ISSN: 1692-7257 - Volumen 1 – Número 47 - 2026

ORIGINAL

RCTA
Revista Colombiana de Tecnologías de Avanzada
UNIPAMPLONA

Comparación de modelos de inteligencia artificial para la detección de cardiopatías usando imágenes de resonancia magnética

Comparison of artificial intelligence models for heart disease detection using magnetic resonance imaging

Jorge Eliecer Gómez Gómez[1]
Juan José Gúzman Pineda[1]
Velssy Liliana Hernández Riaño[1]

[1] **Universidad de Córdoba, Departamento de Ingeniería de sistemas, grupo Sócrates, Montería, Colombia.**

Recibido: 21 julio 2025. Revisado: 21 septiembre 2025. Aceptado: 20 diciembre 2025. Publicado: 01 enero 2026.
Autor de correspondencia: Jorge Eliecer Gómez Gómez

Cómo citar: J. E. Gómez Gómez, J. J. Gúzman Pineda, and V. L. Hernández Riaño, "Comparación de modelos de inteligencia artificial para la detección de cardiopatías usando imágenes de resonancia magnética", *Revista Colombiana de Tecnologías de Avanzada (RCTA)*, vol. 1, no. 47, 2026, art. no. 4285, doi: 10.24054/rcta.v1i47.4285

Esta obra está bajo una licencia internacional Creative Commons Atribución-NoComercial 4.0.

Resumen:

En este trabajo se compararon tres técnicas de inteligencia artificial para la detección binaria (sano/enfermo) de cardiopatías mediante cortes axiales de resonancia magnética cardíaca. Usando un conjunto de 150 pacientes en formato NIfTI, las imágenes se preprocesaron (normalización, reescalado a 128 × 128, conversión a RGB y aumento de datos) y se dividieron en una proporción 80/20 por paciente. Se evaluaron Random Forest con descriptores radiómicos GLCM y estadísticos de primer orden, una red neuronal convolucional (CNN) y un modelo basado en YOLOv8 adaptado a clasificación binaria. Los modelos se compararon mediante accuracy, precision, recall, F1-score y AUC, y se aplicaron técnicas de explicabilidad (SHAP, Grad-CAM, Integrated Gradients y sensibilidad a la oclusión) para validar la coherencia anatómica de las predicciones. En conjunto, los resultados indican que los enfoques de aprendizaje profundo preentrenado, como YOLOv8, ofrecen ventajas sustanciales en términos de precisión e interpretabilidad, posicionándose como una alternativa prometedora para el desarrollo de sistemas inteligentes de apoyo al diagnóstico de cardiopatías estructurales.

Palabras clave: cardiopatías, resonancia magnética cardíaca, YOLOv8, Random Forest, CNN, inteligencia artificial.

Abstract:

In this study, three artificial intelligence techniques were compared for the binary detection (healthy/diseased) of cardiopathies using axial slices from cardiac magnetic resonance imaging. Using a dataset of 150 patients in NIfTI format, the images were preprocessed (normalization, rescaling to 128 × 128, RGB conversion, and data augmentation) and split on a per-patient basis using an 80/20 ratio. A Random Forest model with GLCM radiomic descriptors and first-order statistical features, a convolutional neural network (CNN), and a YOLOv8-based model adapted for binary classification were evaluated. The models were compared using accuracy, precision, recall, F1-score, and AUC, and explainability techniques (SHAP, Grad-CAM, Integrated Gradients, and occlusion sensitivity) were applied to validate the anatomical coherence of the predictions. Overall, the results indicate that pretrained deep learning approaches, such as YOLOv8, offer substantial advantages in terms of accuracy and interpretability, positioning them as a promising alternative for the development of intelligent decision-support systems for the diagnosis of structural cardiopathies.

Keywords: heart disease, MRI, YOLOv8, Random Forest, CNN, artificial intelligence.

INTRODUCCIÓN

Las enfermedades cardiovasculares siguen siendo la principal causa de muerte en el mundo, representando aproximadamente el 32% de todas las defunciones globales cada año, según la Organización Mundial de la Salud (OMS)[1]. Se estima que más de 17.9 millones de personas fallecen anualmente por causas relacionadas con el corazón y los vasos sanguíneos, una cifra que continúa en ascenso debido al envejecimiento poblacional, los cambios en los estilos de vida y la limitada cobertura de estrategias preventivas. La OMS ha enfatizado la necesidad de intervenciones tempranas, especialmente en países de ingresos medios y bajos, donde la carga de enfermedad cardiovascular ha mostrado una tendencia alarmante al alza.

En Colombia, esta situación no es ajena. El Ministerio de Salud y Protección Social ha reportado que, durante los últimos años, las enfermedades del sistema circulatorio han ocupado consistentemente los primeros lugares como causa de muerte en el país, siendo la enfermedad isquémica del corazón, los accidentes cerebrovasculares y las enfermedades hipertensivas las más prevalentes [2]. En 2022, se estimó una tasa de mortalidad por enfermedades cardiovasculares cercana a 176 muertes por cada 100 000 habitantes, con una carga mayor en adultos mayores y zonas rurales [3]. A pesar de avances en infraestructura hospitalaria y cobertura, persisten desafíos estructurales relacionados con el acceso equitativo al diagnóstico oportuno, especialmente en regiones con escasa cobertura especializada. En este escenario, la adopción de tecnologías capaces de facilitar la detección temprana y eficiente de enfermedades cardiovasculares, como soluciones basadas en inteligencia artificial, se perfila como una prioridad de salud pública [4].

Frente a este panorama, la detección precoz y precisa de condiciones cardíacas, incluso antes de que se manifiesten clínicamente, se ha convertido en un objetivo clave de la medicina moderna. Entre las herramientas diagnósticas disponibles, la resonancia magnética cardíaca (IRM) destaca por su capacidad para ofrecer imágenes detalladas de la morfología y función del corazón, sin exposición a radiación ionizante [5], [6]. Sin embargo, el análisis tradicional de estas imágenes depende de la interpretación visual por parte de expertos, una práctica sujeta a variabilidad inter observador, demandante en tiempo y no siempre accesible en todos los entornos clínicos [7].

