

A systematic review on the impact of machine learning on medical diagnostic imaging

Hector Ivan Duque Meza 

Universidad Tecnológica del Perú, Facultad de Ingeniería, Lima- Perú,
U20304053@utp.edu.pe

Abstract - This is a systematic literature review (SLR) that aims to identify and analyze medical image processing techniques using artificial intelligence (AI) and machine learning to improve the accuracy and speed of diagnostic imaging in medical institutions. To this end, SCOPUS selected 2757 articles using inclusion and exclusion criteria, resulting in 53 relevant articles. In addition, the findings are grouped into key themes, highlighting the growing interest in the use of artificial intelligence in medical diagnosis. Therefore, the results indicate that machine learning can overcome the limitations of traditional methods, improve diagnostic accuracy and reduce waiting times; however, some problems were observed, such as dependence on operator expertise and limited technical infrastructure. Therefore, SLR highlights the importance of integrating AI technology into diagnostic imaging, suggesting that its implementation has the potential to transform medical practice, although challenges need to be addressed to maximize its effectiveness.

Keywords -- Artificial Intelligence, Machine Learning, Diagnostic Imaging, Diagnostic Accuracy.

Una revisión sistemática sobre el impacto del aprendizaje automático en el diagnóstico médico por imagen

Hector Ivan Duque Meza
Universidad Tecnológica del Perú, Facultad de Ingeniería, Lima- Perú,
U20304053@utp.edu.pe

Resumen - Esta revisión sistemática de la literatura (RSL) tiene como objetivo analizar las técnicas de procesamiento de imágenes médicas basadas en inteligencia artificial (IA) y aprendizaje automático para mejorar la precisión y eficiencia diagnóstica en instituciones médicas. A partir de una búsqueda en la base de datos SCOPUS, se identificaron 2757 artículos, de los cuales, tras aplicar criterios de inclusión y exclusión, se seleccionaron 53 estudios relevantes. Los resultados evidencian que las técnicas de aprendizaje automático superan limitaciones de los métodos tradicionales al incrementar la precisión del diagnóstico y reducir los tiempos de procesamiento. Entre los principales avances destacan el uso de algoritmos de aprendizaje profundo y la automatización en el análisis de imágenes, aunque persisten desafíos, como la dependencia de operadores capacitados, las restricciones técnicas en ciertas instituciones y la necesidad de validación clínica más amplia. En conclusión, la RSL resalta el potencial transformador de la IA en el diagnóstico médico, con la posibilidad de optimizar procesos y resultados clínicos. Sin embargo, para maximizar su efectividad, es crucial superar las barreras tecnológicas y de implementación identificadas, asegurando así su integración sostenible en la práctica médica.

Palabras Clave -- Inteligencia Artificial, Aprendizaje Automático, Diagnóstico por Imagen, Precisión Diagnóstica.

I. INTRODUCCIÓN

La inteligencia artificial (IA) se ha consolidado como una herramienta fundamental en el diagnóstico médico, especialmente en el análisis de imágenes, al incrementar la precisión y la eficiencia en la detección de diversas condiciones médicas. Un estudio demostró que un sistema de IA mejoró significativamente la precisión diagnóstica de los radiólogos junior en la evaluación de nódulos tiroideos, alcanzando un rendimiento comparable al de radiólogos senior, lo que posiciona a la IA como un recurso valioso en la práctica clínica [1].

El contexto actual del diagnóstico médico enfrenta grandes desafíos debido al creciente volumen de datos clínicos, la necesidad de diagnósticos rápidos y precisos, y la limitada disponibilidad de especialistas en ciertas regiones. En este escenario, la IA surge como una solución prometedora, ofreciendo herramientas capaces de analizar grandes cantidades de datos médicos con alta precisión. Sin embargo, su implementación no está exenta de problemas, como la dependencia de infraestructura tecnológica avanzada y la

necesidad de personal capacitado para interpretar y validar los resultados [2].

En este artículo se abordan también las aplicaciones específicas de la IA en distintas áreas del diagnóstico por imágenes. En el ámbito del ultrasonido mamario, se ha demostrado que los sistemas de IA pueden clasificar imágenes con alta precisión, lo que refuerza su utilidad en la detección temprana del cáncer de mama [3]. Asimismo, un modelo de aprendizaje automático desarrollado para caracterizar la malignidad de nódulos pulmonares solitarios en estudios de PET-CT ilustra la versatilidad de estas tecnologías [4]. Además, la IA ha mostrado resultados prometedores en la detección de fracturas en radiografías, aumentando la sensibilidad y especificidad en los diagnósticos realizados por clínicos de primera línea [2].

La justificación de este estudio radica en la necesidad de integrar la IA en los procesos clínicos para superar las limitaciones de los métodos tradicionales y mejorar la calidad de la atención médica. Aunque ya se han desarrollado numerosas iniciativas, aún persisten desafíos relacionados con la variabilidad en la interpretación debido a la subjetividad del operador [5], así como con las limitaciones en recursos tecnológicos en muchas instituciones de salud [3], [6]. Estos aspectos subrayan la importancia de seguir investigando para maximizar los beneficios de la IA en el ámbito médico.

El procesamiento de imágenes médicas mediante IA también ha mostrado avances significativos en el diagnóstico de enfermedades como el hígado graso [7] o el cáncer de próstata [8]. Estas tecnologías no solo mejoran la precisión diagnóstica, sino que también permiten a los ingenieros de software gestionar grandes volúmenes de datos y desarrollar modelos de entrenamiento que potencien los avances en este campo, contribuyendo al desarrollo de soluciones innovadoras para satisfacer las necesidades del sector salud.

Este documento, una revisión sistemática de la literatura (RSL), tiene como objetivo analizar las técnicas de procesamiento de imágenes médicas basadas en inteligencia artificial (IA) y aprendizaje automático, evaluando su capacidad para mejorar la precisión y eficiencia diagnóstica en instituciones médicas. La estructura del artículo está organizada de la siguiente manera: la metodología (Sección 2) detalla el proceso de selección de estudios relevantes para garantizar un análisis riguroso; los resultados (Sección 3) presentan los hallazgos organizados y enfatizan la precisión de las técnicas de diagnóstico por imagen; la discusión (Sección 4) aborda los desafíos, limitaciones y oportunidades de las tecnologías

basadas en IA; y las conclusiones (Sección 5) sintetizan los principales aportes del estudio y sugieren líneas futuras de investigación.

II. MATERIALES Y MÉTODOS

Se aplicó la técnica PICOC, una herramienta utilizada para descomponer una pregunta de investigación en múltiples interrogantes empíricas. Esta técnica sigue una estructura definida que incluye los siguientes componentes: Problema, Intervención, Comparación, Resultado y Contexto (ver Tablas I y II).

