

INTELIGÊNCIA ARTIFICIAL NA DETECÇÃO DE FRATURAS EM RADIOGRAFIAS: DESEMPENHO DIAGNÓSTICO E POTENCIAL COMO FERRAMENTA DE AUXÍLIO À DECISÃO CLÍNICA

ARTIFICIAL INTELLIGENCE IN FRACTURE DETECTION ON RADIOGRAPHS: DIAGNOSTIC PERFORMANCE AND POTENTIAL AS A CLINICAL DECISION-SUPPORT TOOL

Letícia Santos de Almeida¹; Rubens José da Silva Júnior¹; Vandecia Amanda Silva de Lima²; Edvaldo Barreto Cardoso Junior¹.

1: Curso de Medicina, Unima/Afy – Maceió/AL, Brasil; 2: Fisioterapeuta; Docente do Curso de Medicina, Unima/Afy – Maceió/AL, Brasil.

edivaldojr0514@gmail.com

Eixo Temático: Saúde.

RESUMO

Introdução: A radiografia convencional permanece como um dos principais métodos de investigação inicial das fraturas, porém sua interpretação pode ser dificultada por lesões sutis, sobreposição anatômica e elevada demanda nos serviços de emergência. Nesse contexto, sistemas de inteligência artificial (IA), especialmente aqueles baseados em aprendizado profundo, têm sido desenvolvidos para auxiliar a detecção automatizada de fraturas.

Objetivo: Analisar o desempenho diagnóstico da IA na detecção de fraturas em radiografias e avaliar seu potencial como ferramenta auxiliar à interpretação profissional. **Método:** Foi realizada revisão integrativa da literatura, orientada pela pergunta de pesquisa: “Qual é o desempenho diagnóstico da inteligência artificial na detecção de fraturas em radiografias e qual o impacto de sua utilização como ferramenta auxiliar sobre o desempenho dos profissionais?”. Foram analisados estudos de acurácia diagnóstica, revisões sistemáticas e meta-análises envolvendo sistemas de IA aplicados a radiografias convencionais, considerando principalmente sensibilidade, especificidade e área sob a curva ROC. **Resultados:** Evidências sintetizadas demonstraram elevado desempenho dos algoritmos. Uma meta-análise envolvendo 100 estudos identificou sensibilidade agrupada de 91,43%, especificidade de 92,12% e área sob a curva ROC de 0,968. Em outra meta-análise, envolvendo 26 estudos, a utilização da IA como ferramenta auxiliar aumentou a sensibilidade dos profissionais de 77% para 87% e a especificidade de 88% para 92%, com aumento da área sob a curva de 0,90 para 0,95. Apesar dos resultados favoráveis, foram observados heterogeneidade entre estudos, diferenças relacionadas à localização anatômica, risco de viés e limitações de validação externa. **Conclusão:** A IA apresenta elevado potencial para auxiliar a detecção de fraturas em radiografias e pode melhorar o desempenho diagnóstico dos profissionais, especialmente como ferramenta de segunda leitura ou suporte à decisão. Entretanto, as evidências atuais favorecem sua utilização complementar, não sendo suficientes para sustentar a substituição do julgamento clínico.

Palavras-chave: Inteligência Artificial; Fraturas Ósseas; Radiografia; Diagnóstico por Imagem; Aprendizado Profundo.

ABSTRACT:

