Vulnerabilidades de los sistemas de inteligencia artificial en radiología digital frente a ataques adversariales y su impacto en la confiabilidad diagnóstica y la seguridad del paciente

Avir Sophia Avila Abril

Laura Melissa Cruz Avellaneda

Sergio Andres Diaz Bello

Yeison Vladimir Martínez Olaya

Kely Biviana Tibocha Correa

Asesor

Christian Camilo Rodríguez Castro

Universidad Nacional Abierta y a Distancia- UNAD

Escuela Ciencias de la Salud- ECISA

Tecnología en radiología e imágenes diagnosticas

2026

Resumen

La incorporación de sistemas de inteligencia artificial en la radiología digital ha permitido optimizar los procesos diagnósticos mediante el uso de algoritmos de aprendizaje profundo capaces de identificar patologías con alta precisión. Sin embargo, estos sistemas presentan vulnerabilidades frente a ataques adversariales, los cuales consisten en la manipulación intencional de imágenes médicas mediante perturbaciones imperceptibles que pueden alterar los resultados diagnósticos.

En entornos hospitalarios, donde se emplean infraestructuras como RIS, PACS y el estándar DICOM para la gestión de imágenes, estas amenazas representan un riesgo significativo para la integridad de la información clínica y la seguridad del paciente. La falta de mecanismos robustos de ciberseguridad, junto con factores humanos como la escasa capacitación y la dependencia de la automatización, agravan esta problemática.

El presente estudio tiene como objetivo analizar las vulnerabilidades de los sistemas de inteligencia artificial en radiología digital frente a ataques adversariales, evaluando su impacto en la confiabilidad diagnóstica y en la seguridad del paciente. Asimismo, se abordan las implicaciones éticas, legales y operativas, y se identifican estrategias de mitigación orientadas a fortalecer la resiliencia de los sistemas de salud.

Palabras Clave: inteligencia artificial, radiología digital, ataques adversariales, ciberseguridad, seguridad del paciente

Abstract

The integration of artificial intelligence systems in digital radiology has enhanced diagnostic processes through deep learning algorithms capable of detecting pathologies with high accuracy. However, these systems are vulnerable to adversarial attacks, which involve the intentional manipulation of medical images through imperceptible perturbations that can alter diagnostic outcomes.

In hospital environments, where infrastructures such as RIS, PACS, and the DICOM standard are used for image management, these threats pose a significant risk to the integrity of clinical information and patient safety. The lack of robust cybersecurity mechanisms, along with human factors such as insufficient training and overreliance on automation, further exacerbates this issue.

This study aims to analyze the vulnerabilities of artificial intelligence systems in digital radiology to adversarial attacks, assessing their impact on diagnostic reliability and patient safety. Additionally, ethical, legal, and operational implications are addressed, and mitigation strategies are identified to strengthen the resilience of healthcare systems.

Keywords: artificial intelligence, digital radiology, adversarial attacks, cybersecurity, patient safety

Contenido

Introducción

La evolución tecnológica en el ámbito de la salud ha transformado profundamente la manera en que se realizan los procesos diagnósticos y la gestión de la información clínica. En las últimas décadas, la radiología digital ha experimentado un avance significativo gracias a la incorporación de herramientas computacionales capaces de optimizar la adquisición, almacenamiento y análisis de imágenes médicas. Este desarrollo ha permitido mejorar la eficiencia de los servicios radiológicos y fortalecer la precisión en la detección de múltiples patologías, convirtiendo a la tecnología en un elemento indispensable dentro de los entornos hospitalarios modernos. (Rodríguez, 2025)

En este contexto, la inteligencia artificial (IA) se ha consolidado como una de las innovaciones más importantes dentro de la medicina diagnóstica. Mediante técnicas de aprendizaje profundo y redes neuronales convolucionales, los sistemas de IA tienen la capacidad de reconocer patrones complejos en imágenes médicas y generar análisis automatizados que apoyan el trabajo del profesional de la salud. Su implementación en radiología ha contribuido a agilizar la interpretación de estudios, disminuir la carga operativa y favorecer la detección temprana de enfermedades, especialmente en áreas como oncología, neurología y diagnóstico pulmonar (Perez, 2025)

La integración de estas tecnologías con infraestructuras digitales como RIS, PACS y el estándar DICOM ha permitido consolidar sistemas de información altamente interconectados, facilitando el intercambio de imágenes y datos clínicos entre diferentes servicios médicos. Gracias a esta interoperabilidad, las instituciones de salud han logrado optimizar sus flujos de trabajo y mejorar el acceso oportuno a la información diagnóstica. No obstante, la creciente dependencia de plataformas digitales y sistemas automatizados también ha generado nuevos

desafíos relacionados con la protección de la información médica y la seguridad tecnológica. (Perez, 2025)

Actualmente, uno de los principales retos asociados al uso de inteligencia artificial en radiología corresponde a las vulnerabilidades frente a amenazas cibernéticas y ataques adversariales. Estas alteraciones digitales pueden modificar el comportamiento de los algoritmos de IA sin que los cambios sean perceptibles para el ojo humano, afectando la interpretación de las imágenes médicas y comprometiendo la confiabilidad de los diagnósticos. Esta situación representa una preocupación importante dentro del entorno clínico, debido a que la precisión diagnóstica constituye un elemento fundamental para la toma de decisiones terapéuticas y la seguridad del paciente (López, 2023).

Asimismo, el avance acelerado de la automatización en salud ha superado, en muchos casos, el desarrollo de estrategias de protección y regulación orientadas a garantizar un uso seguro de estas tecnologías. Por esta razón, la ciberseguridad en radiología digital se ha convertido en un tema de creciente interés científico, ético y profesional, ya que involucra aspectos relacionados con la integridad de la información clínica, la confidencialidad de los datos y la responsabilidad en el uso de sistemas inteligentes dentro de la práctica médica (Domínguez & Castaño, 2023).

A partir de esta realidad, la presente investigación se orienta al análisis de las vulnerabilidades de los sistemas de inteligencia artificial en radiología digital frente a ataques adversariales, así como a sus posibles implicaciones en la confiabilidad diagnóstica y la seguridad del paciente. Mediante una revisión documental de literatura científica y fuentes especializadas, se busca comprender el impacto de estas amenazas en los entornos clínicos

actuales y destacar la importancia de fortalecer las estrategias de ciberseguridad y resiliencia tecnológica en el sector salud (Domínguez & Castaño, 2023).

Planteamiento del Problema

En los últimos años, la inteligencia artificial (IA) ha adquirido una gran importancia en el área de la salud, especialmente en la radiología digital, debido a su capacidad para analizar imágenes médicas mediante algoritmos de aprendizaje profundo y redes neuronales convolucionales. Diversas investigaciones han demostrado que estos sistemas pueden detectar patologías con altos niveles de sensibilidad y especificidad, convirtiéndose en herramientas de apoyo para el diagnóstico médico y la toma de decisiones clínicas (Rodríguez, 2025).

De manera paralela, la transformación digital de los servicios de radiología ha permitido la implementación de infraestructuras tecnológicas como RIS, PACS y el estándar DICOM, los cuales facilitan el almacenamiento, procesamiento y transmisión de imágenes médicas dentro de las instituciones de salud. Esta integración tecnológica ha contribuido a optimizar los tiempos de respuesta, mejorar la gestión de la información clínica y fortalecer los procesos diagnósticos (Perez, 2025).

Sin embargo, el creciente uso de sistemas digitales e inteligencia artificial también ha generado nuevas preocupaciones relacionadas con la ciberseguridad y la protección de la información médica. Estudios recientes han evidenciado que los sistemas de inteligencia artificial aplicados a imágenes médicas presentan vulnerabilidades frente a ataques adversariales, los cuales consisten en modificaciones mínimas e imperceptibles realizadas sobre las imágenes diagnósticas con el objetivo de alterar intencionalmente los resultados (López, 2023).

