



Explainable Artificial Intelligence (XAI) en el diagnóstico por imágenes para la detección temprana del cáncer: una revisión sistemática de la literatura

Explainable Artificial Intelligence (XAI) in Diagnostic Imaging for Early Cancer Detection: A Systematic Review of Literature

Walter Fabián Taday Guashpa^{1*} , Jaime David Camacho Castillo¹ 

¹ Escuela Superior Politécnica de Chimborazo. Chimborazo, Ecuador

*Autor para la correspondencia: walter.taday@esPOCH.edu.ec

Cómo citar este artículo

Taday Guashpa WF, Camacho Castillo JD: Explainable Artificial Intelligence (XAI) en el diagnóstico por imágenes para la detección temprana del cáncer: una revisión sistemática de la literatura. Rev haban cienc méd [Internet]. 2026 [citado]; 25. Disponible en: <http://www.revhabanera.sld.cu/index.php/rhab/article/view/6337>

Recibido: 12 de enero de 2026

Aprobado: 13 de abril de 2026

RESUMEN

Introducción: El cáncer constituye una de las principales causas de mortalidad mundial y la detección temprana mediante imágenes médicas representa un factor crítico para mejorar la supervivencia. En este contexto, la Inteligencia Artificial Explicativa ha emergido como una estrategia para superar la opacidad de los modelos de inteligencia artificial diagnósticos.

Objetivo: Analizar la evidencia científica disponible sobre la efectividad de las técnicas de Inteligencia Artificial Explicativa aplicadas al diagnóstico por imágenes para la detección temprana de diversos tipos de cáncer.

Material y Métodos: Se realizó una revisión sistemática siguiendo la guía PRISMA 2020 y los criterios PICO en las bases de datos Scopus y Web of Science, seleccionándose 25 estudios publicados entre 2020 y 2025 tras la evaluación de calidad mediante QUADAS-2 y ROBINS-I.

Resultados: Técnicas como Grad-CAM, SHAP y LIME son las más utilizadas y se aplican principalmente en resonancia magnética, tomografía computarizada, mamografía y dermatoscopia para cáncer de pulmón, mama, cerebro y piel. En múltiples estudios, los modelos explicables igualaron o superaron a los sistemas de "caja negra", se alcanzaron altas métricas de precisión diagnóstica, reduciendo falsos positivos y mejorando la confianza clínica al permitir validar visualmente las decisiones algorítmicas.

Conclusiones: La Inteligencia Artificial Explicativa no solo incrementa la interpretabilidad, sino que fortalece el rendimiento y la aceptabilidad clínica de la inteligencia artificial médica, posicionándose como un componente esencial para la implementación segura y efectiva de sistemas de diagnóstico oncológico asistido por IA en la práctica clínica contemporánea y en futuros entornos sanitarios digitalizados.

Palabras Claves:

Apoyo a la toma de decisiones médicas, detección temprana del cáncer, diagnóstico oncológico, diagnóstico por imágenes, inteligencia artificial explicable, XAI.

ABSTRACT

Introduction: Cancer is one of the main causes of mortality worldwide, and the early detection through medical imaging represents a critical factor to improve survival. In this context, Explanatory Artificial Intelligence has emerged as a strategy to overcome the opacity of diagnostic artificial intelligence models.

Objective: To analyze the available scientific evidence on the effectiveness of Explanatory Artificial Intelligence techniques applied to diagnostic imaging for the early detection of various types of cancer.

Material and Methods: A systematic review was carried out in the Scopus and Web of Science databases, following the PRISMA 2020 guidelines and the PICO criteria. In addition, 25 studies published between 2020 and 2025 were selected after quality assessment using QUADAS-2 and ROBINS-I.

Results: Techniques such as Grad-CAM, SHAP and LIME are the most commonly used and are mainly applied in magnetic resonance imaging, computed tomography, mammography and dermoscopy for lung, breast, brain and skin cancer. In multiple studies, explainable models matched or outperformed "black box" systems, achieving high diagnostic accuracy metrics, reducing false positives, and improving clinical confidence by allowing algorithmic decisions to be visually validated.

Conclusions: Explainable artificial intelligence not only increases interpretability, but also strengthens the performance and clinical acceptability of medical artificial intelligence, positioning itself as an essential component for the safe and effective implementation of AI-assisted cancer diagnostic systems in contemporary clinical practice and in future digitalized healthcare environments.

Keywords:

Medical decision support, early detection of cancer, cancer diagnosis, diagnostic imaging, Explainable artificial intelligence, XAI.



INTRODUCCIÓN

El cáncer constituye una de las principales causas de muerte a nivel mundial, con aproximadamente 20 millones de nuevos casos y 9,7 millones de muertes registradas en 2022. Aproximadamente, 1 de cada 5 personas desarrolla cáncer a lo largo de su vida. Además, 1 de cada 9 hombres y 1 de cada 12 mujeres mueren a causa de esta enfermedad.

⁽¹⁾ La detección temprana del cáncer es fundamental para mejorar el pronóstico y las tasas de supervivencia de los paciente.^(2,3)

En este sentido, la Inteligencia Artificial (IA) está transformando rápidamente el panorama de la investigación oncológica y el avance de las intervenciones clínicas personalizadas.⁽⁴⁾ Los algoritmos de aprendizaje automático (ML) y aprendizaje profundo (DL) han innovado el diagnóstico del cáncer, permitiendo la detección, localización y monitorización de la enfermedad de forma no invasiva, económica y escalable.⁽⁵⁾

En medicina, los ámbitos digitalizados, como la imagenología, se prestan a ser pioneros en la adopción de la IA.⁽⁶⁾ El proceso de adquisición de imágenes, reconstrucción, interpretación, elaboración de informes y comunicación de resultados se gestiona en un entorno digital, lo que facilita la captura y estructuración eficiente de datos relevantes para los sistemas de IA.⁽⁷⁾ En consecuencia, se ha experimentado una creciente adopción para tareas claves, como la mejora de los programas de cribado, la detección y caracterización de lesiones, así como la predicción de la respuesta al tratamiento y el pronóstico.⁽⁸⁾

El análisis tradicional de imágenes médicas presenta limitaciones significativas, que incluyen la subjetividad del diagnóstico, la variabilidad entre observadores y el tiempo considerable requerido para la interpretación.^(9,10) Por ello, los avances de la AI tienen un potencial extraordinario para mejorar la precisión diagnóstica en oncología, alcanzando niveles de exactitud superiores a 90 % en diversos tipos de cáncer.^(9,11,12)

No obstante, modelos como los basados en aprendizaje profundo funcionan como cajas negras que carecen de transparencia e interpretabilidad, lo que constituye una barrera fundamental para su adopción clínica.⁽¹³⁾ La opacidad de estos sistemas genera desconfianza entre los profesionales de la salud y plantea preocupaciones sobre la seguridad del paciente y la responsabilidad médica.⁽¹⁴⁾

Explainable Artificial Intelligence (XAI) emerge como una solución prometedora para abordar esta problemática al proporcionar interpretabilidad y transparencia a los modelos sin comprometer significativamente su rendimiento.^(15,16) La XAI comprende un conjunto de técnicas que permiten a las partes interesadas comprender mejor un algoritmo de IA y las decisiones que estos generan.⁽¹⁷⁾

Diversos estudios recientes han explorado el papel de XAI en aplicaciones específicas de diagnóstico por imágenes para cáncer. Ghasemi *et al.*⁽¹⁸⁾ revisaron estudios que utilizan XAI en la detección y predicción de riesgo de cáncer de mama, destacando que técnicas como SHAP son las más implementadas para explicar los resultados de clasificación y predicción de biomarcadores. De manera similar, investigaciones como las de Muhammad & Bendeche⁽¹⁹⁾ y Koutoulakis *et al.*,⁽²⁰⁾ enfocadas en el análisis de imágenes médicas, han identificado que los métodos XAI pueden abordar la necesidad de interpretabilidad en modelos complejos para mejorar su aplicabilidad clínica, destacando la importancia de explicar de manera efectiva la toma de decisiones de los algoritmos.