Introduction: Conventional radiography remains one of the main methods for the initial investigation of fractures; however, its interpretation can be challenging due to subtle lesions, anatomical overlap, and the high workload in emergency services. In this context, artificial intelligence (AI) systems, particularly those based on deep learning, have been developed to assist in the automated detection of fractures. **Objective:** To analyze the diagnostic performance of AI in fracture detection on radiographs and evaluate its potential as an adjunctive tool for professional interpretation. **Methods:** An integrative literature review was conducted based on the following research question: “What is the diagnostic performance of artificial intelligence in detecting fractures on radiographs, and what is the impact of its use as an assistive tool on professionals’ performance?” Diagnostic accuracy studies, systematic reviews, and meta-analyses involving AI systems applied to conventional radiographs were analyzed, primarily considering sensitivity, specificity, and area under the ROC curve. **Results:** The synthesized evidence demonstrated high algorithmic performance. A meta-analysis involving 100 studies identified a pooled sensitivity of 91.43%, specificity of 92.12%, and an area under the ROC curve of 0.968. In another meta-analysis involving 26 studies, the use of AI as an assistive tool increased professionals’ sensitivity from 77% to 87% and specificity from 88% to 92%, with an increase in the area under the curve from 0.90 to 0.95. Despite these favorable results, heterogeneity among studies, differences related to anatomical location, risk of bias, and limitations in external validation were observed. **Conclusion:** AI demonstrates considerable potential to assist in fracture detection on radiographs and may improve the diagnostic performance of professionals, particularly as a second-reader or decision-support tool. However, current evidence supports its complementary use, as it is not yet sufficient to justify replacing clinical judgment.

Keywords: Artificial Intelligence; Bone Fractures; Radiography; Diagnostic Imaging; Deep Learning.

INTRODUÇÃO

A radiografia convencional permanece como método fundamental na investigação inicial de pacientes com suspeita de fratura, devido à sua ampla disponibilidade, rapidez e menor custo. Entretanto, a interpretação pode ser desafiadora diante de lesões sutis, sobreposição anatômica e elevada demanda assistencial.

A incorporação de técnicas de aprendizado profundo possibilitou o desenvolvimento de algoritmos capazes de identificar padrões associados às fraturas, podendo atuar na triagem, como segunda leitura ou como suporte à interpretação profissional.

Evidências recentes demonstram um elevado desempenho diagnóstico¹. Uma revisão sistemática e meta-análise com 100 estudos encontrou sensibilidade agrupada de 91,43%, especificidade de 92,12% e área sob a curva ROC de 0,968, embora tenham sido observadas diferenças conforme a localização da fratura e a arquitetura dos algoritmos.¹ Entretanto, uma

revisão com 42 estudos identificou resultados semelhantes entre IA e profissionais, mas apontou alto risco de viés em 22 estudos (52%), destacando limitações metodológicas.²

Ademais, recentemente, o foco tem se voltado à IA como ferramenta de colaboração. Meta-análise demonstrou aumento da sensibilidade dos profissionais de 77% para 87% e da especificidade de 88% para 92% com o auxílio da IA.³ Dessa forma, torna-se relevante sintetizar as evidências sobre seu desempenho diagnóstico e discutir seu potencial como ferramenta complementar à prática clínica.

MÉTODO

Trata-se de uma revisão integrativa da literatura, elaborada a partir da seguinte pergunta norteadora: “Qual é o desempenho diagnóstico da inteligência artificial na detecção de fraturas em radiografias e qual o impacto de sua utilização como ferramenta auxiliar sobre o desempenho dos profissionais?”

A pergunta de pesquisa foi estruturada segundo adaptação do modelo PICO, sendo: P (população) — pacientes submetidos a radiografias com suspeita de fratura; I (intervenção) — sistemas de inteligência artificial aplicados à detecção de fraturas; C (comparação) — interpretação humana sem auxílio de IA ou comparação direta entre desempenho humano e algorítmico; O (desfechos) — sensibilidade, especificidade, acurácia, área sob a curva ROC (AUC) e desempenho dos profissionais com assistência de IA.

A busca bibliográfica foi realizada nas bases de dados PubMed/MEDLINE, Scopus e Google Scholar, utilizando a combinação dos descritores (DeCS/MeSH) “Artificial Intelligence”, “Deep Learning”, “Fractures, Bone”, “Radiography” e “Diagnostic Imaging”, associados pelos operadores booleanos AND e OR, sem restrição de idioma, priorizando publicações dos últimos cinco anos. Foram incluídos estudos de acurácia diagnóstica, revisões sistemáticas e meta-análises que avaliaram algoritmos de inteligência artificial na detecção de fraturas em radiografias convencionais, com apresentação de medidas quantitativas de desempenho (sensibilidade, especificidade e/ou AUC). Foram excluídos estudos que avaliaram exclusivamente outras modalidades de imagem (tomografia computadorizada, ressonância magnética), cartas ao editor, resumos de congresso, estudos sem dados quantitativos extraíveis e publicações duplicadas.