Estas alteraciones pueden provocar falsos positivos o falsos negativos, afectando directamente la confiabilidad diagnóstica y aumentando el riesgo de errores médicos. En el entorno hospitalario, donde las imágenes diagnósticas son utilizadas para definir tratamientos, realizar seguimientos clínicos y tomar decisiones terapéuticas, cualquier manipulación maliciosa

de la información puede comprometer la seguridad del paciente y la calidad de la atención médica (López, 2023).

El problema se agrava debido a que muchos sistemas de inteligencia artificial en radiología han sido desarrollados priorizando la automatización y el rendimiento diagnóstico, sin incorporar mecanismos robustos de protección frente a ataques adversariales, validación de integridad de imágenes o estrategias de detección de manipulación digital. Asimismo, factores como la limitada capacitación del personal en ciberseguridad clínica, la interconectividad de los sistemas hospitalarios y la creciente dependencia de herramientas automatizadas incrementan el riesgo de vulnerabilidad dentro de las instituciones de salud (Resühr, 2025).

Además de las consecuencias clínicas, esta problemática también plantea implicaciones éticas, legales y operativas relacionadas con la confidencialidad de la información médica, la responsabilidad profesional y la seguridad de los sistemas digitales de salud.

Por esta razón, surge la necesidad de analizar las vulnerabilidades de los sistemas de inteligencia artificial en radiología digital frente a ataques adversariales, con el fin de comprender su impacto en la confiabilidad diagnóstica y en la seguridad del paciente (Resühr, 2025).

Justificación

La presente investigación se justifica debido al creciente uso de sistemas de inteligencia artificial (IA) en la radiología digital y a la necesidad de garantizar que estas tecnologías sean seguras, confiables y éticamente responsables dentro de los entornos clínicos. Aunque la inteligencia artificial ha permitido optimizar los procesos diagnósticos mediante herramientas capaces de analizar imágenes médicas con altos niveles de precisión, también han surgido nuevas vulnerabilidades relacionadas con ataques adversariales y riesgos de ciberseguridad que pueden comprometer la confiabilidad diagnóstica y la seguridad del paciente (Resühr, 2025).

Desde la perspectiva clínica, esta investigación es relevante porque una alteración maliciosa en imágenes médicas o en sistemas de inteligencia artificial puede generar diagnósticos erróneos, falsos positivos o falsos negativos, ocasionando retrasos terapéuticos, tratamientos innecesarios o decisiones médicas inadecuadas. Debido a que la radiología constituye una herramienta fundamental para el diagnóstico y seguimiento de múltiples enfermedades, cualquier vulnerabilidad que afecte la integridad de la información médica puede tener consecuencias directas sobre la calidad de la atención en salud y la seguridad del paciente (López, 2023).

Desde el ámbito tecnológico, el estudio se justifica porque los sistemas basados en aprendizaje profundo y redes neuronales convolucionales han demostrado ser sensibles a perturbaciones mínimas en las imágenes médicas. Esta situación representa un desafío importante para la implementación segura de tecnologías de inteligencia artificial en los sistemas hospitalarios modernos. En este sentido, la investigación permite comprender los riesgos asociados a la integración de IA con infraestructuras digitales como RIS, PACS y DICOM,

aportando conocimiento sobre la necesidad de fortalecer las estrategias de ciberseguridad y resiliencia tecnológica en el sector salud (Ramanathan, 2022).

Asimismo, esta investigación posee relevancia ética y legal, ya que el uso de inteligencia artificial en medicina implica responsabilidades relacionadas con la protección de la información clínica, la confidencialidad de los datos y la seguridad del paciente. Si los sistemas pueden ser manipulados sin mecanismos adecuados de detección y control, se vulneran principios fundamentales de la atención en salud, como la beneficencia, la no maleficencia y la responsabilidad profesional (Resühr, 2025).

De igual manera, este trabajo beneficia a diferentes actores del entorno sanitario y académico. Los profesionales de la salud podrán comprender mejor las vulnerabilidades asociadas al uso de inteligencia artificial en radiología y la importancia de implementar medidas de seguridad informática en la práctica clínica. Las instituciones hospitalarias podrán reconocer la necesidad de fortalecer sus sistemas digitales y establecer protocolos de protección frente a amenazas cibernéticas. Del mismo modo, estudiantes e investigadores del área de radiología, informática médica y ciberseguridad podrán utilizar esta investigación como base teórica para futuros estudios relacionados con inteligencia artificial aplicada a la salud.

A nivel social, la investigación contribuye a generar conciencia sobre la importancia de proteger la información médica digital y promover el desarrollo de tecnologías más seguras y confiables para los pacientes. Además, fomenta la discusión sobre el impacto de la inteligencia artificial en la atención médica y la necesidad de equilibrar el avance tecnológico con la seguridad clínica y la ética profesional.

Finalmente, este estudio puede servir como punto de partida para futuras investigaciones orientadas al desarrollo de mecanismos de detección de ataques adversariales, estrategias de

protección de imágenes médicas, modelos de inteligencia artificial más robustos y protocolos de ciberseguridad aplicados a sistemas de radiología digital. De esta manera, se espera contribuir al fortalecimiento de la seguridad tecnológica y a la mejora continua de la calidad de los servicios de salud.

Objetivos

Objetivo General

Analizar las vulnerabilidades de los sistemas de inteligencia artificial en radiología digital frente a ataques adversariales y su impacto en la confiabilidad diagnóstica y la seguridad del paciente, con base en el análisis documental y la revisión de literatura científica.

Objetivos Específicos

Describir el funcionamiento de los sistemas de inteligencia artificial en radiología digital y su integración con infraestructuras clínicas como RIS, PACS y DICOM, a partir de la revisión bibliográfica.

Reconocer los principales tipos de ataques adversariales que afectan los sistemas de inteligencia artificial en radiología digital, mediante el análisis de literatura científica y casos documentados.

Examinar las implicaciones de las vulnerabilidades de la inteligencia artificial en la confiabilidad diagnóstica y la seguridad del paciente, con base en la evidencia científica consultada.

Analizar la relación entre la seguridad de la información y la calidad del diagnóstico en entornos digitales de salud, a partir del análisis documental.

Marco Teórico

Infraestructura de la Radiología Digital Moderna

Sistema de Información Radiológica (RIS)

Un Sistema de Información Radiológica es un sistema que utilizan los departamentos de imágenes para la administración electrónica. Un RIS es una solución destinada a obtener, almacenar y compartir datos de imágenes médicas, y se utiliza ampliamente en todo el sector de la salud.

Un sistema RIS permite optimizar múltiples procesos dentro de los servicios de radiología, ya que facilita la programación y seguimiento de pacientes mediante el acceso a historiales médicos, el control del estado de los tratamientos, la gestión de registros clínicos y la organización de citas médicas. Asimismo, contribuye a la administración eficiente de documentos, garantizando la actualización constante y la integridad de los datos de los pacientes, lo que permite que la información clínica esté disponible cuando el personal médico la requiera. De igual manera, el RIS asegura la correcta identificación de las imágenes médicas al asociarlas adecuadamente con los datos de cada paciente, reduciendo el riesgo de errores o confusiones en los estudios diagnósticos. Además, este sistema favorece la optimización de procesos administrativos relacionados con facturación e informes radiológicos, gracias al almacenamiento de registros financieros, la automatización de pagos electrónicos y el análisis de datos. Finalmente, la intercomunicación entre modalidades y el uso de listas de trabajo dentro del RIS permiten agilizar el procesamiento de imágenes médicas y mejorar el control de calidad en los servicios de radiología (Ramanathan, 2022)

Los Beneficios de un RIS

En general, todos los beneficios inmediatos de utilizar un RIS culminan en una forma más rápida y fiable de gestionar la información de los pacientes se mejora el proceso de diagnóstico, se reduce el riesgo de escasez de personal y los errores de entrada de datos son tremendamente limitados para incluir solo unos pocos.