Sin embargo, persiste una limitada sistematización de la evidencia comparativa entre modelos explicables y no explicables aplicados al diagnóstico por imágenes para la detección temprana del cáncer. Esta situación dificulta la comprensión integral y una evaluación crítica de las fortalezas y debilidades de XAI en oncología, dificultando su implementación clínica efectiva. Por lo tanto, resulta necesario y pertinente una revisión sistemática de la literatura que identifique, evalúe y sintetice la evidencia disponible sobre el uso de XAI en el diagnóstico por imágenes oncológicas, con el fin de establecer una base científica sólida que oriente futuras investigaciones y favorezca su adopción en la práctica clínica.

El presente estudio tiene como objetivo analizar la evidencia bibliográfica disponible sobre la efectividad en el uso de las técnicas de Explainable Artificial Intelligence (XAI) aplicadas al diagnóstico por imágenes para la detección temprana de varios tipos de cáncer, mediante una revisión sistemática de literatura en casos de estudio encontrados en Scopus y Web of Science.

Las preguntas de investigación planteadas son: (1) ¿Cuáles son las principales técnicas de XAI utilizadas en el diagnóstico por imágenes para la detección temprana del cáncer?; (2) ¿Qué modalidades de imagen médica son frecuentes y en qué tipos de cáncer se han implementado las técnicas de XAI?; y (3) ¿Cuál es la efectividad de las técnicas de XAI en comparación con otros modelos de IA sin explicabilidad aplicadas al diagnóstico por imágenes para la detección temprana de varios tipos de cáncer?.

MATERIAL Y MÉTODOS

El presente estudio corresponde a una investigación documental con un diseño no experimental, transversal y descriptivo-analítico, caracterizado por la no manipulación directa de las variables. Se basó en la recopilación, evaluación crítica y síntesis de estudios primarios relacionados con el uso de XAI en el diagnóstico por imágenes para la detección temprana de varios tipos de cáncer.

Se adoptó un enfoque metodológico mixto, combinando un análisis cualitativo para examinar las características, aplicabilidad clínica y limitaciones de las técnicas de XAI, y un enfoque cuantitativo, para sistematizar y comparar las métricas diagnósticas reportadas en los estudios incluidos.

Para garantizar rigurosidad, transparencia, trazabilidad y reproducibilidad, se siguió el método de revisión sistemática de literatura conforme a la guía PRISMA 2020.⁽²¹⁾ La definición de los criterios de selección y el alcance de la investigación se basó en la estrategia PICO. (Tabla 1).

Tabla 1. Criterios PICO para la selección de los estudios	
Característica	Descripción
P (Población)	Pacientes sometidos a diagnóstico por imágenes médicas para la detección temprana de varios tipos de cáncer
I (Intervención)	Técnicas de XAI integradas en modelos de IA para análisis de imágenes
C (Comparador)	Modelos de IA sin explicabilidad o con técnicas convencionales (“caja negra”)
O (Resultados)	Mejoras en métricas diagnósticas, interpretabilidad y aceptación clínica

Criterios de elegibilidad

Los criterios de inclusión consideraron artículos científicos originales publicados entre 2020 y la fecha de búsqueda (07/01/2026), en los que se abordaran explícitamente técnicas de XAI en el análisis de imágenes médicas para la detección temprana del cáncer, y que reportaran resultados cuantitativos o cualitativos relacionados con el desempeño diagnóstico, la interpretabilidad del modelo o su comparación con enfoques de IA tradicionales. Además, no se planteó restricción de idioma.

Por otro lado, se excluyeron estudios duplicados, retractados, editoriales, cartas al editor, capítulos de libro o libros completos, actas de conferencias, revisiones y estudios sin acceso a texto completo. Asimismo, se excluyeron aquellos estudios que, tras la evaluación de la calidad metodológica mediante las herramientas QUADAS-2 y ROBINS-I, presentaron un riesgo de sesgo moderado-alto, o cuyo enfoque principal no contribuyera a responder las preguntas de investigación planteadas.

Estrategia de búsqueda

Las fuentes de información seleccionadas para esta revisión sistemática fueron las bases de datos bibliográficas *Scopus* y *Web of Science*, debido a su amplia cobertura de literatura científica de alto impacto en los campos de inteligencia artificial, Ciencias de la Computación y Ciencias de la Salud. Estas bases de datos garantizan el acceso a estudios revisados por pares y con alto rigor académico.

La estrategia de búsqueda se diseñó conforme a cuatro grupos conceptuales principales: “*explainable artificial intelligence*”; “*medical imaging*”; “*early diagnosis*”; “*cáncer*”. Para adaptar la sintaxis a las particularidades de cada base de datos, se emplearon términos con operadores booleanos (“AND”, “OR”), enfocándose la búsqueda en títulos, resúmenes y palabras claves. En la Tabla 2 se detalla la cadena de búsqueda que se aplicó en cada base de datos y los resultados obtenidos, la cual muestra un total de 85 estudios.

Tabla 2. Cadena de búsqueda para cada base de datos seleccionada		
Base de datos	Cadena de búsqueda	Estudios
Scopus	(TITLE-ABS-KEY ("explainable artificial intelligence" OR "explainable AI" OR "XAI" OR "interpretable AI" OR "model interpretability" OR "transparent AI") AND TITLE-ABS-KEY ("medical imaging" OR "diagnostic imaging" OR "radiology") AND TITLE-ABS-KEY ("early detection" OR "early diagnosis") AND TITLE-ABS-KEY ("cancer" OR "oncology" OR "tumor")) AND PUBYEAR > 2019 AND PUBYEAR < 2026 AND (LIMIT-TO (DOCTYPE, "ar"))	71
Web of Science	TS=("explainable artificial intelligence" OR "explainable AI" OR "XAI" OR "interpretable AI" OR "model interpretability" OR "transparent AI") AND TS=("medical imaging" OR "diagnostic imaging" OR "radiology") AND TS=("early detection" OR "early diagnosis") AND TS=("cancer" OR "oncology" OR "tumor")	14
Total		85

Proceso de selección de estudios

El proceso de selección de estudios se llevó a cabo siguiendo rigurosamente las fases establecidas por el protocolo PRISMA.⁽²¹⁾ En la fase de identificación, se recuperaron 71 registros desde Scopus y 14 desde Web of Science, sumando un total de 85 estudios. Posteriormente, se eliminaron 9 registros duplicados. En la fase de cribado, se examinaron 76 estudios mediante la revisión de títulos, resúmenes y palabras clave, excluyéndose 11 por no contribuir a los objetivos del estudio ni responder a las preguntas de investigación.

En la fase de elegibilidad, se solicitaron 65 estudios para su recuperación completa, de los cuales 6 no pudieron ser recuperados. Finalmente, se evaluaron 59 estudios a texto completo, excluyéndose 29 por presentar un alto riesgo de sesgo metodológico y 5 por no cumplir los criterios de inclusión establecidos para esta revisión. Como resultado del proceso de inclusión, un total de 25 estudios fueron seleccionados para la revisión sistemática final. La Figura 1 muestra el diagrama PRISMA, así como los estudios resultantes de cada etapa..