En este contexto, la inteligencia artificial (IA) ha emergido como una solución prometedora para automatizar procesos diagnósticos, reducir la carga clínica y mejorar la consistencia del análisis [8]. Se han desarrollado múltiples modelos basados en aprendizaje profundo que permiten segmentar cavidades cardíacas, cuantificar volúmenes o clasificar patologías específicas a partir de imágenes médicas [9], [10]. Sin embargo, en la práctica clínica diaria, muchas decisiones iniciales se reducen a una pregunta fundamental: ¿existe o no una condición cardíaca que amerite atención?

Abordar esta pregunta desde un enfoque binario presencia o ausencia de enfermedad resulta especialmente relevante en escenarios como cribado poblacional, triaje en urgencias o telemedicina [11], [12]. En tales contextos, modelos ligeros, interpretables y eficientes, capaces de ofrecer respuestas confiables a partir de imágenes básicas, podrían ser determinantes para reducir tiempos diagnósticos y priorizar adecuadamente los recursos clínicos disponibles [13].

Más allá del rendimiento técnico, la confiabilidad de estos sistemas juega un papel fundamental en su adopción. Diversos estudios han evidenciado que los modelos de IA pueden alcanzar niveles de precisión comparables o superiores a los de expertos humanos, especialmente cuando se entrenan con conjuntos de datos bien etiquetados y heterogéneos. No obstante, su implementación requiere no solo validación cuantitativa, sino también mecanismos que permitan explicar y justificar cada decisión clínica, con el fin de generar confianza entre los profesionales de la salud y pacientes por igual.

La incorporación de técnicas de explicabilidad ha demostrado ser una estrategia efectiva para mejorar la transparencia de estos sistemas, permitiendo identificar qué regiones de la imagen influyeron más en la predicción realizada[12]. Esto no solo facilita auditorías médicas y segundas opiniones, sino que además respalda la integración ética y legal de la IA en sistemas de salud estructurados, particularmente en procesos como tamizajes automáticos, diagnósticos asistidos y segundas lecturas remotas.

Este artículo se enfoca en el desarrollo y validación de modelos de detección binaria de enfermedades cardíacas basado en IRM y aprendizaje profundo. A través de la implementación de aprendizaje automático y redes neuronales, se busca ofrecer un mecanismo eficaz, reproducible y clínicamente viable para apoyar la toma de decisiones tempranas en cardiología, con especial atención a entornos de alta demanda o recursos limitados.

ESTADO DEL ARTE

En esta sección se analizan de manera crítica los estudios más relevantes, agrupados en cinco líneas temáticas que representan las principales tendencias actuales de investigación.

Segmentación automática de estructuras cardíacas

Una de las líneas más consolidadas en la literatura es la segmentación automática de estructuras cardíacas a partir de IRM. En este campo, el trabajo de[7]marcó un hito al presentar los resultados del Automatic Cardiac Diagnosis Challenge (ACDC), donde se evaluaron diversos métodos de aprendizaje profundo aplicados a la segmentación del ventrículo izquierdo (LV),

ventrículo derecho (RV) y el miocardio (MYO). Este benchmark estableció un estándar en el área, logrando coeficientes Dice promedio de 0.95 para LV, 0.89 para RV y 0.89 para MYO, utilizando principalmente arquitecturas basadas en U-Net. A pesar de los avances, los autores señalaron limitaciones en cuanto a la generalización de los modelos fuera del conjunto ACDC, lo que impulsó investigaciones posteriores enfocadas en robustez y transferibilidad.

En respuesta a estas limitaciones, en[14]propusieron una red convolucional que integra mapas de características direccionales (Directional Feature Maps, DFM), diseñados para capturar patrones espaciales orientados propios de las estructuras cardíacas, como bordes curvos y límites entre el miocardio y las cavidades ventriculares. Este enfoque permitió superar las limitaciones de arquitecturas isotrópicas como U-Net, alcanzando coeficientes Dice de 0.955, 0.901 y 0.888 para LV, RV y MYO respectivamente, sobre el conjunto ACDC. El valor añadido de esta arquitectura radica en su capacidad para abordar tanto la indistinción interclase como la variabilidad intraclase, aspectos críticos para un diagnóstico clínico confiable.

Por su parte, en[15]desarrollaron Heart-Net, una arquitectura multimodal que combina imágenes IRM con señales fisiológicas (ECG y presión arterial), integrando una CNN y una LSTM en un enfoque de fusión tardía. Esta aproximación no solo mejora la precisión segmentando las estructuras cardíacas, sino que también permite clasificar patologías cardíacas con una precisión superior al 98%. La combinación de información morfológica y funcional simula el razonamiento diagnóstico médico y ha demostrado ser altamente robusta frente a variaciones anatómicas y ruido clínico.

Complementando estas propuestas, en[16]introdujeron un sistema híbrido de diagnóstico asistido por computador (CAD) que emplea ResU-Net para la segmentación de las estructuras cardíacas y un Vision Transformer (ViT) para la clasificación de regiones infartadas. Este enfoque fue particularmente efectivo en la identificación de tejido miocárdico dañado, logrando un AUC del 0.98 y una precisión diagnóstica global del 97.3%. La capacidad de los transformers de capturar relaciones espaciales de largo alcance lo hace ideal para detectar patrones diseminados y sutiles, lo cual resulta fundamental en el diagnóstico temprano del infarto de miocardio.

Clasificación automática de enfermedades cardíacas

La clasificación directa de patologías cardíacas a partir de IRM representa una evolución significativa respecto a enfoques tradicionales que dependen de múltiples etapas intermedias. En [11] presentaron SA-YOLO, una versión adaptada de YOLOv8 con módulos de atención espacial que permite la detección de patologías directamente desde las imágenes, sin necesidad de segmentaciones previas. Este modelo alcanzó una precisión global del 98.1% y un tiempo de inferencia de 23 ms por imagen, haciéndolo viable para su uso en tiempo real en entornos clínicos con limitaciones computacionales.

Mientras tanto, en[17]introdujeron un enfoque innovador al utilizar campos de deformación entre las fases diastólica y sistólica para codificar la dinámica contráctil del corazón; usando una arquitectura basada en autoencoders y un clasificador SVM, lograron una precisión media del 94.5%. Este tipo de entrada funcional permite identificar disfunciones contráctiles propias de enfermedades como la miocardiopatía dilatada (DCM) o hipertrófica (HCM), proporcionando una alternativa más informativa que las imágenes estáticas.

