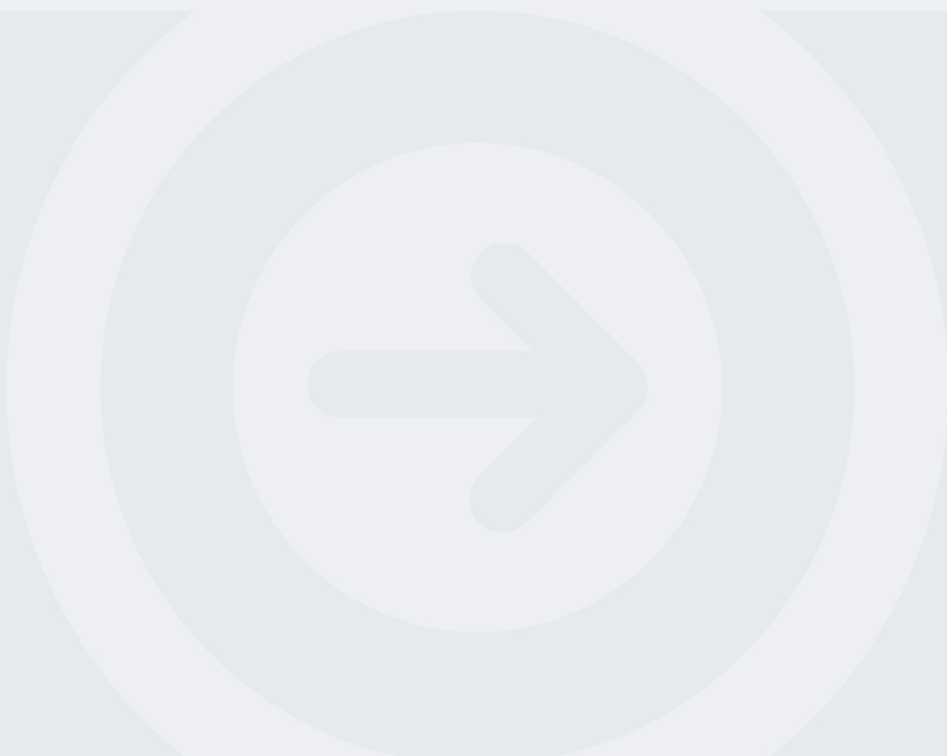




A inteligência artificial no apoio ao diagnóstico do cancro da mama

MARIA CAROLINA CAMPOS BRANDÃO DO COUTO

novembro de 2025



A inteligência artificial no apoio ao diagnóstico do cancro da mama

Maria Carolina Campos Brandão do Couto

**Dissertação para obtenção do Grau de
Mestre em Engenharia Biomédica**

Orientador: Professor Carlos Augusto Xavier Ramos

Júri:

Presidente:

Joaquim Alves, Prof. Adjunto, ISEP-IPP

Vogais:

Carlos A. X. Ramos, Prof. Ajunto, ISEP-IPP

Jorge António Silva Mendes, Prof. Ajunto, ISEP-IPP

Resumo

A Inteligência Artificial (IA) tem vindo a transformar o setor da saúde, particularmente no que diz respeito ao diagnóstico médico. O diagnóstico do cancro da mama, uma das principais causas de mortalidade feminina, pode beneficiar significativamente da aplicação destas tecnologias, que permitem uma deteção precoce e mais precisa da doença. Este trabalho tem como objetivo analisar o estado da arte na utilização da IA no diagnóstico do cancro da mama e avaliar a utilidade de sistemas de Diagnóstico Assistido por Computador (CAD), analisando a sua eficácia, fiabilidade e limitações. Foram escolhidos e analisados três sistemas CAD comerciais – *Lunit INSIGHT MMG*, *Transpara* e *ProFound AI*. Os resultados da revisão indicam que a IA possui um elevado potencial para melhorar a precisão diagnóstica, reduzir falsos positivos e otimizar o tempo de análise. No entanto, desafios relacionados com bases de dados mais diversificadas, explicabilidade dos algoritmos, aceitação por parte dos profissionais de saúde, bem como questões éticas e regulamentares, ainda limitam a sua integração plena na prática clínica. Esta revisão pretende, assim, fornecer uma visão abrangente do papel da IA no diagnóstico do cancro da mama e perceber como melhorar a sua integração em contexto clínico.

Palavras-chave: Inteligência Artificial, Cancro da mama, Diagnóstico, Deteção precoce, Diagnóstico Assistido por Computador

Abstract

Artificial Intelligence (AI) has been transforming the healthcare sector, particularly in medical diagnosis. Breast cancer diagnosis, one of the leading causes of female mortality, can significantly benefit from the application of these technologies, which enable early and more accurate disease detection. This study aims to review the state of the art in the use of AI in breast cancer diagnosis and evaluate the usefulness of Computer-Aided Diagnosis (CAD) systems, analyzing their effectiveness, reliability, and limitations. Three commercial CAD systems were selected and analyzed – Lunit INSIGHT MMG, Transpara, and ProFound AI. The results of the review indicate that AI has significant potential to improve diagnostic accuracy, reduce false positives, and optimize analysis time. However, challenges related to more diverse databases, algorithm explainability, acceptance by healthcare professionals, as well as ethical and regulatory issues, still limit its full integration into clinical practice. This review, therefore, aims to provide a comprehensive overview of the role of AI in breast cancer diagnosis and understand how to improve its integration in a clinical context.

Keywords: Artificial Intelligence, Breast Cancer, Diagnosis, Early Detection, Computer-Aided Diagnosis

Agradecimentos

A conclusão deste trabalho representa o culminar de uma etapa importante da minha vida académica e pessoal, que não teria sido possível sem o apoio de várias pessoas a quem quero expressar a minha sincera gratidão.

Em primeiro lugar, agradeço ao meu orientador, Professor Carlos Augusto Xavier Ramos, pela orientação, disponibilidade e paciência ao longo de todo o desenvolvimento desta dissertação. Agradeço também aos meus colegas e amigos, pelo companheirismo, incentivo e pelas conversas que tornaram este percurso mais leve e enriquecedor.

Um agradecimento muito especial à minha família, em particular aos meus pais, irmão e namorado, pelo amor incondicional, apoio constante e por acreditarem sempre em mim.

Por fim, agradeço a todos os que, de forma direta ou indireta, contribuíram para a realização deste projeto e para o meu crescimento académico e pessoal.

A todos, o meu mais sincero obrigado.

Índice

1.	Introdução	1
1.1.	Apresentação do Tema	1
1.2.	Objetivos Gerais e Específicos	2
2.	Metodologia	3
2.1.	Pesquisa Bibliográfica	3
2.2.	Seleção dos Sistemas CAD	5
3.	Revisão da Literatura	7
3.1.	Cancro da Mama	7
3.2.	Inteligência Artificial no Cancro da Mama	8
3.3.	Modalidades de Imagem e Integração com IA	9
3.3.1.	Mamografia e Tomossíntese Digital da Mama	10
3.3.2.	Ressonância Magnética	11
3.3.3.	Ultrassonografia	12
3.3.4.	Patologia Digital e <i>Whole Slide Imaging (WSI)</i>	13
3.3.5.	Modalidades Emergentes e Fusão Multimodal	14
3.4.	Resultados Quantitativos e Análise	14
4.	Algoritmos e Sistemas CAD	20
4.1.	Evolução dos Sistemas CAD	20
4.2.	Algoritmos de IA Aplicados ao Diagnóstico por Imagem	21
5.	Análise dos Sistemas CAD	25
5.1.	ProFound AI	26
5.2.	Lunit INSIGHT MMG	30
5.3.	Transpara	33
5.4.	Dados Recolhidos e Análise	36
6.	Discussão	43
6.1.	Considerações Gerais	43
6.2.	Implicações Clínicas e Científicas	43
6.3.	Desafios e Barreiras Identificadas	44
7.	Conclusão	47
7.1.	Limitações do Estudo	47
7.2.	Perspetivas Futuras e Recomendações	48

Lista de Gráficos

Gráfico 1 – Comparação de métricas entre a IA e o método tradicional	18
Gráfico 2 – Comparação das métricas do método tradicional com o sistema ProFound AI	39
Gráfico 3 – Comparação das métricas do método tradicional com o sistema Lunit INSIGHT MMG	40
Gráfico 4 – Comparação das métricas do método tradicional com o sistema Transpara	41

Lista de Tabelas

Tabela 1 – Resumo da Pesquisa Bibliográfica.....	4
Tabela 2 – Sistemas CAD selecionados.....	6
Tabela 3 – Principais algoritmos de IA utilizados no diagnóstico médico	22
Tabela 4 – Análise geral dos dados recolhidos.....	38

Acrónimos

Lista de Acrónimos

ABUS	<i>Automated Breast Ultrasound</i> (Mamografias Ultrassonográficas Automatizadas da Mama)
AUC	<i>Area Under the Curve</i> (Área Sob a Curva)
BI-RADS	<i>Breast Imaging Reporting and Data System</i> (Sistema de Relatório e Dados de Imagem da Mama)
BRCA	<i>Breast Cancer Gene</i> (Gene do cancro da mama)
CAD	<i>Computer-aided Diagnosis/Detection</i> (Diagnóstico/Deteção Assistida por Computador)
CADe	<i>Computer-aided Detection</i> (Deteção Assistida por Computador)
CADx	<i>Computer-aided Diagnosis</i> (Diagnóstico Assistido por Computador)
CDR	<i>Cancer Detection Rate</i> (Taxa de Deteção de Cancro)
CNN	<i>Convolutional Neural Network</i> (Rede Neuronal Convolucional)
DBT	<i>Digital Breast Tomosynthesis</i> (Tomossíntese Digital Mamária)
DCIS	<i>Ductal carcinoma in situ</i> (Carcinoma ductal in situ)
DNN	<i>Deep Neural Network</i> (Rede Neuronal Profunda)
FDA	<i>Food and Drug Administration</i> (Administração de Alimentos e Medicamentos)
FFDM	<i>Full Field Digital Mammography</i> (Mamografia Digital de Campo Total)
GAN	<i>Generative Adversarial Network</i> (Rede Generativa Adversarial)
HTA	<i>Health Technology Assessment</i> (Avaliação de Tecnologias da Saúde)

IA	Inteligência Artificial
KNN	<i>K-Nearest Neighbors</i> (K-Vizinhos mais próximos)
LPCC	Liga Portuguesa Contra o Cancro
MMG	Mamografia
MWI	<i>Microwave Image</i> (Imagem por Microondas)
NHS	<i>National Health Service</i> (Serviço Nacional de Saúde do Reino Unido)
RM	Ressonância Magnética
ROC	<i>Receiver Operating Characteristic</i> (Característica de Funcionamento do Receptor)
SVM	<i>Support Vector Machine</i> (Máquina de Vetores de Suporte)
TC	Tomografia Computorizada
US	Ultrassonografia
WSI	<i>Whole Slide Imaging</i> (Imagem de lâmina inteira)
XAI	<i>Explainable Artificial Intelligence</i> (Inteligência Artificial Explicável)

1. Introdução

1.1. Apresentação do Tema

O cancro da mama é uma das doenças mais incidentes entre mulheres em todo o mundo, sendo responsável por um número significativo de óbitos anualmente. A deteção precoce é essencial para melhorar as taxas de sobrevivência, permitindo intervenções médicas mais eficazes e menos invasivas. No entanto, os métodos tradicionais de diagnóstico, como mamografias e biópsias, apresentam desafios significativos, incluindo a subjetividade na interpretação das imagens, a variabilidade nos diagnósticos entre especialistas e o tempo necessário para a execução e análise dos exames.

Neste contexto, a Inteligência Artificial (IA) tem emergido como uma ferramenta promissora, e potencialmente revolucionária, no apoio ao diagnóstico médico, oferecendo soluções avançadas para a deteção e classificação de padrões em imagens médicas. Diferentes algoritmos de IA têm sido desenvolvidos para aumentar a precisão e a rapidez dos diagnósticos, reduzindo a subjetividade e o excesso de tempo associados à análise tradicional.

Ao utilizar algoritmos avançados de *machine learning* (aprendizagem automática) e *deep learning* (aprendizagem profunda), a IA pode analisar grandes volumes de dados e identificar padrões com alta precisão, auxiliando os profissionais de saúde na deteção precoce do cancro. A sua aplicação pode contribuir para reduzir a taxa de falsos positivos e falsos negativos, otimizando o tempo de diagnóstico (Hasan, Khalid and Abid, 2025).

No entanto, apesar do seu potencial, a implementação da IA na prática clínica ainda enfrenta desafios, como a necessidade de validação científica robusta, a aceitação por parte dos profissionais de saúde, para além de questões éticas relacionadas ao uso de dados sensíveis dos pacientes (Goh *et al.*, 2025).

Desta forma, esta pesquisa é relevante tanto do ponto de vista científico como clínico, pois pretende avaliar a eficácia de alguns destes algoritmos de IA no apoio à deteção do cancro da mama, enquanto investiga os desafios e limitações que impedem a sua adoção em larga escala. Os resultados poderão fornecer perceções valiosas para o desenvolvimento de tecnologias mais adaptadas e percutíveis, contribuir para a otimização dos protocolos de diagnóstico e ajudar a moldar estratégias para a implementação da IA na área da saúde.

1.2. Objetivos Gerais e Específicos

O presente trabalho tem como objetivo rever e analisar criticamente a utilização da inteligência artificial no apoio ao diagnóstico do cancro da mama, com especial foco nos sistemas de Diagnóstico Assistido por Computador (CAD). Pretende-se compreender de que forma a IA contribui para a melhoria da precisão diagnóstica, para a redução da variabilidade interobservador e para o aumento da eficiência no rastreio e diagnóstico desta patologia, que permanece uma das principais causas de mortalidade feminina a nível mundial.

Através de uma revisão sistematizada da literatura, de acordo com uma metodologia que será apresentada posteriormente, será avaliada a evidência científica disponível relativamente ao desempenho de diferentes algoritmos e abordagens de IA aplicadas a modalidades de imagem mamária, principalmente, mamografia digital e tomossíntese mamária. Para além disso, será realizada uma análise complementar de sistemas CAD já disponíveis no mercado, com o intuito de analisar o desempenho e eficácia dos mesmos, e identificar os que apresentam maior validação clínica, certificação regulatória e aplicabilidade prática em contexto clínico.

Com esta abordagem, este trabalho procura sintetizar os benefícios, limitações e desafios associados à implementação da IA na prática clínica, destacando tanto os avanços já alcançados como as barreiras que ainda persistem, sejam elas de ordem técnica, regulatória ou ética. O objetivo último é fornecer uma visão abrangente e fundamentada que permita compreender o real valor acrescentado da IA no diagnóstico do cancro da mama e orientar futuras investigações e aplicações clínicas nesta área.

2. Metodologia

Esta pesquisa será de caráter descritivo e exploratório, utilizando uma abordagem qualitativa e quantitativa.

2.1. Pesquisa Bibliográfica

A presente revisão foi conduzida de forma estruturada, com o objetivo de identificar, analisar e sintetizar os principais avanços da aplicação da inteligência artificial no diagnóstico do cancro da mama.

No que se refere à fundamentação teórica, a revisão bibliográfica foi realizada na base de dados *PubMed*, utilizando as palavras-chave “*breast cancer*”, “*artificial intelligence*” e “*diagnosis*” (“cancro da mama”, “inteligência artificial” e “diagnóstico”). Aplicou-se um filtro temporal que restringiu a seleção a publicações entre Janeiro de 2024 e Setembro de 2025, de forma a assegurar a atualidade da informação. A pesquisa foi conduzida em língua inglesa, uma vez que a maioria da literatura científica relevante se encontra publicada neste idioma, o que permite assegurar maior diversidade e representatividade dos artigos incluídos.

Complementarmente, efetuaram-se pesquisas manuais nas plataformas *Google Scholar* e *ResearchGate*, de forma a garantir uma maior abrangência dos resultados.

Foram então definidos os seguintes critérios de inclusão:

- Foram considerados elegíveis os artigos que contemplassem as palavras-chave previamente definidas para este estudo ou que estabelecessem uma relação temática clara e relevante entre elas;
- Foram incluídos artigos com acesso livre ou disponibilizados pela instituição;
- Foram incluídos artigos publicados recentemente, com data de publicação não superior a um ano.

