CardioRadNet: Cardiac Mass Diagnosis Through Integrated Segmentation and Radiomic Analysis

Meri Ferretti , Michele Pagliaccia , Andrea Baggiano ,
Luigi Lovato , Francesco Angeli , Matteo Armillotta ,
Luca Bergamaschi , Carmine Pizzi , Gianluca Pontone ,
Valentina D.A. Corino

PII: S0169-2607(26)00202-6
DOI: https://doi.org/10.1016/j.cmpb.2026.109448
Reference: COMM 109448

To appear in: *Computer Methods and Programs in Biomedicine*

Received date: 5 January 2026
Revised date: 29 April 2026
Accepted date: 14 May 2026

Please cite this article as: Meri Ferretti , Michele Pagliaccia , Andrea Baggiano , Luigi Lovato , Francesco Angeli , Matteo Armillotta , Luca Bergamaschi , Carmine Pizzi , Gianluca Pontone , Valentina D.A. Corino , CardioRadNet: Cardiac Mass Diagnosis Through Integrated Segmentation and Radiomic Analysis, *Computer Methods and Programs in Biomedicine* (2026), doi: https://doi.org/10.1016/j.cmpb.2026.109448

Highlights

- An end-to-end deep learning and radiomics pipeline enables contrast-free cardiac mass classification
- Cardiac masses are accurately segmented on T1-weighted CMR using minimal user input
- Radiomics-based models reliably distinguish benign from malignant cardiac masses
- Semiautomated analysis achieves performance comparable to manual segmentation
- Malignancy-prioritized classification supports clinically acceptable decision-making in cardio-oncology

CardioRadNet: Cardiac Mass Diagnosis Through Integrated Segmentation and Radiomic Analysis

Meri Ferretti[a,b], Michele Pagliaccia[c], Andrea Baggiano[d], Luigi Lovato[f], Francesco Angeli[g,h], Matteo Armillotta[g,h], Luca Bergamaschi[g,h], Carmine Pizzi[g,h], Gianluca Pontone[d,e], Valentina D. A. Corino[a,b]

[a]*Cardio Tech-Lab, Centro Cardiologico Monzino IRCCS, Via Carlo Parea 4, 20138, Milan, Italy*
[b]*Department of Electronics, Information and Bioengineering, Politecnico di Milano, Via Golgi 39, 20133, Milan, Italy*
[c]*Cardiology Department, San Giovanni Battista Hospital, Via Massimo Arcamone, Foligno, Italy*
[d]*Department of Perioperative Cardiology and Cardiovascular Imaging, Centro Cardiologico Monzino IRCCS, Milan, Italy*
[e]*Department of Biomedical, Surgical and Dental Sciences, University of Milan, Milan, Italy*
[f]*Pediatric and Adult CardioThoracic and Vascular, Oncohematologic and Emergency Radiology Unit, UOC di Radiologia, Radiologia Cardiovascolare, IRCCS Azienda Ospedaliero-Universitaria di Bologna, 40138 Bologna, Italy*
[g]*Department of Medical and Surgical Sciences, Alma Mater Studiorum, University of Bologna, 40138 Bologna, Italy*
[h]*Cardiovascular Division, Morgagni-Pierantoni University Hospital, Forlì, Italy*

Abstract

Background and objective: Cardiac masses (CMs), though rare, include a wide spectrum of benign and malignant lesions that require distinct therapeutic strategies. Prior studies typically address narrow classification tasks (e.g., thrombus vs tumor) and rely on manually delineated regions. Building on these gaps, we introduce CardioRadNet: the first integrated framework combining deep learning segmentation and radiomics-based classification on contrast-free T1-weighted cardiac MR, designed to differentiate benign from malignant CMs.

Methods: A total of 127 patients with pathologically confirmed CMs (62 malignant, 65 benign) were included. A segmentation network incorporating point-based guidance was developed for mass delineation. Radiomic features were extracted from both manually and semiautomatically segmented volumes, and two separate radiomics-based classification models were developed accordingly. Feature selection and classifier performance were optimized using a 5-fold cross-validation.

Results: The segmentation network achieved a Dice score of 0.78, with 88% of the radiomic features extracted from the semiautomatic ROIs showing good reproducibility (ICC > 0.6) when compared with those derived from manual ROIs. The two models achieved identical balanced accuracy (0.85) and the same number of misclassifications (both used 10 features). Notably, the semiautomated ROI model yielded one fewer false negative, thereby reducing missed malignant cases.

Conclusion: CardioRadNet offers a novel, accurate and contrast-free solution for comprehensive CM classification. Unlike prior studies, it includes the full spectrum of CMs and uses semiautomated segmentation for broader clinical applicability. Overall, this approach supports scalable integration into routine CMR workflows as a decision-support tool for early risk stratification and, in turn, improved patient management.

Keywords: Cardiac masses, Classification, Deep Learning, Radiomics, Segmentation Network.

1. Introduction

Although rare, cardiac masses (CMs) represent a significant concern in the field of cardio-oncology. They encompass a heterogeneous group of lesions, which may be either neoplastic or non-neoplastic in nature. Among neoplastic masses, primary tumors, that are originating from within the heart, represent only 0.3% to 0.7% of all cardiac neoplasms and can be either benign or malignant [1]. Approximately 75% of primary lesions are benign, while 25% are malignant [2], [3]. Secondary neoplastic lesions, meaning those originating outside of the heart, are always malignant and 30-40 times more common than primary lesions [2], [3], [4]. Eventually, non-neoplastic masses, often referred to as “pseudo tumors”, are benign and can include thrombi, vegetations, and normal anatomical variants [5].

