análise da imagem. Na incidência CC, destacam-se o posicionamento do mamilo e a adequada representação do tecido mamário medial e lateral (Gupta et al., 2020; Tanyel et al., 2024).

Figura 1: Marcos anatômicos nas incidências MLO e CC



Fonte: Elaborada pelos autores (2026)

2.3. IA APLICADA À QUALIDADE TÉCNICA DA MAMOGRAFIA

6

O aprendizado profundo tem ampliado as possibilidades de análise automatizada de imagens médicas ao permitir que modelos computacionais aprendam características diretamente dos dados. Na mamografia, essas técnicas vêm sendo aplicadas não apenas à detecção de alterações suspeitas, mas também à avaliação da própria qualidade técnica das imagens, incluindo classificação, segmentação de estruturas anatômicas, localização de marcos de posicionamento e sistemas de *feedback* durante a aquisição. Entre as abordagens descritas na literatura, destacam-se as CNNs, amplamente utilizadas em tarefas de visão computacional por sua capacidade de extrair padrões espaciais das imagens. Diferentes arquiteturas têm sido empregadas conforme a tarefa proposta, incluindo modelos voltados à classificação da qualidade técnica e arquiteturas destinadas à segmentação ou localização de estruturas anatômicas (Gupta et al., 2020; Hejduk et al., 2023; Watanabe et al., 2023; Tanyel et al., 2024; Sexauer et al., 2025).

Além do desempenho computacional, a aplicação clínica dessas tecnologias envolve desafios relacionados à diversidade das imagens, diferenças anatômicas entre pacientes, equipamentos de aquisição, critérios utilizados como referência e concordância entre

especialistas. Por esse motivo, a análise da IA aplicada à qualidade técnica da mamografia não deve se limitar à acurácia dos modelos, mas considerar também sua capacidade de generalização, reprodutibilidade e integração ao fluxo clínico.

3. METODOLOGIA

Esta RSL foi conduzida com base nas diretrizes propostas por Kitchenham e Charters (2007) e nas recomendações do PRISMA 2020 (*Preferred Reporting Items for Systematic Reviews and Meta-Analyses*), buscando assegurar transparência, rastreabilidade e reprodutibilidade ao processo de revisão (Page et al., 2021).

O processo metodológico foi organizado em três etapas principais: planejamento, envolvendo a definição do protocolo, das questões de pesquisa, da estratégia de busca e dos critérios de elegibilidade; condução, compreendendo a busca nas bases de dados, triagem, seleção dos estudos e extração das informações; e síntese, destinada à avaliação metodológica e à análise das evidências obtidas. A ferramenta *Parsifal* foi utilizada como apoio ao gerenciamento da revisão, permitindo organizar os estudos recuperados e registrar as diferentes etapas do processo.

7

3.1. QUESTÕES DE PESQUISA

As questões de pesquisa foram definidas com o propósito de orientar a busca, a extração e a análise das evidências relacionadas ao uso da IA na avaliação da qualidade técnica da mamografia. Considerando o objetivo desta revisão, buscou-se compreender como essas tecnologias vêm sendo empregadas, quais abordagens predominam, como seu desempenho tem sido avaliado e quais limitações ainda dificultam sua incorporação à prática clínica. A Tabela 1 apresenta as questões de pesquisa definidas para esta revisão.

Tabela 1 – Questões de Pesquisa

ID	Questão de Pesquisa
QP ₁	Como a IA tem sido aplicada na avaliação automatizada da qualidade técnica da mamografia?
QP ₂	Quais arquiteturas de IA e estratégias de processamento de imagens predominam nos estudos analisados?
QP ₃	Quais critérios de qualidade, métricas de desempenho e resultados são reportados na avaliação automatizada da qualidade técnica da mamografia?
QP ₄	Quais são os principais desafios, limitações e oportunidades para a aplicação da IA no controle da qualidade técnica da mamografia?

3.2. ESTRATÉGIA DE BUSCA

A busca bibliográfica foi realizada nas bases ACM Digital Library, IEEE Xplore, PubMed, ScienceDirect, Scopus, SpringerLink e Web of Science, selecionadas por sua relevância para estudos nas áreas de IA, processamento de imagens médicas, radiologia e engenharia biomédica. A utilização de diferentes fontes teve como finalidade ampliar a cobertura da literatura e contemplar tanto estudos provenientes da área computacional quanto pesquisas desenvolvidas no contexto clínico.

