ISSN: 1692-7257 - Volume 1 – Issue 47 - 2026

ORIGINAL

RCTA
Revista Colombiana de Tecnologías de Avanzada
UNIPAMPLONA

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

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

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.**

Received: 21 julio 2025. Revised: 21 septiembre 2025. Accepted: 20 diciembre 2025. Published: 01 enero 2026.
Corresponding Author: Jorge Eliecer Gómez Gómez

How to cite: 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

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

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.

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.

INTRODUCTION

Cardiovascular diseases remain the leading cause of death worldwide, accounting for approximately 32% of all global deaths each year, according to the World Health Organization (WHO)[1]. More than 17.9 million people are estimated to die annually from causes related to the heart and blood vessels, a figure that continues to rise because of population aging, changes in lifestyle, and the limited coverage of preventive strategies. The WHO has emphasized the need for early interventions, especially in low- and middle-income countries, where the burden of cardiovascular disease has shown an alarming upward trend.

In Colombia, this situation is no different. The Ministry of Health and Social Protection has reported that, in recent years, diseases of the circulatory system have consistently ranked first as a cause of death in the country, with ischemic heart disease, cerebrovascular accidents, and hypertensive diseases being the most prevalent [2]. In 2022, the cardiovascular disease mortality rate was estimated at close to 176 deaths per 100 000 inhabitants, with a greater burden in older adults and rural areas [3]. Despite advances in hospital infrastructure and coverage, structural challenges related to equitable access to timely diagnosis persist, especially in regions with limited specialized coverage. In this scenario, the adoption of technologies capable of facilitating the early and efficient detection of cardiovascular diseases, such as artificial intelligence-based solutions, is emerging as a public health priority [4].

Against this backdrop, the early and accurate detection of cardiac conditions, even before they manifest clinically, has become a key goal of modern medicine. Among the available diagnostic tools, cardiac magnetic resonance imaging (MRI) stands out for its ability to provide detailed images of the morphology and function of the heart without exposure to ionizing radiation [5], [6]. However, the traditional analysis of these images depends on visual interpretation by experts, a practice subject to interobserver variability, time-consuming, and not always accessible in all clinical settings [7].

In this context, artificial intelligence (AI) has emerged as a promising solution to automate diagnostic processes, reduce the clinical burden, and improve the consistency of analysis [8]. Multiple deep learning-based models have been developed that allow the segmentation of cardiac chambers, the quantification of volumes, or the classification of specific pathologies from medical images [9], [10]. However, in daily clinical practice, many initial decisions come down to a fundamental question: is there or is there not a cardiac condition that warrants attention?

Addressing this question as a binary detection of the presence or absence of disease is particularly relevant in scenarios such as population screening, emergency department triage or telemedicine [11], [12]. In such contexts, lightweight, interpretable and efficient models, capable of providing reliable answers from basic images, could be decisive for reducing diagnostic times and appropriately prioritizing available clinical resources [13].

Beyond technical performance, the reliability of these systems plays a fundamental role in their adoption. Several studies have shown that AI models can achieve levels of accuracy comparable to or higher than those of human experts, especially when trained with well-labeled and heterogeneous datasets. However, their implementation requires not only quantitative validation, but also mechanisms that allow each clinical decision to be explained and justified, in order to build trust among health professionals and patients alike.

The incorporation of explainability techniques has proven to be an effective strategy for improving the transparency of these systems, allowing identification of which regions of the image most influenced the prediction made[12]. This not only facilitates medical audits and second opinions, but also supports the ethical and legal integration of AI into structured health systems, particularly in processes such as automated screening, assisted diagnosis and remote second readings.

This article focuses on the development and validation of binary detection models for heart disease based on MRI and deep learning. Through the implementation of machine learning and neural networks, the aim is to offer an effective, reproducible and clinically viable mechanism to support early decision-making in cardiology, with special attention to high-demand or resource-limited settings.

STATE OF THE ART

This section critically analyzes the most relevant studies, grouped into five thematic lines that represent the main current research trends.

Automatic segmentation of cardiac structures

One of the most consolidated lines in the literature is the automatic segmentation of cardiac structures from MRI. In this field, the work of[7]marked a milestone by presenting the results of the Automatic Cardiac Diagnosis Challenge (ACDC), where various deep learning methods applied to the segmentation of the left ventricle (LV), right ventricle (RV) and the myocardium (MYO) were evaluated. This benchmark established a standard in the area, achieving average Dice coefficients of 0.95 for LV, 0.89 for RV and 0.89 for MYO, mainly using architectures based on U-Net. Despite the advances, the authors

noted limitations regarding the generalization of the models outside the ACDC dataset, which prompted subsequent research focused on robustness and transferability.