TABLA I
PREGUNTA ORIENTADORA PICO Y
SUBPREGUNTAS

RQ:	¿Cómo influye el uso de aprendizaje automático en la precisión y el tiempo de generación del diagnóstico por imagen en comparación con los métodos tradicionales en instituciones de salud?
RQ1:	¿Cuáles son las limitaciones de precisión y tiempo en los diagnósticos por imagen utilizando métodos tradicionales en instituciones de salud?
RQ2:	¿Qué técnicas específicas de aprendizaje automático se están utilizando para mejorar los diagnósticos por imagen en el ámbito de la salud?
RQ3:	¿Qué diferencias se observan en la precisión y el tiempo de diagnóstico entre el uso de aprendizaje automático y los métodos tradicionales?
RQ4:	¿Qué métricas se emplean para evaluar la precisión y el tiempo de respuesta en diagnósticos por imagen en ambas metodologías?
RQ5:	¿En qué centros, lugares o instituciones se realiza diagnóstico por imagen?

TABLA II
PREGUNTA ORIENTADORA PICO Y
SUBPREGUNTAS

P	Limitaciones de precisión y tiempo en los diagnósticos por imagen	imaging OR image OR picture OR scan OR film OR screening
I	Técnicas de aprendizaje automático usadas en diagnóstico por imagen	"Artificial Intelligence" OR "machine learning" OR "Integrative Analytics" OR "Image Segmentation"
C	Tipo de exámenes usados en el diagnóstico por imagen	"ct scan" OR ultrasound OR "Magnetic resonance imaging" OR "positron emission tomography" OR x-ray OR electroencephalogram OR "Magnetic resonance angiography"
O	Métricas que miden la precisión y el tiempo de respuesta en diagnósticos por imagen	precision OR accuracy OR specificity OR sensitivity OR validity OR "Diagnostic time" OR "Response time"
C	Lugares donde realiza diagnóstico por imagen	hospital OR clinic OR "medical center" OR "imaging center"

A. Estrategia de búsqueda

Con base en las palabras y términos clave señalados en la tabla anterior, se elaboró la siguiente ecuación de búsqueda para su aplicación en SCOPUS: (TITLE-ABS-KEY (imaging OR image OR picture OR scan OR film OR screening) AND TITLE-ABS-KEY ("Artificial Intelligence" OR "machine learning" OR "Integrative Analytics" OR "Image Segmentation") AND TITLE-ABS-KEY ("ct scan" OR ultrasound OR "Magnetic resonance imaging" OR "positron emission tomography" OR x-ray OR electroencephalogram OR "Magnetic resonance angiography") AND TITLE-ABS-KEY (precision OR accuracy OR specificity OR sensitivity OR validity OR "Diagnostic time" OR "Response time") AND TITLE-ABS-KEY (hospital OR clinic OR "medical center" OR "imaging center"))

B. Criterios de elegibilidad

Se aplicaron los siguientes criterios de inclusión:

- Los documentos reportaron las limitaciones de precisión y tiempo en los diagnósticos por imagen al utilizar métodos tradicionales en instituciones de salud.
- Los documentos informaron sobre las técnicas específicas de aprendizaje automático empleadas para mejorar los diagnósticos por imagen en el ámbito de la salud.
- Los estudios abarcaron las diferencias observadas en la precisión y el tiempo de diagnóstico entre el uso de aprendizaje automático y los métodos tradicionales.
- Los estudios presentaron las métricas utilizadas para evaluar la precisión y el tiempo de respuesta en diagnósticos por imagen en ambas metodologías.
- Se consideraron únicamente estudios realizados en centros de salud que incluyeron diagnósticos por imagen.

Así mismo, los criterios de exclusión considerados fueron:

- No se consideraron estudios publicados antes del año 2019.
- No se incluyeron documentos que no fueran artículos originales.
- No se tomaron en cuenta documentos redactados en idiomas distintos al inglés

C. Criterios de elegibilidad

Se identificaron 2,757 artículos en SCOPUS, los cuales fueron filtrados de acuerdo con los criterios establecidos. Como resultado, la muestra final se redujo a 53 artículos, utilizando la técnica PRISMA (Figura 1).

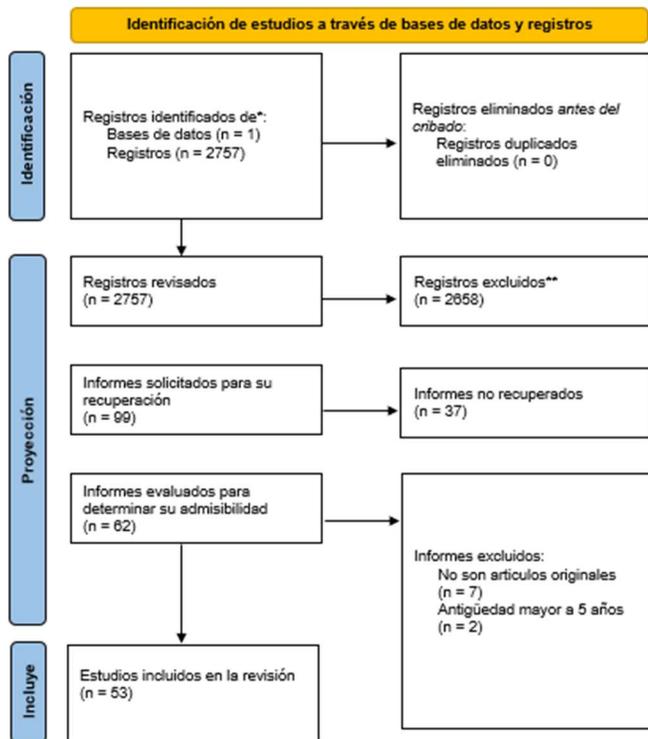


Fig. 1 Diagrama de flujo PRISMA

II. RESULTADOS

El análisis de las palabras clave en los artículos citados reveló tendencias significativas en el campo. El término más frecuente fue **"imagen"**, lo que refleja el interés predominante en el procesamiento de imágenes médicas. **"Inteligencia artificial"** se identificó en 27 artículos, destacando su creciente incorporación en este ámbito. Asimismo, 17 artículos mencionaron el uso de **"ultrasonido"** como método diagnóstico, mientras que 40 exploraron aspectos relacionados con la **"precisión"**, enfatizando la relevancia de obtener resultados exactos. Por último, el término **"hospital"** apareció en 43 artículos, evidenciando el enfoque en la aplicación de estas tecnologías en entornos de atención médica.

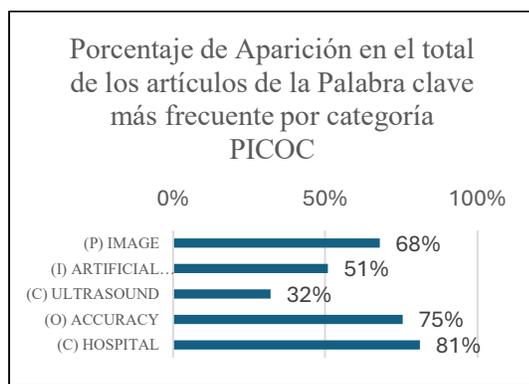


Fig. 2 Palabras clave PICOC más frecuentes

Por otro lado, la Figura 3 presenta la distribución de los artículos según su año de publicación, abarcando el período 2020-2024. Este rango fue seleccionado debido a que, a partir de 2020, se observa un aumento gradual en el número de estudios, alcanzando su punto máximo de publicación en 2024, tal como se ilustra en la figura.