Un RIS puede ser esencial para proporcionar un modelo de servicio completo y sin fisuras para la atención del paciente. Las imágenes del RIS pueden ser una gran ventaja para descentralizar la prestación de servicios de radiología, ya que permiten agilizar la labor del médico y aumentar las comunicaciones entre los médicos y permitir el seguimiento de los pacientes en tiempo real y gestionar mejor el flujo de trabajo, lo que se traduce en diagnósticos más rápidos (Ramanathan, 2022)

Sistema de Archivado y Comunicación de Imágenes (PACS)

PACS es una tecnología de imágenes médicas que proporciona almacenamiento de datos y acceso cómodo a las imágenes desde múltiples modalidades. Que es utilizada principalmente por hospitales y organizaciones de atención médica. Creado para almacenar y transmitir de forma segura imágenes y datos electrónicos de los pacientes, un Sistema PACS elimina la necesidad de metodologías tradicionales que implican la administración manual de archivos (por ejemplo, carátulas de películas) y la entrega.

Si bien los radiólogos han sido los principales usuarios de Sistemas PACS, además de ser uno de los principales generadores de imágenes radiográficas, el PACS también se ha implementado en otros campos relacionados con la salud, como la cardiología, la oncología, la dermatología, la patología y las imágenes de medicina nuclear. Debido a su uso prolífico, el PACS está diseñado para procesar formatos de imagen producidos a partir de una amplia

variedad de modalidades, como las mamografías (MG), la resonancia magnética (RM), la ecografía (EE. UU.), la tomografía computarizada (TC) y la radiografía digital.

Al aprovechar un PACS, los médicos pueden acceder fácilmente a la información de sus pacientes de forma digital. El acceso digital significa acelerar y mejorar la atención, minimizar las posibilidades de errores en el tratamiento y la prescripción, además de evitar la realización de pruebas innecesarias (Ramanathan, 2022)

Imágenes Digitales y Comunicaciones en Medicina (DICOM)

DICOM es el estándar de comunicación y gestión reconocido a nivel mundial para imágenes médicas y otros datos de pacientes. Se utiliza con frecuencia en el campo de la medicina para almacenar y transmitir imágenes médicas, lo que facilita la integración con los dispositivos médicos y los sistemas PACS (Ramanathan, 2022)

DICOM es el estándar mundial para la transmisión y el almacenamiento de imágenes médicas. Desde su creación en 1993, el estándar DICOM ha sido indispensable en la transformación de la práctica radiológica, que pasó de ser una película de rayos X a un flujo de trabajo digitalizado (Ramanathan, 2022)

DICOM existe como protocolo de comunicación y como tipo de formato. Al guardar las imágenes médicas en este formato, se garantiza que todos los datos relacionados con un estudio en particular (es decir, la información del paciente, la imagen médica, etc.) permanezcan juntos, lo que permite una transmisión fluida entre dispositivos compatibles con DICOM, traducidos a compartir datos de pacientes de forma más sencilla entre los médicos y un proceso de diagnóstico más rápido (Ramanathan, 2022)

Un archivo es una imagen guardada en formato DICOM. En su interior, se incluye una imagen médica generada a partir de una modalidad (por ejemplo, una resonancia magnética, un escáner, una ecografía, etc.), junto con etiquetas de metadatos (Ramanathan, 2022)

PACS Contra DICOM

PACS es el sistema de archivo y gestión que almacena y distribuye imágenes médicas. DICOM es el estándar técnico y formato de archivo universal que permite a esas imágenes ser visualizadas y transmitidas entre equipos. PACS utiliza el estándar DICOM para funcionar (Ramanathan, 2022)

La integración entre sistemas RIS, PACS y el estándar DICOM ha permitido optimizar significativamente la gestión de imágenes médicas y la comunicación entre diferentes servicios hospitalarios. Sin embargo, esta alta interconectividad también incrementa la superficie de exposición frente a amenazas cibernéticas, ya que una vulnerabilidad en cualquiera de estos sistemas puede comprometer la integridad de la información diagnóstica. En el contexto de la inteligencia artificial aplicada a radiología, esta situación adquiere mayor relevancia debido a que los algoritmos dependen directamente de la autenticidad y calidad de las imágenes médicas para generar resultados confiables.

Sistema de Información Clínica (CIS)

Un CIS implica soluciones de software en red que trabajan juntas dentro de una práctica de radiología, como un RIS y un sistema de historia clínica electrónica. Orientado específicamente al uso de la atención clínica (por ejemplo, en las unidades de cuidados intensivos), un CIS es un sistema de información en red con sistemas informáticos que se utiliza en varios departamentos de los hospitales actuales. Estos departamentos incluyen cardiología,

radiología y patología. El CIS recopila los datos de los pacientes y los transfiere a un registro electrónico al que el médico tratante puede acceder mientras visita la cama del paciente.

A través de ella, las organizaciones de atención médica reciben un apoyo adecuado en sus operaciones, implementación de políticas y administración de los datos de los pacientes. Para entender por qué el CIS es una herramienta esencial para brindar una atención óptima al paciente, examinemos cómo beneficia tanto al paciente como a los médicos. El aprovechamiento de un sistema de información clínica (CIS) permite mejorar la comunicación entre los pacientes y los equipos de atención médica, así como entre los mismos profesionales de la salud, favoreciendo una interacción más eficiente dentro de los entornos clínicos. Además, facilita la transferencia integral de conocimientos, lo que contribuye a que los médicos puedan tomar decisiones informadas y acertadas en relación con el diagnóstico y tratamiento de los pacientes. De igual manera, un CIS permite un acceso más ágil a radiografías, escaneos y otros estudios diagnósticos cuando son necesarios, optimizando la atención médica. Asimismo, contribuye a mejorar la calidad de la administración y la prestación de los cuidados en salud, al tiempo que fortalece las posibilidades de desarrollo e investigación clínica avanzada.

Dado que un CIS captura y procesa electrónicamente toda la información médica relevante, se reduce la necesidad de documentación en papel y disminuyen los errores de ingreso de datos, lo que en última instancia se traduce en un tiempo (¡y costo!) considerable ahorros para las organizaciones de atención médica. De este modo, proporcionando impactos positivos en la calidad de la atención al paciente siendo entregado (Ramanathan, 2022)

Interconectividad

La radiología, por su naturaleza, está intrínsecamente conectada a Internet y se encuentra a la vanguardia de la tecnología en medicina. En los últimos años se ha observado un aumento

drástico de la tecnología basada en Internet en la atención médica, con la imagenología como aplicación central. Numerosas aplicaciones y tecnologías basadas en Internet han incursionado en la medicina, y para la radiología la integración es más fluida que en otras especialidades clínicas. Muchas aplicaciones en la práctica de la radiología se basan en Internet y cada día se agregan más. La introducción de dispositivos móviles y su integración en el flujo de trabajo de imagenología ha reforzado el papel que desempeña Internet en la radiología. Debido a la rápida proliferación de dispositivos portátiles y teléfonos inteligentes, la tecnología habilitada para IoT está transformando la atención médica, pasando de sistemas convencionales basados en centros a sistemas de atención médica más personalizados. Este artículo analiza brevemente cómo el IoT desempeña un papel útil en el flujo de trabajo diario de imagenología y sus aplicaciones actuales y potenciales futuras, cómo se pueden integrar los dispositivos móviles en los flujos de trabajo de radiología y el impacto del IoT en la educación de residentes y estudiantes de medicina, la investigación y la participación del paciente en radiología (Johnson, 2020)

Amenazas Cibernéticas en el Entorno Clínico

Ransomware en Salud

Los ataques de secuestro de datos, o ransomware, son programas maliciosos que cifran los sistemas digitales de un hospital, bloqueando el acceso a historiales médicos y equipos críticos para exigir un rescate económico. Son objetivos críticos porque manejan información confidencial de alto valor, utilizan sistemas obsoletos vulnerables y necesitan continuidad 24/7, lo que presiona a pagar para evitar poner en riesgo la vida de los pacientes (Reed, 2024)

El ransomware en radiología cifra sistemas PACS e imágenes DICOM, paralizando el diagnóstico y tratamiento al dejar inaccesibles los estudios de los pacientes. Constituye una amenaza crítica que busca extorsión económica, provocando cancelación de citas, pérdida de

datos y alto riesgo asistencial. La protección requiere copias de seguridad robustas y ciberseguridad (Reed, 2024)

Aunque tradicionalmente los ataques cibernéticos en salud se han asociado con pérdida de acceso a la información o interrupción de servicios, actualmente el riesgo es mayor debido a la incorporación de sistemas de inteligencia artificial en los procesos diagnósticos. Un ataque que altere imágenes médicas o afecte plataformas digitales puede influir directamente en las decisiones clínicas automatizadas, generando diagnósticos erróneos y aumentando el riesgo para los pacientes. Esto demuestra que la ciberseguridad en radiología no solo representa un problema tecnológico, sino también un desafío clínico y ético.