O *Google Scholar* não foi incluído na estratégia de busca devido às limitações relacionadas à reprodutibilidade dos resultados e ao controle dos critérios de indexação.

Essa decisão foi adotada visando preservar a consistência e a rastreabilidade do processo de recuperação dos estudos.

3.3. CONSTRUÇÃO E REFINAMENTO DA ESTRATÉGIA DE BUSCA

A estratégia de busca foi construída e refinada de forma iterativa por meio de buscas piloto. Esse processo permitiu ajustar os descritores, sinônimos e operadores booleanos até alcançar uma combinação compatível com o escopo da revisão. Foram considerados termos relacionados à mamografia, IA, qualidade técnica, posicionamento mamográfico e repetição de exames.

A estratégia principal foi posteriormente adaptada às particularidades sintáticas de cada base de dados, preservando-se, sempre que possível, a equivalência semântica entre as consultas. A expressão de busca apresentada no protocolo foi: ("Mammography"OR "Breast Imaging") AND ("Artificial Intelligence"OR "Deep Learning"OR "Machine Learning") AND ("Positioning Quality"OR "Image Quality"OR "Technical Quality") AND ("Repeat Rate"OR "Retake"OR "Technical Repeat").

3.4. CRITÉRIOS DE ELEGIBILIDADE

Os critérios de inclusão e exclusão foram definidos previamente para assegurar que os estudos selecionados estivessem diretamente relacionados ao objetivo da revisão. Foram considerados estudos primários publicados entre 2019 e 2025, realizados em contexto clínico ou utilizando bases de imagens clínicas, que aplicassem técnicas de IA à avaliação ou ao apoio à qualidade técnica da mamografia. Foram excluídos estudos duplicados, revisões da literatura, pesquisas sem acesso ao texto completo, estudos baseados exclusivamente em *phantoms*,

simulações ou dados sintéticos e trabalhos sem relação direta com a qualidade técnica ou o posicionamento mamográfico. Os critérios adotados são apresentados na Tabela 2.

Tabela 2 – Critérios de elegibilidade

Cod	Critério de Inclusão (CI)	Cod	Critério de Exclusão (CE)
CI ₁	Estudos realizados em contexto clínico real ou com bases clínicas reais.	CE ₁	Artigos duplicados.
CI ₂	Estudos primários retrospectivos, prospectivos ou de implementação clínica.	CE ₂	Artigos sem acesso ao texto completo.
CI ₃	Estudos que aplicam IA na avaliação ou orientação da qualidade técnica da mamografia.	CE ₃	Estudos baseados apenas em <i>phantom</i> , simulações ou dados sintéticos.
CI ₄	Estudos que avaliam critérios técnicos de posicionamento mamográfico ou frameworks como PGMI, MQSA e FDA.	CE ₄	Estudos sem relação direta com a qualidade técnica ou o posicionamento mamográfico.
CI ₅	Estudos que reportam métricas quantitativas relacionadas à avaliação da qualidade técnica.	CE ₅	Estudos publicados antes de 2019.
CI ₆	Estudos com imagens clínicas de mamografia de rastreamento ou diagnóstico.	CE ₆	Revisões sistemáticas, revisões narrativas, meta-análises.

3.5. EXTRAÇÃO E SÍNTESE DE DADOS

Após a aplicação dos critérios de elegibilidade, os estudos incluídos foram submetidos à extração sistemática dos dados. Para padronizar esse processo e permitir a comparação entre os trabalhos, foi utilizado um formulário estruturado contemplando características bibliográficas, metodológicas, clínicas e computacionais.

Foram extraídas informações referentes aos autores e ano de publicação, desenho do estudo, modalidade de aquisição mamográfica, características das bases de dados, arquiteturas de IA, tarefas realizadas, critérios de qualidade avaliados, método utilizado como referência (*ground truth*), métricas de desempenho, principais resultados, diversidade dos dados, aspectos relacionados à interpretabilidade e reprodutibilidade e limitações reportadas.

A organização dessas informações foi definida de modo a estabelecer uma correspondência direta entre os dados extraídos e as questões de pesquisa, conforme apresentado na Tabela 3.