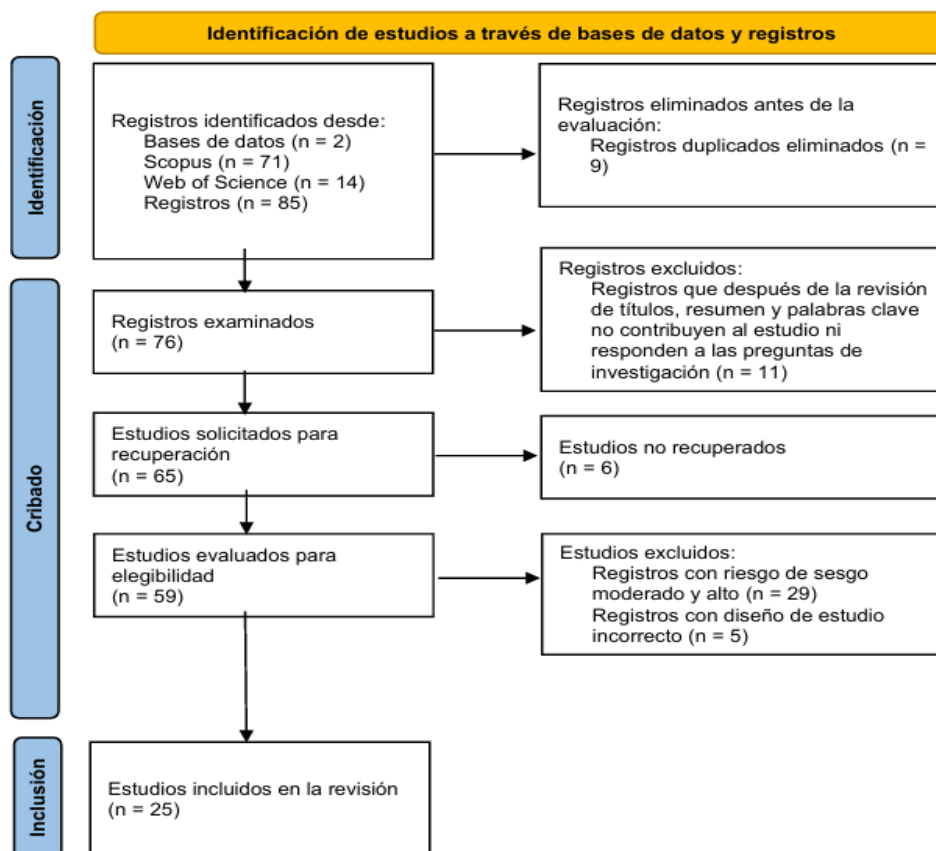


Figura 1: Diagrama PRISMA para la selección y descarte de los artículos Evaluación de la calidad de los estudios

La evaluación de la calidad metodológica y el riesgo de sesgo de los estudios identificados constituyó una fase fundamental de esta revisión sistemática, con el fin de garantizar la validez interna, la confiabilidad de los resultados y la solidez de las conclusiones. Para los estudios de validación diagnóstica, se aplicó la herramienta QUADAS-2, que reveló un riesgo de sesgo bajo en 25 estudios y moderado en 27. En los estudios de tipo observacional y experimental, se utilizó ROBINS-I presentando un riesgo de sesgo moderado en ambos estudios. Como criterio metodológico final, se incluyeron exclusivamente los 25 estudios con riesgo bajo, como se muestra en la Figura 2, descartando aquellos con riesgo moderado para preservar el rigor y la confiabilidad de la síntesis.

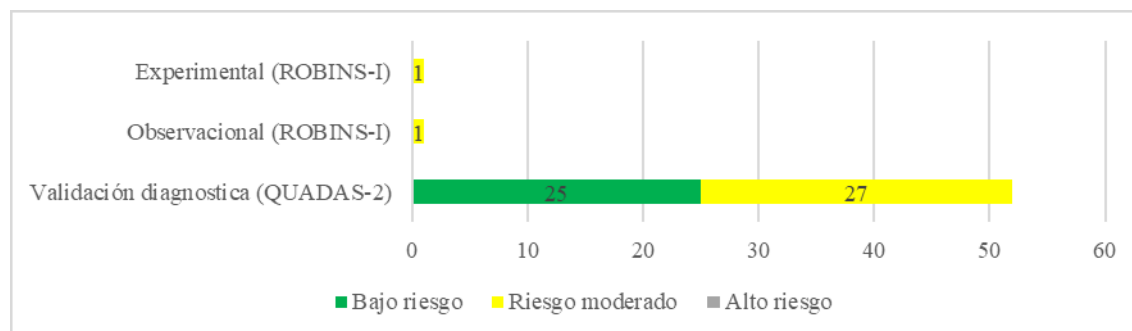


Figura 2. Evaluación de riesgo de sesgo

Métodos de síntesis

Para facilitar la organización y trazabilidad del proceso, se empleó los siguientes instrumentos y herramientas: Excel para la gestión de datos y Zotero para la gestión y verificación de referencias bibliográficas. Asimismo, se diseñó una matriz de extracción de datos que incluye información relevante como: autor, año, país, tipo de estudio, tipo de modelo XAI, tipo de cáncer, métricas de desempeño, interpretabilidad, comparación y principales hallazgos. Las respuestas se codificaron temáticamente de acuerdo con las preguntas de investigación y se agruparon en categorías analíticas que faciliten la sistematización y comparación transversal de la evidencia. Los resultados fueron sintetizados en forma narrativa y tabular

RESULTADOS

Los resultados de esta revisión sistemática se sintetizan mediante una matriz estructurada (Tabla 3), la cual permite comparar de manera sistemática las características metodológicas, las técnicas de XAI empleadas, los modelos de inteligencia artificial subyacentes y la efectividad diagnóstica reportada en los estudios incluidos. La matriz de resultados resume la evidencia empírica proveniente de los 25 estudios primarios publicados entre 2020 y 2025, abarcando distintos tipos de cáncer, modalidades de imagen médica y enfoques de explicabilidad. (Tabla 3).^(22,23,24,25,26,27,28,29,30,31,32,33,34,35,36,37,38,39,40,41,42,43,44,45,46)