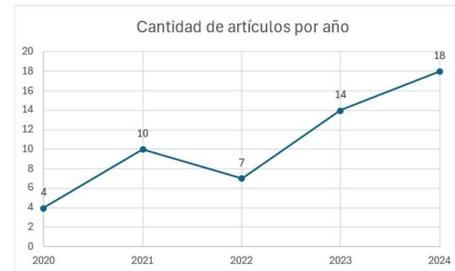


Fig. 3 Artículos según el año de publicación

La Figura 4 muestra la distribución de los artículos según su país de publicación. En primer lugar, China se destaca con el mayor número de artículos publicados (27), seguida por Estados Unidos con 7 artículos y el Reino Unido con 5. Además, otros países como Irán, Australia, Corea del Sur e India también han contribuido al tema, aunque con una menor cantidad de publicaciones.

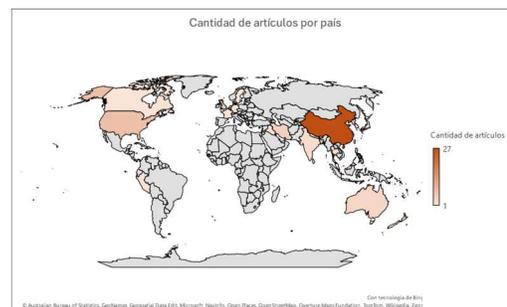


Fig. 4 Artículos según su país de publicación

RQ1.1 ¿Cuáles son las principales causas que afectan la precisión de los métodos tradicionales?

A partir de la revisión sistemática de literatura, se identificaron múltiples limitaciones asociadas a los métodos tradicionales de diagnóstico por imagen, las cuales afectan significativamente su precisión. Una de las principales causas reportadas en los estudios revisados es la alta variabilidad en la interpretación de los resultados, atribuida a la subjetividad y experiencia del operador. Esta problemática ha generado inconsistencias notables en el diagnóstico de enfermedades como la espondiloartritis axial (axSpA) y la tuberculosis [3], [9].

Además, se constató que la dependencia de la pericia del profesional y la infraestructura tecnológica insuficiente en ciertas instituciones de salud dificultan tanto la consistencia como la escalabilidad de los diagnósticos, especialmente en patologías de alta complejidad [5], [6], [10], [11]. Asimismo, varios estudios señalaron que los métodos estadísticos tradicionales presentan baja sensibilidad y especificidad, afectando su precisión, en particular en enfermedades emergentes como la COVID-19 [1], [12], [13].

Finalmente, se observó que la similitud de síntomas y la presencia de comorbilidades entre los pacientes representan un desafío adicional, aumentando el riesgo de diagnósticos incorrectos y comprometiendo la efectividad del tratamiento [14], [15].



Fig. 5 Precisión de métodos tradicionales

RQ1.2 ¿Qué factores limitan la rapidez de diagnóstico en los métodos tradicionales?

En la revisión, se identificó diversos factores que limitan la rapidez en la obtención de resultados en los métodos tradicionales de diagnóstico médico. Entre las principales limitaciones reportadas, destacan los problemas asociados a los recursos y la tecnología, como la falta de hardware adecuado y la escasez de kits de prueba, que prolongan significativamente los tiempos de espera [3].

Asimismo, los estudios revisados resaltan que la dependencia de procedimientos confirmatorios, como cultivos adicionales, incrementa de manera considerable el tiempo necesario para establecer un diagnóstico definitivo [4]. Otro factor crítico identificado es la variabilidad en la interpretación de los resultados y la alta dependencia del operador, lo que introduce subjetividad y ralentiza el proceso diagnóstico [5].

Además, la heterogeneidad de los casos clínicos y la creciente demanda de servicios de salud representan desafíos adicionales, generando retrasos tanto en la emisión de informes como en la atención efectiva de los pacientes [16], [17].

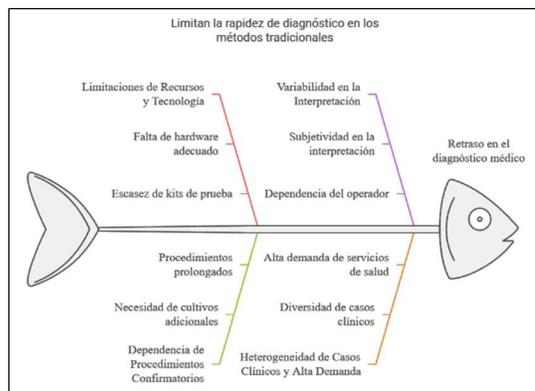


Fig. 6 Diagrama de Ishikawa sobre los Factores que Limitan la rapidez de diagnóstico

RQ2.1 ¿Cuáles son las técnicas de aprendizaje automático más usadas en diagnóstico por imagen?

A partir de la revisión de literatura, se identificó que las redes neuronales convolucionales (CNN) y profundas (DNN) son las técnicas más empleadas en la medicina diagnóstica, representando el 41% de las aplicaciones analizadas [3], [6], [7], [8], [14], [18] - [22]. Estas arquitecturas han sido aplicadas de manera destacada en la detección de diversas patologías, como el cáncer de mama [3] y el cáncer de próstata [8]. Entre las arquitecturas más utilizadas se encuentran ResNet [22] y DenseNet121 [5], ampliamente empleadas para la identificación de patologías médicas.

Los algoritmos de boosting, como XGBoost y LightGBM, comprenden el 12% de las aplicaciones estudiadas, siendo aplicados en la clasificación de tumores pancreáticos [23] y en la detección de enfermedades respiratorias [11]. Por su parte, las máquinas de soporte vectorial (SVM) representan el 6% de las aplicaciones y se han empleado en la clasificación de nódulos tiroideos [23] y en la detección de enfermedades óseas [13].

Los árboles de decisión, incluidos Decision Tree [23] y Random Forest [13], [16], [23], constituyen el 18% de las aplicaciones revisadas y se utilizan en la clasificación de tumores de tejidos blandos [14] y en la detección de enfermedades respiratorias [15].

Finalmente, un 23% de las aplicaciones estudiadas corresponde a otros algoritmos y técnicas avanzadas, como YOLOv3 [3], Faster R-CNN [18], [24], WSDAN [5], atención guiada por máscaras [25], y aprendizaje federado [26]. Estas metodologías han demostrado utilidad en la detección de patologías como el cáncer de próstata [8] y la enfermedad de Alzheimer [21].

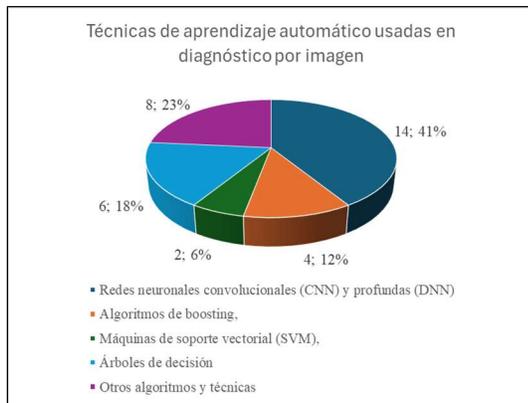


Fig. 7 Técnicas de aprendizaje automático más usadas en diagnóstico por imagen

RQ2.2 ¿Para qué tipos de diagnóstico por imagen se utilizan principalmente estas técnicas?