A seleção dos estudos ocorreu em duas etapas: leitura de títulos e resumos, seguida de leitura na íntegra dos artigos pré-selecionados, realizada de forma independente pelos autores,

com resolução de discordâncias por consenso. Foram identificados 347 registros, dos quais 5 foram incluídos após aplicação dos critérios de elegibilidade.

A análise dos estudos selecionados considerou o desempenho diagnóstico dos algoritmos, sua comparação com profissionais, o impacto da assistência por IA sobre o desempenho humano e as principais limitações metodológicas relatadas para aplicação clínica.

Declaração sobre o uso ético de inteligência artificial

A inteligência artificial generativa foi utilizada apenas como auxílio na organização textual e revisão linguística. A seleção das referências, interpretação das evidências e conteúdo científico foram realizados e conferidos pelos autores. Não foram utilizadas ferramentas de IA para criar ou simular dados, resultados, participantes ou referências inexistentes.

RESULTADOS

As evidências sintetizadas demonstram desempenho diagnóstico elevado dos sistemas de IA na identificação de fraturas radiográficas.

A maior síntese identificada, publicada em 2024, incluiu 100 estudos. Entre os estudos que apresentaram dados suficientes para análise quantitativa, a sensibilidade agrupada foi de 91,43%, a especificidade de 92,12% e a área sob a curva ROC sumarizada de 0,968.¹

Quadro 1: Principais resultados quantitativos da literatura sobre IA na detecção de fraturas

Evidência	Nº de estudos	Sensibilidade	Especificidade	AUC
Nowroozi et al. ¹	100	91,43%	92,12%	0,968
Kuo et al. ² – IA, validação externa	42*	91%	91%	-
Qin et al. ³ – profissionais sem IA	26	77%	88%	0,90
Qin et al. ³ – profissionais com IA	26	87%	92%	0,95

Fonte: Elaborado pelos autores (2026), com base nas referências 1-3 desta revisão.

*O estudo de Kuo et al. incluiu 42 estudos, sendo 37 relacionados à radiografia e 5 à tomografia computadorizada.

A comparação entre IA e profissionais demonstra resultados próximos. No que se refere à validação externa, a IA apresentou sensibilidade de 91% e especificidade de 91%, enquanto os profissionais apresentaram 94% para ambos os parâmetros. Além disso, não foram identificadas diferenças estatisticamente significativas no desempenho global.²

Entretanto, a utilização da IA como ferramenta complementar apresentou resultados particularmente relevantes. Em meta-análise de 26 estudos, a sensibilidade dos profissionais

aumentou de 77% para 87% e a especificidade de 88% para 92% após a introdução da assistência algorítmica, com aumento da AUC de 0,90 para 0,95.³

Quadro 2 — Efeito da assistência por inteligência artificial na interpretação profissional

Desfecho	Sem IA	Com IA	Diferença
Sensibilidade	77%	87%	+ 10 p.p.
Especificidade	88%	92%	+ 4 p.p.
AUC	0,90	0,95	+ 0,05

Fonte: Elaborado pelos autores (2026), com base nas referências 1-3 desta revisão.

A análise por regiões anatômicas demonstra desempenho variável da IA. Em fraturas do membro superior, foi observada sensibilidade de 97,65% e especificidade de 91,38%, apesar da elevada heterogeneidade entre os estudos.⁴ Para fraturas do colo femoral, uma meta-análise de 11 estudos encontrou sensibilidade de 87%, especificidade de 91% e acurácia de 91%, também com heterogeneidade elevada.⁵

Assim, embora os algoritmos apresentem resultados promissores, seu desempenho pode variar conforme a região anatômica, o modelo utilizado, a população estudada e as características das imagens analisadas.