Após a seleção inicial, foram revistos todos os artigos selecionados de forma a perceber a sua utilidade para o tema em estudo. Foram aplicados alguns critérios de exclusão:

- Foram excluídos os artigos que não cumpriam os critérios de inclusão previamente estabelecidos;

- Foram excluídos os artigos que não incluíam as palavras-chave definidas ou termos relacionados;
- Foram excluídos os artigos que não se revelaram pertinentes para o tema em questão;
- Foram excluídas publicações em duplicado ou sem acesso ao texto integral.

A Tabela 1, apresenta um resumo do processo de seleção dos artigos.

Tabela 1 – Resumo da Pesquisa Bibliográfica

Termos-chave utilizados	Ferramenta de pesquisa	Filtros aplicados	Número de artigos selecionados	Número de artigos finais
<i>"breast cancer" AND "artificial intelligence" AND "diagnosis"</i>	<i>PubMed</i>	<i>Publication date: 1 year Text Availability: Free Full Text</i>	64	43

Após a seleção final, procedeu-se à análise global de todos os artigos incluídos. Sempre que possível, foram extraídas informações relativas aos seguintes aspetos:

- Modalidade de imagem (mamografia, tomossíntese mamária, ressonância magnética, ultrassonografia, etc.).
- Abordagem ou sistema de IA.
- Métricas relatadas (sensibilidade, especificidade, precisão, AUC, etc.).
- Comparação com método tradicional.
- Principais limitações e desafios identificados pelos autores.

Em síntese, esta abordagem metodológica permitiu identificar, organizar e analisar de forma sistemática os artigos mais relevantes para a presente revisão.

Esta fundamentação metodológica fornece uma base sólida para compreender qual o estado atual da IA no apoio ao diagnóstico do cancro da mama, estabelecendo o enquadramento necessário.

2.2. Seleção dos Sistemas CAD

Para além da revisão de literatura já mencionada, este trabalho inclui também a análise de sistemas de Diagnóstico Assistido por Computador que incorporem a IA no diagnóstico do cancro da mama, com o objetivo de compreender a aplicabilidade real desta tecnologia em ambiente clínico.

A seleção destes sistemas baseou-se nos seguintes critérios:

1. Certificação regulatória (marcação CE, aprovação FDA).
2. Relevância na literatura científica.
3. Adoção clínica em hospitais ou centros de rastreio.
4. Disponibilidade de documentação técnica e relatórios de desempenho.

Com base nestes critérios, foram selecionados os seguintes sistemas CAD:

- **Lunit INSIGHT MMG** – Software de deteção/diagnóstico assistido por computador (CAdE/x) desenvolvido pela *Lunit AI* que identifica e classifica áreas suspeitas de cancro da mama em mamografias através de tecnologia de *deep learning*.
- **Transpara** – Software de IA desenvolvido pela *ScreenPoint Medical* que atua como um segundo leitor para os radiologistas, ajudando-os a detetar o cancro da mama mais precocemente.
- **ProFound AI / iCAD** – Software de IA desenvolvido pela *iCAD, Inc.*, concebido para a deteção de cancro da mama e avaliação de risco a partir de imagens de mamografia.

A análise destes sistemas consistiu na recolha de dados a partir de artigos científicos, relatórios e documentação técnica. A sua avaliação será feita com base nas métricas existentes, de forma a ser possível avaliar o desempenho individual dos mesmos, e poder compará-los, tanto entre eles, como com os métodos tradicionais de diagnóstico.

Os sistemas CAD selecionados são apresentados na Tabela 2.

Tabela 2 – Sistemas CAD selecionados

Sistema CAD	Empresa	Aplicação	Pontos fortes
<i>Lunit INSIGHT MMG</i>	<i>Lunit AI</i>	Diagnóstico assistido com detecção de lesões.	Alta sensibilidade; Integração em fluxos hospitalares; Aprovado em vários países.
<i>Transpara</i>	<i>ScreenPoint Medical</i>	Atribuição de pontuação de risco por exame.	Interface intuitiva; Apoio à triagem; Aprovação pela CE e FDA.
<i>ProFound AI (iCAD)</i>	<i>iCAD, Inc.</i>	Deteção de lesões e atribuição de pontuação de risco.	Reduz falsos positivos; Diminui o tempo de leitura; Aprovação pela CE e FDA.

3. Revisão da Literatura

3.1. Cancro da Mama

O cancro da mama é caracterizado pelo crescimento descontrolado de células malignas no tecido mamário, podendo afetar tanto mulheres como, em casos raros, homens. Os fatores de risco para este cancro incluem histórico familiar, idade, alterações genéticas (como as mutações BRCA1 e BRCA2), estilo de vida e fatores hormonais. O diagnóstico atual baseia-se na combinação de exames físicos, imagiológicos e laboratoriais, sendo a mamografia o método de rastreio mais comum. A deteção precoce é essencial para o sucesso do tratamento, estando associada a melhores prognósticos e menores taxas de mortalidade.

No que diz respeito ao rastreio do cancro da mama em Portugal, este é feito pela Liga Portuguesa Contra o Cancro (LPCC) através do Programa Nacional de Rastreio do Cancro da Mama. São convocadas anualmente mulheres entre os 50 e os 69 anos para uma mamografia que geralmente é feita em unidades móveis. A análise das imagens é feita por dupla leitura, onde são precisos dois radiologistas que interpretam o mesmo exame de forma independente, de forma a aumentar a sensibilidade do diagnóstico. Após esta análise inicial, caso seja detetada alguma área anormal, podem ser pedidos exames complementares, como ultrassonografia ou ressonância magnética. A biópsia é o único método que confirma com certeza a existência de cancro.

Segundo dados da LPCC, anualmente são detetados 9000 novos casos de cancro da mama e mais de 2000 mulheres morrem com esta doença (“Cancro da Mama _ Liga Portuguesa Contra o Cancro,” 2025). Mesmo com o elevado conhecimento científico e biológico sobre o cancro da mama, este continua a ser uma das principais causas de mortalidade feminina a nível mundial, representando um enorme desafio para a saúde pública. Apesar dos progressos na prevenção, rastreio e tratamento, esta patologia mantém elevada incidência e prevalência, especialmente em países desenvolvidos, onde os programas de rastreio populacional são amplamente utilizados (Uwimana, Gnecco and Riccaboni, 2025).

A deteção precoce do cancro da mama desempenha um papel crucial na redução da mortalidade, permitindo o diagnóstico em estados iniciais e aumentando a eficácia terapêutica (Marinovich *et al.*, 2025). Entre as modalidades utilizadas, a mamografia continua a ser o exame

de referência para o rastreio populacional, sendo responsável pela redução de 20% a 30% da mortalidade por cancro da mama. Contudo, as limitações como baixa sensibilidade em mamas densas, falsos positivos, necessidade de leituras repetidas e dor ou desconforto durante o exame permanecem obstáculos de relevo (Uwimana, Gnecco and Riccaboni, 2025).

3.2. Inteligência Artificial no Cancro da Mama

Neste contexto, a inteligência artificial (IA), e mais especificamente o *machine learning*, surgem como ferramentas inovadoras com potencial transformador na prática clínica. O aparecimento da tecnologia *deep learning*, particularmente das redes neurais convolucionais (CNNs), impulsionou avanços notáveis na imagiologia médica, com desempenhos comparáveis ou mesmo superiores aos de especialistas humanos em tarefas específicas (Datwani *et al.*, 2025). A aplicação da IA no diagnóstico mamário não se limita à mamografia, estendendo-se também a modalidades como a Tomossíntese Digital Mamária (DBT), Ultrassonografia (US), Ressonância Magnética (RM) e patologia digital, cobrindo assim todo o espectro do diagnóstico e caracterização do cancro da mama (Abdullah *et al.*, 2025; Wu *et al.*, 2025).

Uma revisão abrangente de Uwimana, Gnecco and Riccaboni (2025), que analisou 104 estudos publicados entre 2015 e 2023, concluiu que a maioria dos trabalhos avaliou a eficácia clínica, e que esta obteve um valor geral de 78,84%, mostrando que sistemas de IA ou radiologistas assistidos por IA superaram a prática convencional em várias métricas. No caso da métrica relativa à especificidade houve uma melhoria de 76,1% da prática convencional para 80,9% com a integração da IA. Também a sensibilidade mostrou melhores resultados, com a IA a obter 82,7% comparativamente aos 80,7% da prática convencional, assim como a precisão, com um crescimento de 80,8% para 85,1%.

Alguns sistemas CAD, como o *ProFound AI* e o *Lunit INSIGHT MMG*, destacaram-se como exemplos de soluções comerciais já aplicadas em cenários clínicos, reforçando a relevância prática destas tecnologias. Ainda assim, os autores salientam limitações importantes: apenas 10,78% dos estudos demonstraram generalização para diferentes populações e apenas 14,7% partilharam código ou dados, revelando barreiras significativas de reprodutibilidade e transparência destes sistemas.

A integração da IA nas modalidades de imagem acrescenta benefícios específicos. Na mamografia, onde a densidade mamária pode comprometer a acuidade diagnóstica, os sistemas de IA têm contribuído para reduzir falsos positivos e melhorar a detecção de lesões subtis (Singh, Singh and Bhattacharya, 2024). Na ressonância magnética, meta-análises relataram valores da Área Sob a Curva (AUC) – que mede a capacidade de um modelo em diferenciar corretamente casos positivos de casos negativos – próximos de 0,90, com sensibilidade média de 88% e especificidade de 90% para algoritmos de *deep learning* (Abdullah *et al.*, 2025).

Na ecografia, estudos japoneses já validaram comercialmente algoritmos como *YOLOv3-tiny*, que alcançou uma AUC de 0,88 e sensibilidade de até 92%, em cenários de rastreio populacional (Wu *et al.*, 2025). Já na patologia digital, algoritmos treinados em milhões de imagens histopatológicas, como o *Galen Breast*, demonstraram desempenhos superiores a 0,95 de AUC na detecção de carcinoma invasor e DCIS (carcinoma ductal in situ, que representa uma lesão não invasiva da mama), ilustrando o potencial da IA para revolucionar também o diagnóstico anatomopatológico (Datwani *et al.*, 2025).

No próximo capítulo será feita uma breve análise das modalidades de imagem mais utilizadas para a detecção e diagnóstico do cancro da mama.

3.3. Modalidades de Imagem e Integração com IA

A evolução da aplicação da inteligência artificial no diagnóstico do cancro da mama tem sido notória, impulsionada pela diversidade de modalidades de imagem disponíveis na prática clínica. Desde a mamografia, consolidada como exame padrão de rastreio da população, até modalidades emergentes como a imagem por micro-ondas, a IA tem-se mostrado uma ferramenta poderosa para aumentar a precisão diagnóstica, reduzir a variabilidade interobservador e otimizar o fluxo de trabalho. Esta secção explora em detalhe as principais modalidades de imagem e o papel da IA na sua integração, destacando avanços técnicos, resultados quantitativos e desafios futuros.

3.3.1. Mamografia e Tomossíntese Digital da Mama

A mamografia representa o exame de referência no rastreio do cancro da mama, recorrendo a doses reduzidas de raios-X para a obtenção de imagens bidimensionais que permitem visualizar alterações estruturais do tecido mamário, como nódulos, microcalcificações ou distorções arquiteturais. Trata-se do único exame de imagem com evidência robusta de redução da mortalidade por cancro da mama em programas de rastreio populacional, apresentando valores entre 20% a 30% (Uwimana, Gnecco and Riccaboni, 2025). No entanto, a sua principal limitação reside na sobreposição dos tecidos, que pode dificultar a deteção de lesões, sobretudo em mamas densas. A elevada taxa de falsos positivos e o desconforto físico para as pacientes são também limitações de peso. Além disso, a interpretação da mamografia exige grande experiência e encontra-se sujeita à fadiga dos radiologistas, especialmente em programas de rastreio populacional que envolvem a leitura de milhares de exames por ano (Singh, Singh and Bhattacharya, 2024).

A Tomossíntese Digital da Mama (DBT), também designada por “mamografia 3D”, surgiu como uma evolução tecnológica da mamografia convencional. A técnica baseia-se na aquisição de múltiplas imagens em diferentes ângulos que, após reconstrução computacional, originam cortes seccionais da mama. Este processo reduz significativamente a sobreposição de tecidos mamários, aumentando a sensibilidade na identificação de lesões pequenas ou ocultas em mamografia bidimensional. A DBT melhora, assim, a avaliação da estrutura mamária, proporcionando uma caracterização mais precisa de massas, assimetrias e microcalcificações. Demonstrou também aumentar a taxa de deteção de cancro quando comparada à mamografia tradicional, especialmente em mulheres com elevada densidade mamária (Freitas *et al.*, 2025).

Estudos comparativos mostraram que a DBT aumenta a taxa de deteção de cancro em 27% relativamente à mamografia convencional, embora também aumente a quantidade de dados a interpretar, reforçando a necessidade de ferramentas automáticas de apoio (Marinovich *et al.*, 2025).

É neste cenário que os sistemas CAD-AI têm ganho destaque. Entre os mais referenciados encontram-se o *ProFound AI*, aprovado pela FDA (Administração de Alimentos e Medicamentos, dos Estado Unidos da América), o *MammoScreen*, avaliado em contexto de custo-efetividade no NHS (Sistema Nacional de Saúde, do Reino Unido), e o *Lunit INSIGHT MMG*, amplamente validado para mamografia digital (Uwimana, Gnecco and Riccaboni, 2025). Estes

algoritmos têm demonstrado melhorias consistentes em métricas diagnósticas: sensibilidade média de 82,7% contra 80,7% do método tradicional, especificidade de 80,9% contra 76,1%, e precisão de 85,1% contra 80,8% (Uwimana, Gnecco and Riccaboni, 2025).

Um estudo populacional de Marinovich et al. (2025) modelou a introdução da IA em programas de rastreio mamográfico e observou que, embora a taxa de detecção aumentasse, também se verificava um acréscimo na taxa de reconvocações. Este resultado revela o dilema central da integração da IA na mamografia: o equilíbrio entre ganhos em sensibilidade e potenciais custos acrescidos decorrentes de falsos positivos.

Para além dos sistemas comerciais, diversos modelos académicos têm sido desenvolvidos. Por exemplo, um estudo de Bahl et al., (2025) comparou algoritmos de *deep learning* aplicados à DBT com sistemas CAde (Detecção Assistida por Computador) tradicionais — a primeira geração de softwares de apoio à interpretação de exames de mamografia 2D. Os resultados demonstraram uma AUC significativamente superior para os algoritmos de IA (0,873 vs. 0,693), evidenciando o potencial desta tecnologia para melhorar a interpretação de imagens tridimensionais da mama.

Outro ponto relevante é a classificação automática da densidade mamária, um parâmetro essencial no risco individual de cancro e na interpretação da imagem. Algoritmos como o *Visage Breast Density* e o *intelliMammo* foram desenvolvidos especificamente para este fim, demonstrando uma boa correlação com a avaliação humana e uma maior consistência interobservador (Tan et al., 2025).

Assim, na mamografia e DBT, a IA já não se limita a ser uma promessa: existem sistemas comerciais validados e integrados em fluxos clínicos. Contudo, os desafios prendem-se com a regulamentação, a aceitação pelos profissionais e o impacto real em programas populacionais, que ainda necessita de avaliação em ensaios prospetivos multicêntricos.

3.3.2. Ressonância Magnética

A Ressonância Magnética (RM) mamária é uma técnica de imagem que recorre a campos magnéticos e agentes de contraste para produzir imagens detalhadas dos tecidos mamários. É reconhecida como uma das modalidades de imagem com maior sensibilidade. A ressonância

magnética da mama, em particular, é considerada a técnica mais sensível para a deteção e caracterização de lesões mamárias, sendo especialmente útil em pacientes de alto risco ou em contextos pré-operatórios (Abdullah *et al.*, 2025). Diversos estudos mencionam uma sensibilidade superior a 90%, embora a especificidade possa ser mais baixa, resultando numa maior incidência de falsos positivos. Entre as limitações associadas à RM destacam-se o elevado custo e a disponibilidade reduzida, fatores que restringem o seu uso a situações clínicas específicas.

A integração da IA na RM mamária visa, sobretudo, melhorar a caracterização de lesões e reduzir falsos positivos. Um estudo (Abdullah *et al.*, 2025) relatou uma AUC agrupada próxima de 0,90, com sensibilidade média de 88% e especificidade de 90% para algoritmos de *deep learning* aplicados à RM. Estes valores são comparáveis, ou mesmo superiores, ao desempenho médio de radiologistas.