In response to these limitations,[14]proposed a convolutional network that integrates Directional Feature Maps (DFM), designed to capture oriented spatial patterns characteristic of cardiac structures, such as curved edges and boundaries between the myocardium and the ventricular cavities. This approach made it possible to overcome the limitations of isotropic architectures such as U-Net, achieving Dice coefficients of 0.955, 0.901 and 0.888 for LV, RV and MYO respectively, on the ACDC dataset. The added value of this architecture lies in its ability to address both interclass indistinction and intraclass variability, aspects that are critical for reliable clinical diagnosis.

For their part,[15]developed Heart-Net, a multimodal architecture that combines MRI images with physiological signals (ECG and blood pressure), integrating a CNN and an LSTM in a late-fusion approach. This approach not only improves accuracy by segmenting the cardiac structures, but also makes it possible to classify cardiac pathologies with an accuracy above 98%. The combination of morphological and functional information mimics medical diagnostic reasoning and has proven to be highly robust to anatomical variations and clinical noise.

Complementing these proposals,[16]introduced a hybrid computer-aided diagnosis (CAD) system that employs ResU-Net for the segmentation of the cardiac structures and a Vision Transformer (ViT) for the classification of infarcted regions. This approach was particularly effective in identifying damaged myocardial tissue, achieving an AUC of 0.98 and an overall diagnostic accuracy of 97.3%. The ability of transformers to capture long-range spatial relationships makes them ideal for detecting disseminated and subtle patterns, which is essential in the early diagnosis of myocardial infarction.

Automatic classification of cardiac diseases

The direct classification of cardiac pathologies from MRI represents a significant evolution with respect to traditional approaches that depend on multiple intermediate stages. In [11] they presented SA-YOLO, an adapted version of YOLOv8 with spatial attention modules that enables the detection of pathologies directly from the images, without the need for prior segmentations. This model reached an overall accuracy of 98.1% and an inference time of 23 ms per image, making it viable for real-time use in clinical settings with computational limitations.

Meanwhile,[17]introduced an innovative approach by using deformation fields between the diastolic and systolic phases to encode the contractile dynamics of the heart; using an architecture based on autoencoders and an SVM classifier, they achieved a mean accuracy of 94.5%. This type of functional input makes it possible to identify contractile dysfunctions characteristic of diseases such as dilated (DCM) or hypertrophic (HCM) cardiomyopathy, providing a more informative alternative than static images.

In another notable study,[18]applied Vision Transformers combined with explainability techniques such as Grad-CAM to detect myocarditis. Their model reached an F1-score of 0.9741 and an AUC of 0.9616, standing out both for its accuracy and for its interpretative capacity. This work represents an important advance in that it explicitly integrates visualization of the active regions that drive the classification, improving the transparency of the automated diagnostic process.

Finally,[12]adopted a complementary approach by integrating algorithms such as XGBoost and LightGBM on multi-source clinical variables, without relying exclusively on MRI. Although they do not segment images, they achieve an average AUC of 0.989, demonstrating that combining structured data with high-performance models can offer reliable predictions even in the absence of detailed imaging.

Explainability and interpretability of models

The acceptability of artificial intelligence models in clinical contexts depends not only on their accuracy but also on their capacity to be interpreted. In this sense,[19]proposed a joint learning architecture that simultaneously segments the myocardium and detects post-infarction scar tissue; the model incorporates an attention mechanism that highlights the relevant regions in the predictions, achieving a sensitivity of 91.8% and a diagnostic accuracy of 94.6%. The spatial coherence between the segmentation and functional detection tasks improves clinical confidence in the results.

Likewise, the work of[18]stands out for incorporating Grad-CAM as an explainability tool in a sensitive context such as the diagnosis of myocarditis. These techniques allow the model's decisions to be validated visually and quantitatively, showing which specific regions influenced the prediction, which strengthens transparency and medical confidence.