The clinical presentation of CMs is highly variable, ranging from incidental findings on imaging performed for unrelated reasons to cases presenting with shock or aborted sudden cardiac death [6]. Many CMs remain asymptomatic, while others are initially detected due to symptoms that lead to further imaging. The clinical manifestations largely depend on the CM’s size and anatomical location [7]. Many CMs are discovered on transthoracic echocardiogram (TTE) which is generally the first imaging exam performed due to its portability and widespread accessibility. However, more advanced cardiac imaging techniques, such as cardiac magnetic resonance (CMR), cardiac computed tomography (CCT), and positron-emission tomography/computed tomography (PET/CT), are often needed to produce a differential diagnosis [4]. Indeed, TTE has several limitations in imaging cardiac tumors, including a limited field of view, reduced visualization in patients with unfavorable body habitus, and limited capability for tissue characterization [3]. CCT is valuable for evaluating calcified masses and for preoperative planning, as it provides detailed visualization of the tumor’s relationship to adjacent vascular and extracardiac structures [8]. However, it involves exposure to ionizing radiation and the use of contrast agents, and it lacks the temporal resolution and tissue characterization capabilities of CMR, which has the advantage of avoiding radiation altogether [9]. PET/CT evaluates the metabolic activity of masses, helping to distinguish benign from malignant lesions, and is widely used in oncology. Its role in cardiac tumor imaging may expand, though challenges include dietary preparation and additional radiation exposure, especially when performed solely for cardiac assessment rather than alongside routine oncologic scans [4]. Moreover, some benign cardiac tumors, such as myxomas or hemangiomas, can exhibit high metabolic activity on PET, potentially mimicking malignant lesions and complicating diagnostic interpretation [10], [11].

In this context, CMR is likely the most proper imaging modality for cardiac tumors, thanks to its excellent spatial and contrast resolution without radiation exposure, its wide field of view, and its capability for multiplanar imaging [3]. Recent evidence further supports its diagnostic value: a systematic review and meta-analysis reported pooled sensitivity and specificity of 93% and 94%, respectively, for differentiating benign from malignant cardiac tumors, with an AUC of 0.98 [12]. However, most of the included studies relied on qualitative, expert-based interpretation of CMR scans, and despite these promising results, agreement with histopathological diagnosis remains imperfect. Moreover, the ability to correctly identify specific tumor subtypes is still limited, highlighting the need for more advanced and standardized diagnostic approaches.

Early and accurate discrimination between benign and malignant CMs is crucial for determining the appropriate clinical management and significantly impacts patient prognosis. Indeed, malignant tumors often require prompt surgical intervention and/or oncologic treatment, while benign masses may be managed conservatively or with elective surgery [7]. Delayed or incorrect diagnosis can lead to rapid disease progression, especially in malignant cases, reducing the chances of effective treatment and long-term survival.

Currently, no objective and standardized imaging-based procedures exist to definitively diagnose CMs. Clinical interpretation of imaging data still depends largely on qualitative assessments and basic geometric measurements, leaving much of the rich quantitative information contained within the scans underutilized.

Radiomics offers a transformative approach by extracting a large number of quantitative features from medical images, enabling a more detailed and objective analysis of tissue characteristics [13]. This data-driven approach has already shown significant promise in oncologic imaging, particularly for tasks like tumor classification and phenotyping [14], [15]. Although still in its early stages in cardiac imaging, radiomics is attracting growing interest, supported by an expanding body of evidence demonstrating its feasibility, reproducibility, and diagnostic value [16], [17], [18].

Due to the extreme rarity of CMs, objective differentiation remains relatively underexplored in the scientific literature, with most existing studies limited by small sample sizes [19], [20], [21], need for contrast administration [20], [21], [22], [23], [24], [25] or reliance on visually interpreted radiological features [21], [22], [23], [24]. Additionally, prior studies often adopt a narrow scope, focusing on partial classification problems (e.g., tumor vs thrombus or benign vs malignant neoplasms), without encompassing the entire spectrum of CMs [19], [20], [24], [25], [26].

To overcome these limitations, the aim of this study is to develop a robust, non-invasive and reproducible diagnostic tool that can assist clinicians in the accurate classification of all types of CMs based on imaging data alone. Specifically, we built CardioRadNet, a comprehensive end-to-end pipeline that integrates a deep learning-based segmentation network for CMs delineation with a radiomics-driven framework for their classification into benign or malignant (see Figure 1). Importantly, this approach relies exclusively on contrast-free T1-weighted CMR, making it suitable even for patients with contraindications to gadolinium-based contrast agents. To the best of our knowledge, no prior study has implemented such an end-to-end framework for CM classification, as existing radiomics-based approaches rely on manual segmentation [19], [25], [26], thereby limiting scalability and reproducibility.

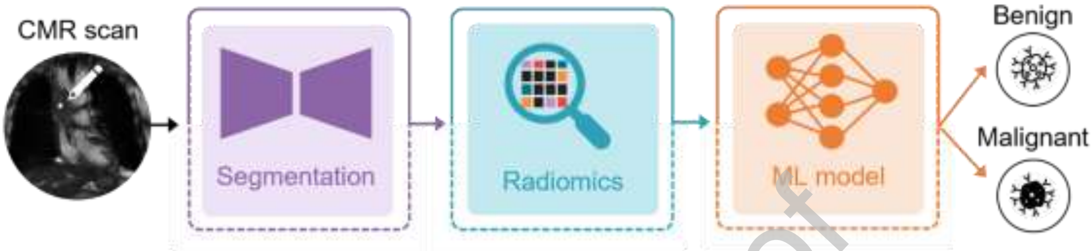


Figure 1: Schematic representation of CardioRadNet. CMR: Cardiac Magnetic Resonance, ML: Machine Learning.

2. Materials and methods

2.1. Dataset

The dataset comprises 127 patients from University Hospital Policlinico Sant'Orsola Malpighi (Bologna, Italy) who underwent CMR for suspected CMs. Definitive histological diagnoses were obtained via biopsy or surgical specimens. For cases involving cardiac thrombi, diagnoses were confirmed either by surgical findings or radiological resolution after appropriate anticoagulation therapy. Following the diagnostic work-up, CMs were classified as benign or malignant according to the World Health Organization 2021 Classification of Tumors of the Heart and Pericardium [27]. Specifically, 62 patients had malignant lesions (21 primary malignant tumors and 41 metastatic tumors), while the remaining 65 had benign lesions (38 primary benign tumors and 27 cases of pseudo tumors). Detailed information on tumor subtypes and their respective frequencies is provided in Table S1 of the Supplementary material. The study was conducted according to the guidelines of the Declaration of Helsinki. The study protocol was approved by the local Ethics Committee (Registration No. 102/2017/Oss/AOUBo) and all the patients signed the informed consent.