A RM apresenta ainda uma oportunidade única para a fusão multimodal de dados. Alguns estudos exploraram a integração de características radiómicas, que são propriedades quantitativas extraídas de imagens médicas, com algoritmos de IA, obtendo desempenhos acrescidos na diferenciação entre lesões benignas e malignas. Apesar destes avanços, as revisões apontam limitações metodológicas, como amostras pequenas, falta de validação externa e heterogeneidade de protocolos de aquisição (Abdullah *et al.*, 2025).

3.3.3. Ultrassonografia

A ultrassonografia, também conhecida como ecografia, é uma técnica de imagem que utiliza ondas sonoras de alta frequência para produzir imagens dos tecidos internos. Desempenha um papel fundamental como exame complementar na avaliação da mama, sendo especialmente útil em mulheres jovens e em casos de mamas densas, onde a sensibilidade da mamografia pode ser reduzida (Wu *et al.*, 2025). Esta modalidade apresenta várias vantagens, incluindo acessibilidade, segurança por não utilizar radiação ionizante e boa tolerância pelos pacientes. No entanto, a qualidade das imagens e a acuidade diagnóstica dependem fortemente da experiência e habilidade do operador, o que pode promover variabilidade na interpretação dos resultados. Além disso, a ultrassonografia é eficaz na diferenciação de lesões sólidas e císticas e pode orientar procedimentos de intervenção, como biópsias e aspirações.

O uso da IA na ultrassonografia mamária tem-se expandido rapidamente, com sistemas comerciais como o *Koios DS* e o *S-Detect* (Wu *et al.*, 2025). Em particular, algoritmos baseados no modelo *YOLOv3-tiny* já foram aprovados para o uso clínico em rastreamento, demonstrando uma AUC de 0,88 e sensibilidade até 92% em ensaios clínicos (Wu *et al.*, 2025).

A principal aplicação da IA na ultrassonografia é a categorização BI-RADS de nódulos sólidos, auxiliando na diferenciação entre lesões benignas e malignas, podendo reduzir o número de biópsias desnecessárias. Além disso, sistemas de detecção em tempo real permitem a triagem de forma automática durante o exame, otimizando a eficiência clínica.

Ainda assim, as limitações incluem a variabilidade de qualidade das imagens, dependência de protocolos de aquisição e escassez de estudos multicêntricos com grandes amostras.

3.3.4. Patologia Digital e *Whole Slide Imaging* (WSI)

A patologia digital envolve a digitalização de lâminas histopatológicas, criando imagens de alta resolução que podem ser visualizadas, armazenadas e analisadas digitalmente. O *Whole Slide Imaging* (WSI) é a tecnologia que permite essa digitalização completa das lâminas, produzindo imagens digitais que reproduzem com fidelidade toda a amostra do tecido, incluindo detalhes microscópicos.

A utilização do WSI permite a aplicação da IA a grandes volumes de imagens histológicas, acelerando a análise de tecidos, reduzindo a variabilidade interobservador e aumentando a consistência diagnóstica. Esta abordagem abriu caminho para o uso de algoritmos de *deep learning* em anatomia patológica, área particularmente adequada para IA devido à quantidade elevada de dados e à complexidade dos padrões presentes nas imagens (Datwani *et al.*, 2025).

Sistemas como o *Galen Breast (IBEX)* foram treinados com milhões de imagens e demonstraram um desempenho notável, com uma AUC superior a 0,95 na detecção de carcinoma invasivo e DCIS (*Carcinoma Ductal In Situ*). Estudos adicionais relataram uma sensibilidade entre 91% e 94% e especificidade entre 95% e 96% para várias análises de imagens (Datwani *et al.*, 2025).

Para além da detecção de neoplasias, a IA tem sido utilizada na quantificação de biomarcadores do cancro da mama – que são substâncias que ajudam a identificar, caracterizar e orientar o

tratamento da doença – com resultados comparáveis aos de patologistas experientes. Esta aplicação é particularmente relevante em oncologia de precisão, onde a correta avaliação de biomarcadores determina a escolha terapêutica (Datwani *et al.*, 2025).

Apesar dos avanços identificados, permanecem vários desafios, incluindo a necessidade de validação multicêntrica, a padronização de *scanners* e algoritmos, o elevado custo dos aparelhos digitais e a integração eficiente nos fluxos de trabalho laboratoriais. Além disso, questões éticas relacionadas com a utilização de bases de dados de tecidos humanos exigem a definição de regulamentação clara e robusta.

3.3.5. Modalidades Emergentes e Fusão Multimodal

Além das modalidades já mencionadas, a investigação recente tem explorado abordagens emergentes. Um exemplo é a imagem por micro-ondas (MWI), investigada como técnica não ionizante, não invasiva, económica, rápida e promissora em contextos de baixo recurso. A MWI é uma técnica experimental que usa micro-ondas para gerar imagens da mama. Sistemas como o *MammoWave* estão em fase de avaliação multicêntrica, com estudos preliminares a sugerirem sensibilidade promissora, embora ainda com necessidade de validação clínica robusta (Uwimana, Gnecco and Riccaboni, 2025).

Outra linha de investigação é a fusão multimodal, que combina imagens de diferentes modalidades, por exemplo, a junção de mamografia com ultrassonografia e RM, com dados clínicos e genómicos. Esta abordagem tem potencial para fornecer um diagnóstico mais holístico e personalizado.

3.4. Resultados Quantitativos e Análise

A análise quantitativa baseada nas métricas de desempenho dos sistemas de inteligência artificial aplicados ao diagnóstico do cancro da mama revela avanços expressivos em múltiplas modalidades, frequentemente superando ou igualando o desempenho humano em tarefas específicas. Este capítulo sintetiza os principais resultados relatados nas revisões e estudos

mencionados. O objetivo não é apresentar dados originais, mas sim sintetizar e comparar métricas-chave, tais como a taxa de deteção do cancro (CDR), sensibilidade, especificidade e AUC, conforme mencionados na literatura.

A compilação destes dados permite identificar tendências, avaliar a consistência dos resultados entre diferentes métodos e algoritmos, e fornecer um panorama geral do potencial clínico da IA no contexto do rastreio mamográfico. A partir da revisão efetuada, foi possível extrair valores agregados e exemplos representativos do desempenho dos diferentes métodos e algoritmos analisados, já mencionados anteriormente de forma breve. A seguir apresenta-se uma síntese dos valores mais frequentemente relatados nos estudos selecionados.

A mamografia assistida por IA apresentou uma sensibilidade média de, aproximadamente, 82,7%, em comparação com a sensibilidade obtida pelos radiologistas sem auxílio da IA, de 80,7%, tendo obtido um incremento de 2%. Já a especificidade apresentou um valor de 80,9% com o auxílio de IA, comparativamente ao valor de 76,1%, obtido pelo método tradicional, apresentando neste caso um incremento perto dos 5% (Uwimana, Gnecco and Riccaboni, 2025).

Noutro estudo (Hashim *et al.*, 2025) foram analisadas mamografias de mais de 120 000 pacientes, e os resultados sugerem que a IA tem uma sensibilidade geralmente maior, de 85% em comparação com o valor dos radiologistas de 77%. O valor de AUC obtido pela IA foi de 0,89, comparativamente ao dos radiologistas, de 0,82. Já a especificidade apresentou melhores valores no caso dos radiologistas, de 90%, comparativamente à sensibilidade da IA que apresentou uma descida de 1% (89%).

Foi realizada uma comparação entre o uso de um método tradicional de deteção assistida por computador (CADe) em mamografias 2D sintéticas com um algoritmo de IA baseado em tomossíntese digital mamária (Bahl *et al.*, 2025). O estudo envolveu 764 pacientes, sendo 106 com cancro da mama confirmado por biópsia. Os resultados mostraram que a IA aplicada à DBT apresentou um desempenho significativamente superior ao CAdE em mamografias 2D, apresentando maior sensibilidade e especificidade na deteção do cancro da mama. O valor de AUC no caso da DBT com IA foi de 0,873, comparativamente ao valor de 0,693 obtido pelo sistema CAdE com mamografia 2D.

No caso da ressonância magnética, uma meta-análise (Abdullah *et al.*, 2025) que avaliou o uso da IA, especificamente com tecnologia *deep learning*, na deteção e classificação do cancro da mama, apresentou valores bastante promissores. Na análise geral, a AUC agrupada foi de, aproximadamente, 0,90, indicando excelente capacidade discriminativa dos modelos. A sensibilidade média foi de cerca de 88%, e a especificidade média foi de, aproximadamente, 90%. Além disso, o tamanho do conjunto de treino foi um fator bastante significativo na performance dos algoritmos, sendo que conjuntos maiores (mais de 600 amostras) tendem a proporcionar melhores resultados, embora a heterogeneidade ainda esteja presente.

Numa revisão de Guo *et al.* (2025) que incluiu o uso da IA combinada com radiómica – método que estuda as informações contidas nas imagens radiográficas – em mamografias ultrassonográficas automatizadas da mama (ABUS) para o diagnóstico de tumores mamários, mostrou valores promissores. Esta revisão destaca como o ABUS supera as limitações da ultrassonografia convencional, oferecendo uma avaliação mais abrangente do tecido mamário. A integração da IA nesta tecnologia melhora a análise das imagens e consequentemente o diagnóstico precoce. Alguns exemplos isolados de estudos incluídos na revisão (Guo *et al.*, 2025) mostraram relatórios de sensibilidade elevada, com valores de 92% e valores de AUC entre 0,88 e 0,90.

No caso da integração da IA na patologia digital, também foram apresentados valores promissores, mostrando a utilidade da sua integração em diversas modalidades. Foram relatados valores de AUC bastante elevados, como por exemplo 0,976 para carcinoma invasivo, DCIS e outras lesões. Os valores de sensibilidade e especificidade foram apresentados em alguns estudos, onde a primeira métrica esteve entre 91% e 94% e a segunda entre 95% e 96% (Datwani *et al.*, 2025).

Numa avaliação de Yu *et al.* (2024) acerca do uso da IA no apoio ao diagnóstico do cancro da mama, num estudo realizado em Hong Kong, foram analisadas quase 400 pacientes com diferentes tipos de cancro. A IA conseguiu detetar cerca de 91,5% dos casos e identificar corretamente casos sem cancro em aproximadamente 96,3% das vezes. Neste caso, o sistema mostrou ser especialmente eficaz em identificar tumores invasivos, palpáveis e que apresentam massas ou distorções arquiteturais, mas apresentou dificuldades em detetar cancros menores ou aqueles que se manifestam de forma mais subtil. Este comportamento pode ser justificado pelo facto de os algoritmos de IA serem geralmente treinados com conjuntos de dados em que

os tumores invasivos, de maiores dimensões e com características visuais mais evidentes estão mais bem representados. Em contraste, lesões menores ou manifestações mais subtis tendem a ser menos frequentes ou mais difíceis de anotar nos conjuntos de treino, o que limita a sua deteção.

Outro estudo, de Qu et al. (2024), analisou o uso da IA baseada em ultrassonografia dinâmica em tempo real para diagnosticar lesões mamárias mínimas (até 10 milímetros). Os resultados mostram que a integração da IA ajuda a melhorar a precisão na diferenciação entre lesões benignas e malignas. Mostrou também que, ao usar limites de BI-RADS – classificação utilizada para padronizar os resultados de exames de imagens mamográficas – (categorias superiores a 3 ou 4A) é possível aumentar a sensibilidade e valor preditivo negativo, reduzindo o risco de perder diagnósticos malignos. Contudo, o estudo apresenta algumas dificuldades, como a baixa especificidade ao usar limites BI-RADS mais baixos, levando a mais diagnósticos falso-positivos. Ou seja, embora os limites do BI-RADS sejam úteis para padronizar a avaliação diagnóstica com IA, há a necessidade de equilibrar a sensibilidade e especificidade, para garantir melhores resultados, minimizando falsos positivos e negativos.

De forma a sintetizar os resultados apresentados nos vários estudos mencionados, o Gráfico 1 apresenta uma média dos valores descritos para as métricas de sensibilidade, especificidade e AUC, com o intuito de comparar o uso da IA com o método tradicional, de leitura apenas pelos radiologistas. A barra de cor azul representa as métricas com a integração da IA, e a barra de cor verde representa as métricas do método tradicional.

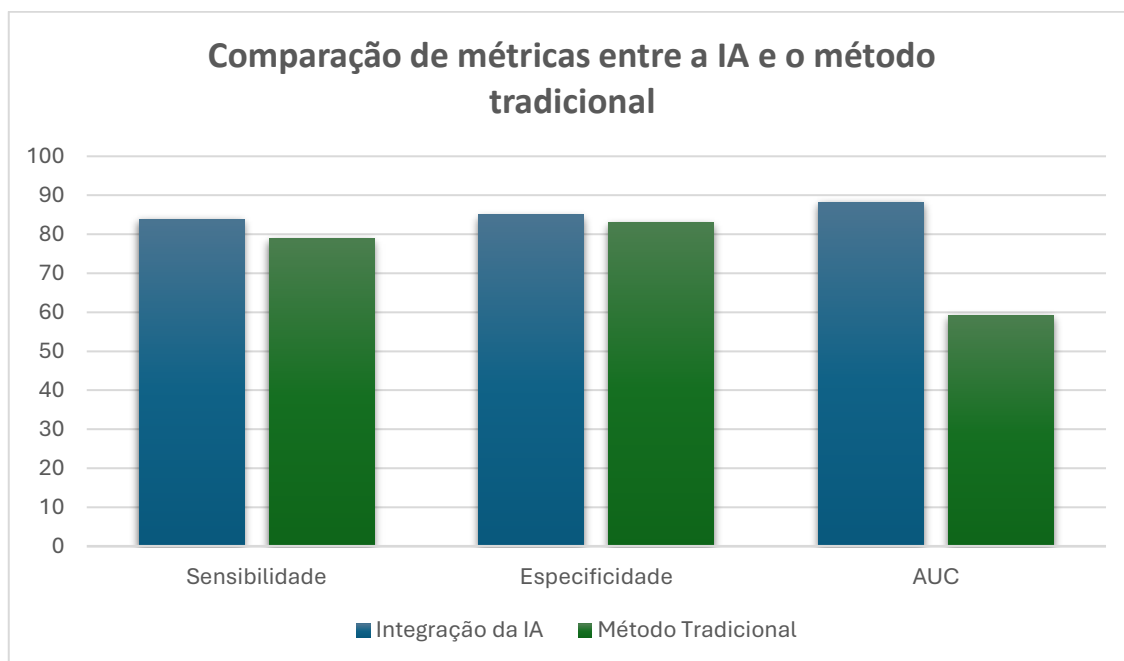


Gráfico 1 – Comparação de métricas entre a IA e o método tradicional

Nas três situações, como é possível verificar no gráfico, a integração da IA apresenta melhores resultados. Os valores mais elevados de sensibilidade e AUC observados podem ser justificados pela maior capacidade destes algoritmos em reconhecer padrões complexos e subtis nas imagens médicas, frequentemente impercetíveis à observação humana. A utilização de técnicas de *deep learning* permite extrair um número muito superior de características relevantes, melhorando a distinção entre tecidos normais e suspeitos. Esta capacidade traduz-se num melhor desempenho global, representado pela AUC superior, e numa maior taxa de deteção de casos verdadeiros de cancro, como se verifica pela elevada sensibilidade.

Já os valores de especificidade mais similares entre as duas situações indicam que, embora a IA detete mais casos positivos, mantém um desempenho comparável ao método tradicional na identificação de exames normais, ou seja, não aumenta de forma significativa o número de falsos positivos. Isto sugere que a IA consegue melhorar a sensibilidade sem comprometer a precisão diagnóstica em casos benignos, alcançando um equilíbrio desejável entre deteção precoce e estabilidade no número de reconvocações, isto é, o número de vezes que um paciente é chamado para repetir um exame. Assim, o comportamento observado reforça o

potencial clínico da IA como ferramenta de apoio, capaz de aumentar a eficácia do rastreio mamográfico sem sobrecarregar o processo com falsos alarmes.

Apesar de os resultados apresentados serem muito promissores, é importante reconhecer que ainda existem algumas lacunas nos estudos existentes. Muitos destes estudos têm caráter retrospectivo ou recorrem a bases de dados enriquecidas com casos positivos, o que pode levar à sobrevalorização das métricas de desempenho. Ainda assim, estas limitações apontam caminhos claros para futuras melhorias, nomeadamente através da validação externa multicêntrica e da utilização de amostras mais representativas.