De acuerdo a la literatura revisada, los tipos de diagnóstico asistido por técnicas avanzadas se corresponden con los distintos sistemas del cuerpo humano. Para el sistema musculoesquelético, se han desarrollado aplicaciones que incluyen el diagnóstico de espondiloartritis axial negativa para HLA-B27 [4], [18], detección de craneosinostosis en niños [18], identificación de fracturas en extremidades [24], diagnóstico de tumores de tejidos blandos [14], displasia del desarrollo de la cadera [6] y escoliosis en adolescentes [22].

En cuanto al sistema digestivo, estas tecnologías han contribuido al diagnóstico de enfermedades como la hepática [7], atresia biliar y cáncer de páncreas [27]. Para el sistema respiratorio, las aplicaciones han permitido clasificar nódulos pulmonares solitarios [16], detectar tuberculosis [9], diagnosticar neumotórax [13], así como identificar casos de COVID-19 [13], [15], [21], [25], [28].

En el sistema endocrino, se ha facilitado la detección de nódulos tiroideos [5] y carcinoma folicular tiroideo [23]. En lo referente al sistema reproductivo, estas herramientas han sido empleadas para diagnosticar masas ováricas [29], cáncer de mama [3], [30] y cáncer de próstata [8].

Finalmente, en el sistema nervioso, estas técnicas han demostrado utilidad en la segmentación y clasificación de tumores cerebrales [19], [10] y en el diagnóstico de lesiones cerebrales traumáticas [12].

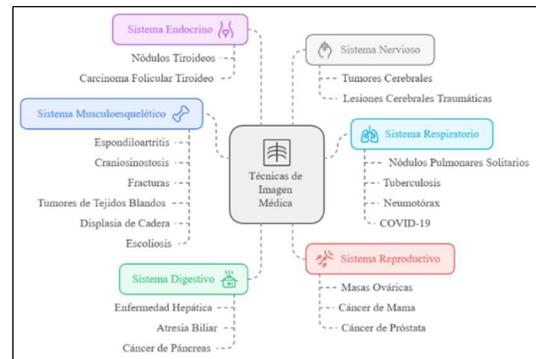


Fig. 8 Diagnósticos por imagen se utilizan principalmente estas técnicas

RQ3.1 ¿Cuánto mejora la precisión con aprendizaje automático en comparación con los métodos tradicionales?

Según la literatura consultada, se observó que la precisión de los sistemas basados en inteligencia artificial (IA) mejora notablemente en comparación con los métodos tradicionales en diversas áreas clínicas. Por ejemplo, en el diagnóstico de cáncer de mama, un sistema de IA con YOLOv3 alcanzó una sensibilidad del 100% para lesiones malignas, superando a los métodos tradicionales [3]. En la clasificación de nódulos pulmonares, la IA logró un área bajo la curva (AUC) de 0.936, superior a los 0.81 y 0.876 obtenidos por los métodos convencionales [8], [16]. Asimismo, en la detección de cáncer de próstata, la IA alcanzó una precisión del 94.3%, mejorando significativamente los resultados de los métodos tradicionales [8].

En el diagnóstico de enfermedades hepáticas, un modelo basado en IA logró una precisión del 90.6% en la detección de hígado graso, superando las tasas de los métodos convencionales [7]. En el diagnóstico del distrés fetal, se observó una mejora significativa en la precisión, ya que solo el 26.7% de los casos fueron identificados mediante monitoreo fetal convencional, mientras que el 73.3% fueron detectados utilizando un enfoque combinado ($P < 0.05$) [11], [31]. Finalmente, en el análisis de imágenes de ultrasonido para la detección de neoplasias tiroideas, la IA alcanzó una precisión del 74.1%, superior al 64.8% logrado por los radiólogos [10].

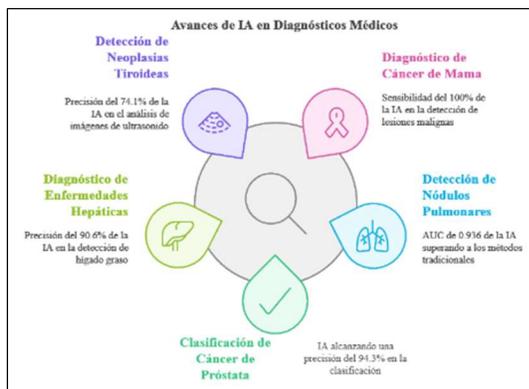


Fig. 9 Avances de IA en Diagnósticos Médicos

RQ3.2 ¿En qué medida se reduce el tiempo de diagnóstico al usar aprendizaje automático comparado con métodos tradicionales?

Los hallazgos derivados de la revisión de la literatura indican que la inteligencia artificial (IA) está generando una transformación significativa en el diagnóstico médico, aportando tres beneficios principales. En primer lugar, se ha observado que las predicciones realizadas por IA son considerablemente más rápidas que las obtenidas mediante métodos tradicionales, generándose en fracciones de segundo, mientras que los métodos convencionales requieren tiempos significativamente más largos. En segundo lugar, las imágenes de diagnóstico asistidas por IA han demostrado ofrecer resultados precisos y rápidos, sin necesidad de procedimientos adicionales, como ocurre con las tomografías computarizadas. Finalmente, se ha destacado que la IA optimiza el flujo de trabajo, reduciendo el tiempo total de diagnóstico a menos de un minuto en algunos casos. Aunque el impacto exacto de estos avances varía según el contexto, los hallazgos de la revisión sugieren que la IA mejora la eficiencia y reduce significativamente los tiempos de diagnóstico [3-6], [9], [17], [19], [21], [32],

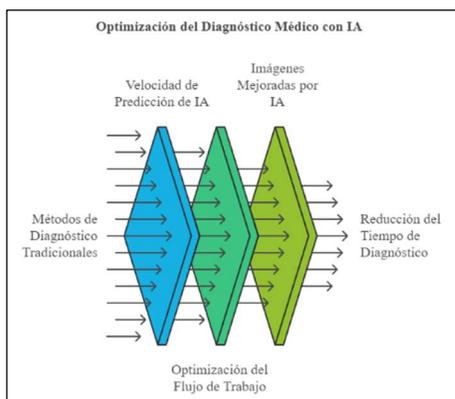


Fig. 10 Optimización del Diagnóstico Médico con IA

RQ4.1 ¿Cuáles son las métricas principales para medir la precisión en diagnósticos por imagen?