In this study only breath-hold steady-state free precession pulse T1-weighted cine sequences (CINE) in axial, long-axis and short-axis orientation have been considered. This choice was motivated by the fact that they are both contrast-free and commonly acquired in standard practice when performing a CMR scan. The examinations were performed on a 1.5 T CMR scanner (Ingenia; Philips Healthcare, Best, the Netherlands). Table S2 of the Supplementary material describes the main acquisition parameters of the employed scanner.

The ground-truth mass segmentations were manually performed by an expert radiologist holding Level III certification from the European Association of Cardiovascular Imaging via the “3D Slicer” software ending up with one volumetric Region Of Interest (ROI) per patient [28]. Subsequently, volumes were preprocessed to reduce noise using a 3×3×3 Gaussian filter and to correct for bias field inhomogeneities with an N4ITK filter. Finally, volumes were standardized via z-score normalization and resampled to 1 mm along all the three scanning axes. This ensured consistent intensity scales and resolution across patients, as recommended by the Image Biomarker Standardisation Initiative (IBSI) for radiomic analysis [29].

2.2. Segmentation network

After preprocessing, the segmentation network was developed. Due to the unfeasibility to train from scratch a deep architecture with this relatively limited dataset (535 slices were derived from the 127 CINE volumes), we opted for the fine-tuning of a segmentation network pre-trained on chondrosarcomas [30]. The architecture is displayed in Figure 2.

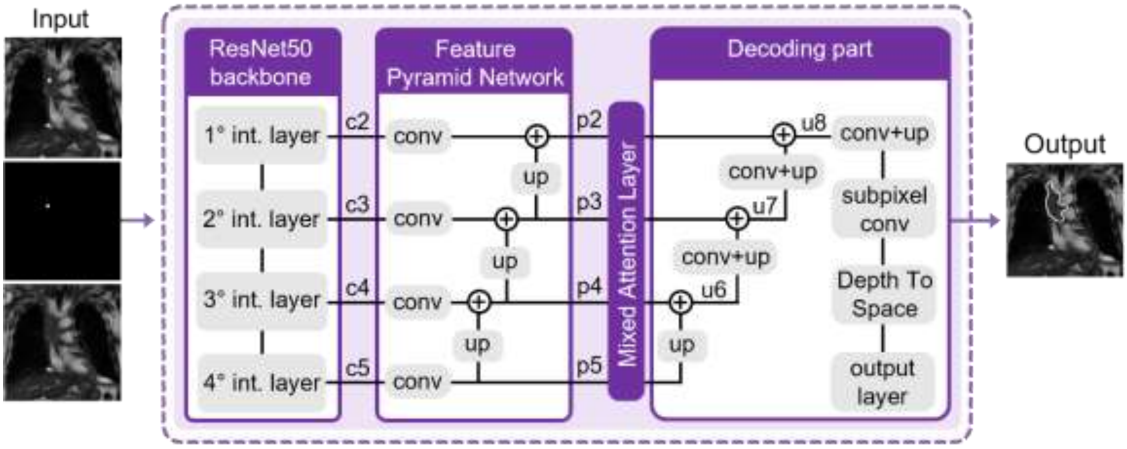


Figure 2: Schematic representation of segmentation architecture. int: intermediate, conv: convolution, up: upsampling.

In particular, the network is composed by a ResNet50 backbone (pretrained on ImageNet) from which output feature maps at four intermediate levels are taken: c2 (64 × 64 × 256), c3 (32 × 32 × 512), c4 (16 × 16 × 1024), and c5 (8 × 8 × 2048). To enable effective cross-scale feature fusion, a Feature Pyramid Network processes these feature maps reducing their channel depth and thus outputting: p2 (64 × 64 × 256), p3 (32 × 32 × 256), p4 (16 × 16 × 256), and p5 (8 × 8 × 256). Each of them is then passed through a mixed attention layer**,** which selectively emphasizes the most diagnostically relevant spatial locations and feature channels at each scale. After that, in the decoder portion, upsampling is carried out: p5 is upsampled and merged with p4 to yield u6 (16 × 16 × 512); similarly, the output is upsampled and merged with p3 to produce u7 (32 × 32 × 512). Eventually, u7 is also upsampled and merged with p2 to give u8 (64 × 64 × 512).

Before being input to the model, all slices were resized to a fixed spatial resolution of 256 × 256 pixels to ensure consistency with the chondrosarcoma images used during the initial training phase.

To provide spatial guidance to the network, a punctual annotation was introduced to indicate the approximate location of the CM within each slice. Specifically, the lesion location was encoded as a single-pixel bright dot placed by the operator at any point inside the mass, represented as a binary image (white dot on a black background). This minimal spatial prior informs the network of the lesion position without imposing any assumptions regarding its size or shape.

This information was incorporated as a dedicated input channel, allowing the model to disentangle anatomical appearance from spatial localization, an aspect particularly relevant given the high variability in CM size, morphology, and position across patients.

Finally, three-channel images were constructed as follows: (i) the CMR scan combined with the punctual input, (ii) the punctual input alone, and (iii) the CMR scan alone (see Figure 3).