Tabla 3. Matriz de resultados

Autor(es)	Conjunto de datos	Tipo de cáncer	Modalidad de imagen médica	Técnica XAI	IA subyacente	Otros modelos AI	Principales resultados
Abdel-Salam <i>et al</i> ⁽²¹⁾	SipakMed	Cáncer cervical	Frotis de Papanicolaou (Pap smear)	LIME	SHO-SVM	LDA, XGBoost, CatBoost	LIME identificó rasgos citomorfológicos clave, mejora la transparencia, detección de sesgos y la confianza clínica.
All <i>et al</i> ⁽²³⁾	7022 imágenes MRI	Tumores cerebrales	MRI cerebral	OSA	CNN	Grad-CAM, Saliency Maps	OSA mostró mayor estabilidad espacial y mejor alineación con expertos clínicos que otros métodos
Arshad <i>et al</i> ⁽²⁴⁾	HAM10000, ISIC2018	Cáncer de piel	Dermatoscopia	LIME, Grad-CAM	Network-level fused Deep architecture: 5-6 block inverted residual network.	Análisis del modelo explicable base	Las técnicas XAI localizaron eficazmente lesiones, validaron predicciones correctas e identificaron fallos de clasificación.
Choi <i>et al</i> ⁽²⁵⁾	97 muestras	Cáncer de próstata	mpMRI	Kernel SHAP	LR	RF, KNN, SVM, Linear SVM, ANN, MLP	SHAP facilitó decisiones clínicas basadas en evidencia y redujo biopsias innecesarias.
Cimino <i>et al</i> ⁽²⁶⁾	318 imágenes	OSCC	Fotografías clínicas (simple plain photographs)	Razonamiento Basado en Casos (CBR)	Faster R-CNN ResNet-50	Análisis del modelo explicable base	La combinación DL-CBR proporcionó alto rendimiento con explicaciones visuales validadas clínicamente.
Dahan <i>et al</i> ⁽²⁷⁾	Kvasir	Cáncer gástrico, colorrectal, de pulmón	Endoscopia	Grad-CAM	Swin Transformer + DCNN	Análisis del modelo explicable base	Mejóro la comprensión clínica y apoyó el diagnóstico temprano.
Gamage <i>et al</i> ⁽²⁸⁾	HAM10000, ISIC	Cáncer de piel	Dermatoscopia	Grad-CAM, Grad-CAM++	CNN, VIT	Grad-CAM vs Grad-CAM++	Grad-CAM++ mostró mejor localización, explicabilidad, cobertura regional y mayor concordancia
Gerbasi <i>et al</i> ⁽²⁹⁾	INbreast, CBIS-DDSM	Cáncer de mama	Mamografía	Grad-CAM, SHAP	ResNet18, VGG16	Análisis del modelo explicable base	XAI incrementó la confiabilidad y consistencia, apoyó casos diagnósticos fronterizos.
Guha <i>et al</i> ⁽³⁰⁾	5725 imágenes	Cáncer de ovario	Tomografía Computarizada (CT)	LIME, Grad-CAM, SHAP	ResNet60	GoogleNet, Inception-v4, ResNeXt, Inception-ResNet, ResNet16, Xception, VGG16, VGG19, ResNet50, EfficientNetB0	Evidenció diferencias por tipo tumoral, fortalece transparencia y utilidad clínica.
Haque <i>et al</i> ⁽³¹⁾	MRI: BraTS, Msoud, BR35H y SARTAJ	Tumores cerebrales	Resonancia magnética (MRI)	Grad-CAM	CNN - censor aware multi-class survival prediction deep learning classifier	Análisis del modelo explicable base	Evidenció diferencias por tipo tumoral, fortalece transparencia y utilidad clínica.
Hermoza <i>et al</i> ⁽³²⁾	NLST	Tumores cerebrales	Radiografía de tórax (CXR)	Grad-CAM	CNN - censor aware multi-class survival prediction deep learning classifier	Análisis del modelo explicable base	Mejora de la localización y predicción a través del aprendizaje.
Kakon <i>et al</i> ⁽³³⁾	BR35H	Tumores cerebrales	Resonancia magnética (MRI)	Grad-CAM++, LIME, SHAP	Deep Ensemble Meta-Learning	YOLOv7, ResNet50 y modelos híbridos CNN	El enfoque explicable alcanzó 99.83% de precisión y validó el razonamiento del modelo.
Kalita <i>et al</i> ⁽³⁴⁾	Cáncer de mama	Cáncer de mama	citología por aspiración con aguja fina (FNAC)	Grad-CAM	DL + ML híbrido	Grad-CAM reveló patrones biológicos relevantes y mejoró la confianza, transparencia y adopción clínica	Grad-CAM reveló patrones biológicos relevantes y mejoró la confianza, transparencia y adopción clínica

Tabla 3 cont. matriz de Resultados

Autor(es)	Conjunto de datos	Tipo de cáncer	Modalidad de imagen médica	Técnica XAI	IA subyacente	Otros modelos AI	Principales resultados
Khare <i>et al</i> (135)	3153 imágenes	Cáncer cervical	Colposcopia	Grad-CAM, LIME, CartoonX, Pixel RDE	Cervix-AID-Net	Métodos XAI: Grad CAM, LIME, CartoonX, Pixel RDE	CartoonX destacó por mayor relevancia clínica al resaltar la zona de transformación. Facilitó el entendimiento de errores del modelo.
Latha <i>et al</i> (136)	BUSI	Cáncer de mama	Ultrasonido (breast ultrasound imaging)	Grad-CAM	EfficientNet-B7	CNNs clásicas	XAI facilitó la validación clínica y la aceptación de modelos asistidos por IA.
Lee <i>et al</i> (137)	652 imágenes	Carcinoma de Células escamosas orales: OSCC	Fotografías de lengua	Grad-CAM, esparcidad (Sparsity analysis)	DCNN	Análisis del modelo explicable base	Las regiones explicadas coincidieron con áreas clínicamente relevantes, mejorando el diagnóstico
Liu <i>et al</i> (38)	LIDC-IDRI	Cáncer de pulmón	Tomografía computarizada torácica (Chest CT)	Grad-CAM	Transformer + Multiple Instance Learning (MIL)	Análisis del modelo explicable base	XAI basada en atributos mejoró la precisión (93,05%) y la interpretabilidad causal.
Oumlaz <i>et al</i> (138)	National Library of Medicine dataset (NLIM)	Cáncer de pulmón	Radiografías de Tórax, Angiografía por Tomografía Computarizada Pulmonar	SHAP	EfficientNet Transformer	Análisis del modelo explicable base	SHAP cerró la brecha entre modelos de caja negra y diagnósticos clínicamente interpretables.
Oviedo <i>et al</i> (140)	RM multicéntrica	Cáncer de mama	Resonancia Magnética de Mama con contraste Dinámico,	Grad-CAM	fully convolutional data description (FCDD)	Binary Cross Entropy (BCE)	El modelo explicable redujo falsos positivos y generó mapas más precisos para radiólogos.
Rajpoot <i>et al</i> (41)	Kaggle MRI, BraTS	Tumores cerebrales	Resonancia Magnética (MRI)	Grad-CAM, Gradient Attention Rollout	BioTransX	CNNs	El modelo híbrido mejoró fidelidad explicativa y localización tumoral
Reaejo <i>et al</i> (42)	1263 pacientes	Cáncer de páncreas	Tomografías Computarizadas (CT) con contraste (CECT)	SHAP y t-SNE	XGBoost, LightGBM, CatBoost, Random Forest, Gradient Boosting	Análisis del modelo explicable base	SHAP identificó biomarcadores radiómicos y mejoró la interpretabilidad a nivel de paciente.
Shariyat <i>et al</i> (43)	NSCLC	Cáncer de pulmón	Tomografía Computarizada, Tomografía por Emisión de Positrones y de Baja Dosis	Grad-CAM	ResNet	VIT, ResNet	XAI vinculó morfología nodular con mutaciones genéticas clínicamente relevantes. Fomenta la confianza y la aceptación clínica
Singh <i>et al</i> (144)	UK Biobank	Hematológicos	Imagen retinal (retinal imaging)	RetHemo	Modelo de regresión de Riesgos Proporcional de Cox	CNNs, Xgboost	XAI vinculó morfología nodular con mutaciones genéticas clínicamente relevantes. Fomenta la confianza y la aceptación clínica
Wang <i>et al</i> (145)	CBIS-DDSM	Cáncer de mama	Mamografía	Grad-CAM	TEBIS	Unet, Res-Basic-Unet, AU-Net, Density-ASP, Swin-Unet, VIT, Inceptionv3	XAI mejoró la interpretación clínica y guía la optimización del modelo.
Wickstrøm <i>et al</i> (146)	EndoScene, MICCAI	Cáncer colorrectal (CRC)	Colonoscopia	Guided Backpropagation	FCNs	SDEM, U-Net, SegNet	La explicabilidad permitió relacionar inercidumbre predictiva con precisión diagnóstica.