Vulnerabilidades del Día Cero

Una vulnerabilidad de día cero es un fallo de seguridad oculto en software o hardware que es desconocido para el proveedor, desarrollador o la comunidad de seguridad. Debido a que no hay un parche o solución oficial disponible en el momento del descubrimiento, los sistemas permanecen expuestos a posibles ataques. Estas vulnerabilidades son particularmente peligrosas porque los ciberdelincuentes pueden explotarlas antes de que el fallo se haga público o se solucione, a menudo causando daños significativos. Las vulnerabilidades de día cero requieren una detección rápida y fuertes medidas de protección contra vulnerabilidades de día cero para minimizar el riesgo y prevenir brechas (Cooper, 2025)

Los exploits de día cero son particularmente peligrosos porque aprovechan los fallos de seguridad antes de que esté disponible un parche. Esto significa que los defensores no tienen tiempo para prepararse, y las medidas de seguridad tradicionales a menudo fallan en detectarlos (Cooper, 2025)

Inteligencia Artificial y Ataques Adversarios

Fundamentos de IA en Radiología

Detectar una patología, por pequeña que sea, puede marcar la diferencia en la vida de un paciente. Aquí es donde las redes neuronales convolucionales (CNN) cambian las cosas ya que son un tipo de inteligencia artificial diseñado específicamente para ver y analizar imágenes.

En radiología esto es muy importante se entrenan estas redes con miles de imágenes médicas, para que cada imagen que procesan aprenda a identificar patrones, texturas, formas y diferencias que podrían indicar la presencia de alguna enfermedad. Las (CNN) pueden ayudarnos a una detección temprana señalando áreas sospechosas en imágenes que el ojo humano cansado después de mucho trabajo podría pasar por alto, y en ese momento es donde ayudarían sirviendo como un asistente (Vega, 2024)

Manipulación de Imágenes

La principal preocupación frente a los ataques adversariales en radiología radica en que las alteraciones realizadas sobre las imágenes suelen ser imperceptibles para el ojo humano, pero suficientes para modificar completamente la interpretación del algoritmo. Esto genera un escenario de riesgo clínico importante, ya que el personal médico podría confiar en resultados aparentemente válidos sin sospechar que la imagen ha sido manipulada. En consecuencia, la dependencia creciente de herramientas automatizadas puede convertir estas vulnerabilidades en una amenaza para la confiabilidad diagnóstica y la seguridad del paciente (Bortsova, 2021)

Imaginemos que tenemos una radiografía de pulmón, para un radiólogo experimentado, es una imagen clara pero aquí es donde entra el concepto de “ruido adversarial” el cual es algo muy sutil, se trata de pequeñas alteraciones en los pixeles de la imagen, tan pequeñas que son completamente imperceptibles al ojo humano.

Estas mínimas modificaciones engañan el algoritmo de la IA y hacen que interprete de una manera diferente ejemplo podría clasificar una imagen de un pulmón sano como si tuviera una patología grave o pasar por alto un tumor claramente visible, este tipo de manipulaciones podría llevar a diagnósticos erróneos, retrasos en los tratamientos y causar falta de confianza.

Además, el envenenamiento de datos evidencia que las vulnerabilidades de la inteligencia artificial no solo se presentan durante el uso clínico de los sistemas, sino también durante su proceso de entrenamiento y desarrollo. Esto demuestra la necesidad de implementar controles de calidad, validación de datos y mecanismos de supervisión continua que permitan garantizar la seguridad y confiabilidad de los modelos utilizados en salud (Bortsova, 2021)

Impacto Operativo y Clínico

El uso cada vez mayor de sistemas digitales en el área de la salud ha permitido mejorar muchos procesos clínicos, organizar mejor la información médica y facilitar el acceso a estudios diagnósticos. Sin embargo, esta gran dependencia de la tecnología también puede traer riesgos operativos cuando los sistemas presentan vulnerabilidades o problemas de ciberseguridad. (Téllez, 2025)

Cuando ocurren ataques cibernéticos o fallas en los sistemas, la continuidad de los servicios de salud puede verse afectada, sobre todo en áreas críticas como radiología, urgencias y oncología. En este sentido, las vulnerabilidades en los sistemas digitales no solo deben verse como un problema tecnológico. También pueden generar consecuencias directas en la atención de los pacientes, influir en la toma de decisiones clínicas y afectar la calidad del servicio que se brinda en las instituciones de salud.

Uno de los efectos más importantes que pueden generar los incidentes de ciberseguridad en los servicios de salud es la interrupción del flujo normal de trabajo clínico. Cuando se

presentan ataques informáticos, los sistemas hospitalarios pueden dejar de funcionar por un tiempo, lo que afecta plataformas utilizadas para la gestión clínica, el almacenamiento de imágenes médicas y los sistemas de información del hospital. (Puentes & Salinas Miranda, 2022)

Cuando ocurre este tiempo de inactividad en sistemas como PACS o RIS, el personal de salud puede perder de manera temporal el acceso a imágenes diagnósticas, informes radiológicos y antecedentes clínicos de los pacientes. Esto provoca retrasos en la interpretación de los estudios, dificulta la priorización de casos urgentes y puede generar desorden en el flujo de trabajo dentro del hospital.

Esta situación es aún más crítica en áreas como urgencias y oncología, donde las decisiones médicas dependen en gran medida de la rapidez con la que se pueda acceder a la información diagnóstica. Cuando los sistemas se interrumpen, la atención puede retrasarse y esto aumenta el riesgo para los pacientes.

Otro aspecto muy importante es la integridad de los datos clínicos. La seguridad de la información en salud se apoya en tres principios básicos, confidencialidad, integridad y disponibilidad. En este caso, la integridad hace referencia a que la información médica debe mantenerse completa y sin modificaciones no autorizadas, es decir, que no haya sido alterada o manipulada.

En los sistemas de imágenes médicas, una posible alteración malintencionada de los registros clínicos o de las imágenes diagnósticas representa un riesgo considerable, ya que podría influir en la interpretación médica y afectar el proceso de diagnóstico. Algunos estudios recientes han evidenciado que los sistemas digitales en salud pueden presentar vulnerabilidades que permiten la manipulación de datos o accesos no autorizados, lo que pone en duda la confiabilidad de la información clínica. (Puentes & Salinas Miranda, 2022)

Además, cuando se pierde la trazabilidad de los registros médicos, es decir, cuando no es posible identificar quién accedió, consultó o modificó cierta información se dificulta comprobar la autenticidad de los datos. Esto también afecta la transparencia y el control dentro de los procesos clínicos.

Finalmente, los incidentes relacionados con la ciberseguridad también pueden generar un deterioro en la calidad del cuidado médico. Cuando los sistemas digitales fallan o dejan de funcionar, el personal de salud muchas veces debe recurrir a procesos manuales o métodos más tradicionales para continuar con la atención. Esto puede hacer que el trabajo sea más lento, menos eficiente y que aumente la posibilidad de cometer errores.