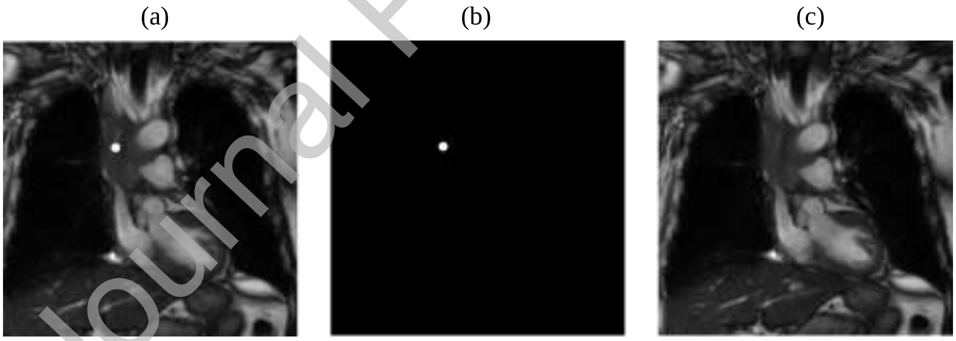


Figure 3: Example of an image used as input to the segmentation network. The first channel (a) displays the punctual input overlaid on the CMR scan, the second channel (b) shows the punctual input alone, and the third channel (c) contains the CMR scan alone.

Prior to the training phase, the dataset was partitioned into a training set (80%) and a test set (20%). The training set was then further divided into learning and validation subsets, with 80% and 20% of the training data, respectively. The split was stratified and enforced at the patient level to ensure an unbiased evaluation of model performance. Learning images were also augmented by randomly applying one of the following transformations: flip, rotation, translation or zoom. As a result, for each original learning image, an augmented version was generated, effectively doubling the size of the learning set. Batches of 32 images were given as input to the segmentation network, trained with the Adam optimizer and a custom loss defined as the mean of the binary cross-entropy (BCE), Dice, and intersection over unit (IoU) losses. This combination leverages the pixel-wise precision of BCE together with the region-level overlap awareness of Dice and IoU, leading to more robust and accurate segmentations. Equation (1) shows the custom loss formula while (2), (3) and (4) show its three components, namely BCE loss, Dice loss and IoU loss:

$$(1)\ L_{custom} = \frac{1}{3}(L_{BCE} + L_{Dice} + L_{IoU})$$

$$(2)\ L_{BCE} = -\frac{1}{N}\sum_{i=1}^{N}(y_i\, log(p_i) + (1 - y_i)log(1 - p_i)$$

$$(3)\ L_{Dice} = 1 - \frac{2\sum_{i=1}^{N} y_i p_i + \varepsilon}{\sum_{i=1}^{N} y_i + \sum_{i=1}^{N} p_i + \varepsilon}$$

$$(4)\ L_{IoU} = 1 - \frac{\sum_{i=1}^{N} y_i p_i + \varepsilon}{\sum_{i=1}^{N} y_i + \sum_{i=1}^{N} p_i - \sum_{i=1}^{N} y_i p_i + \varepsilon}$$

A post-processing step was incorporated to refine the predicted segmentations. Specifically, only the predicted connected components overlapping with the punctual mask were retained.

2.3. Radiomics

A radiomic analysis was performed on the ROIs derived from both the ground-truth and the semiautomated segmentations. A schematic representation of the followed workflow is depicted in Figure 4.

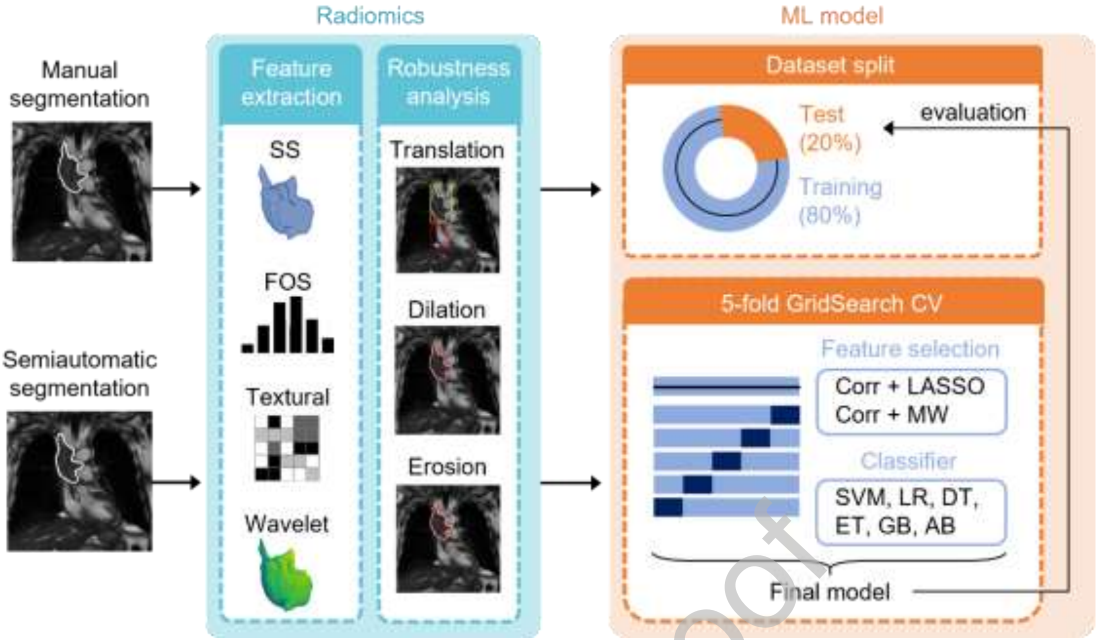


Figure 4: Workflow implemented in parallel for manually segmented and semi-automatically derived ROIs, including radiomic feature extraction, robustness analysis, and model development using a 5-fold cross-validation procedure. SS: shape and size, FOS: first order statistics, CV: cross-validation, LASSO: Least Absolute and Shrinkage Operator, MW: Mann-Whitney U test, SVM: Support Vector Machine, LR: Logistic Regression, DT: Decision Tree, ET: ExtraTrees, GB: Gradient Boosting, AB: AdaBoost.