En otro estudio destacable, en[18]aplicaron Vision Transformers en combinación con técnicas de explicabilidad como Grad-CAM para detectar miocarditis. Su modelo alcanzó un F1-score de 0.9741 un AUC de 09616, destacando tanto por su precisión como por su capacidad interpretativa. Este trabajo representa un avance importante al integrar explícitamente la visualización de las regiones activas que motivan la clasificación, mejorando la transparencia del proceso diagnóstico automatizado.

Finalmente, en[12]adoptaron un enfoque complementario al integrar algoritmos como XGBoost y LightGBM sobre variables clínicas multifuente, sin depender exclusivamente de IRM. Aunque no segmentan imágenes, logran una AUC promedio de 0.989, demostrando que la combinación de datos estructurados con modelos de alto rendimiento puede ofrecer predicciones confiables incluso en ausencia de imagenología detallada.

Explicabilidad e interpretabilidad de modelos

La aceptabilidad de modelos de inteligencia artificial en contextos clínicos depende no solo de su precisión, sino también de su capacidad de ser interpretados. En este sentido, en[19]propusieron una arquitectura de aprendizaje conjunto que simultáneamente segmenta el miocardio y detecta tejido cicatricial post-infarto; el modelo incorpora un mecanismo de atención que resalta las regiones relevantes en las predicciones, logrando una sensibilidad del 91.8% y una precisión diagnóstica del 94.6%. La coherencia espacial entre las tareas de segmentación y detección funcional mejora la confianza clínica en los resultados.

Asimismo, el trabajo de[18]sobresale por incorporar Grad-CAM como herramientas de explicabilidad en un contexto sensible como el diagnóstico de miocarditis. Estas técnicas permiten validar visual y cuantitativamente las decisiones del modelo, mostrando qué regiones específicas influyeron en la predicción, lo que fortalece la transparencia y la confianza médica.

Una de las barreras para la adopción de modelos de IA en medicina es la falta de validación clínica en escenarios reales.

En este aspecto, en[5]realizaron un estudio prospectivo en 61 pacientes comparando secuencias cine convencionales con aquellas reconstruidas mediante aprendizaje profundo. Su modelo demostró una concordancia excelente con las medidas funcionales tradicionales ($r > 0.98$) y redujo el tiempo de adquisición en un 50%, mejorando la calidad visual, especialmente en estructuras subendocárdicas.

Validación clínica y adopción en entornos hospitalarios

Por otro lado, en[20],[6]realizaron revisiones sistemáticas del uso de IA en imagenología cardiovascular, abordando temas como la integración en flujos hospitalarios, la regulación por entidades como la FDA, y la trazabilidad de los modelos. Estas revisiones destacan que, más allá del rendimiento técnico, la aceptación clínica de un modelo depende de su transparencia, tiempo de inferencia aceptable y compatibilidad con sistemas HIS/PACS.

Diagnóstico preventivo y enfoques multimodales

La aplicación de IA en medicina no se limita al diagnóstico actual, sino que se extiende hacia la predicción preventiva del riesgo cardiovascular. En[21]demostraron que redes profundas pueden predecir eventos cardíacos futuros a partir de imágenes hepáticas, con un AUC de 0.87, al combinar IRM abdominal y variables clínicas mediante un transformer multimodal. Esta perspectiva rompe con la lógica tradicional de diagnóstico estructural y permite explorar correlaciones latentes en otras modalidades de imagen.

En el mismo sentido, en[22]entrenaron una CNN tipo ResNet con módulos de atención para clasificar cinco patologías cardíacas directamente desde IRM cine sin segmentación previa. Su modelo logró una precisión del 96.4% y un tiempo de inferencia inferior a un segundo, lo que lo convierte en una opción viable para clínicas con alta carga de estudios y poco personal especializado.

Finalmente, en[13]sintetizaron los avances en IA para enfermedades miocárdicas, destacando el papel emergente de modelos generativos, redes recurrentes y simulaciones virtuales. Estos enfoques, al integrar múltiples dimensiones temporales y morfológicas, amplían el espectro diagnóstico y sugieren nuevas vías para personalizar tratamientos y simular progresiones clínicas.

MATERIALES Y MÉTODOS

Descripción del conjunto de datos

El conjunto de datos utilizado corresponde a estudios de resonancia magnética cardíaca almacenados en formato NIfTI, obtenido de[7], con 150 pacientes, dividido en 5 subgrupos (30 sujetos normales – NOR; 30 pacientes con infarto de miocardio previo; 30 pacientes con miocardiopatía dilatada; 30 pacientes con miocardiopatía hipertrófica; 30 pacientes con ventrículo derecho anormal). Cada volumen incluye cortes axiales de diferentes momentos del ciclo cardíaco (por ejemplo, frame01, frame12), acompañados de máscaras segmentadas manualmente (_gt.nii.gz). Estas máscaras delimitan estructuras como el miocardio, el ventrículo izquierdo (LV) y derecho (RV).

Se aplicó una normalización min-max por paciente (basada en los valores de intensidad de cada volumen). Se interpolaron los cortes axiales a 128×128 o 256×256 píxeles mediante bicúbica, se eliminó ruido de fondo con umbral adaptativo, y las intensidades fueron convertidas a 8 bits. En el caso de YOLOv8, las imágenes fueron duplicadas a tres canales RGB para compatibilidad. El mismo pipeline fue aplicado uniformemente a entrenamiento y validación.

Las imágenes provienen de un repositorio clínico estructurado en subcarpetas por paciente, con variabilidad anatómica y temporal. Esta riqueza permite entrenar modelos robustos para la clasificación binaria (presencia o no de enfermedad), así como experimentar con tareas de segmentación. El dataset fue montado desde Google Drive en un entorno colaborativo utilizando Google Colab.

Flujo General de Procesamiento

Como paso previo a la implementación de cada modelo, se definió un flujo genérico de procesamiento orientado a garantizar la calidad y consistencia de los datos. Este flujo, representado en la Fig. 1, comienza con un proceso de limpieza y normalización de los volúmenes de resonancia magnética cardíaca, seguido por un módulo de preprocesamiento que incluye el recorte de la región de interés, la selección de cortes axiales representativos, el reescalado a resoluciones estándar (128×128 o 256×256 píxeles), y la conversión a formatos compatibles con cada arquitectura (escala de grises o RGB). Posteriormente, se aplicaron técnicas de selección y transformación de características según el enfoque empleado (flattening en Random Forest, convoluciones en CNN, y procesamiento RGB directo en YOLOv8). Finalmente, cada modelo fue entrenado y evaluado de forma independiente usando las mismas métricas de desempeño para permitir una comparación objetiva. Este enfoque estructurado asegura la trazabilidad del pipeline y la replicabilidad de los experimentos.

