

ANÁLISIS DE LA EFICIENCIA DE LA INTELIGENCIA ARTIFICIAL PARA LA DETECCIÓN DE FRACTURAS EN RADIOGRAFÍAS SIMPLES, REVISIÓN DOCUMENTAL 2025

Gilberto, Iglesias¹, Carlos, Tuñón¹, Javier, Vargas¹, Johana Gutiérrez Zehr²

1. Universidad Santander. Licenciatura en radiología e imágenes diagnósticas
2. Universidad de Santander Johana.gutierrez@udes.edu.co. <https://orcid.org/0000-0001-8082-1471>

Resumen

La inteligencia artificial (IA) está en auge en el diagnóstico de imágenes médicas y ha traído esperanzas especialmente altas en la detección de fracturas óseas utilizando radiografías simples. El objetivo de este estudio es examinar la efectividad de la IA en este contexto a través de una revisión documental de artículos publicados entre 2016 y 2024. A pesar del progreso en la tecnología, algunos problemas técnicos, éticos y legales aún deben abordarse, lo que todavía impide su implementación clínica, incluyendo la generalización del modelo en diferentes poblaciones, la variación en la calidad de las imágenes, la transparencia del algoritmo y la responsabilidad médica en caso de errores de diagnóstico. El método adoptado es el desarrollo de una revisión bibliográfica sistemática según las directrices PRISMA en bases de datos académicas (PubMed, Scopus, SciELO, Google Scholar). Proporcionaremos un análisis de parámetros críticos, es decir, sensibilidad, especificidad, pros y contras de diferente software basados en redes neuronales convolucionales.

Los resultados anticipados aclararán la precisión de estas herramientas, la utilidad clínica y el uso potencial como ayudas en entornos con recursos limitados o sobrecargados. Esta investigación conducirá a una mejor apreciación del papel complementario de la IA en el diagnóstico radiológico, así como a una guía para la integración responsable y efectiva de la IA en la práctica clínica.

Palabras clave:

- IA (inteligencia artificial), Sensibilidad, Especificidad, Ventajas de la IA, Desventajas de la IA.

Abstract

Artificial intelligence (AI) is booming in medical imaging diagnostics and has brought particularly high hopes for the detection of bone fractures using simple X-rays. The aim of this study is to examine the effectiveness of AI in this context through a literature review of articles published between 2016 and 2024. Despite advances in technology, some technical, ethical, and legal issues still need to be addressed, which continue to hinder its clinical implementation, including the generalization of the model across different populations, variation in image quality, algorithm transparency, and medical liability in the event of diagnostic errors. The method adopted is the development of a systematic literature review according to PRISMA guidelines in academic databases (PubMed,

Scopus, SciELO, Google Scholar). We will provide an analysis of critical parameters, i.e., sensitivity, specificity, pros, and cons of different software based on convolutional neural networks.

The anticipated results will clarify the accuracy of these tools, their clinical utility, and their potential use as aids in resource-limited or overburdened settings. This research will lead to a better appreciation of the complementary role of AI in radiological diagnosis, as well as guidance for the responsible and effective integration of AI into clinical practice.

Keywords:

- AI (artificial intelligence), Sensitivity, Specificity, Advantages of AI, Disadvantages of AI.

Introducción

En los últimos años, la inteligencia artificial (IA) ha adquirido una nueva relevancia en el sector de la salud, particularmente en el campo de la imagenología diagnóstica. El ejemplo más emocionante de estos sistemas es la detección automática de fracturas óseas en radiografías simples con aprendizaje profundo (especialmente redes neuronales convolucionales, CNN). Cuando se implementa en el entorno clínico, esta tecnología podría optimizar el tiempo de diagnóstico, así como disminuir los errores humanos, mejorar la precisión del diagnóstico y ofrecer asistencia en entornos donde los recursos radiológicos especializados son limitados.

Se llevó a cabo esta investigación, para observar de cerca la precisión diagnóstica de las tecnologías de IA cuando se aplica como herramienta de diagnóstico complementario para identificar fracturas óseas en la práctica clínica. Para ello, se realizó una revisión documental, abordando desde artículos de carácter clínico hasta estudios regulatorios, contemplando enfoques médicos, técnicos, éticos y legales. Esta revisión proporcionó una evaluación integral de la sensibilidad, especificidad, fortalezas y limitaciones de varios modelos de IA en este escenario. Este estudio tuvo como objetivo proporcionar un resumen exhaustivo y actualizado que pueda ser utilizado como referencia para estudios en curso y futuros, la implementación de aplicaciones en hospitales y el establecimiento de marcos regulatorios para el uso ético de la tecnología para un diagnóstico seguro y efectivo de imágenes utilizando inteligencia artificial.