3D radiomic features were extracted using the publicly available PyRadiomics library [31]. A bin width of 0.1 and b-spline interpolator were chosen as parameters for the extractor. In particular, a total of 851 features per volume were extracted, including 14 related to shape and size, 18 first-order statistical features, and 75 textural features. The first-order and textural features were computed both on the original ROIs and on eight additional ROIs obtained by applying a Discrete Wavelet Transform. Then, the radiomic features underwent a robustness analysis [32], whereby robustness was defined as the combination of stability and discriminative capability. This was evaluated by extracting features from perturbated versions of the ROIs as well [33]. Examples of the perturbations applied to the ROIs are shown in Figure 5. Specifically, stability was assessed by evaluating the Intra-class Correlation Coefficient (ICC) of each feature separately when the ROIs were dilatated or eroded by a small amount corresponding to the 15% of the ROI area. Only features that did not vary significantly, i.e., their ICC was bigger than 0.75, were deemed stable [17], [34]. Discriminative capability was assessed by performing a substantial perturbation of the ROIs translating them along the x and y axis of a quantity corresponding to the 80% of the dimension of the mass along the considered axis. Features were considered discriminative if they varied significantly after the perturbation, i.e., if the ICC was smaller than 0.5 [17], [34].

Overlapping the two groups (stable and discriminative features), robust features were derived.

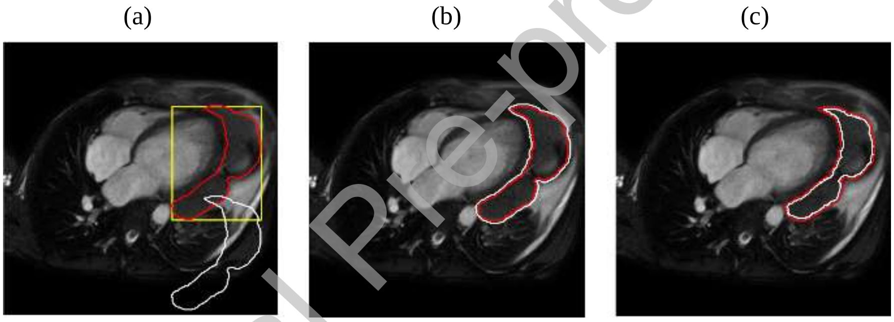


Figure 5: Example of perturbations (in white) applied to the original ROIs (in red) such as (a) translation of -80% of the length of the bounding box (in yellow) along the y axis, (b) dilation of 15% and (c) erosion of 15% of the CM area.

Robustness analysis was performed only on first order and textural classes, while shape and size features were automatically reinserted among the robust ones. To evaluate the similarity between manual and semiautomated segmentations, the ICC was also computed between the robust features derived from the manually segmented ROIs and the same features derived from the semiautomatically segmented ROIs. A high number of stable features in this case would indicate good similarity between manual and semiautomated segmentations.

2.4. Classification model

The same classification pipeline was implemented for features derived from both the ground-truth segmented ROIs and the semiautomatically segmented ROIs.

Specifically, the dataset was split into training (80%) and test set (20%) preserving the class stratification in both sets. The same split used for the segmentation network was adopted here to maintain consistency and prevent data leakage between training and test sets. Then, a 5-fold cross-validation (CV) with a GridSearch was applied on the training set to evaluate the best combination of feature selection process and classifier [35]. Specifically, two feature selection pipelines were tested, both starting with Spearman correlation analysis (CORR) as initial step. The first pipeline then applied the Least Absolute Shrinkage and Selection Operator (LASSO) to shrink some feature coefficients to zero, effectively reducing the feature set in input to the classifier. Alternatively, the second pipeline used the Mann-Whitney U test (MW), a univariate non-parametric statistical test, to assess each feature's ability to discriminate between the two classes, i.e., benign and malignant. Features with statistically significant test results were retained.

Within the GridSearch some hyperparameters were optimized, i.e., the optimal correlation threshold and α parameter for the LASSO selector, along with the best hyperparameters for the tested classifiers: Support Vector Machine (SVM), Decision Tree, ExtraTrees, AdaBoost, Gradient Boosting, and Logistic Regression (LR).

The best hyperparameter combination was determined by the highest mean balanced accuracy across the five validation folds. This process resulted in 12 distinct models, combining the two feature selection pipelines with each of the six classifiers. After selecting the best model, this was retrained on the entire training set and then evaluated on the test set.

3. Results

3.1. Segmentation

As regards the segmentation network development, the best model was obtained at epoch 43 after early stopping was called. The comprehensive quantitative evaluation of this final model yielded the performance metrics reported in Table 1 for individual image slice-level segmentation and Table 2 for complete CM-level segmentation.

The distribution of Dice and IoU scores achieved on all CMs included in the test set is presented in Figure 6. It has to be noted that the two CMs showing metrics below 0.5 are characterized by particularly small volumes (1.9 cm^3 and 0.4 cm^3) compared to the median of the population that is 14.7 cm^3 [5.2 - 40.7] thus increasing the complexity of the segmentation task.

Metric	**Learning set**	**Validation set**	**Test set**
Loss	0.08 ± 0.06	0.20 ± 0.13	0.20 ± 0.11
Accuracy	0.99 ± 0.01	0.99 ± 0.01	0.99 ± 0.01
Dice	0.91 ± 0.08	0.76 ± 0.19	0.75 ± 0.15
IoU	0.88 ± 0.10	0.65 ± 0.20	0.64 ± 0.17

Table 1: Metrics’ values achieved by the segmentation network on learning, validation and test sets per slice.

Metric	**Learning set**	**Validation set**	**Test set**
Loss	0.07 ± 0.06	0.17 ± 0.10	0.19 ± 0.13
Accuracy	0.99 ± 0.01	0.99 ± 0.01	0.99 ± 0.01
Dice	0.93 ± 0.07	0.80 ± 0.14	0.78 ± 0.17
IoU	0.88 ± 0.10	0.70 ± 0.17	0.67 ± 0.20

Table 2: Metrics’ values achieved by the segmentation network on learning, validation and test sets per mass.