Tabela 3 – Síntese dos dados

ID	Dados do Artigo	Descrição	Relevância
1	Caracterização dos estudos	Autores, ano, país, fonte e desenho metodológico	Visão Geral

2	Contexto clínico e tecnológico	Modalidade, incidências e ambiente de aplicação	QP ₁
3	Bases de dados	Origem, tamanho da amostra e validação externa	QP ₁ /QP ₄
4	Abordagem de IA	Arquitetura, estratégia de aprendizado e processamento	QP ₂
5	Tarefa computacional	Classificação, segmentação, detecção/localização, regressão ou <i>feedback</i>	QP ₁ /QP ₂
6	Crítérios de qualidade	PGMI, MQSA/FDA, músculo peitoral, PNL, mamilo, IMF e outros	QP ₃
7	<i>Ground Truth</i>	Especialistas, consenso e outros métodos de comparação	QP ₃
8	Métricas e resultados	Acurácia, Dice, MAE, ICC, kappa, TR e demais resultados quantitativos	QP ₃
9	Diversidade e generalização	Densidade mamária, implantes, idade, equipamentos e validação externa	QP ₄
10	Interpretabilidade e reprodutibilidade	Disponibilidade de código, dados, modelos e descrição metodológica	QP ₄
11	Limitações	Risco de viés, limitações técnicas, clínicas e metodológicas	QP ₄

3.6. AVALIAÇÃO DA QUALIDADE METODOLÓGICA

A qualidade metodológica dos estudos incluídos foi avaliada com base em critérios adaptados das recomendações de Kitchenham e Charters et al. (2007). O objetivo dessa etapa foi verificar a clareza dos objetivos, a descrição dos procedimentos metodológicos, a adequação da avaliação experimental e a consistência entre resultados e conclusões.

Foram consideradas quatro questões de avaliação, conforme Tabela 4.

Tabela 4 – Avaliação da qualidade metodológica

ID	Aspecto Avaliado	Critério
QA ₁	Objetivos	O objetivo da pesquisa está claramente definido?
QA ₂	Metodologia	A metodologia está descrita de forma suficientemente detalhada para permitir sua reprodução?
QA ₃	Avaliação Experimental	A avaliação experimental utiliza métricas apropriadas e claramente descritas?
QA ₄	Resultados	Os resultados e as conclusões são consistentes com os objetivos propostos?

3.7. PROCESSO DE SELEÇÃO DOS ESTUDOS

A seleção dos estudos seguiu as etapas de identificação, triagem, elegibilidade e inclusão recomendadas pelo PRISMA 2020 (Page et al., 2021). Inicialmente, os registros recuperados nas bases pesquisadas foram consolidados e submetidos à identificação e remoção de duplicatas. Em seguida, títulos e resumos foram avaliados de acordo com os critérios de elegibilidade. Os trabalhos potencialmente relevantes foram então analisados em texto completo. Ao final do processo, 13 estudos primários atenderam aos critérios estabelecidos e compuseram o corpus da revisão. O fluxo completo de identificação, exclusão e inclusão dos estudos é apresentado no diagrama PRISMA, Figura 2.

Figura 2: Fluxograma do processo de seleção dos estudos segundo o PRISMA 2020



Fonte: Elaborado pelos autores (2026)

4. RESULTADOS

A aplicação dos critérios de elegibilidade resultou na inclusão de 13 estudos primários, publicados entre 2019 e 2025, que investigaram diferentes aplicações da IA na avaliação da qualidade técnica da mamografia. De modo geral, os estudos mostram uma evolução de abordagens voltadas à identificação automatizada de falhas de posicionamento para soluções capazes de avaliar critérios anatômicos específicos, classificar a qualidade técnica das imagens e fornecer *feedback* durante a aquisição mamográfica.

As evidências foram organizadas de acordo com as questões de pesquisa desta revisão, considerando as aplicações da IA, as principais abordagens computacionais, os critérios e métricas utilizados para avaliação e os desafios relacionados à adoção dessas tecnologias na prática clínica.