4. Algoritmos e Sistemas CAD

O surgimento dos algoritmos de inteligência artificial, em particular das redes neurais profundas, ou *deep learning*, impulsionou uma nova geração de ferramentas capazes de lidar com grandes volumes de dados e padrões complexos. Atualmente, diferentes algoritmos de IA têm sido amplamente aplicados no contexto clínico, cada um com características próprias e adequação a diferentes tipos de análise.

Para contextualizar, *machine learning* e *deep learning* são subáreas da inteligência artificial que permitem aos computadores aprender padrões e realizar previsões a partir de dados. Enquanto o *machine learning* se foca na identificação de relações complexas e classificação de dados sem programação explícita, o *deep learning* utiliza redes neurais profundas capazes de extrair representações hierárquicas e abstratas, reconhecendo padrões complexos em imagens médicas. Estas abordagens são fundamentais para a detecção de lesões subtis e para otimizar a precisão dos sistemas CAD-AI no diagnóstico do cancro da mama.

O desenvolvimento dos sistemas CAD baseados em inteligência artificial representa uma das áreas mais dinâmicas no apoio ao diagnóstico do cancro da mama. Desde os primeiros algoritmos de detecção automática de lesões em mamografia até aos atuais modelos de *deep learning* integrados em fluxos clínicos, a evolução tem sido marcada por avanços significativos em termos de desempenho, capacidade de generalização e integração regulatória.

4.1. Evolução dos Sistemas CAD

Historicamente, os primeiros sistemas CAD foram concebidos sobretudo como ferramentas de detecção assistida por computador (CADE), cujo objetivo principal era identificar potenciais regiões de interesse em exames de imagem, assinalando áreas suspeitas para posterior avaliação por parte do radiologista. Apesar de úteis, estes sistemas apresentavam limitações de relevo: elevada taxa de falsos positivos, baixa especificidade e impacto clínico questionável, pois muitas vezes acrescentavam tempo de leitura em vez de o reduzir (Singh, Singh and Bhattacharya, 2024).

O surgimento do *machine learning* e, mais recentemente, do *deep learning*, proporcionou uma segunda geração de algoritmos, designados CADx (Diagnóstico Assistido por Computador). Estes sistemas não apenas detetam regiões suspeitas, mas também realizam tarefas de classificação, caracterização e previsão do risco, oferecendo informação diagnóstica mais completa. Em alguns casos, os CADx modernos atingem desempenhos equiparáveis aos de especialistas humanos, o que tem fomentado a sua aprovação regulatória e utilização comercial (Uwimana, Gnecco and Riccaboni, 2025). Para esclarecimento, todos os sistemas CAD mencionados neste trabalho são referentes a este último tipo, CADx.

4.2. Algoritmos de IA Aplicados ao Diagnóstico por Imagem

Como já referido, a aplicação de algoritmos de IA no diagnóstico por imagem tem vindo a transformar a prática radiológica, oferecendo novas abordagens para a deteção, caracterização e classificação de lesões. Estes algoritmos, baseados em técnicas de *machine learning* e *deep learning*, permitem analisar grandes volumes de dados de forma rápida e padronizada, reduzindo a variabilidade interobservador e potencializando a precisão diagnóstica.

Com o objetivo de fornecer uma visão geral, esta secção descreve os principais tipos de algoritmos de IA e as suas aplicações no diagnóstico do cancro da mama. A Tabela 3, apresenta a lista dos mesmos, destacando as suas aplicações e contributos para a prática clínica.

Tabela 3 – Principais algoritmos de IA utilizados no diagnóstico médico

Algoritmo de IA	Descrição
Convolutional Neural Network (CNN)	Um dos mais relevantes, bastante aplicado na interpretação de imagens médicas devido à alta precisão na detecção de padrões e anomalias.
Deep Neural Networks (DNN)	Desempenham um papel fundamental na análise de grandes volumes de dados clínicos e genômicos, permitindo a identificação de padrões complexos.
Support Vector Machine (SVM)	Destaca-se na classificação de células cancerígenas e no diagnóstico de doenças como Alzheimer e Parkinson. A eficácia deste tipo de algoritmo é notável em conjuntos de dados pequenos e em problemas de classificação binária, como a distinção entre células saudáveis e doentes.
Random Forest	São modelos bastante utilizados para prever o risco de doenças com base em fatores clínicos e históricos dos pacientes, sendo destacados pela resistência ao overfitting e facilidade de interpretação dos resultados.
Generative Adversarial Networks (GAN)	São utilizados para gerar imagens médicas sintéticas, auxiliando no treino de modelos de IA sem a necessidade de grandes bases de dados reais. Esta abordagem é essencial para aprimorar a precisão dos modelos preditivos, principalmente em cenários onde há escassez de dados.
K-Nearest Neighbors (KNN)	Este algoritmo auxilia na classificação de exames laboratoriais e no prognóstico de doenças com base em sintomas similares, sendo conhecido pela sua simplicidade e facilidade de implementação.
Regressão Logística e Árvores de Decisão	Estes métodos são utilizados para prever a probabilidade de doenças com base em dados clínicos, oferecendo modelos estatísticos de fácil interpretação e integração com sistemas médicos.

Entre os algoritmos apresentados na *Tabela 3*, as redes neurais convolucionais (CNNs) assumem particular destaque, tendo impulsionado avanços significativos na análise de imagens médicas. Graças à sua elevada capacidade de extração e reconhecimento de padrões complexos, as CNNs têm demonstrado um desempenho excepcional em diversas tarefas de visão

computacional. No caso específico da imagiologia mamária, destacam-se as seguintes arquiteturas:

- **AlexNet, VGGNet e ResNet** – São modelos clássicos que impulsionaram o *deep learning* aplicado a imagem médica, com elevada capacidade de extração de características hierárquicas.
- **DenseNet e Inception** – Representam arquiteturas otimizadas para eficiência computacional e complexidade, frequentemente utilizadas em classificadores de mamografia digital.
- **YOLO e Faster R-CNN** – São modelos de detecção em tempo real, utilizados em ultrassonografia mamária para identificar nódulos com elevada sensibilidade.
- **U-Net e variantes** – São arquiteturas amplamente utilizadas em tarefas de segmentação (por exemplo, segmentação de tumores em RM ou em lâminas histológicas digitalizadas).
- **Modelos híbridos multimodais** – Representam combinações de CNNs com *Recurrent Neural Networks* (RNNs) ou com dados clínicos/genómicos, que permitem uma integração mais completa da informação do paciente.

Como já referido, as redes neuronais convolucionais constituem a base para a maioria dos sistemas de inteligência artificial aplicados ao diagnóstico por imagem. No entanto, a sua eficácia está diretamente dependente da disponibilidade de grandes volumes de imagens médicas devidamente anotadas, um requisito que nem sempre é possível cumprir na prática clínica. Para mitigar esta limitação, recorrem-se a estratégias complementares, entre as quais se destacam as técnicas de *transfer learning* e *data augmentation*. Estas abordagens contribuem para mitigar a limitação da escassez de grandes conjuntos de dados anotados em oncologia mamária, permitindo treinar modelos mais robustos e generalizáveis a partir de bases de dados reduzidas (Tan *et al.*, 2025).

O *transfer learning* consiste em aproveitar redes CNN previamente treinadas em grandes bases de dados não médicas, como o *ImageNet*, e adaptá-las a tarefas específicas da imagiologia mamária. Esta abordagem permite reutilizar características já aprendidas nas primeiras camadas convolucionais, que geralmente detetam padrões visuais básicos (como arestas, texturas ou formas), ajustando apenas as camadas finais às particularidades das imagens mamográficas ou tomossintéticas. Deste modo, reduz-se substancialmente o tempo de treino

e a necessidade de grandes conjuntos de dados anotados, mantendo níveis elevados de desempenho.

Já o *data augmentation* procura aumentar artificialmente a diversidade das imagens de treino, através da aplicação de transformações controladas, tais como rotações, inversões, ampliações, variações de contraste ou adição de ruído. Estas alterações não modificam a natureza clínica da imagem, mas permitem que a rede seja exposta a diferentes cenários e variações possíveis na prática, reforçando a sua capacidade de generalização. Ao evitar que a CNN memorize padrões específicos de um conjunto reduzido de imagens, o *data augmentation* contribui para reduzir o sobreajuste (*overfitting*) e melhorar a robustez do modelo.

Desta forma, tanto o *transfer learning* como o *data augmentation* assumem-se como técnicas fundamentais para maximizar o potencial das CNNs no apoio ao diagnóstico do cancro da mama, permitindo ultrapassar uma das principais barreiras deste domínio: a escassez de dados médicos disponíveis para treino e validação dos algoritmos.

Os modelos multimodais são considerados uma das tecnologias mais promissoras da IA aplicada à saúde, especialmente no diagnóstico e tratamento do cancro. São modelos de *deep learning* que processam e integram múltiplos tipos de dados, como imagens médicas, dados clínicos e genéticos e relatórios em forma de texto, oferecendo uma visão geral do paciente. Desta forma permitem uma análise mais robusta e precisa e favorecem decisões personalizadas. Ao integrar diferentes tipos de dados, estes modelos identificam padrões que seriam impercetíveis quando cada modalidade é analisada isoladamente. Esta abordagem reflete o raciocínio clínico humano e tem mostrado potencial para melhorar a deteção precoce do cancro, embora apresentem desafios de explicabilidade, escassez de dados e complexidade computacional.

Em resumo, a diversidade de algoritmos e técnicas demonstra a vasta capacidade de aprendizagem e adaptação dos sistemas de inteligência artificial, constituindo um fundamento essencial para a sua integração em soluções CAD-AI aplicadas ao diagnóstico do cancro da mama.

5. Análise dos Sistemas CAD

Diversos sistemas CAD-AI já alcançaram validação clínica e, em alguns casos, aprovação por entidades regulatórias internacionais, como a FDA e a marcação CE, o que evidencia a maturidade crescente destas tecnologias. Os algoritmos e sistemas de apoio ao diagnóstico baseados em inteligência artificial constituem, assim, uma das maiores inovações no diagnóstico do cancro da mama nas últimas décadas, transformando-se de ferramentas auxiliares rudimentares em soluções robustas e integráveis nos fluxos clínicos.

Inicialmente concebidos apenas como sistemas de deteção assistida por computador (CADE), com capacidade limitada e elevada taxa de falsos positivos, evoluíram para sistemas CADx baseados em *deep learning*, capazes de detetar, classificar e caracterizar lesões com elevada precisão. Atualmente, encontram-se disponíveis soluções comerciais já em utilização clínica, particularmente na mamografia digital, tomossíntese e ecografia, onde têm demonstrado potencial para aumentar a sensibilidade diagnóstica, reduzir a variabilidade interobservador e otimizar o tempo de leitura.

Tendo em conta a diversidade de algoritmos e soluções comerciais atualmente disponíveis, optou-se por selecionar três sistemas de inteligência artificial para análise mais aprofundada: *ProFound AI*, *Lunit INSIGHT MMG* e *Transpara*. Estes três sistemas representam diferentes abordagens dentro do ecossistema de CAD-AI: o *ProFound AI*, amplamente reconhecido como pioneiro na aplicação da IA à tomossíntese digital da mama (DBT); o *Lunit INSIGHT MMG*, validado em múltiplos países e com dados consistentes sobre o seu desempenho em programas de rastreio; e o *Transpara*, um dos primeiros sistemas a ser integrado em programas nacionais de rastreio na Europa, com particular destaque na otimização do fluxo de trabalho e na estratificação do risco.

A análise detalhada destes sistemas permitirá compreender não só as métricas quantitativas relacionadas, mas também os desafios e benefícios da sua implementação prática, contribuindo para uma visão crítica sobre o papel atual e futuro da IA no diagnóstico do cancro da mama.

5.1. ProFound AI

O *ProFound AI* (ou *ProFound Breast Health Suite*) é um sistema de IA desenvolvido pela *iCAD Inc.*, projetado para auxiliar na detecção precoce do cancro da mama. Com a utilização de algoritmos avançados de *deep learning*, o *ProFound AI* analisa imagens de tomossíntese digital (DBT) e mamografias digitais 2D para identificar lesões suspeitas, classificando-as com base na probabilidade de malignidade (Nepute *et al.*, 2025). O *ProFound AI* é compatível com sistemas de mamografia de múltiplos fornecedores, permitindo a integração em diversos fluxos de trabalho clínicos. Além disso, o sistema foi projetado para ser utilizado em tempo real, fornecendo resultados rápidos e precisos que podem melhorar a eficiência do diagnóstico e reduzir a carga de trabalho dos radiologistas.

O *ProFound AI* encontra-se regulamentado para uso clínico, tendo obtido certificações de órgãos internacionais como a FDA nos Estados Unidos, a marca CE na União Europeia e a licença do *Health Canada*, atestando a sua segurança e eficácia.

O *ProFound AI* está dividido em três componentes principais:

- *ProFound Detection* – Ferramenta de detecção de cancros que destaca densidades suspeitas, assimetrias, dilatações arquitetónicas, calcificações, etc.
- *ProFound Density* – Avaliação automática da densidade do tecido mamário.
- *ProFound Risk* – Modelo de risco personalizado, baseado em imagem, que procura prever a probabilidade de cancro nos próximos 1 a 2 anos.

O desenvolvimento do *ProFound AI* baseia-se em algoritmos de *deep learning*, predominantemente redes neuronais convolucionais, concebidos para analisar imagens da mama. Estes algoritmos foram treinados com grandes bases de dados clínicos anotados de diversas instituições e programas de rastreio, contendo imagens positivas, confirmadas histologicamente como malignas, e imagens negativas, sem evidência de doença. Estes conjuntos de dados incluíram uma variedade significativa de perfis demográficos, densidade mamária e diferentes equipamentos de imagem, garantindo a generalização e aplicabilidade do modelo em diversas populações. Cada imagem foi cuidadosamente anotada por radiologistas experientes, que assinalaram lesões suspeitas e as classificaram de acordo com a morfologia, localização e padrão, incluindo massas, microcalcificações, assimetrias e distorções arquitetónicas (Nepute *et al.*, 2025).

O *ProFound AI* foi treinado através de aprendizagem supervisionada, em que o modelo ajusta iterativamente os pesos das redes neurais para minimizar a diferença entre as previsões e as anotações humanas, associando padrões visuais a categorias clínicas. No caso da DBT, o algoritmo analisa cada corte individual, integrando posteriormente as informações ao longo dos múltiplos planos, o que melhora a capacidade de detecção de lesões subtis em mamas densas.

O modelo foi submetido a procedimentos rigorosos de validação interna e externa utilizando subconjuntos de dados independentes não incluídos no treino. Isto permitiu o ajuste fino de parâmetros críticos, como os limites de alerta e a sensibilidade geral, minimizando os falsos positivos e garantindo um desempenho robusto em cenários clínicos reais. Além disso, o *iCAD* atualiza periodicamente o algoritmo, incorporando dados recentes e feedback clínico, para manter um desempenho consistente e adaptável a novas populações, equipamentos e protocolos de rastreio.

Diversos estudos clínicos têm demonstrado que o *ProFound AI* apresenta um desempenho significativo na melhoria da detecção do cancro da mama, quando comparado com a leitura convencional realizada por radiologistas sem o apoio a sistemas de inteligência artificial. Este sistema tem o objetivo de auxiliar o radiologista na identificação de achados suspeitos e na priorização de casos com maior probabilidade de malignidade.

Um estudo de Conant et al., (2019) avaliou o desempenho geral do sistema *ProFound AI*, mostrando que a utilização da IA na leitura de exames mamográficos promoveu melhorias significativas no desempenho dos radiologistas. O desempenho médio apresentou um aumento de 8,0% na sensibilidade, de 77,0% sem IA para 85,0% com IA. Este aumento de sensibilidade não se traduziu num aumento proporcional das taxas de falsos positivos, demonstrando que o sistema mantém uma boa especificidade diagnóstica. A especificidade apresentou um incremento 6,9%, de 62,7% para 69,6%. Relativamente ao valor de AUC houve um incremento de 0,057, tendo sido a média geral da AUC sem AI de aproximadamente 0,795, enquanto com a AI aumentou para cerca de 0,852, o que reflete uma capacidade robusta de distinguir entre casos benignos e malignos.

Estes resultados mantêm-se consistentes em análises estratificadas por densidade mamária, sugerindo que o sistema é capaz de manter um bom desempenho mesmo em mamas densas — uma limitação conhecida da mamografia tradicional (Eriksson et al., 2023).