A partir de la síntesis de los resultados, se identificaron las principales técnicas XAI utilizadas en el diagnóstico por imágenes para la detección del cáncer. La Tabla 4 sintetiza las más empleadas, los modelos de IA subyacentes, las métricas utilizadas para evaluar tanto la explicabilidad como el desempeño predictivo, así como los autores que reportan su aplicación. Se destaca que la mayoría de los estudios evalúan las técnicas XAI principalmente mediante análisis cualitativo y validación por expertos clínicos, complementados en menor medida por métricas cuantitativas de fidelidad y estabilidad de las explicaciones. (Tabla 4)^(22,23,24,25,26,27,28,29,30,31,32,33,34,35,36,37,38,39,40,41,42,43,44,45,46).

Tabla 4. Técnicas XAI para el diagnóstico por imágenes				
Técnica XAI	Evaluación técnica XAI	Modelo de IA subyacente	Métricas de desempeño del modelo	Referencias
Grad-CAM / Grad-CAM++	Evaluación cualitativa por expertos, coherencia espacial, localización de regiones relevantes	CNN, ResNet, EfficientNet, ViT, modelos híbridos	Accuracy, AUC, Sensitivity, Specificity, F1-score	(24, 27, 28, 29, 31, 32, 33, 34, 35, 36, 37, 38, 40, 41, 43, 45, 46)
LIME	Fidelidad local, interpretabilidad de superpíxeles, concordancia clínica	CNN, SVM, XGBoost, modelos híbridos	Accuracy, Precision, Recall, AUC	(22,24,30,33,35)
SHAP / Kernel SHAP	Importancia de características, consistencia global, estabilidad de explicaciones	Random Forest, XGBoost, CNN, Transformers	Accuracy, AUC, ROC, Log-loss	(25,29,30,33,38,39,42)
Guided Backpropagation	Nitidez de mapas de relevancia, alineación visual	FCN, CNN	Accuracy, Dice, IoU	(46)
Gradient Rollout / Attention-based XAI	Distribución de atención, coherencia semántica	Transformers, modelos híbridos CNN-Transformer	Accuracy, AUC	(38,41)
Case-Based Reasoning (CBR)	Validación clínica de casos similares, trazabilidad de decisiones	Faster R-CNN, ResNet	Accuracy, Precision, Recall	(26)
CartoonX	Relevancia clínica estructural, reducción de ruido explicativo	CNN especializados	Accuracy, AUC	(35)
OSA / métodos propios	Estabilidad espacial, consistencia intercasos	CNN personalizados	Accuracy, Sensitivity	(23)
RetHemo / XAI específica de dominio	Interpretabilidad biológica, identificación de patrones clínicos	CNN, XGBoost	AUC, Accuracy	(44)

Aunque muchos estudios analizados utilizan la XAI de forma visual y cualitativa, (30,33) existen métricas específicas diseñadas para evaluar el rendimiento, la fidelidad y la exactitud espacial. Estas métricas son principalmente cuantitativas, cualitativas o centradas en el usuario. (Tabla 5).^(23,28,32,33,34,40,41).

Tabla 5. Métricas de evaluación para técnicas XAI			
Enfoque	Descripción	Métricas	Referencias
Métricas cuantitativas de Localización Espacial	Evalúa qué tan bien se alinean los mapas de calor generados por la técnica XAI con las anotaciones reales (ground truth)	Intersección sobre Unión (IoU); Coeficiente Dice Continuo (cDice); AUC a nivel de pixel (Pixelwise AUC); AUC de precisión por eliminación de píxeles (AUC de perturbación)	(28,32,33,40,41)
Métricas Cualitativas y Centradas en el Usuario	Dado que el objetivo final de la XAI es generar confianza médica, se aplican métricas enfocadas en la usabilidad	Tasa de concordancia o validación por expertos; Escala de Usabilidad del Sistema (SUS - System Usability Scale); Cobertura, precisión y recall de las explicaciones	(23,28,34)

Por lo tanto, las técnicas de XAI han demostrado no solo igualar, sino a menudo superar la efectividad de los modelos de IA tradicionales ("caja negra") al proporcionar un equilibrio crítico entre precisión diagnóstica y confianza clínica.^(36,38) Mientras que los modelos convencionales suelen carecer de transparencia, lo que limita su adopción en entornos médicos reales, las herramientas de XAI permiten a los clínicos validar las decisiones del modelo, reduciendo errores y mejorando la detección temprana de diversas patologías oncológicas.^(23,35)

Se destaca la efectividad de estas tecnologías en comparación con modelos sin explicabilidad desde varios enfoques. La mejora en la precisión y reducción de falsos positivos, muchas superando arquitecturas tradicionales como VGG y ResNet que suelen tener dificultades con los desbalances de clases.⁽³⁶⁾ Además, los marcos de trabajo que integran capas de XAI han reportado precisiones de hasta 99,85 %.^(22,35)

En el cáncer de pulmón, los modelos de XAI entrenados con supervisión débil pueden localizar tumores antes que sean clínicamente detectables en radiografías de tórax de pacientes asintomáticos, algo que los modelos de regresión simples no logran con la misma eficacia.⁽³²⁾ Técnicas como CartoonX han demostrado ser más meticulosas que LIME o RDE al extraer partes suaves de la imagen (como texturas de la zona de transformación cervical), revelando patrones que no son visibles con métodos estándar de mapeo de características.⁽³⁵⁾

En el adenocarcinoma pancreático, el uso de XAI (SHAP) permitió identificar biomarcadores en tejidos sanos adyacentes, logrando un AUC de 97,98 % sin depender de la visibilidad directa del tumor, lo cual supera el paradigma centrado únicamente en la lesión.⁽⁴²⁾ La efectividad de las técnicas de XAI reside en su capacidad para desmitificar la "caja negra", permitiendo que modelos de alto rendimiento técnico sean clínicamente procesables al facilitar la detección temprana, reducir el agotamiento de los especialistas y priorizar la seguridad del paciente mediante la transparencia diagnóstica.^(23,33,36)

Además, las técnicas de XAI no solo sirven para interpretar resultados, sino que actúan como un mecanismo de auditoría que ayuda a identificar y rectificar sesgos potenciales en el modelo.⁽²⁸⁾ Los conjuntos de entrenamiento que no son representativos pueden sesgar la arquitectura de aprendizaje profundo y provocar clasificaciones inexactas.⁽²⁶⁾ Al hacer transparentes los procesos internos, las explicaciones visuales permiten a los desarrolladores y médicos detectar derivas del modelo o sesgos de entrenamiento inesperados.⁽²⁶⁾

Esto permite refinar los modelos para que sean más robustos, reduzcan falsos positivos o negativos y puedan generalizarse mejor en diversas poblaciones de pacientes.⁽²²⁾ A su vez, se advierte que la integración de estos modelos en el flujo de trabajo debe ser cuidadosa para evitar que los propios médicos desarrollen un sesgo o dependencia excesiva hacia los mapas de calor generados por la IA.⁽⁴⁰⁾

DISCUSIÓN

La presente revisión sistemática tuvo como objetivo analizar de manera integral el papel de la *Explainable Artificial Intelligence* (XAI) en el diagnóstico por imágenes para la detección temprana del cáncer. Los resultados muestran que las técnicas XAI mejoran sistemáticamente la transparencia, la confianza clínica y el rendimiento diagnóstico. Los métodos basados en mapas de activación, especialmente Grad-CAM y sus variantes, fueron los más utilizados y demostraron alta capacidad para localizar regiones tumorales relevantes. Se observó también que la integración de explicabilidad no reduce la precisión de los modelos de AI, sino que frecuentemente la incrementa y facilita su validación clínica. Finalmente, se evidenció que la interpretabilidad favorece la aceptación por parte de especialistas, particularmente en casos limítrofes o diagnósticos complejos.