4.1. APLICAÇÕES DA IA NA AVALIAÇÃO DA QUALIDADE TÉCNICA

Os estudos analisados mostram que a IA tem sido empregada principalmente em três tipos de tarefa: classificação da qualidade técnica, segmentação de estruturas anatômicas e detecção ou localização de marcos anatômicos, relacionados ao posicionamento mamográfico.

Na classificação, os modelos são utilizados para identificar imagens adequadas ou inadequadas e avaliar sua conformidade com critérios técnicos de posicionamento. Essa abordagem aparece desde os primeiros estudos incluídos na revisão e evolui para sistemas capazes de comparar avaliações automatizadas com critérios empregados por especialistas (Gupta et al., 2020; Brahim et al., 2022; Santner et al., 2025).

Na segmentação e localização, a IA é utilizada para identificar estruturas e referenciais anatômicos relevantes, como o músculo peitoral, o mamilo, a PNL e o sulco inframamário (*Inframammary Fold* - IMF). A automatização desses critérios permite transformar aspectos tradicionalmente avaliados de maneira visual em informações quantitativas relacionadas ao posicionamento da mama (Hejduk et al., 2023; Watanabe et al., 2023; Tanyel et al., 2024; Kawasaki et al., 2025).

Outra aplicação identificada foi a incorporação da IA ao próprio fluxo de aquisição, nesses casos, sistemas automatizados são utilizados para fornecer *feedback* sobre a qualidade técnica das imagens, permitindo identificar possíveis inadequações e orientar ações de melhoria no serviço de mamografia. Estudos conduzidos em ambiente clínico indicam que essa aplicação pode contribuir para reduzir imagens inadequadas e repetições de exames (Eby et al., 2023; Gennaro et al., 2024; Sexauer et al., 2025).

Assim, as evidências indicam que a IA vem sendo aplicada tanto para avaliação retrospectiva da qualidade técnica quanto ao apoio ao controle da qualidade durante o fluxo de aquisição, abrangendo desde critérios anatômicos individuais até avaliações mais amplas do posicionamento mamográfico.

4.2. ARQUITETURAS E ESTRATÉGIAS PREDOMINANTES

As CNNs constituem a principal base computacional das soluções identificadas. Os estudos utilizam diferentes arquiteturas e estratégias conforme a tarefa proposta, incluindo modelos destinados à classificação, segmentação e localização de estruturas anatômicas (Gupta et al., 2020; Brahim et al., 2022; Hejduk et al., 2023; Watanabe et al., 2023).

Arquiteturas derivadas da U-Net destacam-se especialmente nas tarefas que exigem maior precisão espacial, como segmentação e localização de marcos anatômicos. Tanyel et al. (2024) investigaram a localização do mamilo e do músculo peitoral em imagens MLO, enquanto Kawasaki et al. (2025) utilizaram estratégias de segmentação associadas à avaliação objetiva do posicionamento.

Também foram identificados sistemas comerciais ou destinados à implementação clínica, nos quais os detalhes completos da arquitetura nem sempre constituem o foco principal da publicação. Nesses estudos, a atenção se concentra no desempenho do sistema em condições reais, no *feedback* oferecido aos profissionais e no impacto sobre indicadores de qualidade (Eby et al., 2023; Gennaro et al., 2024; Margolies et al., 2025; Sexauer et al., 2025).

Observa-se, portanto, predominância de abordagens baseadas em CNNs, com utilização de arquiteturas especializadas de acordo com a natureza da tarefa e crescente presença de soluções avaliadas diretamente no ambiente clínico.

4.3. CRITÉRIOS, MÉTRICAS E DESEMPENHO DOS MODELOS

Os estudos utilizaram diferentes critérios para avaliar a qualidade técnica da mamografia, incluindo posicionamento do mamilo, representação do músculo peitoral, PNL, IMF e avaliações globais baseadas em referenciais como PGMI e MQSA. Essa diversidade também se refletiu nas métricas utilizadas, entre elas acurácia, coeficiente de correlação intraclasse (ICC), coeficiente kappa e indicadores relacionados à TR.

Os melhores resultados foram observados principalmente em critérios anatômicos e geométricos mais objetivos. Hejduk et al. (2023) reportaram acurácia de 98,5% na avaliação relacionada ao músculo peitoral, enquanto Waade et al. (2021) encontraram ICC superior a 0,92 para a PNL. Nos critérios estruturais analisados pelos estudos, os valores de concordância kappa variaram entre 0,69 e 0,95.