Outro aspeto amplamente documentado refere-se à eficiência do processo de leitura. A IA também contribuiu para uma redução substancial no tempo de leitura por exame, com uma diminuição aproximada de 52,7%, de 64,1 segundos para 30,4 segundos, sem comprometer a precisão diagnóstica (Conant *et al.*, 2019). Tal redução traduz-se em benefícios significativos em termos de fluxo de trabalho, especialmente em programas de rastreio com elevado volume de exames, em que o tempo de interpretação constitui um fator crítico de produtividade.

Paralelamente, a taxa de reconvocações — isto é, das chamadas para reavaliação após o rastreio — de casos benignos reduziu em aproximadamente 7,2%, de 38,0% para 30,9% (Conant *et al.*, 2019). Destaca-se, ainda, uma melhoria específica na sensibilidade para casos com microcalcificações. As alterações na sensibilidade variaram entre diferentes subgrupos de casos, evidenciando melhorias mais acentuadas em determinados tipos de lesões.

Para além da deteção imediata, há também evidência de que o *ProFound AI* tem potencial para identificar retrospectivamente um número relevante de cancros intervalares, ou seja, tumores diagnosticados entre rastreios periódicos (Letter *et al.*, 2023). Tal capacidade sugere que o sistema pode contribuir para reduzir o número de casos não detetados em rastreios anteriores, reforçando a sua utilidade como ferramenta de apoio à decisão clínica.

Um outro estudo de (Resch *et al.*, 2024) analisou o desempenho deste sistema, relatando também bons resultados. O sistema apresentou uma AUC de 0,93 em toda a amostra, destacando-se particularmente em mamas não densas. Em mamas densas apresentou uma AUC de 0,92. Na análise de critérios de classificação por risco, um ponto de corte de baixo risco (inferior a 51) na pontuação de caso proporcionou uma sensibilidade de 100%, garantindo que nenhum caso maligno fosse erroneamente categorizado como de baixo risco. Para o critério de alto risco (superior a 69), o sistema alcançou sensibilidade de 93,1% e especificidade de 82%. O *ProFound AI* classificou a maioria dos casos de cancro detetados (54 de 58) com pontuações elevadas (70–100), enquanto apenas 3% dos casos malignos foram atribuídos a pontuações intermediárias (30–69), demonstrando uma elevada capacidade de estratificação do risco de malignidade.

Apesar dos bons resultados (Resch *et al.*, 2024), também foram destacadas algumas limitações na análise deste sistema. A amostra utilizada foi relativamente pequena, o que restringe a generalização dos resultados e indica a necessidade de estudos maiores para confirmar as

descobertas. O algoritmo não acedeu a informações prévias do paciente, como mamografias anteriores, o que poderia levar a uma melhoria do desempenho. Outro ponto importante é que as pontuações de detecção variam de 0 a 100, mas há baixa representatividade de câncros em mamas extremamente densas, dificultando a avaliação do desempenho em populações com alta densidade mamária, que é um fator de risco relevante. Por fim, há uma necessidade de maior diversificação nas bases de dados de treino destes algoritmos para garantir a aplicabilidade em grupos socioeconomicamente desfavorecidos, que frequentemente estão subdiagnosticados.

O *ProFound AI 3D*, uma versão otimizada da plataforma *ProFound AI*, especificamente concebida para a análise de imagens de DBT em três dimensões, mostra também resultados promissores. Esta ferramenta permite uma visualização mais detalhada da anatomia mamária, facilitando a detecção mais precisa de lesões em fases iniciais, através da análise de cada fatia da imagem tridimensional de forma individualizada, promovendo uma avaliação mais abrangente.

O sistema *ProFound AI 3D*, apresenta resultados que demonstram uma elevada capacidade de detecção do cancro da mama, com desempenho comparável à dupla leitura realizada por radiologistas (Graewingholt and Duffy, 2021). No estudo analisado, o sistema atingiu taxas de detecção de 97,5% (157 em 161 casos) para um limiar de sensibilidade mais baixo e 99,4% (160 em 161 casos) para um limiar mais alto, evidenciando uma sensibilidade notavelmente elevada na identificação de lesões malignas.

Embora o estudo não tenha relatado taxas de falsos positivos, salientou-se, mais uma vez, o potencial do *ProFound AI* para reduzir a carga de trabalho dos radiologistas e melhorar a detecção de tumores difíceis, sobretudo em casos de diagnóstico desafiante ou em contextos de rastreio populacional.

Globalmente, a integração do *ProFound AI* na prática clínica traduz-se numa melhoria mensurável dos indicadores de qualidade do rastreio mamográfico. A sua integração na prática clínica permite aumentar significativamente a sensibilidade e a taxa de detecção de malignidades, reduzir falsos positivos e reconvocações, otimizar o tempo de leitura e apoiar a estratificação de risco, contribuindo assim para uma tomada de decisão mais informada e eficiente pelos radiologistas. No entanto, o sistema também apresenta algumas limitações,

incluindo a falta de diversidade nas bases de dados utilizadas, a pouca representatividade de mamas extremamente densas e a falta de acesso a informações prévias do paciente.

Não obstante, os resultados gerais obtidos reforçam a aplicabilidade da inteligência artificial como ferramenta de suporte ao diagnóstico precoce do cancro da mama.

5.2. Lunit INSIGHT MMG

O *Lunit INSIGHT MMG* é uma solução avançada de IA desenvolvida para apoiar a deteção precoce do cancro de mama através da análise automática de mamografias digitais. Com a utilização de algoritmos de *deep learning* treinados com um vasto conjunto de dados clínicos, o sistema identifica e caracteriza áreas suspeitas de malignidade, atribuindo-lhes pontuações de anomalia que refletem a probabilidade de presença de cancro.

Além disso, o *Lunit INSIGHT MMG* tem a capacidade de gerar mapas de calor e contornos que destacam as lesões suspeitas, facilitando a interpretação das imagens pelos radiologistas. A solução é compatível com sistemas de mamografia digital de campo total (FFDM) – tipo de mamografia digital que captura imagens de todo o campo mamário de uma vez – e complementa-se aos fluxos de trabalho existentes, permitindo uma análise eficiente e adicional às leituras tradicionais. Ao contrário dos outros dois sistemas CAD analisados, que são integrados diretamente no fluxo de trabalho do radiologista, o *Lunit INSIGHT MMG* é utilizado como uma ferramenta de leitura adicional.

Estudos clínicos demonstraram que este sistema pode melhorar a sensibilidade na deteção do cancro da mama, especialmente em tecidos mamários densos, e reduzir a carga de trabalho dos radiologistas, sem comprometer a precisão diagnóstica. Em termos regulatórios, o *Lunit INSIGHT MMG* obteve aprovação da FDA nos Estados Unidos, e possui certificação CE na União Europeia, permitindo a sua utilização em diversos mercados internacionais. A solução está aprovada em diversas regiões, incluindo União Europeia, Estados Unidos da América, Canadá, Reino Unido, Austrália e Nova Zelândia, estando disponível para implementação em instituições de saúde que utilizem mamografia digital.

O *Lunit INSIGHT MMG* foi desenvolvido utilizando técnicas de *deep learning* aplicadas a um vasto conjunto de dados clínicos. De acordo com informações da *Lunit*, o sistema foi treinado

com um conjunto de dados abrangente, incluindo mais de 240 000 casos de mamografias, dos quais aproximadamente 50 000 foram confirmados por biópsia como positivos para cancro da mama. O treino destes algoritmos envolveu a análise de imagens de mamografia digital de alta qualidade, provenientes de sistemas de mamografia digital de campo total (FFDM), permitindo que o modelo aprendesse a identificar e caracterizar áreas suspeitas de malignidade.

O *Lunit INSIGHT MMG* tem demonstrado um desempenho clínico robusto e consistente em múltiplos estudos internacionais, evidenciando o seu potencial como ferramenta de apoio à decisão diagnóstica em mamografia digital. De acordo com um estudo de Kim *et al.* (2020), que analisou este sistema, foi apresentado um valor de AUC superior a 0,95, indicando uma precisão diagnóstica comparável ou superior à de radiologistas experientes. A utilização do algoritmo permitiu um aumento significativo da sensibilidade, especialmente em casos de mamas densas, em que a sobreposição tecidular tende a dificultar a identificação de lesões suspeitas.

Os resultados obtidos num outro estudo, de Dembrower *et al.* (2023), indicaram que a utilização de IA na leitura de mamografias, dentro do fluxo de trabalho avaliado, resultou num aumento da deteção de cancro da mama em comparação com o método tradicional, especialmente quando combinada com a leitura por radiologistas. Especificamente, houve um aumento de 4% na deteção de cancro através da estratégia de dupla leitura por um radiologista com o apoio da IA, que corresponde a 11 casos adicionais detetados numa amostra de 2 500 exames.

Além disso, o sistema demonstrou capacidade para reduzir substancialmente a carga de trabalho dos radiologistas, uma vez que consegue realizar a triagem automática de uma proporção considerável de exames — em alguns casos, até 60% das mamografias negativas — sem perda de deteções de casos malignos confirmados (Dembrower *et al.*, 2023). Este desempenho reflete não apenas a precisão do modelo, mas também a sua utilidade prática na melhoria da eficiência dos fluxos de trabalho clínicos.

Os resultados obtidos noutra estudo evidenciaram, também, que o desempenho do sistema *Lunit INSIGHT MMG* foi estatisticamente comparável ao de radiologistas experientes na deteção do cancro da mama por mamografia digital (Yilmaz *et al.*, 2025). De forma específica, o algoritmo foi capaz de identificar 82,75% (24 em 29) dos casos de cancro confirmados, enquanto a leitura por votação maioritária de radiologistas — método em que as avaliações de múltiplos radiologistas são combinadas para determinar a classificação final do exame —

alcançou uma taxa de detecção ligeiramente superior, de 89,65% (26 em 29). Relativamente aos valores de especificidade, o valor obtido pela avaliação inicial dos radiologistas foi de 92%, tendo aumentado para 99% com a reavaliação da IA.

O mesmo estudo evidenciou também valores relativos à taxa de reconvocações, que já não foram tão favoráveis, onde o sistema de IA apresentou um valor mais baixo, de 2,5%, quando comparado com o método tradicional, que apresentou uma taxa de 8,9% (Yilmaz *et al.*, 2025).

A análise da curva ROC (*Receiver Operating Characteristic*) revelou uma AUC de 94,4% para a IA, valor que não diferiu significativamente da AUC obtida pelos radiologistas, de 94,7%, valor que poderá ser justificado pela elevada experiência dos especialistas. No entanto é importante salientar que, após a reavaliação das imagens pelo algoritmo, a AUC da IA aumentou para 94,8%, diferença considerada estatisticamente significativa em relação à avaliação inicial. Este resultado demonstra uma elevada precisão e consistência do sistema na classificação de casos positivos, reforçando a capacidade do modelo em identificar padrões radiográficos associados à malignidade.

Os resultados de Brancato *et al.* (2025) demonstraram um aumento substancial da sensibilidade diagnóstica quando o sistema de IA foi utilizado como ferramenta de apoio à leitura radiológica, atingindo uma média de 84,5%, em comparação com a sensibilidade obtida pelos radiologistas na leitura isolada (69,2%). Este efeito foi particularmente expressivo entre os radiologistas com desempenho inicial mais reduzido, cuja sensibilidade aumentou de aproximadamente 56,7% para 79,7% com o suporte do sistema de IA.

Além disso, verificou-se que a combinação entre a leitura tradicional e a análise automática do algoritmo produziu resultados superiores aos obtidos pelo algoritmo de forma independente, com uma sensibilidade de 83,3% face aos 76,67% registados na leitura exclusiva pelo sistema de IA. Tal sinergia evidencia o carácter complementar da interação entre radiologista e IA reforçando o potencial da tecnologia como instrumento de apoio e não de substituição.

Por outro lado, a introdução do sistema implicou uma ligeira redução na especificidade, com um decréscimo de cerca de 0,6% na taxa de falsos positivos, o que reflete o inevitável equilíbrio entre o ganho de sensibilidade e o aumento marginal de falsos positivos. De forma global, os valores obtidos sustentam que o *Lunit INSIGHT MMG* contribui para uma maior precisão diagnóstica e homogeneidade de desempenho entre leitores.

Yoen *et al.* (2024) demonstrou que o *Lunit INSIGHT MMG* detetou corretamente 72% dos casos de cancro da mama, com melhor desempenho em tumores de maior dimensão, HER2 (biomarcador de cancro da mama) positivos e com achados mamográficos evidentes. No entanto, neste caso o sistema apresentou limitações na deteção de lesões pequenas, isodensas ou em mamas muito densas, o que resultou num aumento de falsos negativos. Conclui-se que, embora eficaz, o desempenho do algoritmo é influenciado por características tumorais e radiológicas, sendo essencial a interpretação conjunta com o radiologista para uma utilização clínica segura e otimizada.

Estes estudos mostram que o software *Lunit INSIGHT MMG* manteve resultados consistentes, demonstrando generalização e reprodutibilidade elevadas. Estes resultados sustentam a sua eficácia como ferramenta de apoio diagnóstico, capaz de otimizar a sensibilidade e a acuidade diagnóstica global, contribuindo para uma deteção mais precoce do cancro da mama e para a redução de erros de interpretação humana. O sistema demonstra um desempenho comparável ao de radiologistas experientes, especialmente útil em casos de mamas densas ou em profissionais com menor experiência. Contudo, a sua eficácia pode ser limitada na identificação de lesões pequenas ou de baixa densidade, motivo pelo qual deve ser encarado como um suporte complementar, e não como substituto da leitura humana.

Concluindo, o *Lunit INSIGHT MMG* representa uma ferramenta de apoio diagnóstico robusta e clinicamente validada, no entanto também é essencial compreender as limitações existentes em certos cenários, para promover uma maior eficiência e ainda melhores resultados.

5.3. Transpara

O *Transpara*, desenvolvido pela *ScreenPoint Medical*, constitui também um sistema de auxílio ao diagnóstico por imagem para a deteção do cancro da mama. Destinado a complementar a avaliação dos radiologistas, o sistema analisa mamografias digitais 2D e DBT, identificando padrões sugestivos de malignidade, incluindo massas e microcalcificações, e atribuindo uma pontuação de risco global. Com base na pontuação máxima de risco atribuída a cada exame, o *Transpara* gera uma pontuação global de 1 a 10, refletindo a probabilidade de malignidade. Pontuações entre 1 e 7 correspondem a exames de baixo risco, valores de 8 a 9 indicam risco intermédio, e a pontuação 10 representa alto risco para cancro da mama (Lyth *et al.*, 2025).

Assim como o sistema *ProFound IA*, o *Transpara* integra-se de forma fluída no fluxo de trabalho clínico, permitindo a comparação automática com exames anteriores, apoiando a deteção precoce de alterações mamográficas e contribuindo para a melhoria da precisão diagnóstica.

Assim como os sistemas mencionados anteriormente, o sistema *Transpara* foi desenvolvido com a utilização de técnicas avançadas de *deep learning*, aplicadas a um vasto conjunto de dados de mamografias digitais.

O treino deste modelo envolveu um extenso conjunto de imagens mamográficas, incluindo mais de 9 000 casos com diagnóstico de cancro da mama e aproximadamente 180 000 casos sem anomalias, provenientes de múltiplas regiões geográficas (Resch *et al.*, 2024). Esta diversidade de dados contribuiu para assegurar a robustez e a generalização do desempenho do sistema em diferentes populações clínicas. As imagens utilizadas foram cuidadosamente anotadas por especialistas, abrangendo casos de cancro da mama confirmado por biópsia, lesões benignas e tecidos mamários normais. O modelo foi treinado para identificar características específicas nas mamografias, como microcalcificações e massas, e para atribuir uma pontuação de risco associada a cada achado, facilitando a avaliação da probabilidade de malignidade.

O sistema *Transpara* é certificado com o marco CE, indicando conformidade com os requisitos da União Europeia para dispositivos médicos, e possui também aprovação da FDA. Estas certificações mostram que o *Transpara* cumpre os padrões de segurança e eficácia exigidos para dispositivos médicos na União Europeia e nos Estados Unidos da América.

Um estudo de van Winkel *et al.* (2025) evidenciou que a utilização do sistema *Transpara* contribui significativamente para a melhoria da deteção do cancro da mama em programas de rastreio populacional. Especificamente, o sistema demonstrou uma sensibilidade de aproximadamente 74%, valor comparável ao obtido na dupla leitura realizada por radiologistas e superior à leitura única, cuja sensibilidade situou-se entre 63% e 65%, respetivamente. Os valores de especificidade situaram-se acima de 97%, mas mostraram poder diminuir um pouco na combinação da leitura de um radiologista com AI, especialmente em mamografias de alta densidade. O sistema foi capaz de identificar um número considerável de casos de cancro inicialmente não reconhecidos pelos radiologistas, incluindo tumores clinicamente relevantes que posteriormente evoluíram para estados mais avançados. As taxas de reconvocação já não apresentaram valores tão favoráveis, tendo subido de 2,9% para as leituras humanas tradicionais para 5,0% com a introdução da IA.