Clinical validation and adoption in hospital settings

One of the barriers to the adoption of AI models in medicine is the lack of clinical validation in real-world scenarios. In this regard,[5]conducted a prospective study in 61 patients comparing conventional cine sequences with those reconstructed by means of deep learning. Their model demonstrated excellent agreement with traditional functional measurements ($r > 0.98$) and reduced acquisition time by 50%, improving visual quality, especially in subendocardial structures.

On the other hand, in[20],[6]conducted systematic reviews of the use of AI in cardiovascular imaging, addressing topics such as integration into hospital workflows, regulation by entities such as the FDA, and the traceability of models. These reviews highlight that, beyond technical performance, the clinical acceptance of a model depends on its transparency, acceptable inference time and compatibility with HIS/PACS systems.

Preventive diagnosis and multimodal approaches

The application of AI in medicine is not limited to current diagnosis; it extends toward preventive prediction of cardiovascular risk. In[21], the authors demonstrated that deep networks can predict future cardiac events from liver images, with an AUC of 0.87, by combining abdominal MRI and clinical variables through a multimodal transformer. This perspective breaks with the traditional logic of structural diagnosis and allows latent correlations in other imaging modalities to be explored.

Along the same lines, in[22]the authors trained a ResNet-type CNN with attention modules to classify five cardiac pathologies directly from cine MRI without prior segmentation. Their model achieved an accuracy of 96.4% and an inference time of less than one second, which makes it a viable option for clinics with a high volume of studies and few specialized staff.

Finally, in[13]the authors synthesized advances in AI for myocardial diseases, highlighting the emerging role of generative models, recurrent networks and virtual simulations. By integrating multiple temporal and morphological dimensions, these approaches broaden the diagnostic spectrum and suggest new avenues for personalizing treatments and simulating clinical progressions.

MATERIALS AND METHODS

Description of the dataset

The dataset used corresponds to cardiac magnetic resonance imaging studies stored in NIfTI format, obtained from[7], with 150 patients, divided into 5 subgroups (30 normal subjects – NOR; 30 patients with prior myocardial infarction; 30 patients with dilated cardiomyopathy; 30 patients with hypertrophic cardiomyopathy; 30 patients with abnormal right ventricle). Each volume includes axial slices from different moments of the cardiac cycle (for example, frame01, frame12), accompanied by manually segmented masks (_gt.nii.gz). These masks delimit structures such as the myocardium, the left ventricle (LV) and the right ventricle (RV).

A per-patient min-max normalization was applied (based on the intensity values of each volume). The axial slices were interpolated to 128×128 or 256×256 pixels using bicubic interpolation, background noise was removed with an adaptive threshold, and the intensities were converted to 8 bits. In the case of YOLOv8, the images were duplicated to three RGB channels for compatibility. The same pipeline was applied uniformly to training and validation.

The images come from a clinical repository structured in subfolders by patient, with anatomical and temporal variability. This richness allows robust models to be trained for binary classification (presence or absence of disease), as well as experimentation with segmentation tasks. The dataset was mounted from Google Drive in a collaborative environment using Google Colab.

General Processing Workflow

Before implementing each model, a generic processing pipeline was defined to ensure data quality and consistency. This pipeline, shown in Fig. 1, starts with a cleaning and normalization step for the cardiac magnetic resonance imaging volumes, followed by a preprocessing module that includes cropping of the region of interest, selection of representative axial slices, rescaling to standard resolutions (128×128 or 256×256 pixels), and conversion to formats compatible with each architecture (grayscale or RGB). Feature selection and transformation techniques were then applied according to the approach used (flattening in Random Forest, convolutions in CNN, and direct RGB processing in YOLOv8). Finally, each model was trained and evaluated independently using the same performance metrics, so that the comparison would be objective. This structured approach ensures the traceability of the pipeline and the replicability of the experiments.

Data preparation

The dataset used consists of three-dimensional volumes in .nii.gz format. These were split 80/20 at the patient level, grouping all slices of each subject into the same set, with stratification by class (healthy/diseased). This prevented information leakage and guaranteed independence between training and validation. The volumes were preprocessed to obtain representative 2D axial slices of each study.