Método. Se llevó a cabo una revisión documental, donde se destaca la recolección de datos a partir de referentes teóricos para la identificar el uso de la IA, las medidas de sensibilidad, especificidad, ventajas y limitaciones, bajo metodología Prisma, junto al instrumento de matriz Bibliográfica, cumpliendo los aspectos éticos y las buenas prácticas en investigación. Fuentes: Bases de datos como son Google Académico, Science Direct y Scielo, Pubmed de las cuales se obtendrán artículos y documentos disponibles en línea o en Internet y accesibles a través de páginas de la Universidad Santander.

Palabras claves para consensuar los estudios: Inteligencia artificial, Sensibilidad, Especificidad

Ventajas de la IA, Limitaciones de la IA, Radiografía Simple

Resultados

La información obtenida de la revisión de 30 artículos científicos incluidos en el Análisis de la eficiencia de la inteligencia artificial para la detección de fracturas en radiografías simples desde 2017 hasta 2024. Los datos fueron recopilados y analizados en función del año de publicación, país de origen, campo de estudio, tipo de estudio, metodología, resultados y limitaciones. Manteniéndonos en un período estudiado e identifica una tendencia de aumento continuo de la producción científica sobre IA en el diagnóstico radiológico de fracturas. Más del 70% de las publicaciones se concentran entre 2020 y 2024, lo que corresponde a la proliferación de soluciones de aprendizaje profundo y la liberación de bases de datos abiertas desde su popularización, estas herramientas han dado un crecimiento constante en el ámbito de la medicina.

Algunos antecedentes destacan que se ha demostrado que el modelo CheXNet logró superar en algunas métricas a radiólogos humanos en la detección de patologías torácicas, de otra parte, el Estudio de Kim et al (1) mostraron que la IA aumenta la eficiencia diagnóstica, pero también puede generar falsos positivos si no se integra con juicio clínico. Igualmente el Informe de la FDA del año 2022 indican la aprobación de software de IA para diagnóstico médico, indicando que aún son “adjuntos al diagnóstico clínico” y no reemplazos del juicio humano y por último se encontraron normativas de la OMS sobre IA en salud en el año 2021, donde subrayan la necesidad de validación clínica antes de implementaciones masivas

En la investigación realizada por Duron et al en el año 2021, se incluyeron retrospectivamente 600 pacientes adultos en los que se habían obtenido radiografías multivista después de un trauma reciente, con o sin una o más fracturas de hombro, brazo, mano, pelvis, pierna y pie, de 17 centros médicos franceses, donde se destaca una metodología sólida basada en la comparación entre diagnósticos humanos con y sin asistencia de IA, abarcando múltiples tipos de fracturas en diferentes regiones del cuerpo. Al involucrar tanto radiólogos como médicos de urgencias, el estudio refleja un enfoque realista y práctico. Además, la inclusión de métricas como sensibilidad, especificidad y tiempo de lectura permite medir no solo la precisión diagnóstica, sino también el impacto operativo del uso de la IA en contextos clínicos reales (2)

La inteligencia artificial, comparable al rendimiento médico” realizada por Médica (2022), destaca que la IA funcionaba con un alto grado de precisión, comparable al rendimiento de los médicos”, añadiendo que “Los resultados del estudio apuntan a varias aplicaciones educativas y clínicas prometedoras para la IA en la detección de fracturas”. Además, señala Kuo que “podría reducir la tasa de diagnósticos erróneos tempranos en circunstancias difíciles en el ámbito de las urgencias, incluidos los casos en los que los pacientes pueden sufrir múltiples fracturas. Tiene potencial como herramienta educativa para los médicos noveles”. (3)