O mesmo estudo relata também que, embora o *Transpara* seja capaz de identificar casos de cancro adicionais ao longo do tempo, o seu desempenho permanece inferior à dupla leitura realizada por radiologistas (método tradicional) na deteção inicial de tumores (van Winkel *et al.*, 2025). Adicionalmente, a IA não reconhece todos os cancros detetados pelos radiologistas, particularmente tumores menores ou localizados em mamas densas. O sistema também enfrenta restrições na análise de determinados tipos de mamografias, incluindo exames com implantes mamários ou de qualidade inferior. É importante mencionar que a disponibilidade limitada de dados detalhados sobre os tumores dificultou a análise mais aprofundada sobre a influência de características tumorais no desempenho do sistema.

Num outro estudo, de Resch *et al.* (2024), onde o sistema *Transpara* foi analisado, foi apresentada uma AUC de 0,86 na deteção do cancro da mama, evidenciando um desempenho satisfatório, embora ainda inferior ao observado na dupla leitura por radiologistas, que registou uma AUC de 0,98. No entanto, os dados incluídos no estudo indicam que o sistema é capaz de eliminar até 70% das leituras de pacientes considerados normais, sem comprometer a deteção de casos de cancro, o que se traduz numa redução da carga de trabalho dos radiologistas. Adicionalmente, o *Transpara* classifica as imagens em categorias de risco, sendo que a categoria de menor pontuação (1–7) corresponde, na maioria dos casos, a exames negativos, enquanto os casos de cancro detetados apresentam predominantemente pontuações mais elevadas, refletindo a capacidade do sistema de estratificar o risco de forma consistente.

Neste estudo a avaliação do sistema *Transpara* foi realizada em contexto retrospectivo e numa única instituição, o que restringe a generalização dos resultados a outras populações e cenários clínicos (Resch *et al.*, 2024). A dimensão relativamente reduzida da amostra constitui outra limitação, podendo afetar o poder estatístico das conclusões obtidas. Importa salientar que o sistema não incorpora informações prévias do paciente, como mamografias anteriores, o que pode restringir o contexto clínico na interpretação dos exames. Apesar desta limitação, o *Transpara* demonstrou eficácia na redução do índice de reconvocações em programas de triagem, particularmente em mamografias digitais, mantendo uma elevada sensibilidade e contribuindo para uma maior eficiência do fluxo de trabalho radiológico.

A capacidade demonstrada pelo sistema em identificar casos adicionais de cancro ao longo do seguimento do rastreio sugere que a IA pode auxiliar na deteção de tumores que, de outro modo, poderiam permanecer não identificados na leitura convencional (van Winkel *et al.*,

2025). Tal desempenho indica um potencial para reduzir a taxa de cancros inicialmente não detetados nos programas de rastreio, reforçando, mais uma vez, o seu potencial para otimizar o diagnóstico precoce e, conseqüentemente, melhorar o prognóstico clínico das pacientes.

Em conclusão, o sistema *Transpara* demonstra ser uma ferramenta promissora de apoio ao diagnóstico do cancro da mama, com capacidade de melhorar a sensibilidade da deteção, reduzir o índice de reconvocações e otimizar o fluxo de trabalho em programas de rastreio populacional. A sua utilização permite identificar casos que poderiam passar despercebidos na leitura convencional, complementando a avaliação dos radiologistas e potencialmente contribuindo para diagnósticos mais precoces. Contudo, o sistema ainda apresenta bastantes limitações, incluindo a ausência de histórico prévio do paciente e restrições em mamografias complexas ou de baixa qualidade, o que evidencia a necessidade de validações adicionais e aperfeiçoamentos contínuos. A escassa quantidade de estudos que avaliem as métricas específicas deste sistema é outro fator limitante. No entanto, no seu conjunto, o *Transpara* representa um avanço significativo no apoio à decisão clínica em mamografia.

5.4. Dados Recolhidos e Análise

De forma a resumir e simplificar os dados recolhidos relativos aos três sistemas analisados – *ProFound AI*, *Lunit INSIGHT MMG* e *Transpara* – foi criada uma tabela de análise geral dos mesmos. O objetivo deste capítulo é, então, apresentar, comparar e analisar os dados recolhidos. Todos os valores apresentados foram retirados dos artigos lidos, que se encontram mencionados.

A Tabela 4, apresenta as métricas mais relevantes para esta revisão, incluindo, sensibilidade, especificidade, AUC, tempos de leitura, taxa de reconvocação e taxa de deteção, com o propósito de fornecer uma visão abrangente e mais condensada das capacidades e limitações de cada sistema. Esta análise comparativa permite identificar padrões de eficácia, vantagens específicas e restrições inerentes a cada tecnologia, fornecendo entendimentos para a avaliação crítica da utilização clínica de ferramentas de CAD-AI em programas de rastreio e diagnóstico do cancro mamário.

De forma a facilitar a interpretação dos valores apresentados, são descritas abaixo as métricas mencionadas:

- Sensibilidade – Representa a capacidade de um teste ou sistema de deteção em identificar corretamente os casos positivos. Corresponde à proporção de verdadeiros positivos entre todos os indivíduos que efetivamente apresentam a doença, refletindo o desempenho na deteção de casos reais.
- Especificidade – Indica a capacidade do teste ou sistema em identificar corretamente os casos negativos, correspondendo à proporção de verdadeiros negativos entre todos os indivíduos sem a doença. Valores elevados de especificidade reduzem a ocorrência de falsos positivos.
- AUC (*Area Under the Curve – ROC*) – Quantifica a capacidade discriminativa global de um modelo, medindo a probabilidade de este atribuir uma pontuação mais alta a um caso positivo do que a um caso negativo. Valores próximos de 1 indicam excelente desempenho na distinção entre exames normais e anormais.
- Tempo de leitura – Como o nome indica, corresponde à duração média necessária para a interpretação de cada exame. Esta métrica avalia a eficiência do processo de leitura, permitindo comparar o impacto da utilização de sistemas de apoio à decisão no fluxo de trabalho clínico.
- Taxa de reconvocação (*recall*) – Representa a proporção de exames de rastreio em que é solicitada avaliação adicional, devido a achados médicos considerados suspeitos ou inconclusivos. Esta métrica reflete o equilíbrio entre deteção precoce e falsos positivos, sendo desejável mantê-la em níveis baixos sem comprometer a sensibilidade.
- Taxa de deteção – Representa a proporção de casos positivos corretamente identificados num total de exames analisados, correspondendo à capacidade do sistema em detetar efetivamente a presença da doença. Valores elevados de taxa de deteção indicam maior eficácia na identificação de casos de cancro.

Tabela 4 – Análise geral dos dados recolhidos

Métricas	ProFound AI	Lunit INSIGHT MMG	Transpara
Sensibilidade	77,0% (sem IA) vs. 85,0% (com IA) - Incremento de 8% (Conant <i>et al.</i> , 2019)	69,2% (sem IA) vs. 84,5% (com IA) - Incremento de 15,3% (Brancato <i>et al.</i> , 2025)	63% (sem IA) vs. 74% (com IA) - Incremento de 11% (van Winkel <i>et al.</i> , 2025)
Especificidade	62,7% (sem IA) vs. 69,6% (com IA) - Incremento de 6,9% (Conant <i>et al.</i> , 2019)	92% (sem IA) vs. 99% (com IA) - Incremento de 7% (Yilmaz <i>et al.</i> , 2025)	97,7% (sem IA) vs. 97,8% (com IA) - Incremento de 0,1% (van Winkel <i>et al.</i> , 2025)
AUC	0,795 (sem IA) vs. 0,852 (com IA) – Incremento de 0,057 (Conant <i>et al.</i> , 2019)	0,810 (sem IA) vs. 0,959 (com IA) - Incremento de 0,149 (Yilmaz <i>et al.</i> , 2025)	0,98 (sem IA) vs. 0,86 (com IA) – Redução de 0,12 (Resch <i>et al.</i> , 2024)
Tempo de leitura	Redução de 52,7% (Conant <i>et al.</i> , 2019)	Sem valores	Sem valores
Taxa de reconvocação	38,0% (sem IA) vs. 30,9% (com IA) - Redução de 7,2% (Conant <i>et al.</i> , 2019)	8,9% (sem IA) vs. 2,5% (com IA) - Redução de 6,4% (Yilmaz <i>et al.</i> , 2025)	2,9% (sem IA) vs. 5,0% (com IA) - Incremento de 2,1% (van Winkel <i>et al.</i> , 2025)
Taxa de deteção	Entre 97,5% e 99,4% (para diferentes limiares de sensibilidade) (Graewingholt and Duffy, 2021)	Incremento de 4% (Dembrower <i>et al.</i> , 2023) Redução de 6,9% (Yilmaz <i>et al.</i> , 2025)	Entre 96% e 97% (Guo <i>et al.</i> , 2025)

A análise das métricas de desempenho dos três sistemas CAD demonstra melhorias significativas em vários parâmetros quando comparados à leitura convencional, sem o apoio da

IA. De forma a facilitar a leitura e interpretação dos dados são apresentados três gráficos, um para cada sistema, onde são comparadas as métricas de sensibilidade, especificidade e AUC, com o método tradicional.

O *ProFound AI* apresentou um incremento de 8% na sensibilidade e de 6,9% na especificidade, além de um aumento na AUC de 0,057. O Gráfico 2 apresenta a comparação destas três métricas. As barras do lado esquerdo representam os valores referentes ao método tradicional e as do lado direito representam o sistema.

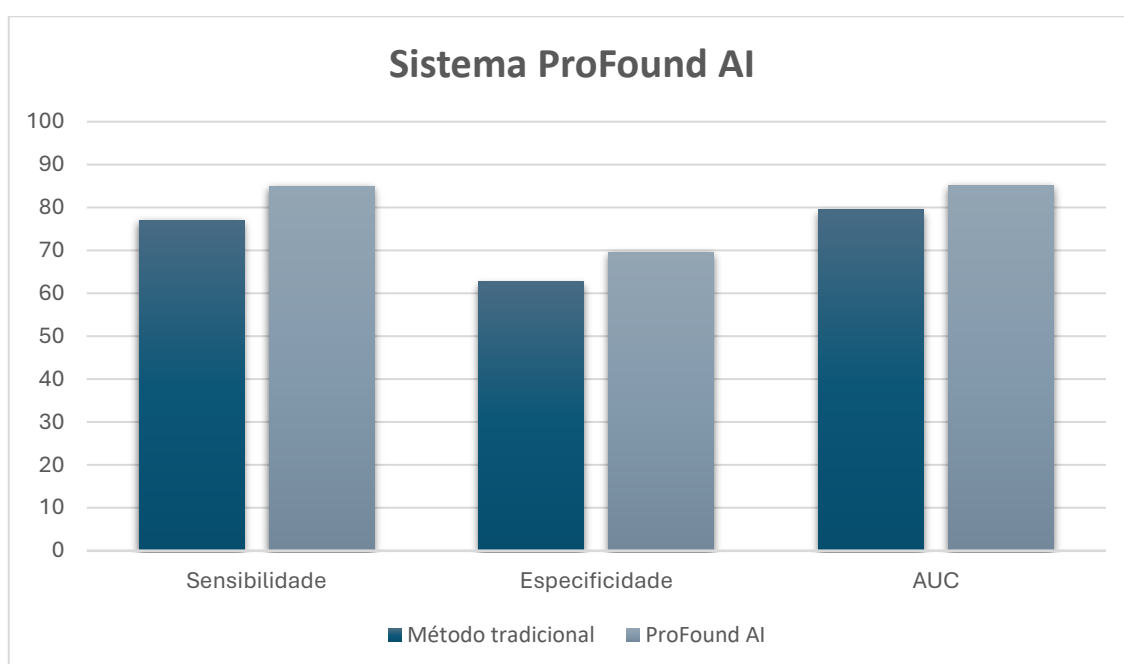


Gráfico 2 – Comparação das métricas do método tradicional com o sistema ProFound AI

O sistema *ProFound AI* destaca-se ainda pela redução do tempo de leitura em 52,7% e pela diminuição da taxa de reconvocação em 7,2%, sugerindo maior eficiência e precisão na interpretação mamográfica (Conant *et al.*, 2019).

O *Lunit INSIGHT MMG* demonstrou uma melhoria expressiva na sensibilidade (15,3%) e na especificidade (7%), e o maior ganho em AUC entre os três sistemas, tendo obtido um incremento de 0,149. O Gráfico 3 apresenta a comparação destas três métricas. As barras do lado esquerdo representam os valores referentes ao método tradicional e as do lado direito representam os valores obtidos pelo sistema.

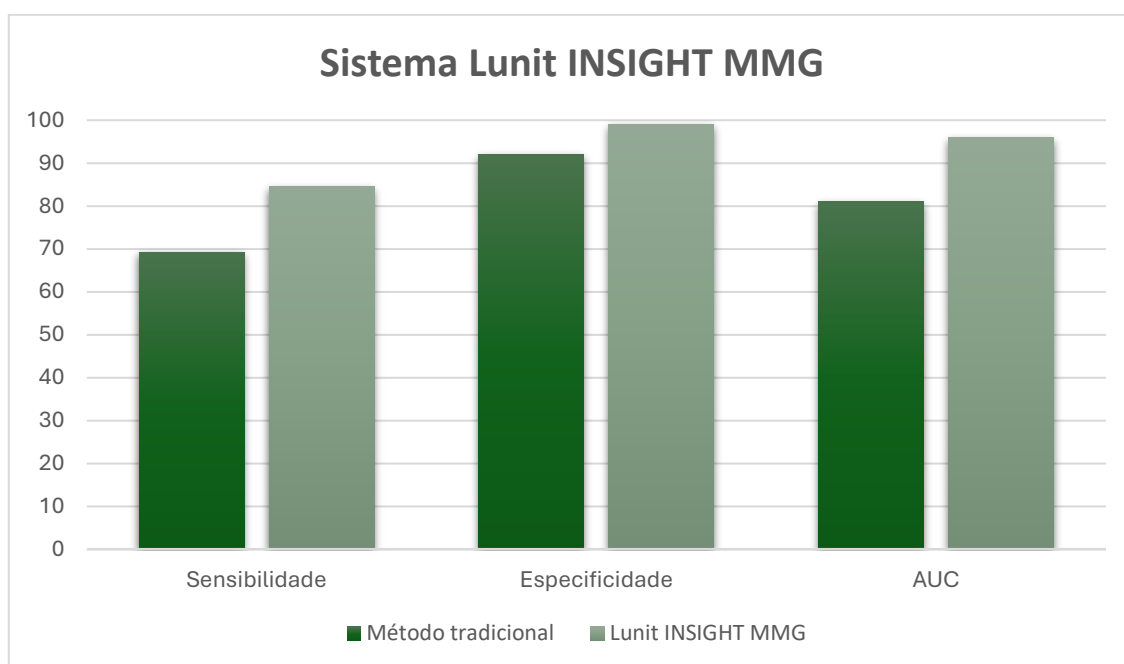


Gráfico 3 – Comparação das métricas do método tradicional com o sistema Lunit INSIGHT MMG

Apesar de não haver dados relativos ao tempo de leitura, observou-se uma redução de 6,4% na taxa de reconvoção, o que indica boa performance. Já a taxa de detecção apresentou variações, com incremento de 4% relatado num estudo (Dembrower *et al.*, 2023) e, uma redução de 6,9% descrita noutro (Yilmaz *et al.*, 2025). Como já mencionado esta discrepância pode dever-se a diferenças metodológicas entre os estudos, nomeadamente a população avaliada, o tipo de rastreio e os limiares de decisão aplicados pelo algoritmo para classificar exames como suspeitos. Além disso, a utilização de versões distintas do software, as bases de dados utilizadas, e a divergência na experiência dos radiologistas envolvidos podem igualmente ter influenciado os resultados.