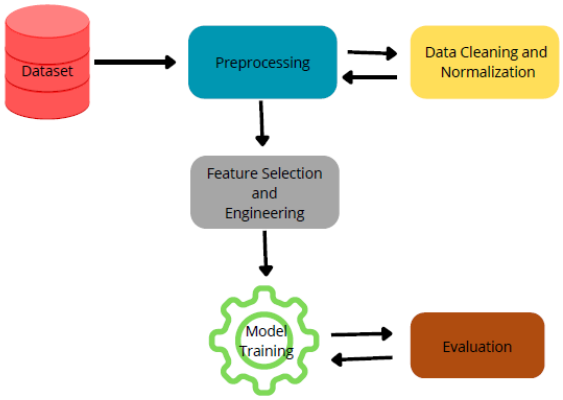

Fig. 1. *General processing pipeline*

The preprocessing stages included:

- 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, with the aim of reducing the variability derived from heterogeneous acquisition conditions.

- 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.

This transformation allowed each image to be represented as a point in a high-dimensional feature space, together with a binary label corresponding to the patient's clinical condition (healthy = 0, diseased = 1).

To improve the discriminative capacity, radiomic features based on GLCM texture (contrast, dissimilarity, homogeneity, energy, correlation) and first-order statistical features (mean, variance, kurtosis, skewness) were extracted. These descriptors were computed per slice using the scikit-image library and served as input to the Random Forest model

Data augmentation

Given the small size of the dataset and its slightly imbalanced distribution, a data augmentation strategy was applied using ImageDataGenerator. This technique increases the diversity of the training set through random transformations, without the need to collect new images. The transformations included random rotations (±10°), random zoom (±10%), horizontal and vertical shifts (±10%), and horizontal flipping. These operations preserve the relevant anatomical features and improve the generalization capacity of the model.

RESULTS

Random Forest model performance

The Random Forest model was trained and evaluated using flattened images from axial slices of cardiac magnetic resonance imaging. After training, its performance was evaluated on a previously separated validation dataset (20% of the total samples), yielding the following quantitative results:

Overall accuracy (Accuracy): 0.8000

Precision (Precision): 0.8000

Sensitivity (Recall): 1.0000

F1-score: 0.8889

Area under the ROC curve (AUC): 0.6424

Figure 2 presents the ROC curve for the model. The area under the curve (AUC = 0.6424) indicates a moderate discrimination capacity of the model to distinguish between healthy patients and diseased patients. Although sensitivity is high, the presence of false positives reduces the overall accuracy of the classifier.

The performance of the Random Forest model is characterized by a high capacity to detect positive cases (diseased), which makes it a potentially useful tool in contexts where missing critical cases is unacceptable (high sensitivity). However, the

number of false positives recorded suggests that it may generate unnecessary alerts in some healthy patients, which must be considered in real clinical applications. This behavior can also be attributed to the loss of spatial information during the image flattening process, which limits the model's ability to differentiate subtle anatomical patterns.

CNN model performance

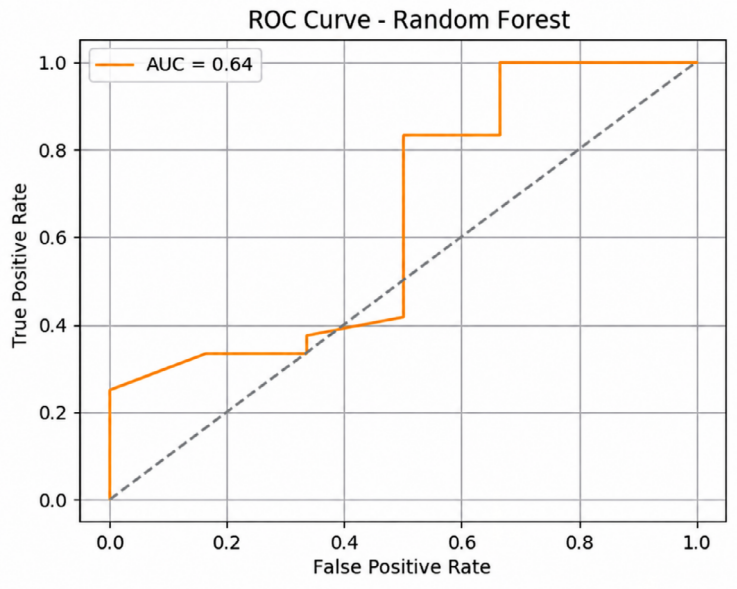

Fig. 2. *ROC curve of the Random Forest model.*

The model was trained and evaluated on a dataset composed of 30 samples. The results obtained by the convolutional network are the following:

Overall accuracy (Accuracy): 0.7000

Precision CHD class: 0.8000

Recall CHD class: 0.8333

F1-score CHD class: 0.8163

Fig. 3 shows the confusion matrix of the CNN model. It is observed that the neural network correctly identified 20 diseased cases (CHD) and 1 healthy case (No CHD). However, it made 5 false positives, that is, it incorrectly classified healthy patients as diseased, and 4 false negatives, corresponding to diseased patients who were classified as healthy.