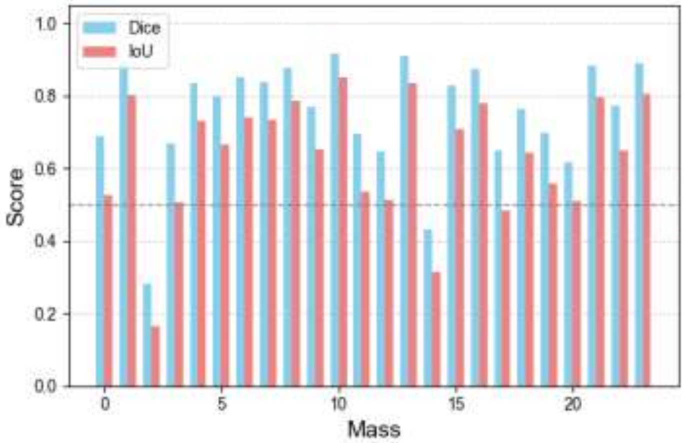


Figure 6: Dice scores and IoU scores achieved on cardiac masses included in the test set. IoU: Intersection over Unit.

Representative examples of CM segmentations are visually illustrated in Figure 7, showcasing the model’s performance across diverse CM morphologies, anatomical presentations, and varying imaging conditions.

A post-processing check was applied on all segmentations to retain only the predicted connected components overlapping with the punctual mask. Figure 8 displays an example in which the correction was employed. In this case, the correction made the Dice score improve from 0.83 to 0.86 and the IoU score from 0.71 to 0.76.

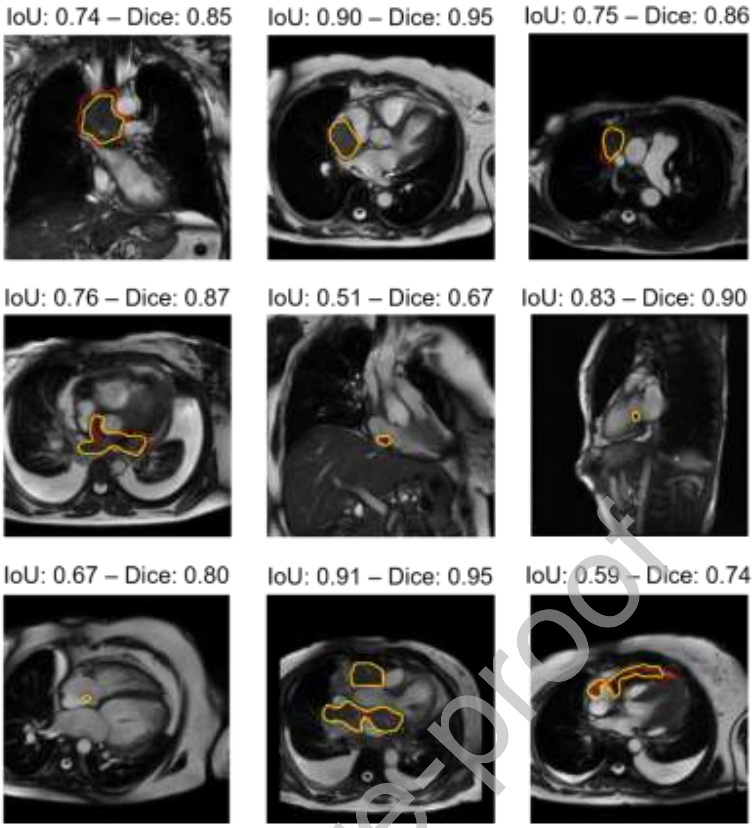


Figure 7: Comparison of semiautomated segmentations (in yellow) with the manual segmentations performed by the clinician (in red).

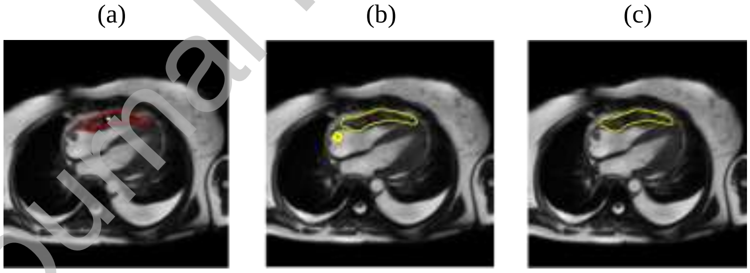


Figure 8: Example of (a) manual segmentation, (b) semiautomated segmentation and (c) corrected semiautomated segmentation for a slice in which the network produced an additional component not overlapping with any punctual input provided (in white).

3.2. Radiomics

After CMs' segmentation, radiomic analysis was performed on both manually- and semiautomatically-derived ROIs.

Robustness analysis retained 240 features for the manually segmented ROIs and 219 features for the semiautomatically segmented ones. As a measure of similarity between the two versions of the segmented CMs, i.e., manual and semiautomated, the ICC between manual ROIs' robust features and the same features derived from the semiautomated ROIs was computed. Almost all the original features (95%) and most of the wavelet ones (72%) resulted to be stable thus similar (ICC>0.75).

As depicted in Figure 9, the mean and median ICC values are 0.84 and 0.88, respectively, with 88% of features exhibiting good reproducibility (ICC > 0.6). However, few features display notably low ICC values. Those correspond to the Total Energy radiomic features extracted from the wavelet-transformed ROIs which quantify the overall magnitude of voxel intensities within a ROI by summing the squared voxel values and scaling by the volume of the ROI. As such, it is strongly affected by the number of voxels included within the segmentation, making it highly sensitive to even slight variations in ROI delineation.

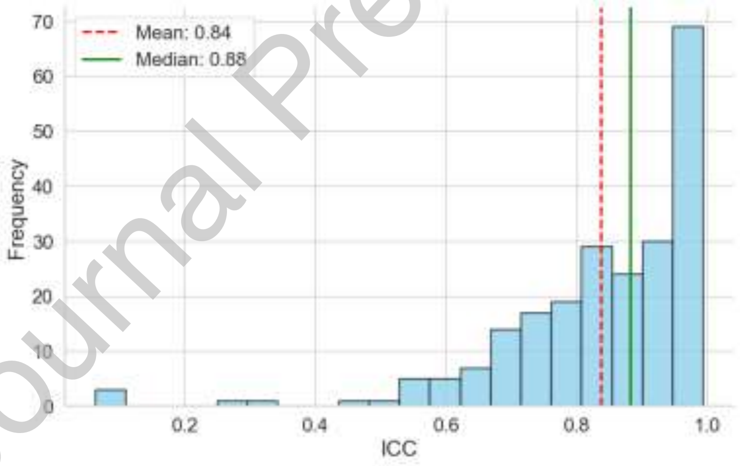


Figure 9: Frequency distribution of Intra-class Correlation Coefficient (ICC) indicating the features' similarity between the two versions of the segmented cardiac masses, i.e., manual and semiautomated. Mean (dashed red line) and median (green line) are indicated in the plot.