Por fim, o *Transpara* registou um incremento de 11% na sensibilidade, e a especificidade manteve-se elevada também, entre 97,2% e 98,6%. Já a AUC sofreu uma ligeira redução, de 0,12, o que poderá sugerir que o algoritmo, no estudo em específico (Resch *et al.*, 2024), não otimizou a capacidade discriminativa do processo de leitura, reforçando a importância de ajustar e validar os modelos conforme o contexto clínico de aplicação. O Gráfico 4 apresenta a comparação das três métricas para este sistema. As barras do lado esquerdo representam os valores referentes ao método tradicional e as do lado direito representam os valores obtidos pelo sistema.

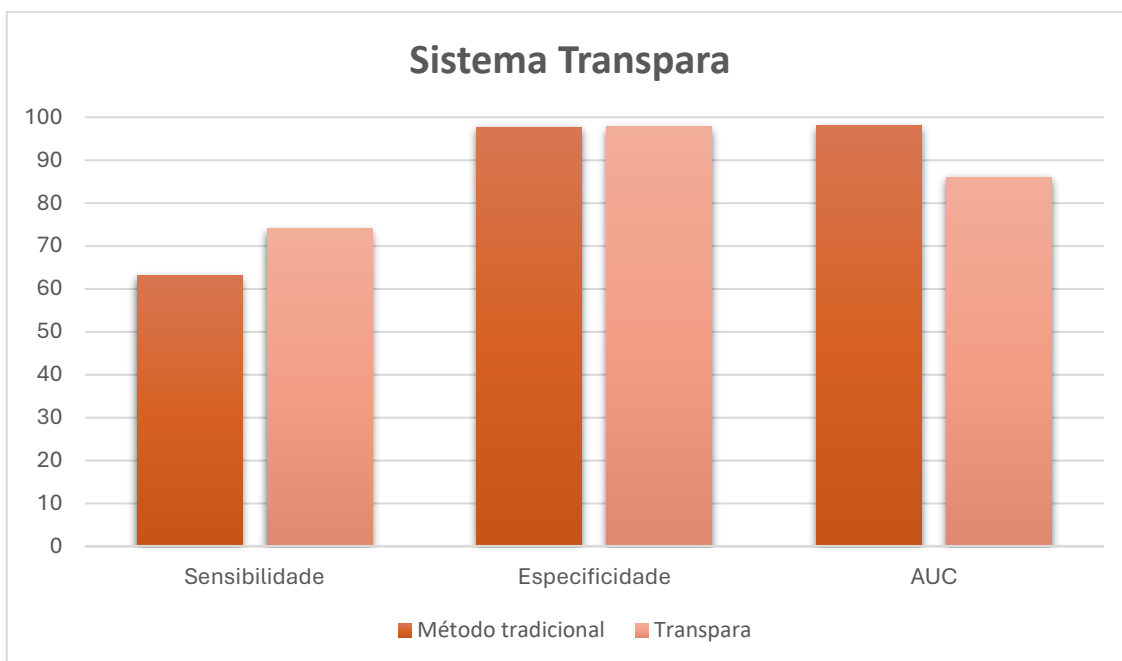


Gráfico 4 – Comparação das métricas do método tradicional com o sistema Transpara

A taxa de reconvocação apresentou um aumento de 2,1%, o que pode refletir maior sensibilidade a falsos positivos. A taxa de detecção também apresentou valores elevados, estando situada entre 96% e 97%.

Nestes dois últimos sistemas, *Lunit INSIGHT MMG* e *Transpara*, não foram apresentados valores relativos ao tempo de leitura. Nenhum dos artigos analisados referia especificamente ou com precisão estes valores, no entanto a maioria destacou uma melhoria geral nesta métrica.

No geral, os valores mais consistentes e mais elevados das métricas de sensibilidade e especificidade evidenciam uma melhoria significativa no desempenho diagnóstico com o apoio da inteligência artificial, demonstrando que os sistemas CAD aumentam simultaneamente a capacidade de detecção de lesões malignas e a precisão na exclusão de casos benignos.

Estes resultados são particularmente relevantes, uma vez que a sensibilidade reflete a eficácia da IA em identificar verdadeiros positivos, enquanto a especificidade indica a redução de falsos positivos, contribuindo para uma leitura mais fiável e consistente. Assim, a combinação de elevados valores de sensibilidade e especificidade indica um equilíbrio desejável entre segurança diagnóstica e eficiência clínica, razão pela qual estas métricas são consideradas mais representativas do verdadeiro impacto da IA no apoio à decisão radiológica.

Apesar da ligeira diferença, é de notar que a IA obteve melhores resultados relativamente à sensibilidade do que à especificidade, indicando que é mais precisa a identificar verdadeiros positivos do que a reduzir os falsos positivos.

Em contrapartida, os valores da AUC e da taxa de reconvocação tendem a apresentar maior instabilidade, podendo variar consoante a base de dados utilizada pelo sistema, o limiar de decisão aplicado, a experiência dos clínicos e o tipo de população avaliada. Pequenas alterações nestes fatores influenciam o equilíbrio entre sensibilidade e especificidade, o que explica oscilações na AUC, bem como flutuações na taxa de reconvocações, que dependem não apenas do desempenho do algoritmo, mas também da estratégia de leitura adotada e da tolerância ao risco de erro dos radiologistas.

De forma global, os resultados apontam para um impacto positivo da utilização de sistemas CAD-AI no apoio ao rastreio mamográfico, com ganhos em sensibilidade, AUC e eficiência no tempo de leitura, ainda que o equilíbrio entre sensibilidade e especificidade varie consoante o algoritmo utilizado. O *ProFound AI* destaca-se pela eficiência e redução do tempo de leitura, o *Lunit INSIGHT MMG* pelo maior ganho em sensibilidade e AUC, e o *Transpara* por manter elevada especificidade. No conjunto, os resultados confirmam o potencial dos sistemas CAD-AI como ferramentas de apoio valiosas na deteção precoce do cancro da mama, promovendo leituras mais rápidas e precisas, embora com variações na taxa de reconvocação entre os diferentes algoritmos.

6. Discussão

6.1. Considerações Gerais

Os resultados obtidos ao longo desta revisão permitem comprovar que a incorporação da inteligência artificial no diagnóstico do cancro da mama representa um avanço significativo na área da imagiologia mamária. De forma geral, os estudos analisados evidenciam melhorias expressivas nos principais indicadores de desempenho, nomeadamente na sensibilidade e especificidade, refletindo uma maior capacidade de deteção de lesões malignas e uma redução do número de falsos positivos. Estes resultados confirmam o potencial da IA como ferramenta de apoio à decisão clínica, capaz de aumentar a precisão diagnóstica e otimizar o fluxo de trabalho dos profissionais de saúde.

Contudo, apesar dos resultados promissores, a integração efetiva da IA na prática clínica continua a exigir uma abordagem prudente e sustentada. A variabilidade entre estudos, a ausência de padronização nos métodos de validação e a necessidade de garantir transparência e interpretabilidade dos algoritmos permanecem como desafios centrais. Embora os sistemas de IA tenham mostrado resultados melhores do que o desempenho coletivo dos radiologistas, ainda tendem a ter um desempenho inferior ao dos melhores radiologistas individuais. Deste modo, as evidências atuais devem ser interpretadas de forma crítica, reconhecendo que a IA, embora demonstre elevado potencial, deve ser entendida como uma ferramenta de apoio à decisão e não como substituto do julgamento clínico humano.

6.2. Implicações Clínicas e Científicas

Os resultados analisados neste estudo sugerem que a utilização de sistemas CAD-AI pode trazer benefícios clínicos significativos, nomeadamente ao melhorar a sensibilidade e especificidade na deteção de lesões mamárias. A aplicação destes algoritmos como suporte à decisão contribui para uma maior precisão diagnóstica, reduzindo a probabilidade de falsos negativos e falsos positivos, e permitindo uma triagem mais eficiente. Consequentemente, a integração da IA na prática clínica tem o potencial de otimizar o tempo de leitura dos radiologistas, reduzir o

número de reconvocações desnecessárias e facilitar o diagnóstico precoce do cancro da mama, com impacto direto na qualidade do cuidado prestado às pacientes.

Do ponto de vista científico, a análise evidencia que os avanços em *deep learning*, e nas técnicas de *transfer learning* e *data augmentation* estão a promover a evolução dos sistemas CAD-AI, tornando-os cada vez mais robustos e capazes de lidar com diferentes tipos de imagem e populações de pacientes. Desta forma, este estudo contribui para consolidar o conhecimento sobre a eficácia destas ferramentas, reforçando a necessidade de estudos futuros que utilizem métricas padronizadas e validação clínica rigorosa. Estes resultados ajudam a estabelecer uma base sólida para a investigação contínua, permitindo identificar tendências emergentes e áreas de desenvolvimento prioritárias na aplicação da IA no diagnóstico do cancro mamário.

Adicionalmente, os achados evidenciam a importância de considerar o papel complementar da IA na prática clínica. Embora os sistemas CAD-AI demonstrem elevada capacidade de apoio à decisão, os profissionais de saúde mantêm um papel central na interpretação dos resultados, na tomada de decisão clínica e na comunicação com as pacientes. Esta complementaridade reforça a necessidade de estratégias de implementação que integrem a IA de forma eficaz e segura, incluindo formação adequada dos radiologistas e adoção gradual de ferramentas que aumentem a confiança e a aceitação da tecnologia no contexto hospitalar.

6.3. Desafios e Barreiras Identificadas

Como já referido, apesar do potencial demonstrado pelos algoritmos e sistemas CAD e dos resultados promissores apresentados, existem desafios substanciais para a adoção generalizada da IA na prática clínica. Um dos principais obstáculos refere-se à qualidade e representatividade dos dados de treino. Bases de dados públicas, como o *Digital Database for Screening Mammography (DDSM)*, *INbreast* e *CBIS-DDSM*, têm sido amplamente utilizadas, mas nem sempre refletem a diversidade de populações ou a realidade clínica atual (Tan *et al.*, 2025). A necessidade de grandes volumes de imagens anotadas com precisão limita o desenvolvimento de modelos robustos, sendo que a variabilidade entre diferentes instituições e equipamentos pode afetar a generalização dos sistemas. Além disso, questões relacionadas com a privacidade e proteção de dados das pacientes impõem restrições adicionais à partilha e utilização de conjuntos de dados clínicos.

As questões éticas e legais, como a responsabilidade em caso de erro diagnóstico, a privacidade dos dados e o viés algorítmico, ainda carecem de regulamentação clara (Akinniyi *et al.*, 2025). Para maior clareza, o viés algorítmico ocorre quando erros sistemáticos nos algoritmos de IA produzem resultados injustos ou discriminatórios, que frequentemente refletem ou reforçam desigualdades socioeconômicas, raciais e de gênero já existentes.

Outro desafio importante é a transparência e explicabilidade dos algoritmos. Muitos sistemas comerciais funcionam como “caixas-pretas”, dificultando a compreensão do processo de decisão por parte dos profissionais de saúde. Esta falta de interpretabilidade pode gerar desconfiança, afetar a aceitação clínica e levantar questões éticas e legais sobre responsabilidade em caso de erro diagnóstico. A necessidade de integrar a IA de forma segura e confiável nos fluxos de trabalho hospitalares torna-se, portanto, um desafio crítico para a adoção efetiva destas tecnologias.

Um estudo de Arravalli *et al.* (2025) mostrou que a utilização de técnicas de Inteligência Artificial Explicável (XAI) fornece transparência e interpretabilidade aos modelos e algoritmos utilizados, ajudando os profissionais de saúde a compreenderem as razões para as decisões tomadas pelo sistema. Estas ferramentas têm a capacidade de identificar quais fatores, como idade, tamanho do tumor, nódulos envolvidos, etc., influenciaram cada diagnóstico, promovendo uma melhor compreensão do modo de funcionamento dos mesmos.

Um outro aspeto crítico prende-se na avaliação económica. Poucos estudos realizaram análises de custo-efetividade robustas, embora alguns exemplos, como a avaliação do *MammoScreen* no NHS (*National Health Service*) do Reino Unido, tenham sugerido a viabilidade com valores de ICER – que representa uma métrica de custo-efetividade na saúde – compatíveis com padrões de decisão em saúde (Uwimana, Gnecco and Riccaboni, 2025). A incorporação da IA em programas de rastreio e diagnóstico requer, assim, modelos da Avaliação de Tecnologias na Saúde (HTA) adaptados, que considerem não apenas métricas técnicas, mas também impacto organizacional, económico e social. A HTA representa conjuntos de processos sistemáticos e metodologias que avaliam as propriedades, efeitos e impactos de diferentes tecnologias na saúde, permitindo dar apoio na tomada de decisões relativas a estas tecnologias.

A aceitação da IA por parte dos pacientes constitui igualmente um fator determinante. Estudos qualitativos demonstram que as mulheres tendem a aceitar soluções baseadas em IA desde que

sejam acompanhadas por supervisão humana, transparência nos resultados e possibilidade de escolha (Gatting *et al.*, 2024), (Carter *et al.*, 2024). As mulheres entrevistadas também mostraram preocupações sobre a perda de envolvimento humano no processo de análise, com receio de que a tecnologia possa aumentar o risco de erros ou afetar negativamente os resultados, o que mostra a falta de confiança na IA. Também foram mencionados receios relativamente à segurança dos dados pessoais e privacidade das pacientes.

Num estudo realizado na Austrália por Woode *et al.* (2025), foi referida também uma predisposição geral das mulheres australianas a aceitar o uso da inteligência artificial em serviços de triagem do cancro da mama quando eram fornecidas informações adicionais sobre as vantagens da IA. Estes dados mostram que os pacientes dão grande valor ao maior conhecimento sobre a IA. As participantes também preferiram uma taxa superior de sensibilidade, que se traduz numa melhoria na taxa de casos perdidos. O tempo necessário para receber os resultados também foi mencionado, indicando uma grande preferência por resultados mais rápidos, outro fator em que IA pode ser útil.

Outra questão bastante relevante é a menor eficácia em casos atípicos ou complexos, como é o caso de lesões ou alterações subtis, lesões pouco representadas ou a presença de implantes mamários, como foi possível verificar na análise dos sistemas CAD escolhidos. Esta limitação pode ser justificada pela falta de dados representativos destes casos nas bases de dados utilizadas para o treino dos algoritmos.

Por fim, existem barreiras relacionadas com formação, adaptação e custos de implementação. A introdução de sistemas CAD-AI exige que os radiologistas e técnicos, e profissionais de saúde no geral, recebam formação específica para utilizar e interpretar corretamente os resultados fornecidos pela IA.

Adicionalmente, os custos associados à aquisição, manutenção e integração destes sistemas nos ambientes clínicos podem limitar a sua reprodução, sobretudo em contextos com recursos restritos. Superar estes desafios será essencial para garantir que a inteligência artificial cumpra o seu papel de apoio à decisão clínica de forma eficiente, segura e equitativa.

7. Conclusão

Em síntese, a presente revisão evidencia o impacto crescente da inteligência artificial no apoio ao diagnóstico do cancro da mama, demonstrando que os sistemas CAD-AI têm potencial para melhorar a precisão e eficiência da interpretação de imagens da mama. A análise dos estudos mais recentes mostra avanços significativos na sensibilidade e especificidade dos algoritmos, reforçando o papel da IA como ferramenta complementar à prática radiológica. A combinação de IA com a *expertise* humana é vista como uma forma de melhorar o diagnóstico precoce do cancro.

No entanto, a sua integração clínica requer uma abordagem cautelosa e multidisciplinar, outro fator muito importante, que assegure a fiabilidade, a explicabilidade e a validação contínua destes sistemas em contextos reais. Assim, conclui-se que, embora os resultados atuais sejam promissores, o verdadeiro valor da IA na prática médica dependerá da capacidade de equilibrar inovação tecnológica com rigor científico e ética clínica, garantindo que estas ferramentas sirvam o propósito último da medicina: melhorar o diagnóstico e o cuidado ao paciente.

7.1. Limitações do Estudo

O presente estudo apresenta algumas limitações que devem ser consideradas na interpretação dos resultados. Em primeiro lugar, o número reduzido de artigos incluídos na análise pode limitar a representatividade da amostra e, conseqüentemente, a generalização das conclusões obtidas. Apesar de terem sido definidos critérios específicos de inclusão e exclusão, a disponibilidade restrita de estudos que cumprissem tais requisitos, bem como a exclusão de publicações sem acesso ao texto integral, pode ter condicionado a abrangência da revisão.