La figura ilustra las métricas clave utilizadas para evaluar las imágenes de diagnóstico, organizadas jerárquicamente desde los indicadores más generales hasta los más específicos, según los hallazgos de la literatura revisada. En la capa externa de la jerarquía se encuentran los indicadores de segmentación, como el coeficiente de similitud de los cubos (DSC) y la distancia de Hausdorff (HD), que se emplean para comparar las áreas segmentadas previstas y reales. En el siguiente nivel, los valores predictivos positivo (VPP) y negativo (VPN) son utilizados para evaluar la probabilidad de diagnósticos correctos. Más cerca del centro, se destacan métricas como el área bajo la curva (AUC) de la curva ROC y la precisión general, las cuales miden la capacidad del modelo para distinguir entre diferentes clases. También se incluyen la sensibilidad y la especificidad, que evalúan la capacidad del modelo para identificar correctamente los casos positivos y negativos, respectivamente. En el núcleo de esta jerarquía se encuentra el objetivo principal: lograr un diagnóstico preciso integrando todas estas métricas, como se evidencia en diversas investigaciones [1-9], [12-25], [27-30], [32-51]

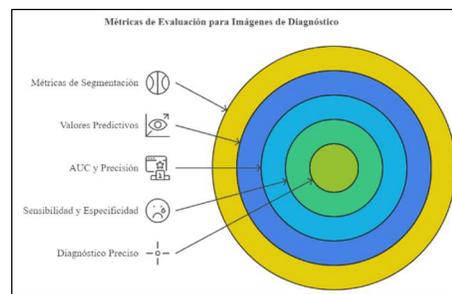


Fig. 11 Métricas de Evaluación para Imágenes de Diagnóstico

RQ4.2 ¿Qué indicadores se usan para medir el tiempo de respuesta en métodos tradicionales y con aprendizaje automático?

La figura compara el rendimiento diagnóstico entre los métodos tradicionales y aquellos basados en aprendizaje automático, utilizando métricas clave de tiempo y precisión, tal como se observó en la literatura revisada. En los métodos tradicionales, los principales indicadores son el tiempo medio hasta el diagnóstico (MTD), la duración de la evaluación de imágenes y el tiempo de espera para diagnóstico o tratamiento, con un enfoque predominante en el tiempo como medida de efectividad.

Por otro lado, los métodos basados en aprendizaje automático destacan por métricas clave como el área bajo la curva ROC (AUC-ROC), la tasa de recuperación y la tasa de falsos positivos, además del tiempo de procesamiento y la velocidad

de segmentación, lo que subraya su capacidad para gestionar grandes volúmenes de datos de manera eficiente. Ambos enfoques consideran, además, métricas como el tiempo promedio de lectura de imágenes y la duración del informe diagnóstico. Sin embargo, en el caso del aprendizaje automático, también se evalúa el tiempo invertido en procesamiento, lo que permite identificar áreas de mejora tanto en precisión como en eficiencia del diagnóstico, especialmente cuando se implementan técnicas avanzadas [4], [5], [6], [8], [9], [16], [17], [19], [24], [33], [41].

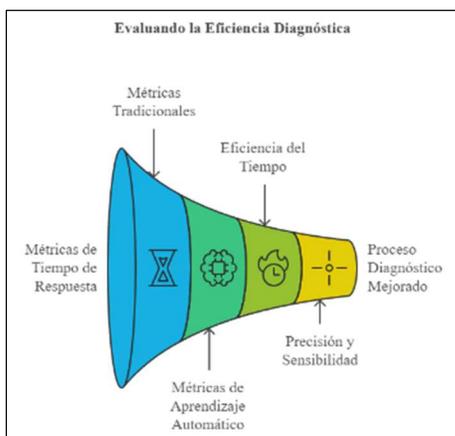


Fig. 12 Evaluando la Eficiencia Diagnóstica

RQ5.1 ¿Qué tipos de centros de salud realizan diagnósticos por imagen?

La figura clasifica los centros de imágenes médicas en cuatro categorías: académicos y no académicos, así como generales y especializados, de acuerdo con los hallazgos observados en la literatura revisada. Los hospitales universitarios combinan atención al paciente, investigación y enseñanza clínica, lo que les permite realizar diagnósticos avanzados, aprovechando su enfoque integral. Por otro lado, las instituciones oncológicas terciarias se centran en el diagnóstico y tratamiento de enfermedades específicas, como el cáncer, lo que les otorga especialización en esta área.

En el ámbito no académico, los hospitales públicos y privados ofrecen servicios generales de diagnóstico por imágenes, mientras que las clínicas de atención primaria están orientadas a diagnósticos básicos de enfermedades comunes, proporcionando atención más accesible para condiciones prevalentes. Esta red de instituciones médicas, que varía en términos de especialización y alcance, integra tecnología avanzada, especialización y una amplia cobertura clínica, subrayando la relevancia de cada tipo de institución en el contexto del diagnóstico médico [1-9], [12-30], [33-35], [45], [50], [52], [53].

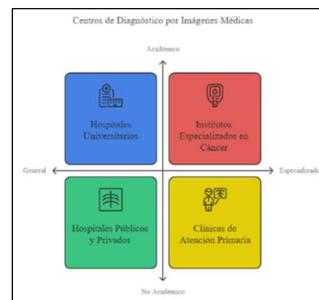


Fig. 13 Centros de Diagnóstico por Imágenes Médicas

RQ5.2 ¿En qué regiones o países es más común el uso de aprendizaje automático en diagnósticos por imagen?

El uso del aprendizaje automático en el diagnóstico por imágenes es más común en regiones con una infraestructura sanitaria avanzada, lideradas por países como Estados Unidos y aquellos de Europa, con ejemplos destacados como Italia y Rusia. Japón y China también están experimentando una creciente adopción de esta tecnología, enfocándose en mejorar la precisión y la rapidez de los diagnósticos médicos. En las regiones en desarrollo, el aprendizaje automático se utiliza principalmente para abordar enfermedades prevalentes, como la tuberculosis, y ha demostrado ser crucial en países como Tailandia y Perú, donde desempeña un papel fundamental en la optimización de diagnósticos y tratamientos [3], [4], [5], [8], [9], [15-19], [21], [32], [52].

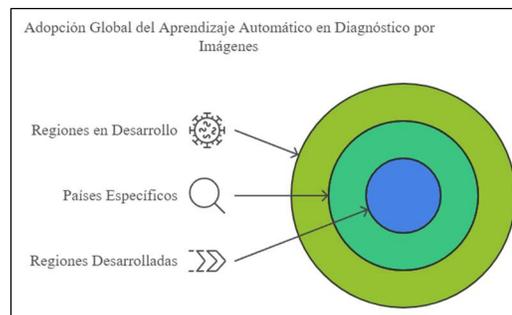


Fig. 14 Adopción Global del Aprendizaje Automático en Diagnóstico por Imágenes.

III. DISCUSIONES

El impacto del aprendizaje automático (AA) en la precisión y rapidez del diagnóstico por imágenes ha demostrado ser significativo en la literatura revisada, con una prominente adopción de técnicas avanzadas como las redes neuronales convolucionales (CNN) y las redes neuronales profundas (DNN), que representaron el 41% de las aplicaciones analizadas. Estos métodos han logrado avances notables en diversas áreas, superando los enfoques tradicionales. Un ejemplo destacado es el diagnóstico de cáncer de mama utilizando YOLOv3, que alcanzó una sensibilidad del 100%, un

rendimiento superior al de los métodos convencionales [3]. Además, en el diagnóstico de nódulos pulmonares, se observó un notable aumento en el área bajo la curva (AUC), pasando de 0.81 a 0.936 con el uso de modelos basados en AA, lo que refleja una mejora considerable en la capacidad de detección [8], [9]. Innovaciones como UNet mejorado y Attention-ResNet también han mostrado un impacto importante al optimizar diagnósticos complejos, reduciendo los tiempos de respuesta a tan solo unos segundos en algunos casos específicos, lo que mejora la eficiencia del diagnóstico [19-21].