Martin (2020) en la investigación titulada Prioridad a las radiografías que muestran fracturas gracias a un algoritmo, indica que “Las fracturas pasadas por alto son una causa común de discrepancia diagnóstica entre la interpretación radiográfica inicial y la lectura final realizada por los radiólogos. Por esto, un grupo de investigadores de la Facultad de Medicina de la Universidad de Boston realizó un estudio en el que identificaron que la inteligencia artificial (IA) puede ayudar a los médicos a detectar fracturas en las radiografías.” Y declara “la IA puede ser una herramienta poderosa para ayudar a los

radiólogos y otros médicos a mejorar el rendimiento del diagnóstico y aumentar la eficiencia. Además, la tecnología tiene el potencial de mejorar la experiencia del paciente en el hospital o la clínica". (4)”

Respecto a los resultados que arrojaron los documentos de las regiones anatómicas más analizadas bajo la IA en radiografías simples fue la multirregional seguida de las extremidades superiores.

Tabla 1. Región Anatómica

Región Anatómica	N° De Artículos
Extremidades superiores (muñeca, húmero).	9
Extremidades inferiores (fémur, tibia).	6
Columna vertebral (fracturas osteoporóticas).	5
Generalizado o multirregional	10

Para la detección de fracturas en las radiografías se identificaron varios sistemas que ya están en funcionamiento en entornos controlados los cuales son:

Tabla 2 Software IA Fracturas simples

Softwares identificados
BoneView (Gleamer)
FracNet
InceptionV3
ResNet-50
Faster R-CNN
YOLOv3
Modelos basados en CNN
OsteoDetect (Imagen Technologies)
FractureDetect (Imagen Technologies)
BriefCase (Aidoc Medical)
HealthVCF (Zebra Medical Vision)
uAI EasyTriage-Rib (Shanghai United Imaging Intelligence)
XRAIT
CheXNet (Stanford University)
GPT-4V (OpenAI)
RoentGen (Stable Diffusion / Stability AI)
CAAI-FDS (conjunto de soluciones comerciales aprobadas)

Se destaca que el uso de herramientas como el BoneView™ y YOLOv5 fueron aplicados con gran éxito, resaltando en regiones como la muñeca y el fémur donde se muestra una gran precisión diagnóstica.

Respecto a la sensibilidad y especificidad, los datos observados en esta revisión arrojaron los siguientes intervalos de sensibilidad y especificidad estuvieron entre rangos del 70% al 99% con promedio entre 90% al 92%.

Tabla 3 Sensibilidad vs Especificidad

	Rango observado	Promedio aproximado
Sensibilidad	70% - 98.35%	91% - 92%
Especificidad	77% - 99.2%	90% - 91%

Discusión

La interpretación de los estudios en esta revisión apunta al desarrollo sustancial de la IA para el diagnóstico radiológico de fracturas. Está respaldada por su efectividad en múltiples contextos clínicos. Los estudios revisados muestran resultados consistentes en cuanto a su sensibilidad y especificidad. Particularmente en regiones con alto flujo de estudios como las articulaciones de miembro superior, esto demostrado en los casos de uso más prometedores donde ya hemos encontrado que dando una segunda lectura/triage utilizando IA, particularmente en el contexto de radiólogos con recursos limitados y sobrecargados de trabajo. Varios estudios clínicos han demostrado que el rendimiento de estos sistemas puede rivalizar — e incluso superar — al de los médicos generales en tareas de lectura primaria, aunque no reemplazan la interpretación de expertos en casos difíciles.

En los aportes más relevantes de esta revisión es el impacto de la inteligencia artificial como herramienta complementaria para los médicos generales y residentes en ambiente con limitaciones de tiempo y recursos, como destaca Adams (5). Sin embargo, aún quedan varios límites para su aplicación rutinaria: la variabilidad de los datos de entrenamiento, los problemas éticos de responsabilidad diagnóstica y se requieren regulaciones claras, como se ha informado recientemente en estudios dedicados a las implicaciones ético-legales estos mencionados en los artículos de Che y Lee Park (6)

Los hallazgos son clínicamente prometedores, el protocolo de Shelmerdine et al. Ghosh R, (7), resalta la importancia del uso de estas herramientas en entornos reales con participaciones clínicas de primera línea donde se contemplen demografías de pacientes más amplias y pruebas en tiempo real son obligatorios para la validación de la seguridad y efectividad de estos modelos en el uso rutinario en la práctica clínica, en lo que podemos analizar el sesgo geográfico en la producción de conocimiento también es notable, ya que se origina principalmente en países con una infraestructura altamente tecnificada, lo que contrasta con la necesidad de aplicar y evaluar estos sistemas en lugares como América Latina o África como en los estudios realizados por Romero Alk, Villegas Caval. (8). Estos indican que es posible adaptar estas herramientas a regiones locales si se cuentan con datos representativos y orientación adecuada al personal.