This behavior reflects a slight bias toward the positive class, favoring the detection of pathological cases over the correct identification of healthy subjects. Such a tendency may be attributed to the imbalance in the dataset (a greater number of images of diseased patients) or to a limited generalization capacity of the model to learn patterns representative of the normal myocardium.

The ROC curve of the model shows an AUC of 0.4167, indicating a discriminative power below the chance threshold, which is not acceptable in the health field. Although the network achieved acceptable sensitivity, the low AUC and the inability to correctly classify negative cases limit its diagnostic usefulness.

Despite its high sensitivity for detecting CHD cases, the CNN model showed significant deficiencies in identifying healthy patients, which could result in a high number of false positives in a clinical setting. This behavior may be related to the imbalance of the dataset and to the possible need for additional adjustments in the architecture or in the regularization methods. In future versions of the model, we recommend applying techniques such as focal loss, class weighting or specific augmentation for the minority class.

Performance of the YOLOv8 model

The performance achieved by this model was as follows:

Overall accuracy (Accuracy): 0.8000

Precision (Precision): 0.8000

Sensitivity (Recall): 1.000

F1-score: 0.8900

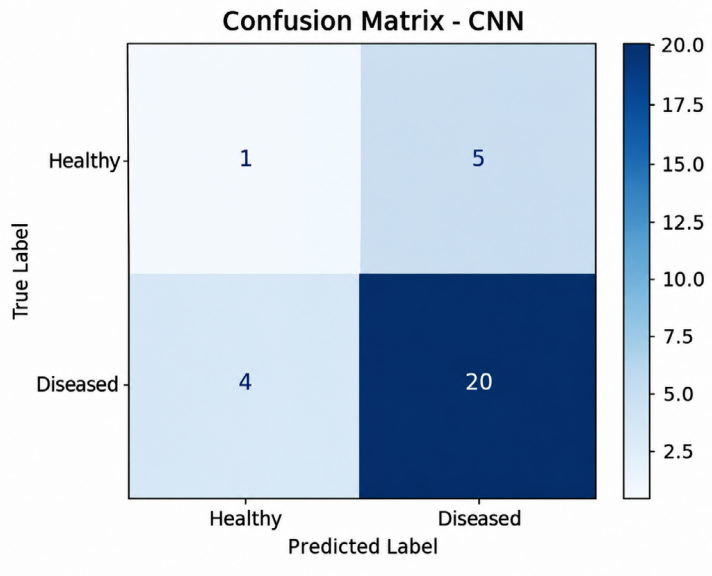

Fig. 3. *Confusion matrix of the CNN model.*

Fig. 4 shows the ROC curve obtained from the classification probabilities of the YOLOv8 model for the positive class (diseased). El area under the curve (AUC) was 0.83, which reflects a good discriminative capacity, although not outstanding. La curva presenta una forma ascendente y se aproxima parcialmente al vértice superior izquierdo, indicando que el modelo logra una high true positive rate with a moderate number of false positives.

This result confirms that YOLOv8 can effectively distinguish between healthy and diseased subjects, although its performance could be improved by more precise calibration of the decision threshold or a better balance between classes.

The model showed superior performance compared with the other approaches evaluated, both in terms of precision and in terms of balance between classes. Its ability to generalize correctly on a diverse validation dataset, together with its low error rate and high AUC score, positions it as a highly viable alternative for the automated detection of congenital heart disease from medical images. The stability of the model also suggests effective learning of discriminative patterns without overfitting, even in a clinical setting with moderate resources.