Furthermore, because this feature is derived from the wavelet transformation, its sensitivity is amplified: wavelets emphasize high-frequency components, i.e., local intensity variations that typically occur along the boundaries of structures. Since discrepancies between manual and semiautomated segmentations predominantly arise at the edges, these differences are further accentuated by the wavelet transform, resulting in substantial variability in the Total Energy feature.

Overall radiomic methodological quality was assessed using the METhodological RadiomICs Score (METRICS) [36], a standardized tool designed to evaluate and improve the methodological quality of radiomics studies. The present work showed good compliance with the framework (70%).

3.3. Classification

Features retained from the robustness analysis were used for subsequent classification of CMs as benign or malignant.

Specifically, a 5-fold CV with GridSearch was employed to identify the optimal hyperparameters for the correlation analysis, LASSO feature selection and the classifiers. The results of this validation, showing performance across all classifiers and feature selection methods for both the semiautomatic and manual segmentation models, are summarized in Figure S3 in the Supplementary material.

For manually segmented CMs, the best feature selection method was CORR+LASSO, and the best classifier was LR, whose combination attained the highest mean balanced validation accuracy (0.85 ± 0.04), leading to the selection of the CORR+LASSO+LR combination as the final model. After retraining on the full training set, this model achieved an AUC of 0.95, a balanced accuracy of 0.85, F1-score of 0.83, sensitivity of 0.77, specificity of 0.92 on the test set. The confusion matrix referred to the test set is reported in Fig. 10(a). The final model was based on a subset of 10 selected features.

Similarly, for semiautomatically segmented CMs, the best feature selection method was CORR+LASSO while the top-performing classifier was AdaBoost, achieving a mean balanced accuracy of 0.84 ± 0.03. Therefore, CORR+LASSO+AdaBoost was selected as the final model for semiautomatically segmented CMs. After retraining on the full training set, the model achieved an AUC of 0.82 and balanced accuracy, F1-score, sensitivity, and specificity all equal to 0.85 on the test set. This model also relied on 10 selected features. A binomial McNemar test has been applied to compare the predictions from the two models which resulted in a p-value of 0.48 indicating no statistically different results.

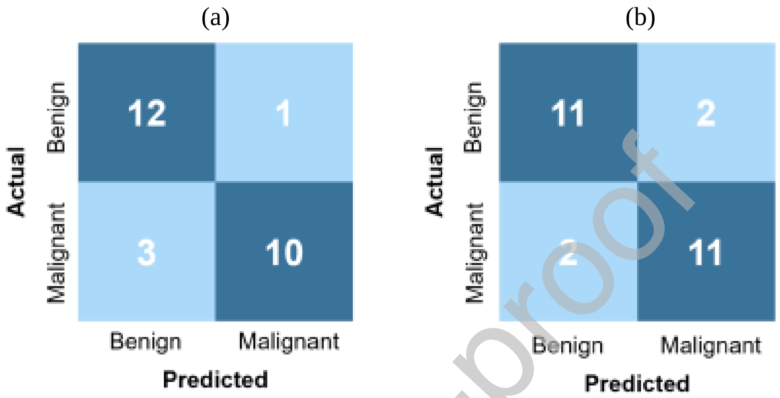


Figure 10: Confusion matrices for the best models chosen to classify (a) manually and (b) semiautomatically segmented CMs evaluated on the test set.

All test-set performance metrics for the two models, including AUC, balanced accuracy, F1 score, sensitivity, specificity and AUC are reported in Table 3 with 95% confidence intervals obtained via bootstrap resampling.

	Manual	**Semiautomated**
AUC	0.95 (CI: 0.91-0.98)	0.82 (CI: 0.75-0.88)
Accuracy	0.85 (CI: 0.80-0.89)	0.85 (CI: 0.80-0.89)
F1 score	0.83 (CI: 0.76-0.89)	0.85 (CI: 0.78-0.90)
Sensitivity	0.77 (CI: 0.67-0.85)	0.85 (CI: 0.77-0.92)
Specificity	0.92 (CI: 0.85-1.00)	0.85 (CI: 0.77-0.92)

Table 3: Test-set performance metrics for the models based on manual and semiautomated ROIs, with 95% confidence intervals obtained via bootstrap resampling. AUC: Area Under the Curve, Acc: balanced accuracy, F1: F1 score, Sens: sensitivity, Spec: specificity.

4. Discussion

Cardiac masses (CMs), though rare, pose a diagnostic challenge due to their variable imaging features and the critical need to distinguish benign from malignant lesions, two entities requiring fundamentally different clinical approaches. Early and precise classification is vital, as malignant tumors often demand urgent surgical or oncologic intervention, while benign masses may be managed conservatively or with elective procedures. Misdiagnosis or delay can lead to rapid disease progression and negatively impact patient outcomes. Despite the clinical importance, no standardized or objective imaging-based diagnostic protocols currently exist. Interpretation still relies heavily on qualitative assessments and basic morphologic features, leaving much of the quantitative information embedded in the scans unused [12]. This issue is compounded by the rarity of CMs, which limits the availability of large, validated datasets and hampers the development of robust tools. Existing studies are often constrained by small sample sizes, dependence on contrast administration, or subjective radiological evaluations.

This work proposes CardioRadNet, a semiautomated end-to-end pipeline that integrates deep learning-based mass segmentation with radiomics-driven classification to enable robust and reproducible differentiation between benign and malignant CMs. By eliminating the need for contrast agents, CardioRadNet is especially suited for patients with contraindications to gadolinium. Ultimately, it aims to support more accurate and timely decision-making, improving patient care in a domain where diagnostic clarity is essential.