Estos datos llegan a la conclusión de que los algoritmos basados en CNN (ResNet y DenseNet) alcanzan picos de sensibilidad superiores al 90% especialmente cuando integran una segunda lectura, como los estudios realizados por Tang A, Tam R, Cadrin-Chênevert (9)

Conclusiones

Después de revisar el estado de la investigación en el campo de la inteligencia artificial (IA) pudimos observar en los 30 artículos científicos que se publicaron entre el 2017 y 2024 permitió desarrollar los objetivos de esta investigación sobre la aplicación de esta herramienta en la detección de fracturas en radiografías simples.

Con la identificación de múltiples softwares que están en uso como herramientas para este propósito como BoneView™, CheXNet, YOLOv3, ResNet-50, entre otros. Estas herramientas han sido probadas en escenarios clínicos reales o ambientes controlados, brindándonos una vista de su creciente incorporación en el campo de la radiología, observando a su vez datos que evidenciaron que estos algoritmos alcanzan niveles muy altos de sensibilidad (10), especialmente cuando fue utilizado como herramienta de apoyo para los residentes y médicos generales también dándonos datos específicos donde el porcentaje fue igual de consistente demostrando que estas herramientas no solo detectan fracturas con buenos resultados, si no que mantienen una baja tasa de falsos positivos (11).

Esta revisión permitió descubrir las ventajas que estas herramientas son capaces de brindar debido a que aporta una mejor eficacia a la hora de ser diagnósticas evitando o reduciendo el error humano en contextos de lugares de alta carga laboral y la capacidad de ser utilizadas en entornos donde el acceso a especialistas es limitado (12). Además, proporciona un gran beneficio a aquellos profesionales que están en formación y residentes. Sin embargo, pudimos identificar limitaciones que están también relacionadas a sus propias ventajas ya que existen modelos de inteligencia artificial que poseen un alto riesgo de “falsos positivos”, además de distintas circunstancias éticas relacionadas también con la responsabilidad legal en compromiso de la privacidad de datos clínicos (13), sin dejar de lado también la desigualdad que existe en el acceso a estas tecnologías en regiones con menos recursos médicos, nuestra región o en lugares como África.

La inteligencia artificial se muestra como una herramienta versátil y poderosa, con un potencial enorme para complementar el diagnóstico en los estudios radiológicos (14) y aunque todavía enfrenta obstáculos técnicos, éticos y de integración, su eficacia demostrada, su utilidad pedagógica y su capacidad de adaptarse a contextos de recursos limitados haciendo hincapié que en Panamá en la caja de seguro social ya existen estas herramientas pero hasta el momento no se han hecho publicaciones sobre su funcionalidad lo cual justifica su promoción para su desarrollo y validación en la práctica clínica.

Referencias

1. Kim, D. H., & MacKinnon, T. (2018). Artificial intelligence in fracture detection: Transfer learning from deep convolutional neural networks. *Clinical Radiology*, 73(5), 439–445. [https://www.clinicalradiologyonline.net/article/S0009-9260\(17\)30535-4/fulltext](https://www.clinicalradiologyonline.net/article/S0009-9260(17)30535-4/fulltext) 15
2. Duron, L., Ducarouge, A., Gillibert, A., Lainé, J., Allouche, C., Cherel, N., Zhang, Nitche, N., Lacave, E., Pourchot, A., Felter, A., Lassalle, L., Regnard, N., & Feydy, A. (2021). Assessment of an AI Aid in Detection of Adult Appendicular Skeletal