Estos hallazgos subrayan el considerable potencial del AA para superar las limitaciones inherentes a los métodos tradicionales de diagnóstico. Por ejemplo, los enfoques tradicionales suelen depender de la experiencia del operador y presentan limitaciones en cuanto a la especificidad, lo que puede llevar a diagnósticos erróneos. En contraste, el AA mejora tanto la precisión como la eficiencia al automatizar tareas y permitir una mayor precisión en la identificación de patrones, lo que se traduce en diagnósticos más rápidos y confiables [5], [15]. Estas ventajas se han reflejado en el diagnóstico de diversas patologías, desde neoplasias tiroideas hasta tumores pancreáticos, donde la precisión del AA ha demostrado ser superior a la de los métodos tradicionales [5], [8], [23]. Sin embargo, no todos los resultados fueron igualmente positivos. En el caso del diagnóstico de tuberculosis, se observó que el AUC alcanzó 0.876, lo que sugiere que existen desafíos específicos relacionados con la naturaleza de la enfermedad o con la calidad de los datos utilizados, lo que podría haber afectado la precisión del modelo en este contexto [9].

El estudio también ha enfrentado varias limitaciones. En primer lugar, la exclusión de artículos previos a 2019 y la restricción a publicaciones en inglés pudieron haber limitado la diversidad de perspectivas y enfoques metodológicos en los estudios analizados [5]. Además, la falta de métricas uniformes entre los estudios revisados dificultó la comparación directa de los resultados, lo que limita la capacidad de establecer conclusiones definitivas sobre el rendimiento general de las tecnologías de AA en diagnóstico médico [7]. También se identificaron resultados inesperados, como una menor precisión en los diagnósticos de COVID-19 asistidos por atención guiada por máscaras, lo que pone de manifiesto la necesidad de ajustar los modelos de AA a contextos clínicos y poblacionales específicos, ya que el rendimiento de los modelos puede variar según el entorno y las características del caso [25].

A partir de los hallazgos de esta revisión, se sugieren varias direcciones para futuras investigaciones. Es necesario explorar la implementación del AA en regiones con recursos limitados, donde la infraestructura sanitaria es menos avanzada, pero donde el potencial de mejora en los diagnósticos puede ser significativo. Además, sería valioso investigar la aplicabilidad de estos modelos en el diagnóstico de enfermedades menos estudiadas y en entornos de atención primaria, donde los diagnósticos rápidos y precisos son cruciales. También es

fundamental examinar cómo la adopción de tecnologías de AA afecta la percepción de los pacientes sobre la atención médica y cómo los profesionales de la salud se adaptan a estas nuevas herramientas, para garantizar una integración ética y equitativa del AA en la práctica clínica. Además, investigar el impacto de estos avances en la formación de médicos y técnicos de salud sería esencial para maximizar los beneficios de la tecnología y asegurar que su implementación sea adecuada a las necesidades del sistema de salud global.

IV. CONCLUSIONES

Esta revisión sistemática de la literatura ofrece una visión completa sobre la precisión diagnóstica y las limitaciones de los métodos tradicionales, destacando el impacto creciente del aprendizaje automático (AA) en el diagnóstico médico. A partir del análisis de 53 artículos relevantes, se evidencia cómo la inteligencia artificial (IA) está optimizando los procesos diagnósticos y reduciendo significativamente los tiempos de espera en la atención sanitaria.

La principal aportación de esta revisión es la organización y síntesis de la información actual sobre técnicas avanzadas en diagnóstico por imagen, proporcionando una base sólida para el avance del campo. Los hallazgos resaltan cómo la IA está transformando los diagnósticos, mejorando tanto la precisión como la eficiencia. Esta recopilación de conocimientos se convierte en una referencia clave para futuras investigaciones, impulsando el desarrollo de mejores prácticas en diagnóstico médico mediante tecnologías emergentes.

Se sugiere que futuras investigaciones profundicen en la adopción de IA en diversas condiciones clínicas, evaluando su eficacia en diferentes patologías y su impacto tanto en la práctica clínica como en la formación de profesionales. Además, es fundamental investigar la percepción de los pacientes sobre el uso de IA en el diagnóstico, abordando los aspectos éticos y de confianza que podrían influir en la aceptación y efectividad de estas tecnologías en la práctica diaria.

V. REFERENCIAS

- [1] D. Xu *et al.*, "The clinical value of artificial intelligence in assisting junior radiologists in thyroid ultrasound: a multicenter prospective study from real clinical practice," *BMC Med*, vol. 22, no. 1, 2024, doi: 10.1186/s12916-024-03510-z.
- [2] A. Novak *et al.*, "Evaluating the impact of artificial intelligence-assisted image analysis on the diagnostic accuracy of front-line clinicians in detecting fractures on plain X-rays (FRACT-AI): protocol for a prospective observational study," *BMJ Open*, vol. 14, no. 9, 2024, doi: 10.1136/bmjopen-2024-086061.
- [3] R. Mori, M. Okawa, Y. Tokumaru, Y. Niwa, N. Matsuhashi, and M. Futamura, "Application of an artificial intelligence-based system in the diagnosis of breast ultrasound images obtained using a smartphone," *World J Surg Oncol*, vol. 22, no. 1, 2024, doi: 10.1186/s12957-023-03286-1.