El conjunto de datos utilizado se compone de volúmenes tridimensionales en formato .nii.gz, los cuales fueron divididos

80/20 a nivel de paciente, agrupando todos los cortes de cada sujeto en el mismo conjunto, con estratificación por clase (sano/enfermo). Esto evitó fuga de información y garantizó independencia entre entrenamiento y validación, y procesados previamente para obtener cortes axiales 2D representativos de cada estudio.

Preparación de datos

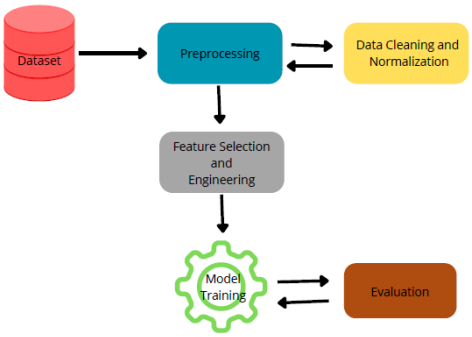

Fig. 1. *Flujo general del procesamiento*

Las etapas del preprocesamiento incluyeron:

- Extracción del corte medio axial del volumen, asumiendo que este plano proporciona una vista representativa del corazón.

- Normalización de las intensidades de píxeles al rango [0, 1] mediante escalado min-max, con el objetivo de reducir la variabilidad derivada de condiciones de adquisición heterogéneas.

- Redimensionamiento uniforme de todas las imágenes a dimensiones fijas de 128×128 píxeles.

- Conversión a formato plano: dado que el modelo Random Forest requiere entradas vectorizadas, cada imagen fue convertida de su forma matricial (128,128,1) a un vector unidimensional de 16,384 características mediante aplanamiento.

Esta transformación permitió representar cada imagen como un punto en un espacio de características de alta dimensionalidad, acompañado de una etiqueta binaria correspondiente a la condición clínica del paciente (sano = 0, enfermo = 1).

Para mejorar la capacidad discriminativa, se extrajeron características radiómicas basadas en textura GLCM (contrast, dissimilarity, homogeneity, energy, correlation) y estadísticos de primer orden (media, varianza, curtosis, asimetría). Estos descriptores se calcularon por corte usando la librería scikit-image y sirvieron como entrada al modelo de Random Forest

Aumento de datos

Dado el tamaño reducido del conjunto de datos y su distribución levemente desbalanceada, se aplicó una estrategia de data augmentation utilizando ImageDataGenerator. Esta técnica permite aumentar la diversidad del conjunto de entrenamiento mediante transformaciones aleatorias, sin necesidad de recolectar nuevas imágenes. Las transformaciones incluyeron: rotaciones aleatorias (±10°), zoom aleatorio (±10%), desplazamientos horizontales y verticales (±10%), y volteo horizontal. Estas operaciones preservan las características anatómicas relevantes y permiten mejorar la capacidad de generalización del modelo.

RESULTADOS

Desempeño del modelo Random Forest

El modelo Random Forest fue entrenado y evaluado utilizando imágenes aplanadas provenientes de cortes axiales de resonancia magnética cardíaca. Tras el entrenamiento, su desempeño fue evaluado sobre un conjunto de validación previamente separado (20% del total de muestras), obteniéndose los siguientes resultados cuantitativos:

Precisión global (Accuracy): 0.8000

Precisión (Precision): 0.8000

Sensibilidad (Recall): 1.0000

F1-score: 0.8889

Área bajo la curva ROC (AUC): 0.6424

En la Figura 2 se presenta la curva ROC correspondiente al modelo. El área bajo la curva (AUC = 0.6424) indica una capacidad de discriminación moderada del modelo para distinguir entre pacientes sanos y enfermos. Si bien la sensibilidad es elevada,

la presencia de falsos positivos reduce la precisión global del clasificador.

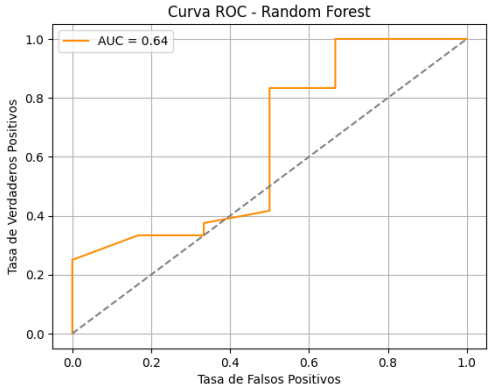

Fig. 2. *Curva ROC del modelo Random Forest.*

El desempeño del modelo Random Forest se caracteriza por una elevada capacidad para detectar casos positivos (enfermos), lo que lo convierte en una herramienta potencialmente útil en contextos donde la omisión de casos críticos es inaceptable (alta sensibilidad). No obstante, el número de falsos positivos registrados sugiere que puede generar alertas innecesarias en algunos pacientes sanos, lo cual debe ser considerado en aplicaciones clínicas reales. Este comportamiento también puede atribuirse a la pérdida de información espacial durante el proceso de aplanamiento de las imágenes, lo que limita la capacidad del modelo para diferenciar patrones anatómicos sutiles.

Desempeño del modelo CNN

El modelo fue entrenado y su evaluación se empleó sobre un conjunto de compuesto por 30 muestras. Los resultados obtenidos por la red convolucional son los siguientes:

Precisión global (Accuracy): 0.7000

Precisión clase CHD: 0.8000

Recall clase CHD: 0.8333

F1-score clase CHD: 0.8163

La Fig. 3 muestra la matriz de confusión del modelo CNN. Se observa que la red neuronal identificó correctamente 20 casos enfermos (CHD) y 1 caso sano (No CHD). Sin embargo, cometió 5 falsos positivos, es decir, clasificó incorrectamente pacientes sanos como enfermos, y 4 falsos negativos, correspondientes a pacientes enfermos que fueron clasificados como sanos.