Em contrapartida, critérios que envolvem maior componente de interpretação apresentaram menor concordância. Para avaliações relacionadas ao IMF e à qualidade global

segundo o PGMI, foram observados valores de kappa entre 0,14 e 0,42. Essa diferença evidencia que o desempenho dos sistemas não é uniforme entre todos os componentes utilizados na avaliação da qualidade mamográfica (Waade et al., 2021; Hejduk et al., 2023; Santner et al., 2025).

Além das métricas de desempenho computacional, alguns estudos avaliaram indicadores relacionados ao fluxo clínico. Eby et al. (2023), em uma análise envolvendo 307.850 imagens, reportaram redução da taxa de repetição de 0,77% para 0,17%, além do aumento da proporção de imagens classificadas como *Perfect/Good* para 59,78%.

Sexauer et al. (2025) observaram redução da proporção de imagens inadequadas de 13,3% para 3,2% após a utilização de *feedback* baseado em IA.

Margolies et al. (2025), em um conjunto de 553.339 imagens, avaliaram a aplicação da IA em larga escala e reportaram valores de kappa entre 0,70 e 0,95 para critérios relacionados à MQSA. Os estudos indicam desempenho mais consistente da IA quando a avaliação envolve critérios anatômicos ou geométricos objetivos, enquanto avaliações globais ou mais dependentes de interpretação apresentam maior variabilidade. Os resultados também mostram que, em determinados cenários clínicos, sistemas de IA foram associados à melhoria de indicadores operacionais, incluindo redução da TR.

4.4. DESAFIOS, LIMITAÇÕES E OPORTUNIDADES

Apesar dos resultados promissores, os estudos apresentam limitações que dificultam a comparação direta e a generalização das evidências. Foram identificadas diferenças entre bases de dados, critérios de qualidade, arquiteturas, métricas e métodos utilizados para estabelecer o *ground truth*. A própria avaliação humana utilizada como referência pode apresentar variabilidade, especialmente em critérios mais subjetivos relacionados à qualidade global do posicionamento.

Outro desafio está relacionado à representatividade dos dados utilizados no desenvolvimento e na validação dos modelos. Situações clínicas específicas, incluindo variações anatômicas, diferentes densidades mamárias, presença de implantes ou dispositivos cirúrgicos e diferenças entre equipamentos de aquisição, podem afetar o desempenho e a capacidade de generalização das soluções.

A literatura também apresenta diferentes níveis de transparência metodológica e validação clínica. Embora alguns estudos utilizem grandes conjuntos de imagens e avaliem sistemas em ambientes reais, outros apresentam bases mais restritas ou menor diversidade de

casos. Além disso, a heterogeneidade das métricas e dos critérios adotados dificulta estabelecer comparações diretas entre os modelos.

Os resultados apontam como principais desafios a generalização dos modelos, heterogeneidade dos protocolos de avaliação, variabilidade do *ground truth*, representatividade das populações estudadas e necessidade de validação clínica mais ampla. Ao mesmo tempo, os estudos que incorporam *feedback* durante a aquisição indicam uma oportunidade relevante de aproximar a IA do controle da qualidade realizado no próprio fluxo de trabalho mamográfico.

4.5. SÍNTESE DO CORPUS

A RSL evidencia o avanço recente da aplicação da IA na avaliação da qualidade técnica da mamografia. Os estudos analisados empregam predominantemente modelos de aprendizado profundo, especialmente CNNs, em tarefas de classificação da qualidade, segmentação de estruturas e localização de marcos anatômicos relacionados ao posicionamento mamográfico. A Tabela 5 apresenta a síntese dos 13 estudos incluídos, destacando seus objetivos, abordagens de IA e principais contribuições. De modo geral, os resultados indicam potencial para tornar a avaliação da qualidade técnica mais objetiva e padronizada, além de apoiar a redução de imagens inadequadas e da taxa de repetição em determinados cenários clínicos. Entretanto, permanecem limitações relacionadas à heterogeneidade das bases de dados, dos critérios de avaliação e das métricas empregadas, bem como à generalização e à validação clínica dos modelos. Essas limitações dificultam a comparação direta entre os estudos e reforçam a necessidade de protocolos mais padronizados e evidências multicêntricas para ampliar a confiabilidade e a aplicabilidade clínica dessas soluções (Gupta et al., 2020; Waade et al., 2021; Brahim et al., 2022; Eby et al., 2023; Hejduk et al., 2023; Tanyel et al., 2024; Santner et al., 2025; Sexauer et al., 2025).