En este tipo de situaciones pueden aparecer retrasos en el diagnóstico, en la programación de exámenes o incluso en la administración de tratamientos. Esto es especialmente delicado en áreas como oncología, donde el seguimiento constante y la precisión de la información clínica son esenciales para tomar decisiones terapéuticas adecuadas. Por eso, cualquier interrupción en los sistemas tecnológicos puede afectar negativamente la atención que reciben los pacientes.

Por esta razón, las vulnerabilidades en ciberseguridad no solo deben verse como un problema relacionado con la protección de la información. También pueden tener un impacto directo en la continuidad de la atención médica, en la seguridad del paciente y en la calidad de los servicios de salud que se brindan (Puentes & Salinas Miranda, 2022)

Lo anterior permite comprender que las vulnerabilidades en los sistemas digitales de radiología tienen repercusiones que van más allá de los aspectos tecnológicos. La afectación de la disponibilidad, integridad o confiabilidad de la información médica puede impactar directamente la toma de decisiones clínicas y la continuidad de la atención en salud. Por esta razón, el fortalecimiento de la ciberseguridad en sistemas de inteligencia artificial aplicados a

radiología se convierte en una necesidad prioritaria para garantizar diagnósticos seguros y servicios médicos de calidad.

Consideraciones Éticas y Legales

Seguridad del Paciente

La seguridad del paciente es uno de los principios fundamentales en los sistemas de salud. En los entornos digitales actuales, los ataques cibernéticos no deben considerarse únicamente como problemas técnicos, sino también como eventos adversos que pueden afectar directamente la atención médica. Cuando un sistema hospitalario o de imágenes médicas es comprometido por un ataque informático, pueden producirse retrasos en los diagnósticos, interrupciones en los tratamientos o pérdida de información clínica crítica. Estas situaciones representan un riesgo real para la vida y la salud de los pacientes, por lo que la ciberseguridad se convierte en un componente esencial para garantizar la seguridad clínica (Resühr, 2025)

Privacidad y Confidencialidad de la Información

Los sistemas de salud manejan grandes volúmenes de información sensible relacionada con los pacientes, como historiales clínicos, diagnósticos e imágenes médicas. Por esta razón, la protección de la privacidad y la confidencialidad de los datos es una obligación ética y legal para las instituciones de salud. Diversas normativas internacionales establecen lineamientos estrictos para la protección de la información médica, entre ellas la Health Insurance Portability and Accountability Act (HIPAA) en Estados Unidos y el Reglamento General de Protección de Datos (RGPD) en Europa. Estas regulaciones exigen la implementación de medidas de seguridad que garanticen el manejo adecuado de los datos personales y prevengan el acceso no autorizado o el robo de información.

La inteligencia artificial (IA) y el aprendizaje automático se están integrando cada vez más en las prácticas radiológicas modernas, facilitando el análisis de imágenes, el diagnóstico y la toma de decisiones. Sin embargo, la integración de la IA introduce nuevas vulnerabilidades, especialmente en lo que respecta a la integridad y la seguridad de los algoritmos y el flujo de datos que procesan. Además de aumentar la vulnerabilidad del proceso de obtención de imágenes, los avances en IA también incrementan la sofisticación de la manipulación de imágenes y aumentan los riesgos, tanto involuntarios como maliciosos, para la integridad de los datos de imagen. Es fundamental garantizar la seguridad de estos sistemas de IA y proteger los flujos de datos que se les proporcionan contra las filtraciones (Wald, 2025)

HIPAA es la piedra angular de la ley de privacidad de la información de salud en los Estados Unidos. Establece requisitos nacionales para salvaguardar la información de salud protegida electrónicamente (ePHI), incluidas las imágenes radiológicas. La Regla de Privacidad de HIPAA garantiza que la ePHI esté debidamente protegida sin comprometer la prestación de atención médica de alta calidad. La Regla de Privacidad también introdujo la Disposición de Puerto Seguro para habilitar un método para anonimizar los datos de atención médica, incluidos los informes e imágenes radiológicas (Wald, 2025)

La Regla de Seguridad de HIPAA, vigente desde el 21 de abril de 2005, exige que los proveedores de atención médica implementen medidas de seguridad administrativas, físicas y técnicas para proteger la confidencialidad , la integridad y la disponibilidad de la información de salud protegida electrónicamente (ePHI).

Los departamentos de radiología deben garantizar que el almacenamiento y el intercambio electrónico de datos confidenciales de pacientes, como imágenes radiológicas, cumplan con las disposiciones de HITECH (Ley de Tecnología de la Información Sanitaria para

la Salud Económica y Clínica). En caso de una violación de datos, los proveedores están obligados a notificar de inmediato a las personas afectadas, al Departamento de Salud y Servicios Humanos y, en algunos casos, a los medios de comunicación. El incumplimiento puede acarrear graves sanciones económicas, así como costos de mitigación posteriores, y dañar la reputación del proveedor. Por lo tanto, los departamentos de radiología deben implementar sólidas medidas de seguridad de datos y mantenerse vigilantes para proteger la información del paciente y mantener la confianza (Wald, 2025)

Responsabilidad Profesional en el Uso de Tecnologías

El uso creciente de tecnologías avanzadas, como sistemas de inteligencia artificial en radiología y diagnóstico médico, plantea nuevos desafíos éticos y legales en relación con la responsabilidad profesional. En situaciones donde un sistema automatizado o una inteligencia artificial pueda ser manipulado o comprometido por un ataque cibernético, surge la interrogante sobre quién asume la responsabilidad si se genera un diagnóstico erróneo. Esta responsabilidad podría involucrar a diferentes actores del proceso, como el tecnólogo en radiología que opera el sistema, el médico radiólogo que interpreta los resultados o el desarrollador del software que diseñó la herramienta tecnológica. Por lo tanto, es necesario establecer marcos regulatorios claros que definan responsabilidades y garanticen la seguridad en el uso de estas tecnologías (Resühr, 2025)

En este contexto, la implementación de inteligencia artificial en radiología plantea la necesidad de establecer marcos éticos y regulatorios más sólidos que permitan definir responsabilidades frente a posibles errores diagnósticos derivados de vulnerabilidades tecnológicas o ataques cibernéticos. La rápida incorporación de sistemas automatizados en salud ha avanzado más rápido que muchas normativas existentes, por lo que resulta necesario

fortalecer políticas de ciberseguridad, supervisión tecnológica y protección de datos clínicos en los entornos hospitalarios.

Estrategias de Mitigación y Resiliencia

Los avances tecnológicos en el sector salud han permitido el desarrollo de sistemas de imágenes médicas digitales, los cuales facilitan el diagnóstico y tratamiento de diversas enfermedades. Equipos como rayos X, tomografía computarizada, resonancia magnética y ecografía generan grandes volúmenes de información que deben ser almacenados, procesados y transmitidos a través de sistemas informáticos especializados. Estos sistemas suelen integrarse con plataformas de gestión hospitalaria y redes digitales, lo que mejora la eficiencia del servicio médico. Sin embargo, también aumenta la exposición a riesgos de seguridad informática, como accesos no autorizados, pérdida de datos o ataques cibernéticos (PROMEDCO, 2023)

En este contexto, las estrategias de mitigación y resiliencia se convierten en elementos fundamentales para proteger la información médica y garantizar la continuidad de los servicios de salud. La mitigación se refiere a las acciones implementadas para reducir la probabilidad de que ocurran incidentes de seguridad, mientras que la resiliencia se relaciona con la capacidad de los sistemas para resistir, adaptarse y recuperarse rápidamente frente a fallos o ataques informáticos.