Este comportamiento refleja que el modelo presenta un ligero sesgo hacia la clase positiva, privilegiando la detección de casos patológicos por encima de la correcta identificación de sujetos sanos. Tal tendencia puede atribuirse al desbalance en el conjunto de datos (mayor cantidad de imágenes de pacientes enfermos) o a una limitada capacidad de generalización del modelo para aprender patrones representativos del miocardio normal.

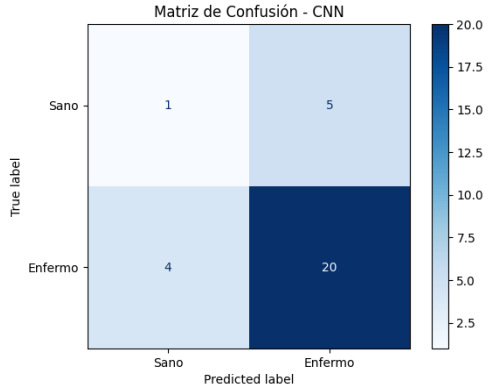

Fig. 3. *Matriz de confusión del modelo CNN.*

La curva ROC del modelo, la cual presenta un AUC de 0.4167, indicando un poder discriminativo inferior al límite del azar, lo cual no es correcto en el área de la salud. A pesar de que la red logró una sensibilidad aceptable, el bajo AUC y la incapacidad

de clasificar correctamente los casos negativos limitan su utilidad diagnóstica.

A pesar de su elevada sensibilidad para detectar casos de CHD, el modelo CNN presentó deficiencias significativas en la identificación de pacientes sanos, lo cual podría resultar en un número elevado de falsos positivos en un escenario clínico. Este comportamiento puede estar relacionado con el desequilibrio del conjunto de datos y la posible necesidad de ajustes adicionales en la arquitectura o en los métodos de regularización. En futuras versiones del modelo se recomienda aplicar técnicas como focal loss, class weighting o aumento específico para la clase minoritaria.

Desempeño del modelo YOLOv8

El desempeño alcanzado por este modelo fue el siguiente:

Precisión global (Accuracy): 0.8000

Precisión (Precision): 0.8000

Sensibilidad (Recall): 1.000

F1-score: 0.8900

La Fig. 4 muestra la curva ROC obtenida a partir de las probabilidades de clasificación del modelo YOLOv8 para la clase positiva (enfermo). El área bajo la curva (AUC) fue de 0.83, lo que refleja una capacidad de discriminación buena, aunque no sobresaliente. La curva presenta una forma ascendente y se aproxima parcialmente al vértice superior izquierdo, indicando que el modelo logra una alta tasa de verdaderos positivos con un número moderado de falsos positivos.

Este resultado confirma que YOLOv8 es capaz de distinguir con eficacia entre sujetos sanos y enfermos, aunque su desempeño podría mejorar mediante una calibración más precisa del umbral de decisión o un mayor equilibrio entre clases.

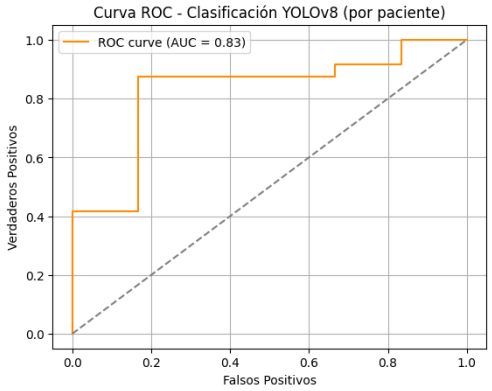

Fig. 4. *Curva ROC del modelo YOLOv8.*

El modelo mostró un rendimiento superior frente a los otros enfoques evaluados, tanto en términos de precisión como de equilibrio entre clases. Su capacidad para generalizar correctamente sobre un conjunto de validación diverso, sumado a su baja tasa de error y su alta puntuación AUC, lo posiciona como una alternativa altamente viable para la detección automatizada de enfermedades cardíacas congénitas a partir de imágenes médicas. La estabilidad del modelo también sugiere un aprendizaje efectivo de patrones discriminativos sin sobreajuste, incluso en un entorno clínico con recursos moderados.

Explicabilidad de los modelos

Random Forest

Con el propósito de comprender cómo el modelo toma sus decisiones y validar la relevancia fisiológica de las características utilizadas, se implementaron dos enfoques complementarios de explicabilidad global y local:

SHAP (SHapley Additive Explanations), basado en teoría de juegos cooperativos. Permutation Importance, que evalúa la variación del desempeño del modelo al alterar aleatoriamente cada variable. Ambos métodos se aplicaron sobre el conjunto de validación, empleando los seis descriptores radiómicos obtenidos a partir de la matriz de co-ocurrencia de niveles de gris (GLCM): Homogeneity, Energy, ASM (Angular Second Moment), Contrast, Dissimilarity y Correlation.

El análisis de valores absolutos medios de SHAP (ver Fig. 5) permitió identificar que las variables con mayor impacto en la predicción fueron:

- GLCM_Homogeneity (0.0449),

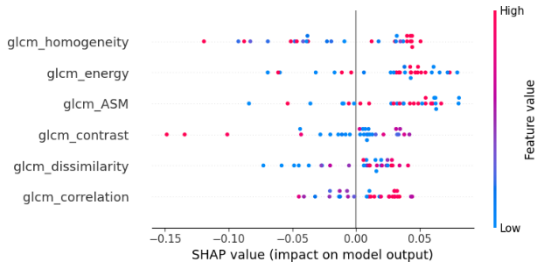

Fig. 5. *Impacto del modelo en las salidas*

- GLCM_Energy (0.0442),
- GLCM_ASM (0.0436).

Estas tres características se asocian con la uniformidad del tejido miocárdico. Los pacientes con valores altos de homogeneidad y energía tendieron a clasificarse como sanos, mientras que texturas más heterogéneas (baja homogeneidad, alto contraste) se vincularon a la clase enfermo. Las variables Contrast, Dissimilarity y Correlation tuvieron menor contribución, aunque su efecto combinado ayudó a distinguir alteraciones estructurales leves. La representación de SHAP individual mostró una tendencia coherente: los valores negativos de homogeneidad disminuyen la probabilidad de ser sano, y los positivos la aumentan, lo que confirma que el modelo internalizó un patrón clínicamente consistente.