Tabela 5 – Síntese do corpus

Autor/Ano	Objetivo	Abordagem de IA	Principal Contribuição
Gupta et al. (2020)	Detectar falhas de posicionamento mamográfico	CNN	Avaliação automatizada de mamografias mal posicionadas
Waade et al. (2021)	Automatizar critérios técnicos de qualidade	Aprendizado profundo	Comparação entre IA e avaliação de profissionais
Brahim et al. (2022)	Classificar qualidade do posicionamento	CNN	Classificação automatizada da qualidade técnica

Hejduk et al. (2023)	Automatizar a garantia da qualidade	dCNN e Regressão	Avaliação padronizada de critérios técnicos
Watanabe et al. (2023)	Avaliar critérios de posicionamento	Aprendizado profundo	Sistema automatizado de controle da qualidade
Alukić et al. (2023)	Avaliar qualidade técnica	Aprendizado profundo	Aplicação de IA ao controle da qualidade
Eby et al. (2023)	Avaliar impacto clínico	Sistema baseado em IA	Redução da taxa de repetição após <i>feedback</i>
Gennaro et al. (2024)	Melhorar o desempenho técnico	Sistema baseado em IA	<i>Feedback</i> automatizado aplicado à aquisição
Tanyel et al. (2024)	Localizar marcos anatômicos	U-Net e variantes	Localização de mamilo e músculo peitoral
Santner et al. (2025)	Avaliar critérios PGMI	IA aplicada ao PGMI	Comparação entre IA e leitores humanos
Kawasaki et al. (2025)	Automatizar critérios técnicos	Segmentação/CNN	Avaliação objetiva do posicionamento
Sexauer et al. (2025)	Avaliar <i>feedback</i> em tempo real	Aprendizado profundo	Redução de imagens inadequadas
Margolies et al. (2025)	Avaliar IA em ambiente clínico	Aprendizado profundo	Avaliação em larga escala da qualidade técnica

5. DISCUSSÃO

Os resultados desta revisão indicam que a IA vem ampliando seu papel no controle da qualidade técnica da mamografia, passando de aplicações isoladas de análise de imagens para soluções mais próximas do fluxo clínico. Sua principal contribuição está na possibilidade de tornar mais objetivos e consistentes critérios tradicionalmente dependentes da avaliação visual. Esse benefício é especialmente relevante no posicionamento mamográfico, em que a variabilidade entre avaliadores pode influenciar a classificação da qualidade das imagens.

As evidências mostram, entretanto, que o desempenho da IA varia conforme a natureza do critério avaliado. Os resultados mais consistentes são observados em tarefas associadas a estruturas anatômicas e relações geométricas objetivas, como a identificação do músculo peitoral e a avaliação da PNL. Em contrapartida, critérios com maior componente interpretativo apresentam menor concordância entre sistemas automatizados e avaliação humana. Esse comportamento sugere que, no estágio atual, a IA apresenta maior maturidade

na quantificação de critérios específicos do que na reprodução integral de avaliações globais, como o PGMI (Waade et al., 2021; Hejduk et al., 2023; Santner et al., 2025; Sexauer et al., 2025).

Além do desempenho algorítmico, alguns estudos demonstram impacto no fluxo de aquisição. Eby et al. (2023) reportaram redução da taxa de repetição de 0,77% para 0,17%, enquanto Sexauer et al. (2025) observaram redução da proporção de imagens tecnicamente inadequadas de 13,3% para aproximadamente 3,2% após a utilização de *feedback* apoiado por IA. Esses achados indicam que a análise automatizada durante a aquisição pode contribuir para identificar inadequações enquanto a paciente ainda se encontra no serviço, permitindo correções oportunas e reduzindo repetições desnecessárias. Contudo, esses resultados devem ser interpretados no contexto dos estudos que os produziram e não permitem assumir o mesmo efeito para diferentes sistemas ou ambientes clínicos.