DISCUSSÃO

Observa-se que a IA apresenta elevado potencial na detecção de fraturas, com sensibilidade e especificidade superiores a 90%.¹ Entretanto, seu principal benefício parece estar no auxílio ao profissional, e não na substituição de sua atuação.

Nesse sentido, a elevação da sensibilidade de 77% para 87% com o auxílio da IA sugere contribuição especialmente importante na redução de resultados falso-negativos.³ Essa aplicação pode ser particularmente relevante em serviços de emergência, caracterizados por elevado volume de exames, necessidade de decisões rápidas e disponibilidade variável de especialistas.

Apesar dos resultados favoráveis, há limitações metodológicas, incluindo alto risco de viés em 52% dos estudos² e elevada heterogeneidade.⁴ Além disso, bons resultados retrospectivos não garantem eficácia clínica em diferentes populações e cenários.

Assim, as evidências favorecem um modelo de colaboração homem-máquina, no qual a IA atua como segunda leitura, mecanismo de triagem ou alerta para exames suspeitos, enquanto o profissional permanece responsável pela integração dos achados de imagem com a história clínica e o exame físico. Sua implementação deve incluir validação local, monitoramento de

desempenho e definição clara das responsabilidades, visando ampliar a capacidade diagnóstica sem substituir o raciocínio clínico da equipe profissional.

CONCLUSÃO:

A inteligência artificial apresenta elevado desempenho na detecção de fraturas em radiografias, com resultados agrupados superiores a 90% para sensibilidade e especificidade em grandes sínteses da literatura. É notório que sua utilização como ferramenta auxiliar também está associada à melhora do desempenho dos profissionais, especialmente da sensibilidade diagnóstica.

Entretanto, a heterogeneidade entre estudos, o risco de viés e as limitações de validação externa impedem que esses resultados sejam interpretados como evidência para substituição da avaliação profissional.

A aplicação mais promissora da IA atualmente demonstra-se como uma ferramenta complementar, atuando na triagem e na segunda leitura das radiografias. Estudos prospectivos e multicêntricos ainda são necessários para determinar seu impacto sobre diagnósticos perdidos, fluxo de trabalho e desfechos clínicos.

REFERÊNCIAS BIBLIOGRÁFICAS:

1. Nowroozi A, Salehi MA, Shobeiri P, Agahi S, Momtazmanesh S, Kaviani P, et al. Artificial intelligence diagnostic accuracy in fracture detection from plain radiographs and comparing it with clinicians: a systematic review and meta-analysis. *Clin Radiol*. 2024;79(8):579-588. doi:10.1016/j.crad.2024.04.009.
2. Kuo RYL, Harrison C, Curran TA, Jones B, Freethy A, Cussons D, et al. Artificial intelligence in fracture detection: a systematic review and meta-analysis. *Radiology*. 2022;304(1):50-62. doi:10.1148/radiol.211785.
3. Qin H, Ding Y, Ju J, Qu Z, Peng L. Enhanced fracture detection on radiographs with AI assistance for clinicians: a systematic review and meta-analysis. *Wien Med Wochenschr*. 2026;176(1):2610079. doi:10.1080/07853890.2025.2610079.
4. Mellon M, Dworsky-Fried J, Rathod P, Shah D, et al. Artificial intelligence in the interpretation of upper extremity trauma radiographs: a systematic review and meta-analysis. *JSES Rev Rep Tech*. 2025;5(3):435-443.
5. Ramadanov N, Lettner J, Hable R, Hakam HT, Prill R, Dimitrov D, et al. Artificial intelligence-guided assessment of femoral neck fractures in radiographs: a systematic review and multilevel meta-analysis. *Orthop Surg*. 2025;17(5):1277-1286. doi:10.1111/os.14250.