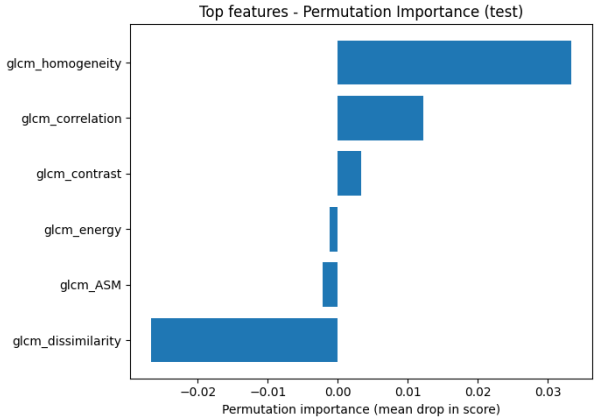

Fig. 6. *Importancia de la permutación.*

En cuanto a importancia de permutación, la Fig. 6 muestra la importancia promedio y la desviación estándar obtenidas mediante el método de permutación. Los resultados coincidieron parcialmente con SHAP: homogeneidad mantuvo el primer lugar en relevancia (0.0333), seguida de Correlación (0.0122) y Contraste (0.0033). No obstante, variables como ASM y Energy mostraron ligeras variaciones negativas, evidenciando cierta redundancia entre ellas, lo cual es esperable dado que ambas reflejan uniformidad y textura suave. La concordancia entre SHAP y Permutación, sugieren una consistencia interna alta en el modelo, la dominancia de descriptores de uniformidad implica que el modelo asocia dispersión de intensidades con patología, coherente con los cambios estructurales típicos de la miocardiopatía (hipertrofia, fibrosis o remodelación). Para finalizar, dado que Ramdon Forest no opera directamente sobre la imagen sino sobre vectores de características globales, su interpretabilidad espacial es limitada; sin embargo, ofrece una explicabilidad conceptual clara basada en propiedades estadísticas del tejido.

4.4.2. CNN

El modelo de red neuronal convolucional (CNN) fue analizado mediante tres técnicas complementarias de explicabilidad visual:

- Grad-CAM (Gradient-weighted Class Activation Mapping)
- Saliency Maps
- Integrated Gradients (IG).

Estas herramientas permiten identificar las regiones de la imagen que más influyen en la predicción final, aportando una interpretación espacial directa sobre la atención del modelo.

La técnica Grad-CAM (ver Fig. 7) generó mapas de calor que indican qué regiones del corte axial activaron más intensamente las capas convolucionales al clasificar un paciente como enfermo.

En la mayoría de los casos positivos, las zonas rojas y amarillas se concentraron en el ventrículo izquierdo y el septo interventricular, áreas comúnmente afectadas en cardiopatías hipertróficas y dilatadas. En pacientes sanos, las activaciones fueron más uniformes y dispersas, lo que refleja menor atención del modelo a regiones específicas.

En la Fig. 8, se aprecia que se calcularon los gradientes de la salida respecto a cada píxel de entrada, evidenciando qué zonas provocan los mayores cambios en la probabilidad de clasificación.

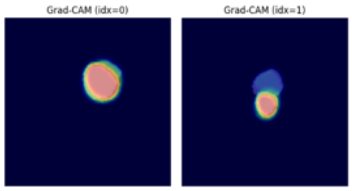

Fig. 7. *Mapas de calor Grad-CAM*

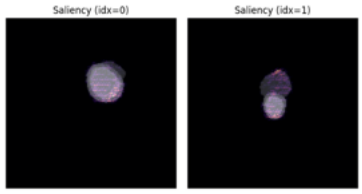

Fig. 8. *Mapas de Saliencia*

Estas visualizaciones resaltaron contornos del corazón y bordes ventriculares, indicando que el modelo basó sus decisiones en la forma y límites del miocardio. El patrón observado confirma que la CNN aprende rasgos de textura y geometría, pero no logra discriminar completamente entre sanos y enfermos, ya que muchos bordes aparecen destacados incluso en cortes sin patología.

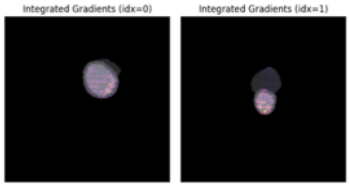

Fig. 9. *Gradientes integrados*

La técnica gradientes integrados como se aprecia en la Fig. 9, ofreció una perspectiva más estable, reduciendo el ruido característico de los gradientes puros.

Los resultados mostraron concentraciones de intensidad moderada sobre el miocardio medio-septal y las paredes ventriculares, lo cual coincide con las zonas que el Grad-CAM identificó como más relevantes. Esta coincidencia entre métodos refuerza la confianza en que la CNN sí está aprendiendo patrones anatómicamente plausibles, aunque su precisión global sea limitada.

4.4.3. YOLOv8

Este modelo, originalmente diseñado para detección de objetos, fue adaptado a la tarea de clasificación binaria de cardiopatías. Debido a su naturaleza completamente convolucional y a la presencia de capas de feature pyramids y self-attention, su comportamiento interno resulta menos intuitivo que el de una CNN tradicional. Por tanto, se aplicaron dos métodos de explicabilidad espacial: Grad-CAM y Sensibilidad a la oclusión, con el objetivo de analizar las regiones anatómicas que contribuyen a la predicción final y validar la coherencia de sus activaciones. Los mapas Grad-CAM generados sobre los cortes axiales más representativos de cada paciente (ver Fig. 10) evidenciaron activaciones intensas en las cavidades ventriculares y el septo interventricular en los casos clasificados como enfermos.

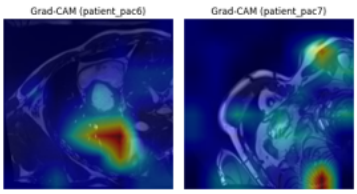

Fig. 10. *Mapas Grad-CAM (pac6 y pac7)*

En particular, los pacientes pac6 y pac7 mostraron zonas calientes (rojas y amarillas) concentradas en el ventrículo izquierdo y en la pared septal, coherentes con la localización típica de alteraciones en miocardiopatías dilatadas o hipertróficas.