Embora os estudos incluídos apresentem metodologias semelhantes e utilizem métricas comparáveis, verificou-se que nem todos disponibilizam de forma completa os indicadores de desempenho, como a sensibilidade, especificidade, AUC, tempo de leitura ou taxa de reconvoções. Esta limitação dificultou a comparação quantitativa rigorosa entre os diferentes sistemas analisados e impediu a obtenção de uma visão totalmente uniforme e precisa do desempenho dos mesmos. Além disso, a forma variável como alguns autores relataram os

resultados — por exemplo, em diferentes unidades ou formatos — representou um desafio adicional à análise consolidada dos dados.

Outra limitação relevante prende-se com a escassez de informação técnica disponibilizada sobre os sistemas de IA analisados. Por se tratarem, em grande parte, de softwares comerciais com algoritmos próprios, não foi possível obter detalhes sobre os conjuntos de dados utilizados no treino, as técnicas de aprendizagem aplicadas ou os processos de validação interna. Esta falta de transparência limita a compreensão da real robustez e generalização dos modelos, condicionando a análise crítica do seu desempenho, e impedindo também uma análise comparativa mais fidedigna.

Apesar deste trabalho ter sido direcionado para estudos recentes, publicados entre o último ano e o mês anterior à realização desta revisão, é importante reconhecer que a natureza altamente dinâmica do campo da inteligência artificial pode rapidamente alterar o panorama atual. Novos algoritmos, atualizações de software e metodologias de validação surgem com frequência, podendo modificar o desempenho e a fiabilidade dos sistemas CAD-AI disponíveis no mercado. Assim, ainda que os resultados aqui apresentados reflitam o estado da arte mais atualizado à data, futuras investigações deverão continuar a acompanhar a evolução tecnológica e clínica destes sistemas, de forma a assegurar a relevância e a precisão das conclusões.

Uma outra limitação de relevo prende-se com a ausência de uma validação empírica própria. Este estudo baseou-se exclusivamente na análise de dados secundários provenientes da literatura publicada, não incluindo a avaliação direta em ambiente clínico ou ensaios com dados originais. Como tal, não foi possível verificar de forma prática o desempenho dos sistemas selecionados nas condições reais de utilização. Esta limitação reforça a necessidade de estudos futuros que integrem validação clínica direta para complementar e corroborar as conclusões obtidas a partir da revisão bibliográfica.

7.2. Perspetivas Futuras e Recomendações

No futuro, será essencial promover estudos clínicos multicêntricos e de larga escala que validem empiricamente o desempenho dos sistemas CAD-AI em diferentes populações e contextos de prática clínica. A transparência dos algoritmos, a normalização das métricas de avaliação e a criação de bases de dados partilhadas e diversificadas constituem etapas fundamentais para

garantir a fiabilidade e a reprodutibilidade dos resultados. Paralelamente, a implementação segura da IA na prática médica exigirá formação contínua dos profissionais de saúde e o desenvolvimento de políticas éticas e regulamentares claras. Apenas através desta integração equilibrada entre ciência, tecnologia e responsabilidade clínica será possível consolidar o papel da inteligência artificial como aliada estratégica no diagnóstico precoce e na melhoria dos cuidados em saúde mamária, e da saúde no geral.

Uma das principais recomendações para pesquisas futuras prende-se com a necessidade de uma validação clínica direta dos algoritmos e sistemas CAD. Estudos que incluam ensaios em hospitais ou ambientes clínicos reais permitirão avaliar o desempenho destes algoritmos em diferentes populações e fluxos de trabalho, garantindo que os resultados obtidos em estudos anteriores se traduzem em benefícios práticos e seguros na rotina clínica diária. Este tipo de investigação é fundamental para consolidar a confiança dos profissionais de saúde e demonstrar o valor clínico da inteligência artificial no diagnóstico do cancro da mama.

Outro aspeto relevante consiste na padronização das métricas de avaliação utilizadas nos estudos que avaliem o desempenho dos algoritmos e sistemas CAD. A utilização consistente e similar de indicadores como sensibilidade, especificidade, AUC, tempo de leitura, taxa de reconvocação e taxa de deteção permitirá comparações mais rigorosas entre os diferentes sistemas e facilitará a realização de revisões sistemáticas e meta-análises. Esta uniformização contribuirá para uma análise mais robusta do impacto da IA na prática clínica e reduzirá a dificuldade associada à interpretação de resultados heterogéneos.

A transparência e o acesso a informações técnicas sobre os algoritmos representam outra área crítica. A inteligência artificial explicável mostra ser uma grande melhoria nesta questão, promovendo mais clareza e compreensão dos algoritmos usados. No que diz respeito à atividade científica é recomendado que os detalhes sobre os treinos, as bases de dados utilizadas, as técnicas de aprendizagem aplicadas e os processos de validação, mesmo que de forma resumida, sejam apresentados. Esta abertura permitirá uma avaliação crítica mais completa, aumentará a replicabilidade dos estudos e fomentará um maior desenvolvimento científico na área, promovendo soluções mais seguras e eficazes.

Relativamente aos utilizadores primários desta tecnologia, como é o caso dos profissionais de saúde, é igualmente importante fornecer mais informação e promover maior abertura sobre a

tomada de decisão destes sistemas, de forma a transmitir maior confiabilidade, e promover um trabalho combinado entre radiologistas e sistemas CAD-AI. A maior abertura e explicação do funcionamento destes sistemas será também benéfica para os pacientes, ao promover maior compreensão e conseqüentemente, maior aceitação na integração da IA na prática clínica.

A formação dos profissionais de saúde é outro ponto a ter em consideração, visto que será essencial para garantir uma interpretação correta dos resultados dos sistemas CAD, uma integração eficiente no fluxo clínico e uma utilização segura. Desta forma é possível maximizar os benefícios da tecnologia sem comprometer a segurança do paciente.

Os modelos multimodais representam outro fator bastante promissor e que poderá impulsionar resultados ainda mais precisos. Estes modelos aproximam o uso da IA da realidade clínica, porque refletem o modo como os médicos tomam decisões. No entanto a escassez de bases de dados multimodais e padronizadas, os requisitos computacionais elevados e a dificuldade de explicabilidade são desafios a ultrapassar. A superação destes desafios requer esforços conjuntos na criação de bases de dados padronizadas e interoperáveis, desenvolvimento de métodos explicáveis e eficientes, e na formação dos profissionais de saúde.

Por fim, a implementação da IA na prática clínica necessita de uma análise do impacto económico e operacional. Investigar métodos de integração no fluxo de trabalho, desenvolver programas de treino para radiologistas e técnicos, e avaliar custos *versus* benefícios, são passos essenciais para garantir a adoção eficiente desta tecnologia.

Resumindo, os resultados quantitativos sugerem que a IA no apoio ao diagnóstico do cancro da mama pode reduzir a carga de trabalho dos profissionais de saúde, aumentar a padronização e melhorar o diagnóstico precoce. No entanto, para que estes ganhos se traduzam em benefícios de saúde pública, é essencial avaliar o custo-efetividade em larga escala, a maior aceitação por profissionais de saúde e pacientes, bases de dados mais diversificadas e a integração em fluxos de trabalho reais.

No panorama atual, pode afirmar-se que a IA no apoio ao diagnóstico do cancro da mama se encontra numa fase de transição entre validações técnicas promissoras e os primeiros passos para integração clínica.

Referências

- Abdullah, K.A. *et al.* (2025) “Deep learning-based breast cancer diagnosis in breast MRI: systematic review and meta-analysis,” *European Radiology*. Springer Science and Business Media Deutschland GmbH, pp. 4474–4489. Available at: <https://doi.org/10.1007/s00330-025-11406-6>.
- Akinniyi, O. *et al.* (2025) “The role of AI for improved management of breast cancer: Enhanced diagnosis and health disparity mitigation,” *Computer Methods and Programs in Biomedicine*. Elsevier Ireland Ltd. Available at: <https://doi.org/10.1016/j.cmpb.2025.109036>.
- Arravalli, T. *et al.* (2025) “Detection of breast cancer using machine learning and explainable artificial intelligence,” *Scientific Reports*, 15(1). Available at: <https://doi.org/10.1038/s41598-025-12644-w>.
- Bahl, M. *et al.* (2025) “Traditional versus modern approaches to screening mammography: a comparison of computer-assisted detection for synthetic 2D mammography versus an artificial intelligence algorithm for digital breast tomosynthesis,” *Breast Cancer Research and Treatment*, 210(3), pp. 529–537. Available at: <https://doi.org/10.1007/s10549-024-07589-z>.
- Brancato, B. *et al.* (2025) “AI-supported approaches for mammography single and double reading: A controlled multireader study,” *European Journal of Radiology*, 187. Available at: <https://doi.org/10.1016/j.ejrad.2025.112101>.
- “Cancro da Mama _ Liga Portuguesa Contra o Cancro” (no date).
- Carter, S.M. *et al.* (2024) “Women’s views on using artificial intelligence in breast cancer screening: A review and qualitative study to guide breast screening services,” *Breast*. Churchill Livingstone. Available at: <https://doi.org/10.1016/j.breast.2024.103783>.
- Conant, E.F. *et al.* (2019) “Improving accuracy and efficiency with concurrent use of artificial intelligence for digital breast tomosynthesis,” *Radiology: Artificial Intelligence*, 1(4). Available at: <https://doi.org/10.1148/ryai.2019180096>.
- Datwani, S. *et al.* (2025) “Artificial intelligence in breast pathology: Overview and recent updates,” *Human Pathology* [Preprint]. W.B. Saunders. Available at: <https://doi.org/10.1016/j.humpath.2025.105819>.
- Dembrower, K. *et al.* (2023) “Artificial intelligence for breast cancer detection in screening mammography in Sweden: a prospective, population-based, paired-reader, non-inferiority study,” *The Lancet Digital Health*, 5(10), pp. e703–e711. Available at: [https://doi.org/10.1016/S2589-7500\(23\)00153-X](https://doi.org/10.1016/S2589-7500(23)00153-X).
- Eriksson, M. *et al.* (2023) “Long-Term Performance of an Image-Based Short-Term Risk Model for Breast Cancer,” *J Clin Oncol*, 41, pp. 2536–2545. Available at: <https://doi.org/10.1200/JCO.22>.
- Freitas, V. *et al.* (2025) “The Transformative Power of Digital Breast Tomosynthesis and Artificial Intelligence in Breast Cancer Diagnosis,” *Canadian Association of Radiologists Journal*.

SAGE Publications Inc., pp. 302–312. Available at:
<https://doi.org/10.1177/08465371241301957>.

Gattig, L. *et al.* (2024) “Acceptability of artificial intelligence in breast screening: focus groups with the screening-eligible population in England,” *BMJ Public Health*, 2(2), p. e000892. Available at: <https://doi.org/10.1136/bmjph-2024-000892>.

Goh, S. *et al.* (2025) “Challenges in Implementing Artificial Intelligence in Breast Cancer Screening Programs: Systematic Review and Framework for Safe Adoption,” *Journal of Medical Internet Research*. JMIR Publications Inc. Available at: <https://doi.org/10.2196/62941>.

Graewingholt, A. and Duffy, S. (2021) “Retrospective comparison between single reading plus an artificial intelligence algorithm and two-view digital tomosynthesis with double reading in breast screening,” *Journal of Medical Screening*. SAGE Publications Ltd, pp. 365–368. Available at: <https://doi.org/10.1177/0969141320984198>.

Guo, Y. *et al.* (2025) “Artificial intelligence-based automated breast ultrasound radiomics for breast tumor diagnosis and treatment: a narrative review,” *Frontiers in Oncology*. Frontiers Media SA. Available at: <https://doi.org/10.3389/fonc.2025.1578991>.

Hasan, A.H., Khalid, U.A.R. and Abid, M.A. (2025) “Leveraging Artificial Intelligence in Breast Cancer Screening and Diagnosis,” *Cureus* [Preprint]. Available at: <https://doi.org/10.7759/cureus.79177>.

Hashim, H.T. *et al.* (2025) “Artificial intelligence versus radiologists in detecting early-stage breast cancer from mammograms: a meta-analysis of paradigm shifts,” *Polish Journal of Radiology*. Termedia Publishing House Ltd., pp. e1–e8. Available at: <https://doi.org/10.5114/pjr/195520>.

Kim, H.E. *et al.* (2020) “Changes in cancer detection and false-positive recall in mammography using artificial intelligence: a retrospective, multireader study,” *The Lancet Digital Health*, 2(3), pp. e138–e148. Available at: [https://doi.org/10.1016/S2589-7500\(20\)30003-0](https://doi.org/10.1016/S2589-7500(20)30003-0).

Letter, H. *et al.* (2023) “Use of Artificial Intelligence for Digital Breast Tomosynthesis Screening: A Preliminary Real-world Experience,” *Journal of Breast Imaging*, 5(3), pp. 258–266. Available at: <https://doi.org/10.1093/jbi/wbad015>.

Lyth, J. *et al.* (2025) “Results from a Swedish model-based analysis of the cost-effectiveness of AI-assisted digital mammography,” *European Radiology* [Preprint]. Available at: <https://doi.org/10.1007/s00330-025-11821-9>.

Marinovich, M.L. *et al.* (2025) “Simulated arbitration of discordance between radiologists and artificial intelligence interpretation of breast cancer screening mammograms,” *Journal of Medical Screening*. SAGE Publications Ltd, pp. 48–52. Available at: <https://doi.org/10.1177/09691413241262960>.

Nepute, J.A. *et al.* (2025) “Improved Breast Cancer Detection with Artificial Intelligence in a Real-World Digital Breast Tomosynthesis Screening Program,” *Clinical Breast Cancer* [Preprint]. Available at: <https://doi.org/10.1016/j.clbc.2025.05.007>.

- Qu, C. *et al.* (2024) "Diagnostic Value of Artificial Intelligence in Minimal Breast Lesions Based on Real-Time Dynamic Ultrasound Imaging," *International Journal of General Medicine*, Volume 17, pp. 4061–4069. Available at: <https://doi.org/10.2147/ijgm.s479969>.
- Resch, D. *et al.* (2024) "AI-enhanced Mammography With Digital Breast Tomosynthesis for Breast Cancer Detection: Clinical Value and Comparison With Human Performance," *Radiology: Imaging Cancer*, 6(4). Available at: <https://doi.org/10.1148/rycan.230149>.
- Singh, Alok, Singh, Akanksha and Bhattacharya, S. (2024) "Research trends on AI in breast cancer diagnosis, and treatment over two decades," *Discover Oncology*. Springer Science and Business Media B.V. Available at: <https://doi.org/10.1007/s12672-024-01671-0>.
- Tan, H. *et al.* (2025) "Mammography-based artificial intelligence for breast cancer detection, diagnosis, and BI-RADS categorization using multi-view and multi-level convolutional neural networks," *Insights into Imaging*, 16(1). Available at: <https://doi.org/10.1186/s13244-025-01983-x>.
- Uwimana, A., Gnecco, G. and Riccaboni, M. (2025) "Artificial intelligence for breast cancer detection and its health technology assessment: A scoping review," *Computers in Biology and Medicine*. Elsevier Ltd. Available at: <https://doi.org/10.1016/j.compbiomed.2024.109391>.
- van Winkel, S.L. *et al.* (2025) "AI as an independent second reader in detection of clinically relevant breast cancers within a population-based screening programme in the Netherlands: a retrospective cohort study," *The Lancet Digital Health* [Preprint]. Available at: <https://doi.org/10.1016/j.landig.2025.100882>.
- Woode, M.E. *et al.* (2025) "Preferences for the Use of Artificial Intelligence for Breast Cancer Screening in Australia: A Discrete Choice Experiment," *Patient* [Preprint]. Available at: <https://doi.org/10.1007/s40271-025-00742-w>.
- Wu, X. *et al.* (2025) "Decoding breast cancer imaging trends: the role of AI and radiomics through bibliometric insights," *Breast Cancer Research*, 27(1). Available at: <https://doi.org/10.1186/s13058-025-01983-1>.
- Yilmaz, E. *et al.* (2025) "Clinical Application of AI in Mammography: Insights from a Prospective Study," *Academic Radiology* [Preprint]. Available at: <https://doi.org/10.1016/j.acra.2025.05.025>.
- Yoen, H. *et al.* (2024) "Artificial Intelligence for Breast Cancer Detection on Mammography: Factors Related to Cancer Detection," *Academic Radiology*, 31(6), pp. 2239–2247. Available at: <https://doi.org/10.1016/j.acra.2023.12.006>.
- Yu, S.M. *et al.* (2024) "Artificial intelligence–based computer-aided diagnosis for breast cancer detection on digital mammography in Hong Kong," *Hong Kong Medical Journal*, 30(6), pp. 468–477. Available at: <https://doi.org/10.12809/hkmj2310920>.