Do ponto de vista tecnológico, predominam abordagens baseadas em CNNs, incluindo arquiteturas destinadas à classificação, segmentação e localização de estruturas anatômicas. Embora os resultados sejam promissores, a heterogeneidade das bases de dados, arquiteturas, métricas e critérios utilizados como referência dificulta comparações diretas entre os estudos. Somam-se a isso diferenças relacionadas à densidade mamária, presença de implantes, histórico cirúrgico, equipamentos de aquisição e variabilidade entre especialistas utilizados na construção do *ground truth*, fatores que podem comprometer a generalização dos modelos (Gupta et al., 2020; Hejduk et al., 2023; Tanyel et al., 2024; Santner et al., 2025; Sexauer et al., 2025).

Em conjunto, as evidências sugerem que a IA apresenta potencial para apoiar o controle da qualidade técnica da mamografia, sobretudo em tarefas objetivas, localizadas e mensuráveis. Sua incorporação ao fluxo de aquisição também pode contribuir para reduzir imagens inadequadas e repetições em determinados contextos. Entretanto, sua adoção clínica mais ampla ainda requer padronização dos critérios e métricas de avaliação, bases de dados mais representativas, maior interpretabilidade e validação multicêntrica. Nesse cenário, a IA mostra-se mais consistente como ferramenta de apoio à avaliação profissional do que como substituta do julgamento humano.

6. CONCLUSÃO

Esta Revisão Sistemática da Literatura evidencia que a IA apresenta potencial relevante para apoiar o controle da qualidade técnica da mamografia, especialmente na avaliação de

critérios anatômicos e geométricos relacionados ao posicionamento. Os estudos analisados demonstram maior consistência em tarefas objetivas e mensuráveis, como a identificação do músculo peitoral e a avaliação da linha posterior do mamilo, enquanto critérios com maior componente interpretativo ainda apresentam maior variabilidade entre os sistemas automatizados e a avaliação humana.

Além do desempenho técnico, alguns estudos apontam benefícios operacionais associados à incorporação da IA ao fluxo de aquisição. Sexauer et al. (2025) observaram redução da proporção de imagens inadequadas de 13,3% para 3,2%, enquanto Eby et al. (2023) reportaram redução da taxa de repetição de 0,77% para 0,17%. Embora esses resultados sejam promissores, devem ser interpretados no contexto específico de cada estudo e não permitem generalizar o mesmo impacto para diferentes serviços, populações ou sistemas de IA.

As evidências também revelam desafios relacionados à heterogeneidade das bases de dados, dos critérios de qualidade e das métricas utilizadas, além da variabilidade do *ground truth* e das diferenças entre populações, características anatômicas e equipamentos de aquisição. Esses fatores limitam a comparação direta entre os estudos e reforçam a necessidade de maior padronização metodológica e validação em diferentes contextos clínicos.

Conclui-se, portanto, que a IA se mostra mais consistente, no estágio atual, como ferramenta de apoio aos profissionais e aos programas de controle da qualidade da mamografia, contribuindo para avaliações mais objetivas e para a identificação de inadequações durante o processo de aquisição. Sua incorporação clínica em maior escala, entretanto, ainda depende de validação multicêntrica, bases de dados mais representativas, maior interpretabilidade dos modelos e padronização dos protocolos de avaliação.

REFERÊNCIAS

ALUKIĆ, E. et al. The impact of subjective image quality evaluation in mammography. *Radiography*, v. 29, n. 3, p. 526–532, 2023.

Disponível em: DOI: <https://doi.org/10.1016/j.radi.2023.02.025>. Acesso em: 01 set. 2026.

BRAHIM, M. et al. Automated assessment of breast positioning quality in screening mammography. *Cancers*, v. 14, n. 19, p. 4704, 2022.

Disponível em: DOI: <https://doi.org/10.3390/CANCERS14194704>. Acesso em: 01 set. 2026.

EBY, P. R. et al. Impact of artificial intelligence-driven quality improvement software on mammography technical repeat and recall rates. *Radiology: Artificial Intelligence*, v. 5, n. 6, e230038, 2023. Disponível em: <https://doi.org/10.1148/ryai.230038>. Acesso em: 01 set. 2026.