En el caso del paciente pac27 (ver Fig. 11), el mapa Grad-CAM presentó activaciones ligeramente desplazadas hacia áreas periféricas, reflejando una predicción con confianza moderada (probabilidad de 0.61).

Esto sugiere que, aunque el modelo reconoce señales patológicas, su atención puede dispersarse en casos con texturas cardíacas menos pronunciadas o con bajo contraste. En pacientes sanos, las activaciones Grad-CAM se distribuyeron de forma difusa, sin concentraciones claras, lo cual coincide con la ausencia de patrones estructurales anómalos.

Para complementar el análisis, se aplicó el método de Sensibilidad a la oclusión (ver Fig. 12), el cual mide la disminución de probabilidad al enmascarar parches locales de la imagen.

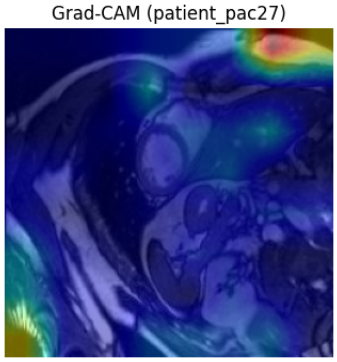

Fig. 11. *Mapas Grad-CAM(pac27)*

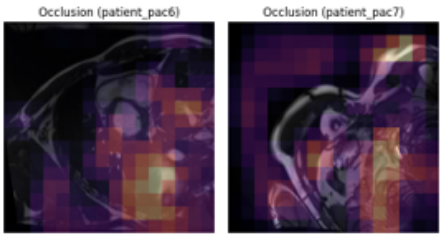

Fig. 12. *Occlusion Sensitivity (pac6, pac7)*

Los resultados mostraron que las regiones cuya oclusión provocó mayor caída de probabilidad (zonas amarillas y rojas) coincidieron con las mismas áreas destacadas por Grad-CAM: paredes ventriculares y septo interventricular.

En los pacientes enfermos, la oclusión del ventrículo izquierdo redujo la probabilidad de la clase “enfermo” en más del 40%, confirmando que estas regiones son críticamente determinantes para la decisión del modelo, en cambio, en los pacientes sanos, la oclusión no produjo cambios significativos en la salida, lo que valida la especificidad de la atención.

CONCLUSIÓN

Este estudio presentó una comparación sistemática entre tres enfoques representativos de la inteligencia artificial aplicados a la detección de cardiopatías mediante imágenes de resonancia magnética cardíaca: aprendizaje computacional clásico (Random Forest), redes neuronales convolucionales (CNN) y modelos de clasificación optimizados (YOLOv8).

La metodología abarcó desde el preprocesamiento de imágenes y la extracción de características hasta la evaluación del rendimiento con métricas estandarizadas (accuracy, precision, recall, F1-score y AUC), complementada con un análisis de explicabilidad (XAI) que permitió interpretar las decisiones de cada modelo.

Los resultados mostraron que YOLOv8 obtuvo el mejor desempeño global, con accuracy y precision del 80%, F1-score de 0.89 y un AUC de 0.83, reflejando una buena capacidad de discriminación entre pacientes sanos y enfermos, su arquitectura permitió conservar información espacial relevante y detectar patrones estructurales complejos en el miocardio, demostrando su potencial para tareas médicas críticas. Por otro lado, el Random Forest también logró un rendimiento competitivo (accuracy = 0.80; recall = 1.00), apoyándose en descriptores texturales derivados de matrices GLCM; mientras que la CNN obtuvo un desempeño inferior (AUC = 0.4167), mostrando dificultad para generalizar correctamente la clase negativa debido al desbalance de datos.

Desde la perspectiva de la inteligencia artificial explicable, los tres enfoques evidenciaron coherencia fisiológica en sus decisiones. El Random Forest basó su clasificación en propiedades texturales homogéneas del tejido miocárdico; la CNN focalizó parcialmente su atención en el ventrículo izquierdo y el septo interventricular; y YOLOv8 concentró sus activaciones en regiones anatómicamente relevantes, confirmando su capacidad de representación espacial. Estos hallazgos subrayan que la explicabilidad no solo aporta transparencia, sino también una herramienta de validación clínica complementaria al

rendimiento cuantitativo.

En el ámbito clínico, la alta sensibilidad observada, especialmente en YOLOv8 y Random Forest, resulta alentadora para el desarrollo de sistemas de apoyo al diagnóstico; sin embargo, la presencia de falsos positivos en los modelos basados en aprendizaje profundo pone de manifiesto la necesidad de recalibrar umbrales de decisión, mejorar el balance de clases e incorporar validaciones cruzadas con especialistas.

Finalmente, en este trabajo se demuestra que la inteligencia artificial es una vía viable y robusta para el diagnóstico asistido de cardiopatías estructurales. Modelos como YOLOv8 combinan rendimiento y explicabilidad, ofreciendo una base sólida para su implementación en flujos de trabajo clínicos reales.

Futuras investigaciones deberán centrarse en aumentar la diversidad de los datos, optimizar la interpretabilidad y evaluar la reproducibilidad multicéntrica de estos modelos, consolidando su papel como herramientas de soporte clínico confiables y transparentes.

REFERENCIAS

[1] World Health Organization, "Cardiovascular diseases (CVDs)," 2023. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds)

[2] Ministerio de Salud y Protección Social, "Enfermedades cardiovasculares." [Online]. Available: https://www.minsalud.gov.co/salud/Paginas/Enfermedades-cardiovasculares.aspx

[3] Ministerio de Salud y Protección Social, "Minsalud conmemora el Día Mundial del Corazón," Sep. 29, 2022. [Online]. Available: https://www.minsalud.gov.co/Paginas/Minsalud-conmemora-el-dia-mundial-del-Corazon.aspx

[4] Ministerio de Salud y Protección Social, "Plan Nacional de Salud Pública 2022–2031." [Online]. Available: https://www.minsalud.gov.co/plandecenal/Paginas/PDSP-2022-2031.aspx

[5] D. De Santis, A. Trivisonno, and E. G. Caiani, "Accelerated deep learning-based function assessment in cardiovascular magnetic resonance: a clinical comparison with conventional cine sequences," La Radiologia Medica, 2025, doi: https://doi.org/10.1007/s11547-025-02019-6

[6] A. Argentiero, G. Muscogiuri, M. G. Rabbat, C. Martini, N. Soldato, P. Basile, et al., "The applications of artificial intelligence in cardiovascular magnetic resonance—A comprehensive review," *Journal of Clinical Medicine*, vol. 11, no. 10, p. 2866, 2022, doi: https://doi.org/10.3390/jcm11102866

[7] O. Bernard et al., "Deep learning techniques for automatic MRI cardiac multi-structures segmentation and diagnosis: Is the problem solved?" *IEEE Trans. Med. Imaging*, vol. 37, no. 11, pp. 2514–2525, 2018. https://doi.org/10.1109/TMI.2018.2837502.