Una de las estrategias más importantes es la ciberhigiene, que comprende un conjunto de prácticas y protocolos de seguridad que deben ser aplicados por el personal técnico y administrativo que interactúa con los sistemas digitales. Estas prácticas incluyen la actualización constante de los sistemas operativos y programas, el uso de contraseñas seguras, la instalación de software antivirus, la restricción de dispositivos externos y la capacitación del personal en el manejo responsable de la información. La ciberhigiene busca reducir los errores humanos y

prevenir vulnerabilidades que puedan ser aprovechadas por atacantes (Domínguez & Castaño, 2023)

Otra estrategia clave es la segmentación de redes, la cual consiste en dividir la infraestructura tecnológica en diferentes subredes o zonas de seguridad. En los entornos de radiología digital, esta técnica permite aislar las modalidades de imagen y los sistemas de almacenamiento de imágenes médicas de la red pública de internet. De esta forma, se limita el acceso a los equipos críticos y se reduce la posibilidad de que un ataque informático se propague a toda la infraestructura hospitalaria (Iqba, 2023)

Asimismo, el cifrado de datos y las firmas digitales son mecanismos fundamentales para proteger la información médica. El cifrado permite transformar los datos en un formato codificado que solo puede ser interpretado por usuarios autorizados, garantizando la confidencialidad de la información del paciente. Por su parte, las firmas digitales permiten verificar la autenticidad e integridad de los archivos, asegurando que las imágenes médicas no hayan sido modificadas o alteradas desde el momento de su captura hasta su análisis por parte del profesional de salud (Navarro, 2025)

La implementación de estas estrategias fortalece la seguridad de los sistemas de imágenes médicas y contribuye a mantener la confiabilidad de los servicios de radiología. Además, permite a las instituciones de salud responder de manera efectiva ante incidentes de seguridad, minimizar el impacto de posibles ataques y garantizar la protección de la información clínica. En consecuencia, la aplicación de medidas de mitigación y resiliencia se considera un componente esencial para asegurar la continuidad de los servicios médicos y preservar la calidad de la atención al paciente.

Sin embargo, la implementación de estrategias de mitigación no debe limitarse únicamente al uso de herramientas tecnológicas. También es fundamental promover procesos de capacitación del personal de salud, cultura de ciberseguridad institucional y actualización constante de protocolos clínicos y tecnológicos. La protección de los sistemas de inteligencia artificial en radiología requiere un enfoque integral que combine seguridad informática, supervisión humana y responsabilidad organizacional.

Marco Metodológico

Tipo de Investigación y Diseño

La presente investigación se clasifica como una investigación de tipo descriptivo, ya que busca identificar y caracterizar las vulnerabilidades de los sistemas de inteligencia artificial en radiología digital frente a ataques adversariales, así como sus implicaciones en la confiabilidad diagnóstica y la seguridad del paciente, sin manipular variables ni establecer relaciones causales de manera experimental.

En este sentido, la investigación corresponde a un diseño no experimental de tipo transversal, debido a que no se manipulan variables ni se realizan intervenciones sobre los sistemas analizados. El estudio se basa en la revisión y análisis de información existente en un momento determinado, con el fin de comprender las vulnerabilidades de los sistemas de inteligencia artificial en radiología digital frente a ataques adversariales y sus implicaciones en el entorno clínico.

Asimismo, la investigación adopta un diseño documental o de revisión bibliográfica, ya que se fundamenta en la recopilación, análisis e interpretación de información proveniente de fuentes secundarias como artículos científicos, informes técnicos, normativas y publicaciones académicas relacionadas con la inteligencia artificial, la radiología digital y la ciberseguridad en salud.

Teniendo en cuenta lo anterior, se llevó a cabo una revisión bibliográfica, ya que según Guirao (2015), tanto en el ámbito clínico como académico, la revisión bibliográfica constituye una etapa fundamental dentro del proceso investigativo, debido a que permite identificar qué se conoce y qué aspectos aún requieren profundización dentro de un tema específico. Asimismo, el autor expone que una revisión bibliográfica reúne diferentes investigaciones y publicaciones

científicas, proporcionando una visión del estado actual del conocimiento y favoreciendo el análisis crítico de la información recopilada.

La investigación se desarrolla bajo un enfoque cualitativo, debido a que se orienta al análisis e interpretación de las vulnerabilidades de los sistemas de inteligencia artificial en radiología digital frente a ataques adversariales. A través de la revisión documental y teórica, se busca comprender las implicaciones clínicas, tecnológicas, éticas y operativas que estas vulnerabilidades pueden generar en la confiabilidad diagnóstica y en la seguridad del paciente.

Recolección y Análisis de la Información

La recolección de la información se realizó mediante la búsqueda, selección y revisión de artículos científicos, documentos académicos, informes técnicos y normativas relacionadas con inteligencia artificial, radiología digital, ciberseguridad y ataques adversariales en el sector salud. Para ello, se consultaron fuentes académicas y científicas confiables como PubMed, ScienceDirect, Scielo, la Organización Mundial de la Salud (OMS), la Food and Drug Administration (FDA), ENISA y otras publicaciones especializadas en tecnología médica y seguridad informática.

Se priorizaron publicaciones comprendidas entre los años 2020 y 2025, considerando criterios de actualidad, relevancia científica, confiabilidad de la fuente y relación directa con los objetivos de investigación. Posteriormente, los documentos seleccionados fueron analizados y organizados de acuerdo con las principales categorías temáticas del estudio.

Las principales categorías de análisis utilizadas en la investigación fueron la inteligencia artificial en radiología digital, la infraestructura tecnológica en sistemas de imágenes médicas, los ataques adversariales y las vulnerabilidades digitales, la ciberseguridad en entornos clínicos,

la seguridad del paciente y las implicaciones éticas y legales relacionadas con el uso de tecnologías digitales en el sector salud.

La población de estudio está constituida por el conjunto de documentos académicos, científicos y técnicos relacionados con la temática de investigación, mientras que la muestra corresponde a una selección de fuentes pertinentes obtenidas bajo criterios de actualidad, relevancia científica y relación directa con el problema de investigación. El tipo de muestreo utilizado fue no probabilístico por criterio, ya que se seleccionaron intencionalmente los documentos considerados más significativos para el análisis.

Finalmente, para el análisis de la información se empleó la técnica de análisis de contenido, la cual permitió clasificar, comparar e interpretar la información recopilada, identificando patrones, relaciones y tendencias en torno a las vulnerabilidades de los sistemas de inteligencia artificial en radiología digital y sus implicaciones en la práctica clínica y la seguridad del paciente.

Consideraciones Éticas

Esta investigación documental se clasifica dentro de la categoría denominada sin riesgo según la Resolución 8430 de 1993 por medio de la cual se establecen las normas científicas, técnicas y administrativas para la investigación en salud, la cual en su artículo No. 11 menciona que:

Dentro de este tipo de investigaciones están los estudios que emplean técnicas y métodos de investigación documental retrospectivos y aquellos en los que no se realiza ninguna intervención o modificación intencionada de las variables biológicas, fisiológicas, psicológicas o sociales de los individuos que participan en el estudio, entre los que se consideran: revisión de

historias clínicas, entrevistas, cuestionarios y otros en los que no se le identifique ni se traten aspectos sensitivos de su conducta (p.3)

Resultados

A partir de la revisión bibliográfica y el análisis documental de artículos científicos, normativas y publicaciones especializadas relacionadas con inteligencia artificial, radiología digital y ciberseguridad en salud, se identificaron diferentes vulnerabilidades tecnológicas y riesgos asociados a los ataques adversariales en sistemas de diagnóstico automatizado.