El análisis detallado de los estudios incluidos revela una correlación consistente entre el tipo de técnica XAI empleada y la mejora de métricas diagnósticas específicas. Abdel-Salam *et al.*⁽²²⁾ evidenciaron que LIME permitió identificar rasgos citomorfológicos clave en citologías cervicales sin comprometer la precisión, mientras que Choi *et al.*⁽²⁵⁾ demostraron que SHAP incrementó el AUC y redujo biopsias innecesarias en cáncer de próstata. De forma similar, Guha *et al.*⁽³⁰⁾ reportaron una precisión de 97.5 % en cáncer de ovario al combinar múltiples técnicas explicables, y Liu *et al.*⁽³⁸⁾ alcanzaron 93,05 % de exactitud en cáncer pulmonar mediante Grad-CAM integrado con aprendizaje múltiple de instancias.

Asimismo, diversos autores observaron que la elección de la técnica XAI influye directamente en la calidad de la interpretación clínica. Ali *et al.*⁽²³⁾ encontraron que la técnica OSA mostró mayor estabilidad espacial que mapas de gradiente convencionales en tumores cerebrales, mientras que Gamage *et al.*⁽²⁸⁾ evidenciaron que Grad-CAM++ superó a Grad-CAM en concordancia con la verdad fundamental. Khare *et al.*⁽³⁵⁾ identificaron que CartoonX resaltó regiones clínicamente relevantes con mayor precisión que otros métodos, y Rezaei *et al.*⁽⁴²⁾ demostraron que SHAP facilitó la detección de biomarcadores radiómicos interpretables a nivel de paciente. Estos hallazgos confirman que la explicabilidad no es homogénea entre técnicas, sino que su efectividad depende del tipo de imagen, del modelo subyacente y de la tarea diagnóstica.

Estos resultados coinciden con investigaciones recientes que destacan el valor clínico de la XAI en imagenología oncológica. Por ejemplo, Ahmed *et al.*⁽⁴⁷⁾ señalan que la explicabilidad mejora la interpretabilidad y la confianza médica al revelar patrones utilizados por los modelos, lo cual es esencial para su integración clínica. Asimismo, Abas *et al.*⁽⁴⁸⁾ reportan que los métodos XAI permiten comprender decisiones diagnósticas complejas y facilitan la validación por especialistas, especialmente en modelos profundos de clasificación médica. En el contexto aplicado, Gurmessa y Jimma⁽⁴⁹⁾ demostraron que técnicas explicables mejoran la exactitud y transparencia en la detección automatizada de cáncer de mama, mientras que Ghasemi *et al.*⁽⁵⁰⁾ documentaron beneficios similares en clasificación de tumores cerebrales mediante aprendizaje profundo interpretable.

En el ámbito regional y comparativo, estudios recientes también respaldan la relevancia de la XAI en sistemas diagnósticos. Ahmed *et al.*⁽⁴⁷⁾ concluyeron que la explicabilidad es un requisito clave para la adopción clínica segura de modelos de aprendizaje automático en salud. Por su parte, Purwono *et al.*⁽⁵¹⁾ demostraron que marcos explicables mejoran la confiabilidad en diagnóstico médico automatizado, mientras que Shifa *et al.*⁽⁵²⁾ evidenciaron que los modelos explicables permiten comprender la influencia de variables clínicas en predicciones, fortaleciendo la toma de decisiones médicas. La convergencia entre estos estudios y los resultados de la presente revisión indica que existe consenso científico respecto a que la interpretabilidad no es solo un atributo deseable, sino un componente esencial para la validación clínica de la inteligencia artificial diagnóstica.

Desde una perspectiva práctica, los hallazgos implican que la incorporación sistemática de XAI en sistemas de diagnóstico por imágenes puede facilitar su implementación hospitalaria, mejorar la comunicación médico-paciente y optimizar procesos de cribado oncológico. La capacidad de visualizar regiones relevantes, identificar biomarcadores y justificar predicciones reduce la resistencia clínica hacia modelos de “caja negra” y promueve su uso como herramientas de apoyo a la decisión, especialmente en entornos con alta carga diagnóstica o escasez de especialistas.

No obstante, el estudio presenta **limitaciones**. La heterogeneidad metodológica entre los trabajos incluidos, en términos de datasets, métricas y arquitecturas, dificulta la comparación directa de resultados y limita la posibilidad de realizar metaanálisis cuantitativos. Además, la mayoría de estudios se basan en conjuntos de datos públicos o retrospectivos, lo que puede afectar la generalización clínica. Como líneas futuras, se recomienda desarrollar evaluaciones multicéntricas prospectivas, estandarizar métricas de explicabilidad y explorar enfoques híbridos que integren interpretabilidad intrínseca y post-hoc, con el fin de consolidar la evidencia y acelerar la transferencia de la XAI al entorno clínico real.

CONCLUSIONES

Las técnicas de XAI más utilizadas en diagnóstico oncológico por imágenes incluyen métodos basados en mapas de activación y atribución de características, como Grad-CAM, LIME y SHAP, debido a su capacidad para proporcionar interpretaciones visuales y comprensibles de las predicciones. Asimismo, técnicas emergentes como CartoonX, métodos de atención y enfoques basados en casos muestran potencial creciente.

Su aplicación abarca diversas modalidades de imagen médica como resonancia magnética, tomografía computarizada, mamografía, dermatoscopia y radiografía de tórax, enfocadas principalmente a cáncer de pulmón, mama, cerebro, piel y cuello uterino

En conjunto, la evidencia indica que los modelos basados en XAI alcanzan un rendimiento igual o superior a los enfoques no explicables, contribuyendo a mejorar la precisión diagnóstica, reducir sesgos y fortalecer la confianza clínica, lo que favorece su adopción en la práctica médica.