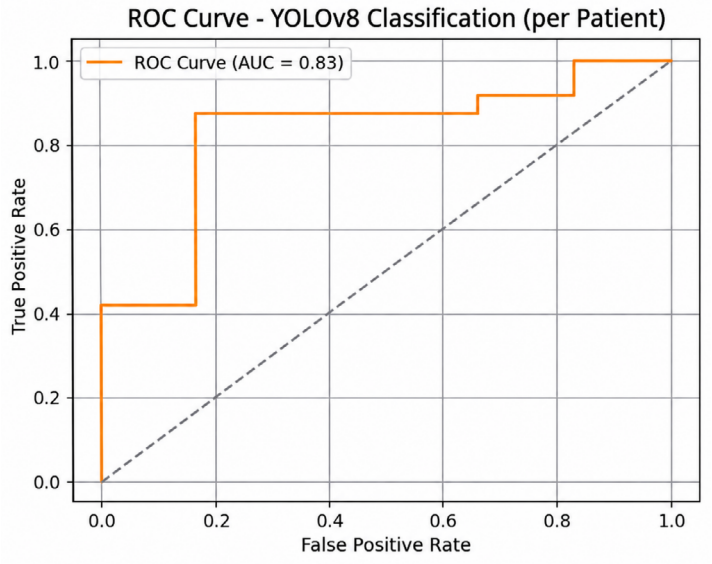

Fig. 4. *ROC curve of the YOLOv8 model.*

Random Forest

In order to understand how the model makes its decisions and to validate the physiological relevance of the features used, two complementary approaches to global and local explainability were implemented:

SHAP (SHapley Additive Explanations), based on cooperative game theory. Permutation Importance, which assesses the variation in model performance when each variable is randomly altered. Both methods were applied to the validation dataset, using the six radiomic descriptors obtained from the gray-level co-occurrence matrix (GLCM): Homogeneity, Energy, ASM (Angular Second Moment), Contrast, Dissimilarity and Correlation.

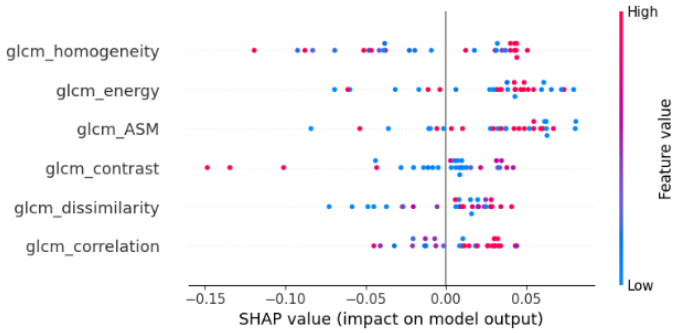

Fig. 5. *Impact of the model on the outputs*

The analysis of mean absolute SHAP values (see Fig. 5) made it possible to identify that the variables with the greatest impact on the prediction were:

- GLCM_Homogeneity (0.0449),
- GLCM_Energy (0.0442),
- GLCM_ASM (0.0436).

These three features are associated with the uniformity of the myocardial tissue. Patients with high homogeneity and energy values tended to be classified as healthy, whereas more heterogeneous textures (low homogeneity, high contrast) were linked to the diseased class. The Contrast, Dissimilarity and Correlation variables had a smaller contribution, although their combined effect helped to distinguish mild structural alterations. The individual SHAP representation showed a consistent trend: negative homogeneity values decrease the probability of being healthy, and positive values increase it, which confirms that the model internalized a clinically consistent pattern.

Regarding permutation importance, Fig. 6 shows the average importance and standard deviation obtained using the permutation method. The results partially agreed with SHAP: homogeneity remained first in relevance (0.0333), followed by Correlation (0.0122) and Contrast (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. The agreement between SHAP and Permutation suggests high internal consistency in the model; the dominance of uniformity descriptors implies that the model associates intensity dispersion with pathology, consistent with the structural changes typical of cardiomyopathy (hypertrophy, fibrosis or remodeling). Finally, since Random Forest does not operate directly on the image but on global feature vectors, its spatial interpretability is limited; however, it offers clear conceptual explainability based on statistical properties of the tissue.

4.4.2. CNN

The convolutional neural network (CNN) model was analyzed using three complementary visual explainability techniques:

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

These tools make it possible to identify the regions of the image that most influence the final prediction, providing a direct spatial interpretation of the model's attention.

The Grad-CAM technique (see Fig. 7) generated heatmaps indicating which regions of the axial slice most intensely activated the convolutional layers when a patient was classified as diseased.

In most positive cases, the red and yellow areas were concentrated in the left ventricle and the interventricular septum, areas commonly affected in hypertrophic and dilated heart disease. In healthy patients, activations were more uniform and dispersed, reflecting less attention of the model to specific regions.