La investigación permitió evidenciar que la integración de sistemas de Inteligencia Artificial (IA) en radiología digital se encuentra estrechamente vinculada con infraestructuras tecnológicas como los sistemas RIS (Radiology Information System), PACS (Picture Archiving and Communication System) y el estándar DICOM, los cuales facilitan el almacenamiento, transmisión y análisis automatizado de imágenes médicas. Esta interoperabilidad ha optimizado significativamente el flujo de trabajo radiológico, reduciendo tiempos de procesamiento y apoyando la detección temprana de patologías mediante algoritmos de deep learning. Sin embargo, también se identificó que dicha integración incrementa la superficie de ataque cibernético dentro de los entornos hospitalarios (Wald, 2025)

La revisión documental también permitió identificar que la implementación de redes neuronales convolucionales (CNN) en radiología ha mejorado la capacidad de detección automatizada de patologías en estudios de tomografía computarizada, resonancia magnética y radiografía digital. Según Vega (2024) y López (2023), estos modelos de aprendizaje profundo logran reconocer patrones complejos en imágenes médicas con altos niveles de sensibilidad diagnóstica, especialmente en patologías pulmonares, neurológicas y oncológicas. No obstante, los autores coinciden en que la precisión de estos sistemas depende directamente de la calidad e integridad de las imágenes procesadas, lo que incrementa la preocupación frente a posibles manipulaciones adversariales.

En relación con el estándar DICOM, se encontró que, aunque representa el principal mecanismo de comunicación entre equipos médicos e imágenes diagnósticas, presenta vulnerabilidades importantes en términos de ciberseguridad. La estructura de sus metadatos puede contener información sensible del paciente y, en muchos entornos clínicos, la transmisión de datos se realiza sin mecanismos robustos de cifrado nativo, lo que facilita posibles accesos no autorizados, manipulación de imágenes o interceptación de información clínica. (Ramanathan, 2022)

Asimismo, el análisis bibliográfico evidenció que la interconectividad entre dispositivos médicos, plataformas RIS, PACS y sistemas CIS favorece la eficiencia operativa de los servicios de radiología, pero simultáneamente incrementa las vulnerabilidades de seguridad informática. Johnson (2020) señala que la incorporación de tecnologías basadas en Internet y dispositivos IoT en radiología ha ampliado la superficie de exposición frente a amenazas cibernéticas, facilitando posibles accesos no autorizados y ataques dirigidos a infraestructuras hospitalarias conectadas.

Dentro de la investigación se logró establecer una taxonomía de los principales ataques adversariales dirigidos a sistemas de IA en radiología. Entre ellos, se destacó el ruido adversarial, consistente en pequeñas modificaciones de píxeles imperceptibles para el ojo humano pero suficiente para alterar el comportamiento de Redes Neuronales Convolucionales (CNN), provocando errores diagnósticos en la clasificación de imágenes médicas. Asimismo, se identificó el envenenamiento de datos (data poisoning), donde conjuntos de entrenamiento son alterados deliberadamente con información corrupta o manipulada, ocasionando que el modelo aprenda patrones incorrectos y disminuya su confiabilidad clínica.

Otro hallazgo importante corresponde a los ataques de inversión y extracción de modelos, mediante los cuales un atacante puede intentar reconstruir información privada de los pacientes a

partir de las respuestas generadas por los algoritmos de IA. Este tipo de amenaza representa un riesgo significativo para la confidencialidad médica y el cumplimiento de normativas de protección de datos en salud.

Los resultados también evidenciaron que la manipulación de imágenes médicas mediante ataques adversariales puede generar falsos positivos y falsos negativos en el diagnóstico radiológico. En diversos estudios analizados, pequeñas alteraciones digitales lograron inducir a los sistemas de IA a detectar patologías inexistentes o, por el contrario, omitir lesiones reales como tumores, hemorragias o fracturas. Estas alteraciones comprometen directamente la integridad de la información clínica, considerada uno de los pilares fundamentales de la seguridad informática en salud.

De igual manera, se encontró que los ataques de ransomware representan una de las amenazas más frecuentes y críticas para los servicios de imágenes diagnósticas. Reed (2024) describe que estos ataques pueden paralizar sistemas PACS y bloquear el acceso a imágenes DICOM, afectando la continuidad operativa de hospitales y retrasando procesos diagnósticos y terapéuticos. Los estudios revisados evidencian que este tipo de incidentes compromete tanto la disponibilidad de la información clínica como la seguridad del paciente.

Otro hallazgo relevante corresponde a la limitada preparación institucional frente a amenazas dirigidas a sistemas de inteligencia artificial en salud. Domínguez y Castaño (2023) y Kelly (2023) destacan que muchas instituciones hospitalarias aún presentan deficiencias en políticas de ciberseguridad, capacitación del personal y protocolos de protección digital. Esto aumenta el riesgo de errores humanos, accesos inseguros y fallas en la detección temprana de incidentes informáticos.

La revisión de literatura también permitió establecer que la confianza excesiva en sistemas automatizados puede convertirse en un factor de riesgo clínico. Resühr (2025) y Sánchez (2022) mencionan que la dependencia creciente de herramientas basadas en IA podría reducir la capacidad crítica de supervisión por parte del personal médico, especialmente cuando los resultados generados por los algoritmos son asumidos como completamente confiables. Esta situación incrementa la probabilidad de aceptar diagnósticos alterados por ataques adversariales sin una validación clínica adecuada.

En relación con la protección de la información médica, Wald (2025) y Navarro (2025) identifican que el uso de mecanismos de cifrado, firmas digitales y autenticación multifactor contribuye significativamente a fortalecer la integridad y confidencialidad de las imágenes diagnósticas. Sin embargo, los autores también señalan que la implementación de estas medidas aún es desigual entre instituciones de salud, especialmente en entornos con limitaciones presupuestales o tecnológicas.

Finalmente, la investigación permitió identificar que la ciberseguridad en radiología aún se encuentra en una etapa de adaptación frente al crecimiento acelerado de la automatización diagnóstica. Aunque existen estrategias emergentes como autenticación multifactor (MFA), segmentación de redes, auditorías digitales y firmas criptográficas para imágenes médicas, su implementación todavía no es homogénea en todas las instituciones de salud, especialmente en entornos con limitaciones tecnológicas o presupuestales. Asimismo, se evidenció que la literatura científica relacionada con ataques adversariales en radiología digital aún se encuentra en desarrollo, especialmente en contextos clínicos reales, lo que demuestra la necesidad de continuar investigando mecanismos de protección, regulación y validación de sistemas de inteligencia artificial en salud.

Por último, la investigación permitió evidenciar que existe una necesidad creciente de fortalecer los marcos regulatorios y éticos relacionados con inteligencia artificial en radiología. Diversos autores, entre ellos Wald (2025), Resühr (2025) y Domínguez y Castaño (2023), coinciden en que el avance acelerado de la automatización diagnóstica ha superado parcialmente el desarrollo de normativas específicas orientadas a regular la responsabilidad profesional, la protección de datos clínicos y la validación segura de sistemas de IA en entornos hospitalarios.

Conclusiones

La investigación permitió concluir que la creciente incorporación de modelos de deep learning en radiología digital ha transformado significativamente los procesos diagnósticos, pero también ha introducido nuevas vulnerabilidades tecnológicas. La alta sensibilidad de los algoritmos de Inteligencia Artificial a perturbaciones mínimas convierte a la radiología digital en un objetivo crítico para ataques adversariales, especialmente porque una alteración aparentemente insignificante puede modificar un diagnóstico médico y afectar directamente la vida del paciente.

Se concluye además que los ataques adversariales no deben considerarse únicamente como fallos técnicos o incidentes informáticos, sino como eventos adversos clínicos con potencial impacto sobre la seguridad del paciente. La manipulación de imágenes médicas puede ocasionar retrasos terapéuticos, tratamientos incorrectos, intervenciones quirúrgicas innecesarias o la omisión de enfermedades graves, comprometiendo la calidad de la atención en salud y la confianza en los sistemas automatizados de diagnóstico.