REFERENCIAS BIBLIOGRÁFICAS

1. World Health Organization. Global cancer burden growing, amidst mounting need for services [Internet]. Ginebra: WHO; 2024. [Citado 26/06/2025]. Disponible en: <https://www.who.int/news/item/01-02-2024-global-cancer-burden-growing--amidst-mounting-need-for-services>
2. Crosby D, Bhatia S, Brindle KM, Coussens LM, Dive C, Emberton M, et al. Early detection of cancer. *Science* [Internet]. 2022;375(6586):eaay9040. Disponible en: <http://doi.org/10.1126/science.aay9040>
3. World Health Organization. Cancer [Internet]. Ginebra: WHO; 2025. [Citado 26/06/2025]. Disponible en: <https://www.who.int/news-room/fact-sheets/detail/cancer>
4. Cheng CH, Shi S sheng. Artificial intelligence in cancer: applications, challenges, and future perspectives. *Mol Cancer* [Internet]. 2025;24(1):274. Disponible en: <http://doi.org/10.1186/s12943-025-02450-3>
5. Swanson K, Wu E, Zhang A, Alizadeh AA, Zou J. From patterns to patients: Advances in clinical machine learning for cancer diagnosis, prognosis, and treatment. *Cell* [Internet]. 2023;186(8):1772–91. Disponible en: <http://doi.org/10.1016/j.cell.2023.01.035>
6. Miao KH, Miao JH, Finkelstein M, Chatterjee A, Oto A. Next-Generation Advances in Prostate Cancer Imaging and Artificial Intelligence Applications. *J Imaging* [Internet]. 2025;11(11):390. Disponible en: <http://doi.org/10.3390/jimaging11110390>
7. Koh DM, Papanikolaou N, Bick U, Illing R, Kahn CE, Kalpathi-Cramer J, et al. Artificial intelligence and machine learning in cancer imaging. *Commun Med* [Internet]. 2022;2(1):133. Disponible en: <http://doi.org/10.1038/s43856-022-00199-0>
8. Cellina M, Cè M, Irmici G, Ascenti V, Khenkina N, Toto-Brocchi M, et al. Artificial Intelligence in Lung Cancer Imaging: Unfolding the Future. *Diagnostics* [Internet]. 2022;12(11):2644. Disponible en: <http://doi.org/10.3390/diagnostics12112644>
9. Hammad M, ElAffendi M, El-Latif AAA, Ateya AA, Ali G, Plawiak P. Explainable AI for lung cancer detection via a custom CNN on CT images. *Sci Rep* [Internet]. 2025;15(1):12707. Disponible en: <http://doi.org/10.1038/s41598-025-97645-5>
10. Marias K. The Constantly Evolving Role of Medical Image Processing in Oncology: From Traditional Medical Image Processing to Imaging Biomarkers and Radiomics. *J Imaging* [Internet]. 2021;7(8):8. Disponible en: <http://doi.org/10.3390/jimaging7080124>
11. Agarwal D. Explainable AI in Cancer Diagnosis: Enhancing Interpretability with SHAP on Benign and Malignant Tumor Detection. *Int J Res Appl Sci Eng Technol* [Internet]. 2025;13(1):1394–402. Disponible en: <http://doi.org/10.22214/ijraset.2025.66580>
12. Alom MR, Farid FA, Rahaman MA, Rahman A, Debnath T, Miah ASM, et al. An explainable AI-driven deep neural network for accurate breast cancer detection from histopathological and ultrasound images. *Sci Rep* [Internet]. 2025;15(1):17531. Disponible en: <http://doi.org/10.1038/s41598-025-97718-5>
13. Hrinivich WT, Wang T, Wang C. Editorial: Interpretable and explainable machine learning models in oncology. *Front Oncol* [Internet]. 2023;13. Disponible en: <http://doi.org/10.3389/fonc.2023.1184428>
14. Tanveer H, Faheem M, Khan AH. Explainable AI in Medical Decision-Making: Challenges and Opportunities [Internet]. 2022;5(12) [Citado 26/06/2025]. Disponible en: <https://www.irejournals.com/formatedpaper/1703509.pdf>
15. Olumuyiwa BI, Han TA, Shamszaman ZU. Enhancing Cancer Diagnosis with Explainable & Trustworthy Deep Learning Models [Internet]. EE UU: arXiv; 2024 [Citado 26/06/2025]. Disponible en: <http://doi.org/10.48550/arXiv.2412.17527>

16. Singhal A, Agrawal KK, Quezada A, Aguiñaga AR, Jiménez S, Yadav SP. Explainable Artificial Intelligence (XAI) Model for Cancer Image Classification. *Comput Model Eng Sci* [Internet]. 2024;141(1):401–41. Disponible en: <http://doi.org/10.32604/cmesci.2024.051363>
17. Wyatt LS, van Kernenbeek LM, Wijkhuizen M, Geldof F, Dashtbozorg B. Explainable Artificial Intelligence (XAI) for Oncological Ultrasound Image Analysis: A Systematic Review. *Appl Sci* [Internet]. 2024;14(18):18. Disponible en: <http://doi.org/10.3390/app14188108>
18. Ghasemi A, Hashtarkhani S, Schwartz DL, Shaban-Nejad A. Explainable artificial intelligence in breast cancer detection and risk prediction: A systematic scoping review. *Cancer Innov* [Internet]. 2024;3(5):e136. Disponible en: <http://doi.org/10.1002/cai2.136>
19. Muhammad D, Bendeche M. Unveiling the black box: A systematic review of Explainable Artificial Intelligence in medical image analysis. *Comput Struct Biotechnol J* [Internet]. 2024;24:542–60. Disponible en: <http://doi.org/10.1016/j.csbj.2024.08.005>
20. Koutoulakis E, Trivizakis E, Markodimitrakis E, Agelaki S, Tsiknakis M, Marias K. A critical review of explainable deep learning in lung cancer diagnosis. *Artif Intell Rev* [Internet]. 2025;59(1):28. Disponible en: <http://doi.org/10.1007/s10462-025-11445-x>
21. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* [Internet]. 2021;372(71). Disponible en: <http://doi.org/10.1136/bmj.n71>
22. Abdel-Salam M, Askr H, Hassanien AE. Revolutionizing cervical cancer detection: a new optimized explainable artificial intelligence model. *Neural Comput Appl* [Internet]. 2025;37(29):23979–4023. Disponible en: <http://doi.org/10.1007/s00521-025-11548-0>
23. Ali SA, Arain KR, Mushtaq NA, Rehman O. Interpretable Deep Learning for Brain Tumor Diagnosis: Occlusion Sensitivity-Driven Explainability in MRI Classification. *VFAST Trans Softw Eng* [Internet]. 2025;13(2):135–46. Disponible en: <http://doi.org/10.21015/vtse.v13i2.2082>
24. Arshad M, Khan MA, Almujaally NA, Alasiry A, Marzougui M, Nam Y. Multiclass skin lesion classification and localization from dermoscopic images using a novel network-level fused deep architecture and explainable artificial intelligence. *BMC Med Inform Decis Mak* [Internet]. 2025;25(1):215. Disponible en: <http://doi.org/10.1186/s12911-025-03051-2>
25. Choi JY, Park S, Shim JS, Park HJ, Kuh SU, Jeong Y, et al. Explainable artificial intelligence-driven prostate cancer screening using exosomal multi-marker based dual-gate FET biosensor. *Biosens Bioelectron* [Internet]. 2025;267:116773. Disponible en: <http://doi.org/10.1016/j.bios.2024.116773>
26. Cimino MGCA, Campisi G, Galatolo FA, Neri P, Tozzo P, Parola M, et al. Explainable screening of oral cancer via deep learning and case-based reasoning. *Smart Health* [Internet]. 2025;35:100538. Disponible en: <http://doi.org/10.1016/j.smhl.2024.100538>
27. Dahan F, Shah JH, Saleem R, Hasnain M, Afzal M, Alfakih TM. A hybrid XAI-driven deep learning framework for robust GI tract disease diagnosis. *Sci Rep* [Internet]. 2025;15(1):21139. Disponible en: <http://doi.org/10.1038/s41598-025-07690-3>
28. Gamage L, Isuranga U, Meedeniya D, De Silva S, Yogarajah P. Melanoma Skin Cancer Identification with Explainability Utilizing Mask Guided Technique. *Electronics* [Internet]. 2024;13(4):680. Disponible en: <http://doi.org/10.3390/electronics13040680>
29. Gerbasi A, Clementi G, Corsi F, Albasini S, Malovini A, Quaglini S, et al. DeepMiCa: Automatic segmentation and classification of breast MicroCalcifications from mammograms. *Comput Methods Programs Biomed* [Internet]. 2023;235:107483. Disponible en: <http://doi.org/10.1016/j.cmpb.2023.107483>
30. Guha S, Kodipalli A, Fernandes SL, Dasar S. Explainable AI for Interpretation of Ovarian Tumor Classification Using Enhanced ResNet50. *Diagnostics* [Internet]. 2024;14(14):1567. Disponible en: <http://doi.org/10.3390/diagnostics14141567>
31. Haque R, Khan MA, Rahman H, Khan S, Siddiqui MIH, Limon ZH, et al. Explainable deep stacking ensemble model for accurate and transparent brain tumor diagnosis. *Comput Biol Med* [Internet]. 2025;191:110166. Disponible en: <http://doi.org/10.1016/j.compbiomed.2025.110166>
32. Hermoza R, Nascimento JC, Carneiro G. Weakly-supervised preclinical tumor localization associated with survival prediction from lung cancer screening Chest X-ray images. *Comput Med Imaging Graph* [Internet]. 2024;115:102395. Disponible en: <http://doi.org/10.1016/j.compmedimag.2024.102395>