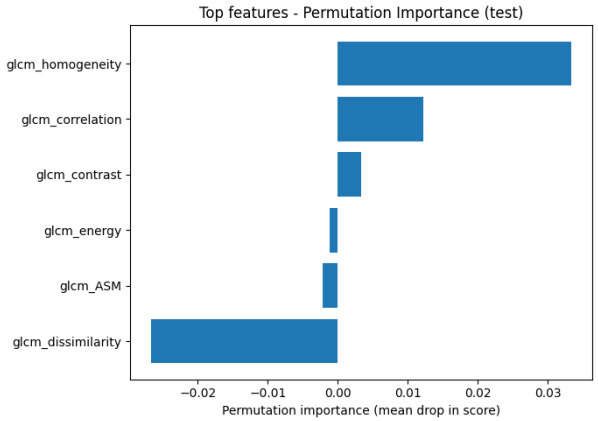

Fig. 6. *Permutation importance.*

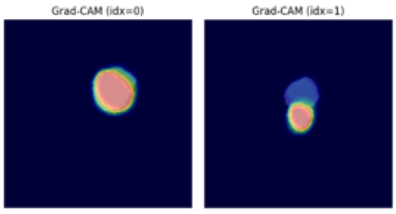

Fig. 7. *Grad-CAM heatmaps*

Fig. 8 shows that the gradients of the output with respect to each input pixel were calculated, revealing which areas cause the greatest changes in classification probability.

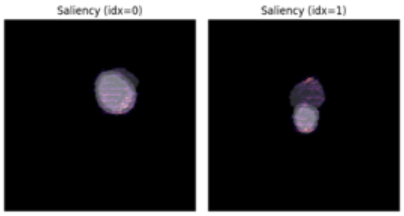

Fig. 8. *Saliency Maps*

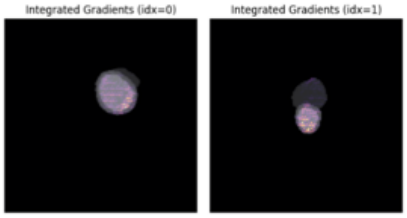

Fig. 9. *Integrated Gradients*

These visualizations highlighted heart contours and ventricular borders, indicating that the model based its decisions on the shape and boundaries of the myocardium. The observed pattern confirms that the CNN learns texture and geometry features, but it fails to fully discriminate between healthy and diseased tissue, as many borders appear prominent even in sections without pathology.

The integrated gradients technique, as shown in Fig. 9, offered a more stable perspective, reducing the noise characteristic of pure gradients.

The results showed moderate intensity concentrations over the mid-septal myocardium and the ventricular walls, which coincides with the regions that Grad-CAM identified as most relevant. This agreement between methods reinforces confidence that the CNN is indeed learning anatomically plausible patterns, even though its overall accuracy is limited.

4.4.3. YOLOv8

This model, originally designed for object detection, was adapted to the binary classification task for heart disease. Because it

is fully convolutional and includes feature pyramid and self-attention layers, its internal behavior is less intuitive than that of a traditional CNN. Therefore, two spatial explainability methods were applied: Grad-CAM and occlusion sensitivity, with the aim of analyzing the anatomical regions that contribute to the final prediction and validating the coherence of its activations. The Grad-CAM maps generated on the most representative axial slices of each patient (see Fig. 10) showed intense activations in the ventricular cavities and the interventricular septum in the cases classified as diseased.

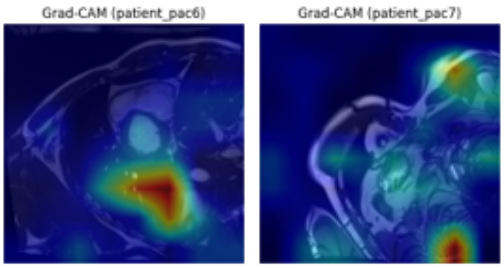

Fig. 10. *Grad-CAM maps (pac6 and pac7)*

In particular, patients pac6 and pac7 showed hot zones (red and yellow) concentrated in the left ventricle and the septal wall, consistent with the typical location of alterations in dilated or hypertrophic cardiomyopathies.

In the case of patient pac27 (see Fig. 11), the Grad-CAM map presented activations slightly displaced toward peripheral areas, reflecting a prediction with moderate confidence (probability of 0.61).