Otro aspecto relevante identificado durante la investigación es la necesidad de fortalecer los marcos éticos y legales relacionados con el uso de IA en medicina. La evolución de normativas internacionales de protección de datos y seguridad sanitaria, como HIPAA, deberá contemplar responsabilidades específicas frente a diagnósticos erróneos derivados de sistemas de Inteligencia Artificial manipulados o vulnerados cibernéticamente. Esto implica definir claramente responsabilidades institucionales, técnicas y profesionales ante posibles daños al paciente.

Asimismo, se concluye que la seguridad en radiología digital debe evolucionar desde modelos tradicionales de protección perimetral hacia estrategias integrales de resiliencia

cibernética. No basta con proteger el acceso a los sistemas; es necesario implementar mecanismos permanentes de ciberhigiene, monitoreo continuo, segmentación de redes hospitalarias, cifrado robusto y firmas digitales que permitan garantizar la autenticidad e integridad de las imágenes médicas desde el momento de su adquisición hasta su interpretación diagnóstica.

Finalmente, la investigación resalta la importancia del factor humano dentro de los entornos clínicos digitalizados. La capacitación constante del personal de salud resulta fundamental para evitar una dependencia excesiva o "confianza ciega" en los resultados generados por la IA. Los profesionales deben mantener una postura crítica y analítica frente a las recomendaciones automatizadas, verificando la coherencia clínica de los hallazgos y fortaleciendo así la seguridad diagnóstica y la protección del paciente frente a posibles manipulaciones tecnológicas.

En resumen, la investigación permitió evidenciar que la incorporación de inteligencia artificial en radiología digital representa un avance significativo para el diagnóstico médico; sin embargo, también introduce desafíos importantes relacionados con la ciberseguridad, la integridad de la información clínica y la seguridad del paciente. Por ello, resulta fundamental continuar desarrollando investigaciones orientadas al fortalecimiento de mecanismos de protección, regulación y supervisión de sistemas de IA en salud, con el fin de garantizar una implementación tecnológica segura, ética y confiable en los entornos hospitalarios.

Bibliografía

ASEFARMA. (24 de 01 de 2022). *¿Cuáles son los PRM en farmacia?* https://www.asefarma.com

Bortsova, G. (2021). *Vulnerabilidad ante ataques adversarios de los sistemas de análisis de imágenes médicas: factores inexplorados*. https://www.sciencedirect.com/science/article/pii/S1361841521001870

Cooper, V. (21 de Octubre de 2025). *Vulnerabilidades de Día Cero: Riesgos Clave y Estrategias de Defensa*. https://www.splashtop.com/es/blog/zero-day-vulnerabilities

Cooper, V. (21 de 10 de 2025). *Vulnerabilidades de Día Cero: Riesgos Clave y Estrategias de Defensa*. Splashtop: https://www.splashtop.com/es/blog/zero-day-vulnerabilities

Domínguez, D., & Castaño, O. (2023). *Ciberseguridad e inteligencia artificial: principales retos y posibles soluciones en el ámbito de la salud*. https://iasalut.cat/es/noticia/ciberseguretat-i-intelligencia-artificial-reptes-principals-i-possibles-solucions-en-lambit-de-la-salut/

Ganen, O. R. (2017). *La dispensación como herramienta para lograr el uso* . https://www.medigraphic.com/pdfs/revcubmedgenint/cmi-2017/cmi174g.pdf

Goris, A. G. (2015). *Utilidad y tipos de revisión de literatura*. https://scielo.isciii.es/scielo.php?script=sci_arttext&pid=S1988-348X2015000200002

Iqba, S. (2023). *Sobre el análisis de imágenes médicas mediante técnicas tradicionales de aprendizaje automático y redes neuronales convolucionales*. https://link.springer.com/article/10.1007/s11831-023-09899-9

Johnson, E. M. (2 de Enero de 2020). *Radiología, dispositivos móviles e Internet de las cosas (IoT)*. https://pmc.ncbi.nlm.nih.gov/articles/PMC7256153/

Kelly, B. S. (2023). *Consideraciones de ciberseguridad para los departamentos de radiología que utilizan inteligencia artificial*. https://pmc.ncbi.nlm.nih.gov/articles/PMC10667413/

López, D. I. (2023). *Impacto de la Inteligencia Artificial en la Radiología*. http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S1684-18592023000100013

Machado, J. (Abril de 2011). *FARMACOVIGILANCIA DE INTERACCIONES MEDICAMENTOSAS EN PACIENTES AFILIADOS AL SISTEMA DE SALUD DE COLOMBIA*. http://www.scielo.org.co/scielo.php?script=sci_arttext&pid=S0124-81462011000100005

Navarro, F. (2025). *Un algoritmo de cifrado de imágenes que mejora la privacidad para proteger las imágenes médicas*. https://www.mdpi.com/2073-8994/17/9/1470

Oñatibia-Astibia, A. (29 de 03 de 2021). *El papel del farmacéutico comunitario en la detección y disminución de los errores de medicación: revisión sistemática exploratoria*. https://scielo.isciii.es/scielo.php?pid=S2340-98942021000100015&script=sci_arttext&tlng=en

Perez, A. (2025). *Transformación de la imagen médica: El papel de la integración de la inteligencia artificial en los sistemas PACS para mejorar la precisión diagnóstica y la eficiencia del flujo de trabajo*. https://www.sciencedirect.com/org/science/article/pii/S1573405625000670

Pérez-Ricart, A. (14 de 10 de 2019). *Integración de la farmacovigilancia en la rutina del servicio de farmacia: nueve años de experiencia*. https://scielo.isciii.es/scielo.php?pid=S1130-63432019000400004&script=sci_arttext&tlng=es

PROMEDCO. (2023). *EL FUTURO DE LA TECNOLOGÍA DE IMÁGENES MÉDICAS EN COLOMBIA*. https://www.promedco.com/noticias/avances-tecnologicos-de-las-imagenes-diagnosticos

Puentes, G., & Salinas Miranda, E. (2022). *Inteligencia artificial y radiología: la disrupción tecnológica en la transformación de un paradigma*. https://revistamedicina.net/articulo/revmedicinacolombia-1552

Ramanathan, V. (02 de Febrero de 2022). *Todas las diferencias entre RIS, PACS, DICOM y CIS*. https://www.ramsoft.com/es/blog/what-is-ris-pacs-dicom-cis

Reed, J. (2024). *Cuando el ransomware mata: ataques a centros sanitarios*. https://www.ibm.com/es-es/think/insights/when-ransomware-kills-attacks-on-healthcare-facilities

Resühr, D. (2025). *Lo bueno, lo malo y lo feo de la IA en la imagenología médica*. https://www.emjreviews.com/en-us/amj/radiology/article/the-good-the-bad-and-the-ugly-of-ai-in-medical-imaging-j140125/

Rodríguez, S. G. (2025). *El futuro de la radiología Digital: transformación en la imagenología y mejora de la precisión diagnóstica*. https://www.revista-portalesmedicos.com/revista-medica/el-futuro-de-la-radiologia-digital-transformacion-en-la-imagenologia-y-mejora-de-la-precision-diagnostica/

Sánchez, J. (2022). *Siete riesgos de la inteligencia artificial en salud*. https://www.telefonicaempresas.es/grandes-empresas/blog/riesgos-inteligencia-artificial-salud/

Téllez, R. I. (2025). *Vulnerabilidad de datos sensibles en sistemas de salud*. https://repository.unad.edu.co/bitstream/handle/10596/74288/Ritellezm10.pdf?sequence=3&isAllowed=y

Vega, J. G. (2024). *Estudio e implementación de redes*. https://gredos.usal.es/bitstream/handle/10366/163859/memoria.pdf?sequence=1

Vijay Ramanathan. (02 de Febrero de 2022). *Todas las diferencias entre RIS, PACS, DICOM y CIS*. https://www.ramsoft.com

Wald, C. (Octubre de 2025). *Consenso, marcos y mejores prácticas en una revisión de los estándares de ciberseguridad y privacidad en imágenes médicas*. https://www.sciencedirect.com/science/article/pii/S1546144025003436