- Fractures by Emergency Physicians and Radiologists: A Multicenter Crosssectional Diagnostic. <https://pubs.rsna.org/doi/full/10.1148/radiol.2021203886>
3. Médica, R. (2022, 29 marzo). Sensibilidad del 92% de IA en fracturas óseas. Redacción Médica. <https://www.redaccionmedica.com/secciones/radiologia/lainteligencia-artificial-logra-detectar-el-92-de-las-fracturas-oseas-7387>
 4. Martín, I. (2022, 3 enero). Algoritmo ahorra trabajo al detectar fracturas al instante. Redacción Médica. <https://www.redaccionmedica.com/secciones/radiologia/prioridad-a-lasradiografias-que-muestran-fracturas-gracias-a-un-algoritmo-2402>
 5. Adams M, Chen W, Holcdorf D, McCusker MW, Howe PD, Gaillard F. Computer-aided diagnosis in radiology: Does it improve diagnostic accuracy and efficiency *Clin Radiol.* 2019;74(5):338–45. Available from: [https://www.clinicalradiologyonline.net/article/S0009-9260\(19\)30073-X/fulltext](https://www.clinicalradiologyonline.net/article/S0009-9260(19)30073-X/fulltext)
 6. Chee; Lee Park H, Zhang Y, Zhang W, Liao P, Li K, Hu W. Application of deep learning to radiographic fracture detection: A systematic review and meta-analysis. *Skeletal Radiol.* 2022;51:1307–20. Available from: <https://pmc.ncbi.nlm.nih.gov/articles/PMC10606060/>
 7. Ghosh R, Ghosh K, Reddy R. Ethical and legal challenges in AI-based diagnostic imaging. *J Digit Imaging.* 2020;33:1323–30. Available from: <https://pmc.ncbi.nlm.nih.gov/articles/PMC7332220/>
 8. Ray Romero ALK, Villegas Caval C. Implicancia de la Inteligencia Artificial en la disminución del error diagnóstico radiológico: Un metanálisis. [Internet]. 2019. Available from: <https://repositoriobibliotecas.uv.cl/serveruv/api/core/bitstreams/6cba71da-c7c9-41d0-b5fd-9225b0885409/content>
 9. Tang A, Tam R, Cadrin-Chênevert A, Guest W, Chong J, Barfett J, et al. Canadian Association of Radiologists white paper on artificial intelligence in radiology. *Can Assoc Radiol J.* 2018;69(2):120–35. Available from: <https://journals.sagepub.com/doi/10.1016/j.carj.2018.02.002>
 10. Wang, S., Wang, R., Su, W., Zhang, M., Li, X., & Zhang, H. (2021). Artificial intelligence in orthopedics: Insights, challenges, and future directions. *Frontiers in Bioengineering and Biotechnology*, 9, 619–626. <https://pmc.ncbi.nlm.nih.gov/articles/PMC12188104/>
 11. Meza, N. G., Mauricio, E. S. R., & Jimbo, J. D. B. (2025). Efectividad de la inteligencia artificial en el diagnóstico médico por imágenes: Una revisión sistemática. *Sapiens in Artificial Intelligence*, 2(2). https://revistasapiensec.com/index.php/Sapiens_in_Artificial_Intelligen/issue/archive • Oakden-Rayner, L. (2020). Exploring large-scale public medical image datasets. *Academic Radiology*, 27(1), 106–112. [https://www.academicradiology.org/article/S1076-6332\(19\)30494-5/abstract](https://www.academicradiology.org/article/S1076-6332(19)30494-5/abstract)
 12. Zhou, S. K., Greenspan, H., Davatzikos, C., Duncan, J. S., Van Ginneken, B., Madabhushi, A., ... Summers, R. M. (2021). A review of deep learning in medical imaging: Imaging traits, technology trends, case studies with progress highlights, and future promises. *Proceedings of the IEEE*, 109(5), 820–838. <https://ieeexplore.ieee.org/document/9363915>

13. Tang, A., Tam, R., Cadrin-Chênevert, A., Guest, W., Chong, J., Barfett, J., ... Gray, B. (2018). Canadian Association of Radiologists white paper on artificial intelligence in radiology. *Canadian Association of Radiologists Journal*, 69(2), 120–135. <https://journals.sagepub.com/doi/10.1016/j.carj.2018.02.002>
14. Lindsey, R., Daluiski, A., Chopra, S., Lachapelle, A., Mozer, M., Sicular, S., ... Hegde, R. (2018). Deep neural network improves fracture detection by clinicians. *Proceedings of the National Academy of Sciences*, 115(45), 11591–11596. <https://www.pnas.org/doi/full/10.1073/pnas.1806905115>