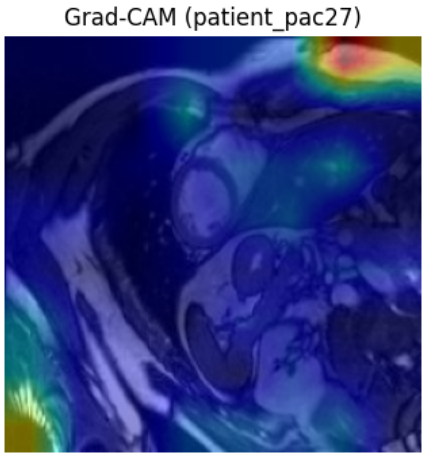

Fig. 11. *Grad-CAM maps (pac27)*

This suggests that, although the model recognizes pathological signals, its attention may become dispersed in cases with less pronounced cardiac textures or low contrast. In healthy patients, the Grad-CAM activations were diffusely distributed, without clear concentrations, which is consistent with the absence of anomalous structural patterns.

To complement the analysis, the occlusion sensitivity method was applied (see Fig. 12), which measures the decrease in probability when local patches of the image are masked.

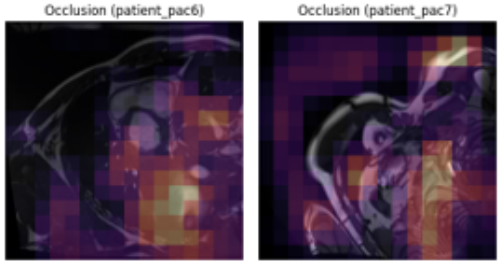

Fig. 12. *Occlusion Sensitivity (pac6, pac7)*

The results showed that the regions whose occlusion caused the greatest drop in probability (yellow and red zones) coincided with the same areas highlighted by Grad-CAM: ventricular walls and interventricular septum.

In diseased patients, occlusion of the left ventricle reduced the probability of the "diseased" class by more than 40%, confirming that these regions are critically determinant for the model's decision; by contrast, in healthy patients, occlusion did not produce significant changes in the output, which validates the specificity of the attention.

CONCLUSION

This study presented a systematic comparison among three representative artificial intelligence approaches applied to the detection of heart disease using cardiac magnetic resonance imaging: classical machine learning (Random Forest), convolutional neural networks (CNN) and optimized classification models (YOLOv8).

The methodology ranged from image preprocessing and feature extraction to performance evaluation with standardized metrics (accuracy, precision, recall, F1-score and AUC), complemented by an explainability (XAI) analysis that made it possible to interpret the decisions of each model.

The results showed that YOLOv8 achieved the best overall performance, with accuracy and precision of 80%, F1-score of 0.89 and an AUC of 0.83, reflecting a good discrimination capacity between healthy patients and diseased patients; its architecture made it possible to preserve relevant spatial information and detect complex structural patterns in the myocardium, demonstrating its potential for critical medical tasks. On the other hand, Random Forest also achieved competitive performance (accuracy = 0.80; recall = 1.00), relying on texture descriptors derived from GLCM matrices; whereas the CNN obtained lower performance (AUC = 0.4167), showing difficulty in correctly generalizing the negative class due to data imbalance.

From the perspective of explainable artificial intelligence, the three approaches evidenced physiological coherence in their decisions.

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 explainability not only provides transparency, but also a clinical validation tool complementary to quantitative performance.

In the clinical setting, the high sensitivity observed, especially in YOLOv8 and Random Forest, is encouraging for the development of decision support systems; however, the presence of false positives in the deep learning-based models highlights the need to recalibrate decision thresholds, improve class balance and incorporate cross-validation with specialists.

Finally, this work demonstrates that artificial intelligence is a viable and robust pathway for the assisted diagnosis of structural cardiopathies. Models such as YOLOv8 combine performance and explainability, offering a solid basis for their implementation in real clinical workflows.

Future research should focus on increasing data diversity, optimizing interpretability and evaluating the multicenter reproducibility of these models, consolidating their role as reliable and transparent clinical support tools.

REFERENCES

[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

FUNDING

This project was funded by Universidad de Córdoba, through internal call resources with project code FI-01-24.

CONFLICT OF INTEREST

None

AUTHORSHIP CONTRIBUTIONS

Drafting – original draft: Jorge Eliecer Gómez Gómez, Juan José Gúzman Pineda, Velssy Liliana Hernández Riaño.

Writing–review and editing: Jorge Eliecer Gómez Gómez, Juan José Gúzman Pineda, Velssy Liliana Hernández Riaño.