- [4] Z. Lu *et al.*, "Artificial intelligence improves the diagnosis of human leukocyte antigen (HLA)-B27-negative axial spondyloarthritis based on multisequence magnetic resonance imaging and clinical features," *Quant Imaging Med Surg*, vol. 14, no. 8, pp. 5845–5860, 2024, doi: 10.21037/qims-24-729.
- [5] P. Namsena *et al.*, "Diagnostic performance of artificial intelligence in interpreting thyroid nodules on ultrasound images: a multicenter retrospective study," *Quant Imaging Med Surg*, vol. 14, no. 5, pp. 3676–3694, 2024, doi: 10.21037/qims-23-1650.
- [6] J. L. Jaremko, A. Hareendranathan, S. E. S. Bolouri, R. F. Frey, S. Dulai, and A. L. Bailey, "AI aided workflow for hip dysplasia screening using ultrasound in primary care clinics," *Sci Rep*, vol. 13, no. 1, 2023, doi: 10.1038/s41598-023-35603-9.
- [7] T. Kim, D. H. Lee, E.-K. Park, and S. Choi, "Deep learning techniques for fatty liver using multi-view ultrasound images scanned by different scanners: development and validation study," *JMIR Med Inform*, vol. 9, no. 11, 2021, doi: 10.2196/30066.
- [8] Z. Sun *et al.*, "Using an artificial intelligence model to detect and localize visible clinically significant prostate cancer in prostate magnetic resonance imaging: a multicenter external validation study," *Quant Imaging Med Surg*, vol. 14, no. 1, pp. 43–60, 2024, doi: 10.21037/qims-23-791.
- [9] A. M. Biewer *et al.*, "Accuracy of digital chest x-ray analysis with artificial intelligence software as a triage and screening tool in hospitalized patients being evaluated for tuberculosis in Lima, Peru," *PLOS Global Public Health*, vol. 4, no. 2 February, 2024, doi: 10.1371/journal.pgph.0002031.
- [10] Y. Pan *et al.*, "A Multi-Atlas-Based [18F]9-Fluoropropyl-(+)-Dihydrotetrabenazine Positron Emission Tomography Image Segmentation Method for Parkinson's Disease Quantification," *Front Aging Neurosci*, vol. 14, 2022, doi: 10.3389/fnagi.2022.902169.
- [11] S. Liu, Y. Sun, and N. Luo, "Doppler Ultrasound Imaging Combined with Fetal Heart Detection in Predicting Fetal Distress in Pregnancy-Induced Hypertension under the Guidance of Artificial Intelligence Algorithm," *J Healthc Eng*, vol. 2021, 2021, doi: 10.1155/2021/4405189.
- [12] C. Marincowitz, L. Paton, F. Lecky, and P. Tiffin, "Predicting need for hospital admission in patients with traumatic brain injury or skull fractures identified on CT imaging: a machine learning approach," *Emergency Medicine Journal*, vol. 39, no. 5, pp. 394–401, 2022, doi: 10.1136/emmermed-2020-210776.
- [13] M. Shanbehzadeh, H. Kazemi-Arpanahi, A. Orooji, S. Mobarak, and S. Jelvay, "Performance evaluation of selected machine learning algorithms for COVID-19 prediction using routine clinical data: With versus without CT scan features," *J Educ Health Promot*, vol. 10, no. 1, 2021, doi: 10.4103/jehp.jehp_1424_20.
- [14] H. Xie *et al.*, "Deep learning driven diagnosis of malignant soft tissue tumors based on dual-modal ultrasound images and clinical indexes," *Front Oncol*, vol. 14, 2024, doi: 10.3389/fonc.2024.1361694.
- [15] J. Ye *et al.*, "Association Between Artificial Intelligence BaseChest Computed Tomography and Clinical/ Laboratory Characteristics with Severity and Mortality in COVID-19 Hospitalized Patients," *J Inflamm Res*, vol. 17, pp. 2977–2989, 2024, doi: 10.2147/JIR.S456440.
- [16] I. D. Apostolopoulos, N. D. Papathanasiou, D. J. Apostolopoulos, N. Papandrianos, and E. I. Papageorgiou, "Integrating Machine Learning in Clinical Practice for Characterizing the Malignancy of Solitary Pulmonary Nodules in PET/CT Screening," *Diseases*, vol. 12, no. 6, 2024, doi: 10.3390/diseases12060115.
- [17] S. R. Blake *et al.*, "Using Artificial Intelligence to Stratify Normal versus Abnormal Chest X-rays: External Validation of a Deep Learning Algorithm at East Kent Hospitals University NHS Foundation Trust," *Diagnostics*, vol. 13, no. 22, 2023, doi: 10.3390/diagnostics13223408.
- [18] M. Sabeti, R. Boostani, B. Taheri, and E. Moradi, "Image processing and machine learning for diagnosis and screening of craniosynostosis in children," *Interdiscip Neurosurg*, vol. 36, 2024, doi: 10.1016/j.inat.2023.101887.
- [19] Y. Yang *et al.*, "Segmentation method of magnetic resonance imaging brain tumor images based on improved UNet network," *Transl Cancer Res*, vol. 13, no. 3, pp. 1567–1583, 2024, doi: 10.21037/tcr-23-1858.
- [20] Y. Liu *et al.*, "Artificial intelligence versus radiologist in the accuracy of fracture detection based on computed tomography images: a multidimensional, multi-region analysis," *Quant Imaging Med Surg*, vol. 13, no. 10, pp. 6424–6433, 2023, doi: 10.21037/qims-23-428.
- [21] M. Ghafoori *et al.*, "Predicting survival of Iranian COVID-19 patients infected by various variants including omicron from CT Scan images and clinical data using deep neural networks," *Heliyon*, vol. 9, no. 11, 2023, doi: 10.1016/j.heliyon.2023.e21965.
- [22] H. Wang *et al.*, "An intelligent composite model incorporating global / regional X-rays and clinical parameters to predict progressive adolescent idiopathic scoliosis curvatures and facilitate population screening," *EBioMedicine*, vol. 95, 2023, doi: 10.1016/j.ebiom.2023.104768.
- [23] Y. Zheng *et al.*, "Diagnostic value of an interpretable machine learning model based on clinical ultrasound features for follicular thyroid carcinoma," *Quant Imaging Med Surg*, vol. 14, no. 9, pp. 6311–6324, 2024, doi: 10.21037/qims-24-601.
- [24] Y. Xie *et al.*, "Artificial intelligence diagnostic model for multi-site fracture X-ray images of extremities based on deep convolutional neural networks," *Quant Imaging Med Surg*, vol. 14, no. 2, pp. 1930–1943, 2024, doi: 10.21037/qims-23-878.
- [25] M. Maftouni *et al.*, "A mask-guided attention deep learning model for COVID-19 diagnosis based on an integrated CT scan images database," *IJSE Trans Healthc Syst Eng*, vol. 13, no. 2, pp. 132–149, 2023, doi: 10.1080/24725579.2022.2142866.
- [26] Y. Zhou *et al.*, "Cloud-magnetic resonance imaging system: In the era of 6G and artificial intelligence," *Magnetic Resonance Letters*, 2024, doi: 10.1016/j.mrl.2024.200138.
- [27] S. Yao, D. Yao, Y. Huang, S. Qin, and Q. Chen, "A machine learning model based on clinical features and ultrasound radiomics features for pancreatic tumor classification," *Front Endocrinol (Lausanne)*, vol. 15, 2024, doi: 10.3389/fendo.2024.1381822.
- [28] Y. Kuroda *et al.*, "Artificial intelligence-based point-of-care lung ultrasound for screening COVID-19 pneumoniae: Comparison with CT scans," *PLoS One*, vol. 18, no. 3 March, 2023, doi: 10.1371/journal.pone.0281127.
- [29] L. Liu *et al.*, "Ultrasound image-based nomogram combining clinical, radiomics, and deep transfer learning features for automatic classification of ovarian masses according to O-RADS," *Front Oncol*, vol. 14, 2024, doi: 10.3389/fonc.2024.1377489.
- [30] L. Zhong, L. Shi, L. Zhou, X. Liu, L. Gu, and W. Bai, "Development of a nomogram-based model combining intraand peritumoral ultrasound radiomics with clinical features for differentiating benign from malignant in Breast Imaging Reporting and Data System category 3-5 nodules," *Quant Imaging Med Surg*, vol. 13, no. 10, pp. 6899–6910, 2023, doi: 10.21037/qims-23-283.
- [31] D. A. Badr, F. De Lucia, A. Carlin, J. C. Jani, and M. M. Cannie, "Chest CT scan predictors of intensive care unit admission in hospitalized pregnant women with COVID-19: a case-control study," *Journal of Maternal-Fetal and Neonatal Medicine*, vol. 36, no. 2, 2023, doi: 10.1080/14767058.2023.2241107.
- [32] S. Sharma, "Drawing insights from COVID-19-infected patients using CT scan images and machine learning techniques: a study on 200 patients," *Environmental Science and Pollution Research*, vol. 27, no. 29, pp. 37155–37163, 2020, doi: 10.1007/s11356-020-10133-3.
- [33] A. Novak *et al.*, "Evaluation of the impact of artificial intelligence-assisted image interpretation on the diagnostic performance of clinicians in identifying pneumothoraces on plain chest X-ray: a multi-case multi-reader study," *Emergency Medicine Journal*, vol. 41, no. 10, pp. 602–609, 2024, doi: 10.1136/emmermed-2023-213620.
- [34] W. Hendrix *et al.*, "Deep learning for the detection of benign and malignant pulmonary nodules in non-screening chest CT scans," *Communications Medicine*, vol. 3, no. 1, 2023, doi: 10.1038/s43856-023-00388-5.