33. Kakon SC, Sazid ZA, Begum IA, Samad MA, Hosen ASMS. Explainable Deep Ensemble Meta-Learning Framework for Brain Tumor Classification Using MRI Images. *Cancers* [Internet]. 2025;17(17):2853. Disponible en: <http://doi.org/10.3390/cancers17172853>
34. Kalita M, Mahanta LB, Das AK, Laskar D. An Interpretable Hybrid AI Model for Breast Fine Needle Aspiration Cytology Image Classification. *J Med Syst* [Internet]. 2025;49(1):182. Disponible en: <http://doi.org/10.1007/s10916-025-02267-z>
35. Khare SK, Booth BB, Blanes-Vidal V, Petersen LK, Nadimi ES. An explainable attention model for cervical precancer risk classification using colposcopic images. *Comput Methods Programs Biomed* [Internet]. 2025;271:108976. Disponible en: <http://doi.org/10.1016/j.cmpb.2025.108976>
36. Latha M, Kumar PS, Chandrika RR, Mahesh TR, Kumar VV, Guluwadi S. Revolutionizing breast ultrasound diagnostics with EfficientNet-B7 and Explainable AI. *BMC Med Imaging* [Internet]. 2024;24(1):230. Disponible en: <http://doi.org/10.1186/s12880-024-01404-3>
37. Lee YH, Jeon S, Jung J, Auh QS, Lee JS, Chaurasia A, et al. DCNN models with post-hoc interpretability for the automated detection of glossitis and OSCC on the tongue. *Sci Rep* [Internet]. 2025;15(1):31940. Disponible en: <http://doi.org/10.1038/s41598-025-16760-5>
38. Liu H, She Q, Lin J, Chen Q, Fang F, Zhang Y. Attribute and Malignancy Analysis of Lung Nodule on Chest CT with Cause-and-Effect Logic. *J Med Biol Eng* [Internet]. 2024;44(5):763–76. Disponible en: <http://doi.org/10.1007/s40846-024-00895-3>
39. Oumlaz M, Oumlaz Y, Oukaira A, Benelhaouare AZ, Lakhssassi A. Advancing Pulmonary Nodule Detection with ARSGNet: EfficientNet and Transformer Synergy. *Electronics* [Internet]. 2024;13(22):4369. Disponible en: <http://doi.org/10.3390/electronics13224369>
40. Oviedo F, Kazerouni AS, Liznerski P, Xu Y, Hirano M, Vandermeulen RA, et al. Cancer Detection in Breast MRI Screening via Explainable AI Anomaly Detection. *Radiology* [Internet]. 2025;316(1):e241629. Disponible en: <http://doi.org/10.1148/radiol.241629>
41. Rajpoot R, Jain S, Semwal VB. BioTransX: A novel bi-former based hybrid model with bi-level routing attention for brain tumor classification with explainable insights. *Comput Biol Med* [Internet]. 2025;195:110515. Disponible en: <http://doi.org/10.1016/j.combiomed.2025.110515>
42. Rezaei SM, Eftekhar A, Rouhi S, Keshavarzi B, Mohammadi Z, Firouzabad LA, et al. Neighboring tissues as diagnostic windows: Neighborhood effects in radiomic detection of pancreatic ductal adenocarcinoma. *Comput Methods Programs Biomed* [Internet]. 2025;272:109056. Disponible en: <http://doi.org/10.1016/j.cmpb.2025.109056>
43. Shariaty F, Pavlov V, Baranov M. AI-Driven Precision Oncology: Integrating Deep Learning, Radiomics, and Genomic Analysis for Enhanced Lung Cancer Diagnosis and Treatment. *Signal Image Video Process* [Internet]. 2025;19(9):693. Disponible en: <http://doi.org/10.1007/s11760-025-04244-y>
44. Singh A, Nooka AK, Modanwal G, Jain N, Dhodapkar MV, Arepalli S, et al. AI-informed retinal biomarkers predict 10-year risk of onset of multiple hematological malignancies. *Eur J Cancer* [Internet]. 2025;229. Disponible en: <http://doi.org/10.1016/j.ejca.2025.115752>
45. Wang H, Wei L, Liu B, Li J, Li J, Fang J, et al. Transformer-Based Explainable Model for Breast Cancer Lesion Segmentation. *Appl Sci* [Internet]. 2025;15(3):1295. Disponible en: <http://doi.org/10.3390/app15031295>
46. Wickstrøm K, Kampffmeyer M, Jenssen R. Uncertainty and interpretability in convolutional neural networks for semantic segmentation of colorectal polyps. *Med Image Anal* [Internet]. 2020;60:101619. Disponible en: <http://doi.org/10.1016/j.media.2019.101619>
47. Ahmed F, Naz NS, Khan S, Rehman AU, Ismael WM, Khan MA. Explainable artificial intelligence (XAI) in medical imaging: a systematic review of techniques, applications, and challenges. *BMC Med Imaging* [Internet]. 2026;26(1):37. Disponible en: <http://doi.org/10.1186/s12880-025-02118-w>
48. Abas Mohamed Y, Ee Khoo B, Shahrimie Mohd Asaari M, Ezane Aziz M, Rahiman Ghazali F. Decoding the black box: Explainable AI (XAI) for cancer diagnosis, prognosis, and treatment planning-A state-of-the art systematic review. *Int J Med Inf* [Internet]. 2025;193:105689. Disponible en: <http://doi.org/10.1016/j.ijmedinf.2024.105689>
49. Gurmessa Dk, Jimma W. Explainable machine learning for breast cancer diagnosis from mammography and ultrasound images: a systematic review. *BMJ Health Care Inform* [Internet]. 2024;31(1):e100954. Disponible en: <http://doi.org/10.1136/bmjhci-2023-100954>
50. Ghasemi A, Hashtarkhani S, Schwartz DL, Shaban-Nejad A. Explainable artificial intelligence in breast cancer detection and risk prediction: A systematic scoping review. *Cancer Innov* [Internet]. 2024 Jul 3;3(5):e136. Disponible en: <http://doi.org/10.1002/cai2.136>

51. Purwono P, Wulandari ANE, Nisa K. Explainable Artificial Intelligence (XAI) in Medical Imaging: Techniques, Applications, Challenges, and Future Directions. Adv Mech Mechatron Syst [Internet]. 2025;1(1):1. Disponible en: <http://doi.org/10.53623/amms.v1i1.692>

52. Shifa N, Saleh M, Akbari Y, Maadeed SA. A review of explainable AI techniques and their evaluation in mammography for breast cancer screening. Clin Imaging [Internet]. 2025;123. Disponible en: <http://doi.org/10.1016/j.clinimag.2025.110492>

Financiamiento:

Los autores declaran no tener ninguna fuente de financiamiento para el desarrollo de esta investigación.

Conflicto de intereses

Los autores declaran que no existe conflicto de intereses.

Contribución de autoría

Walter Fabián Taday Guashpa: Conceptualización, metodología, investigación, análisis formal, curación de datos, software.

Jaime David Camacho Castillo: Recursos, validación, visualización, adquisición de fondos, administración del proyecto, supervisión, redacción (borrador original), redacción (revisión y edición).

Ambos autores participamos en la discusión de los resultados y hemos leído, revisado y aprobado el texto final.