[8] M. Alsharqi and E. R. Edelman, "Artificial Intelligence in Cardiovascular Imaging and Interventional Cardiology: Emerging Trends and Clinical Implications," *J. Soc. Cardiovasc. Angiogr. Interv.*, vol. 4, no. 3, p. 102558, 2025. https://doi.org/10.1016/j.jscai.2024.102558

[9] Y.-R. Wang, K. Yang, Y. Wen, P. Wang, Y. Hu, Y. Lai, Y. Wang, K. Zhao, S. Tang, A. Zhang, H. Zhan, M. Lu, X. Chen, S. Yang, Z. Dong, Y. Wang, H. Liu, L. Zhao, L. Huang, Y. Li, L. Wu, Z. Chen, Y. Luo, D. Liu, et al., "Screening and diagnosis of cardiovascular disease using artificial intelligence-enabled cardiac magnetic resonance imaging," *Nature Medicine*, vol. 30, pp. 1471–1480, 2024, doi: https://doi.org/10.1038/s41591-024-02971-2

[10] R. Jafari, R. Azad, and S. Minaee, "Automatic diagnosis of myocarditis disease in cardiac MRI modality using deep transformers and explainable artificial intelligence," arXiv preprint arXiv:2210.14611, 2022. https://doi.org/10.48550/arXiv.2210.14611

[11] D. Li, H. Zhang, and C. Wu, "An algorithm for cardiac disease detection based on magnetic resonance imaging using an improved SA-YOLO model," *Scientific Reports*, vol. 15, 88567, 2025, doi: https://doi.org/10.1038/s41598-025-88567-3

[12] M. U. Rehman, S. Akbar, and A. Imran, "Predicting coronary heart disease with advanced machine learning classifiers for improved cardiovascular risk assessment," *Scientific Reports*, vol. 15, 96437, 2025, doi: https://doi.org/10.1038/s41598-025-96437-1

[13] T. Tsampras, T. Karamanidou, G. Papanastasiou, and T. G. Stavropoulos, "Deep learning for cardiac imaging: Focus on myocardial diseases: A narrative review," Hellenic Journal of Cardiology, 2024. https://doi.org/10.1016/j.hjc.2024.12.002

[14] F. Cheng et al., "Learning directional feature maps for cardiac MRI segmentation," in Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent. (MICCAI), Cham, Switzerland: Springer, Sep. 2020, pp. 108–117. https://doi.org/10.1007/978-3-030-59719-1_11

[15] D. M. Alsekait et al., "Heart-Net: A multi-modal deep learning approach for diagnosing cardiovascular diseases," *Comput. Mater. Contin.*, vol. 80, no. 3, 2024. https://doi.org/10.32604/cmc.2024.054591

[16] R. K. Sahu, D. R. Sahu, and B. K. Pattanayak, "A hybrid segmentation and classification CAD framework for myocardial infarction detection using cardiac MRI," *Scientific Reports*, vol. 15, 98893, 2025, doi: https://doi.org/10.1038/s41598-025-98893-1

[17] X. Liu, F. Xing, H. K. Gaggin, C. C. J. Kuo, G. El Fakhri, and J. Woo, "Successive subspace learning for cardiac disease classification with two-phase deformation fields from cine MRI," in Proc. 2023 IEEE 20th Int. Symp. Biomed. Imaging (ISBI), Apr. 2023, pp. 1–5. https://doi.org/10.1109/ISBI53787.2023.10230746

[18] J. Xing, S. Wang, K. C. Bilchick, A. R. Patel, and M. Zhang, "Joint deep learning for improved myocardial scar detection from cardiac MRI," in Proc. 2023 IEEE 20th Int. Symp. Biomed. Imaging (ISBI), Apr. 2023, pp. 1–5. https://doi.org/10.1109/ISBI53787.2023.10230541

[19] G. P. Veldhuizen, T. Lenz, D. Cifci, M. van Treeck, J. Clusmann, Y. Chen, C. V. Schneider, T. Luedde, P. W. de Leeuw, A. El-Armouche,

D. Truhn, and J. N. Kather, “Deep learning can predict cardiovascular events from liver imaging,” JHEP Reports, 2025, Art. no. 101427. https://doi.org/10.1016/j.jhepr.2025.101427

[20] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, “Grad-CAM: Visual explanations from deep networks via gradient-based localization,” *International Journal of Computer Vision*, vol. 128, pp. 336–359, 2019, https://doi.org/10.1007/s11263-019-01228-7

[21] J. Liu, A. Wei, L. Cao, X. He, and C. Tang, “Contrastive trustworthy prototype learning for multi-modality myocardial pathology segmentation,” *Applied Soft Computing*, vol. 173, Art. no. 112909, 2025, doi: https://doi.org/10.1016/j.asoc.2025.112909

[22] Z. Dong, Y. Tang, P. Sun, G. Yin, K. Zhao, X. Ma, and S. Zhao, “Early identification of myocardial microstructural alterations in hypertrophic cardiomyopathy with in vivo cardiac diffusion-tensor imaging,” *Radiology: Cardiothoracic Imaging*, vol. 7, no. 1, Art. no. e240009, 2025, doi: https://doi.org/10.1148/ryct.240009

FINANCIACIÓN

Este proyecto fue financiado por la Universidad de Córdoba, mediante recursos de convocatoria interna con el código de proyecto FI-01-24.

CONFLICTOS DE INTERÉS

No existen

CONTRIBUCIÓN DE AUTORÍA

Redacción - borrador original: Jorge Eliecer Gómez Gómez, Juan José Gúzman Pineda, Velssy Liliana Hernández Riaño.

Redacción - revisión y edición: Jorge Eliecer Gómez Gómez, Juan José Gúzman Pineda, Velssy Liliana Hernández Riaño.