- [35] S. Lu *et al.*, "Machine learning applied to HR-pQCT images improves fracture discrimination provided by DXA and clinical risk factors," *Bone*, vol. 168, 2023, doi: 10.1016/j.bone.2022.116653.
- [36] L. Zhixin, L. Gang, J. Zhixian, and P. Silin, "The development and validation of an artificial intelligence-based screening method for atrial septal defect in children's chest x-rays," *Front Pediatr*, vol. 11, 2023, doi: 10.3389/fped.2023.1203933.
- [37] Y. Yang *et al.*, "A prospective multicenter clinical research study validating the effectiveness and safety of a chest X-ray-based pulmonary tuberculosis screening software JF CXR-1 built on a convolutional neural network algorithm," *Front Med (Lausanne)*, vol. 10, 2023, doi: 10.3389/fmed.2023.1195451.
- [38] W. Li *et al.*, "Incorporation of a machine learning pathological diagnosis algorithm into the thyroid ultrasound imaging data improves the diagnosis risk of malignant thyroid nodules," *Front Oncol*, vol. 12, 2022, doi: 10.3389/fonc.2022.968784.
- [39] P. Arunukthavon, D. Songsaeng, C. Keatmanee, S. Klabwong, M. Ekpanyapong, and M. N. Dailey, "Diagnostic Performance of Artificial Intelligence for Interpreting Thyroid Cancer in Ultrasound images," *International Journal of Knowledge and Systems Science*, vol. 13, no. 1, 2022, doi: 10.4018/IJKSS.309431.
- [40] Y. Chen, X. Tian, K. Fan, Y. Zheng, N. Tian, and K. Fan, "The Value of Artificial Intelligence Film Reading System Based on Deep Learning in the Diagnosis of Non-Small-Cell Lung Cancer and the Significance of Efficacy Monitoring: A Retrospective, Clinical, Nonrandomized, Controlled Study," *Comput Math Methods Med*, vol. 2022, 2022, doi: 10.1155/2022/2864170.
- [41] Y. Wu *et al.*, "Ultrasound Lung Image under Artificial Intelligence Algorithm in Diagnosis of Neonatal Respiratory Distress Syndrome," *Comput Math Methods Med*, vol. 2022, 2022, doi: 10.1155/2022/1817341.
- [42] A. Golla *et al.*, "Automated screening for abdominal aortic aneurysm in CT scans under clinical conditions using deep learning," *Diagnostics*, vol. 11, no. 11, 2021, doi: 10.3390/diagnostics11112131.
- [43] Z. Jiao *et al.*, "Prognostication of patients with COVID-19 using artificial intelligence based on chest x-rays and clinical data: a retrospective study," *Lancet Digit Health*, vol. 3, no. 5, pp. e286–e294, 2021, doi: 10.1016/S2589-7500(21)00039-X.
- [44] X. Zhang *et al.*, "Evaluating the Accuracy of Breast Cancer and Molecular Subtype Diagnosis by Ultrasound Image Deep Learning Model," *Front Oncol*, vol. 11, 2021, doi: 10.3389/fonc.2021.623506.
- [45] J. C. Quiroz *et al.*, "Development and validation of a machine learning approach for automated severity assessment of COVID-19 based on clinical and imaging data: Retrospective study," *JMIR Med Inform*, vol. 9, no. 2, 2021, doi: 10.2196/24572.
- [46] M. Y. Kamil, "A deep learning framework to detect Covid-19 disease via chest X-ray and CT scan images," *International Journal of Electrical and Computer Engineering*, vol. 11, no. 1, pp. 844–850, 2021, doi: 10.11591/ijece.v11i1.pp844-850.
- [47] Y. Wang, X. Wang, L. Jin, and X. Wei, "X-Ray Film under Artificial Intelligence Algorithm in the Evaluation for Nursing Effect of Gamma Nail Internal Fixation in Elderly Patients with Intertrochanteric Fracture of Femur," *Comput Math Methods Med*, vol. 2021, 2021, doi: 10.1155/2021/2562575.
- [48] W. Gao, P. Zhang, H. Wang, P. Tuo, and Z. Li, "Magnetic Resonance Imaging Image Feature Analysis Algorithm under Convolutional Neural Network in the Diagnosis and Risk Stratification of Prostate Cancer," *J Healthc Eng*, vol. 2021, 2021, doi: 10.1155/2021/1034661.
- [49] H. Li *et al.*, "Identification of bipolar disorder using a combination of multimodality magnetic resonance imaging and machine learning techniques," *BMC Psychiatry*, vol. 20, no. 1, 2020, doi: 10.1186/s12888-020-02886-5.
- [50] J. P. Kim *et al.*, "Staging and quantification of florbetaben PET images using machine learning: impact of predicted regional cortical tracer uptake and amyloid stage on clinical outcomes," *Eur J Nucl Med Mol Imaging*, vol. 47, no. 8, pp. 1971–1983, 2020, doi: 10.1007/s00259-019-04663-3.
- [51] I. Shin *et al.*, "Application of machine learning to ultrasound images to differentiate follicular neoplasms of the thyroid gland," *Ultrasonography*, vol. 39, no. 3, pp. 257–265, 2020, doi: 10.14366/usg.19069.
- [52] M. F. Mohd Hisham, N. A. Lodz, E. N. Muhammad, F. N. Asari, M. I. Mahmood, and Z. Abu Bakar, "Evaluation of 2 Artificial Intelligence Software for Chest X-Ray Screening and Pulmonary Tuberculosis Diagnosis: Protocol for a Retrospective Case-Control Study," *JMIR Res Protoc*, vol. 12, 2023, doi: 10.2196/36121.
- [53] F.-R. Hsu, S.-T. Dai, C.-M. Chou, and S.-Y. Huang, "The application of artificial intelligence to support biliary atresia screening by ultrasound images: A study based on deep learning models," *PLoS One*, vol. 17, no. 10, October, 2022, doi: 10.1371/journal.pone.0276278.