GENNARO, G. et al. Using automated software evaluation to improve the performance of breast radiographers in tomosynthesis screening. *European Radiology*, v. 34, n. 7, p. 4738-4749, 2024. Disponível em: <https://doi.org/10.1007/s00330-023-10457-x>. Acesso em: 01 set. 2026.

GUPTA, V. et al. Deep learning-based automatic detection of poorly positioned mammograms to minimize patient return visits for repeat imaging: a real-world application. *arXiv*, 2020. arXiv:2009.13580v1.

Disponível em: <https://doi.org/10.48550/arXiv.2009.13580>. Acesso em: 01 set. 2026.

HEJDUK, P. et al. Automatic and standardized quality assurance of digital mammography and tomosynthesis with deep convolutional neural networks. *Insights into Imaging*, v. 14, p. 90, 2023. Disponível em: <https://doi.org/10.1186/s13244-023-01396-8>. Acesso em: 01 set. 2026.

INSTITUTO NACIONAL DE CÂNCER (INCA). Estimativa 2023-2025: incidência de câncer no Brasil. 2024.

Disponível em: <https://rbc.inca.gov.br/index.php/revista/article/view/3700>. Acesso em: 01 set. 2026.

KAWASAKI, A. et al. Development of an AI-based positioning technical assistance system for mammography. *Jpn J Radiol Technol*, v. 81, n. 12, p. 1-11, 2025. Disponível em: <https://doi.org/10.6009/jjrt.25-1574>. Acesso em: 01 set. 2026.

KITCHENHAM, B.; CHARTERS, S. Guidelines for performing systematic literature reviews in software engineering. Technical report. Keele: Keele University, 2007.

MARGOLIES, L. R. et al. Artificial intelligence for assessment of digital mammography positioning reveals persistent challenges. *Journal of Breast Imaging*, v. 7, n. 5, p. 530-540, 2025. Disponível em: <https://doi.org/10.1093/jbi/wbaf025>. Acesso em: 01 set. 2026.

NGUYEN, H. Q. et al. VinDr-Mammo: a large-scale benchmark dataset for computer-aided diagnosis in full-field digital mammography. *Scientific Data*, v. 10, n. 1, p. 277, 2023. Disponível em: <https://doi.org/10.1038/s41597-023-02100-7>. Acesso em: 01 set. 2026.

PAGE, M. J. et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*, v. 372, n. 71, 2021.

RONNEBERGER, O.; FISCHER, P.; BROX, T. U-Net: convolutional networks for biomedical image segmentation. In: *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*. Cham: Springer, 2015. v. 9351, p. 234-241. Disponível em: https://doi.org/10.1007/978-3-319-24574-4_28. Acesso em: 01 set. 2026.

SANTNER, T. et al. PGMI assessment in mammography: AI software versus human readers. *Radiography*, v. 31, 103017, 2025. Disponível em: <https://doi.org/10.1016/j.radi.2025.103017>. Acesso em: 01 set. 2026.

SANTNER, T. et al. AI-based image quality assessment of positioning in mammography: considerations and challenges. *Insights into Imaging*, v. 17, n. 47, 2025. Disponível em: <https://doi.org/10.1186/s13244-025-02191-3>. Acesso em: 01 set. 2026.

SEXAUER, R. et al. Enhancing breast positioning quality through real-time AI feedback. *European Radiology*, 2025. Disponível em: <https://doi.org/10.1007/s00330-025-11812-w>. Acesso em: 01 set. 2026.

TANYEL, T. et al. Mammographic breast positioning assessment via deep learning. *arXiv*, 2024. *arXiv:2407.10796v1*. Disponível em: https://doi.org/10.1007/978-3-031-77789-9_11. Acesso em: 01 set. 2026.

WAADE, G. G. et al. Assessment of breast positioning criteria in mammographic screening: agreement between artificial intelligence software and radiographers. *Journal of Medical Screening*, v. 28, n. 4, p. 448–455, 2021. Disponível em: <https://doi.org/10.1177/0969141321998718>. Acesso em: 01 set. 2026.

WATANABE, H. et al. Quality control system for mammographic breast positioning using deep learning. *Scientific Reports*, v. 13, n. 1, p. 7066, 2023.

Disponível em: <https://doi.org/10.1038/s41598-023-34380-9>. Acesso em: 01 set. 2026.