The main findings of the present study are: i) a deep learning semiautomated segmentation network able to correctly delineate the CMs; ii) a radiomic model able to differentiate malignant from benign CMs, including both neoplastic and non-neoplastic lesions; iii) comparable performance achieved using semiautomated and manually performed segmentations.

While few studies have investigated the use of radiomic and clinical imaging features for characterizing CMs, they have addressed only a subset of the diagnostic challenge, typically focusing either on distinguishing tumors from thrombi [19], [25] or on differentiating benign from malignant neoplasms [26]. This limited scope inherently excludes at least one of the clinically relevant CM categories. Furthermore, all of these studies relied on manual segmentation of the ROIs, which is both time-consuming and prone to inter-observer variability. For instance, Lee et al. [25] developed three classification models, radiomic, clinical, and a combined approach, using contrast-enhanced CT scans to differentiate tumors from thrombi from 192 patients. Their combined model achieved a test set accuracy of 89.4%, with the radiomic-only model reaching 85.3%. However, their study did not distinguish between benign and malignant tumors, focusing instead on the broader separation of neoplastic and thrombotic lesions, and required contrast administration and ionizing radiation exposure. Similarly, Son et al. [19] tackled the same classification task using contrast-free cardiac MRI reaching an accuracy of 0.95. While this aligns more closely with contrast medium concerns, their dataset comprised only 41 patients, and the absence of a training-test split limits the robustness and generalizability of their findings. Bao et al. [26], on the other hand, focused exclusively on differentiating benign from malignant cardiac tumors using TTE. Despite employing a relatively large cohort of 215 patients, the dataset was heavily skewed toward malignant cases (173 out of 215), and no balancing techniques were applied. As a result, the reported test accuracy of 85%, with a sensitivity of 96% but a specificity of only 41%, likely reflects a model biased toward over-predicting malignancy due to the imbalanced data.

Other studies have leveraged clinical and radiological variables to perform either statistical analyses or to develop diagnostic models (not radiomics-based). For example, Kassi et al. [20] proposed a tree-based model using qualitative radiological features extracted from multiple CMR sequences, including contrast-enhanced imaging. However, their dataset was limited (66 patients), and non-neoplastic lesions such as thrombi were excluded. Pazos-Lopez et al. [24] similarly analyzed various CMR-derived features for tumor characterization, but their study also relied heavily on contrast-enhanced imaging and featured a highly unbalanced dataset (only 17 malignant tumors among 126 cases), undermining its diagnostic generalizability.

More recent efforts, including those by Yue et al. [23], Paolisso et al. [22], and Hu et al. [21], aimed to address the core clinical task of differentiating benign from malignant CMs. The first two works proposed scoring systems based on radiological features from contrast-enhanced CMR, while the last one used PET/CT imaging to perform univariate statistical comparisons of metabolic parameters in a small cohort of 41 patients. Despite targeting all the CMs' categories, these studies were based on a small population size and lacked a test set, thereby limiting the reliability and clinical applicability of their results.

In our study, we analyzed a relatively large dataset comprising 127 patients. Importantly, the dataset was balanced in terms of lesion type, including 62 malignant and 65 benign CMs. These encompassed the full spectrum of CM categories, namely, primary benign and malignant neoplasms, secondary (metastatic) tumors, and pseudo tumors. This balanced distribution reflects the study setting, i.e., a tertiary referral "hub" center with a dedicated cardiac imaging and surgery program, which mainly receives complex cases from peripheral "spoke" hospitals, including a higher proportion of suspected or confirmed malignant lesions. Consequently, the cohort composition mirrors a specialized referral case-mix rather than the epidemiological distribution in the general population, where benign lesions are more frequent.

A key novelty of our approach lied in the adoption of a semi-automated segmentation strategy within the proposed pipeline, which, to the best of our knowledge, has not been previously applied to the task of CM classification. While similar approaches have been successfully explored in other clinical domains, such as post-infarction scar detection [37] and musculoskeletal tumor diagnosis [38], prior studies on CMs have relied exclusively on manual ROI delineation. This advancement represents an important step toward improving reproducibility, efficiency, and clinical translation.

Despite the high morphological heterogeneity of CMs, the fine-tuned segmentation network demonstrated good performance, achieving a Dice score of 0.78 on the test set. Segmentation performance appears adequate for subsequent radiomic analysis, with 88% of the radiomic features extracted from the semiautomatic ROIs showing good reproducibility (ICC > 0.6) when compared with those derived from manual ROIs.

From a computational standpoint, the proposed framework is efficient and clinically feasible. Despite being more complex than a standard U-Net, the segmentation network maintains comparable inference times (20–50 ms per image vs. 10–25 ms).

For the classification task, two models were developed: one based on radiomic features extracted from manually segmented CMs, and the other using features derived from semiautomatically segmented volumes. Both models demonstrated highly comparable performance, with identical balanced accuracy (0.85) and the same number of misclassifications. Minor differences were observed in other metrics: the manual segmentation model achieved a higher AUC (0.95 vs. 0.82) but a lower F1-score (0.77 vs. 0.85) compared to the semiautomated model. These results indicate that the two models are effectively interchangeable for the classification task. Notably, despite the same number of misclassifications, the semiautomated model yielded one fewer false negative than the manual model. From a clinical perspective, this aspect is particularly relevant, as reducing false negatives, i.e., missed malignant masses, is generally more critical than misclassifying benign lesions. Indeed, failing to identify a malignant mass may delay diagnosis and treatment, potentially leading to disease progression and worse patient outcomes.

Three misclassifications were shared between the two models: two malignant lesions (neoplastic thrombosis from renal carcinoma and hepatic metastasis) were misclassified as benign, and one benign lesion (ventricular fibroma) was incorrectly classified as malignant. Beyond these common errors, each model presented one additional unique misclassification: the manual segmentation model failed to identify a diffuse large B-cell lymphoma, while the semiautomated model incorrectly assigned a pericardial hemangioma to the malignant class. Radiomic analysis of the common misclassified cases revealed that the lesions exhibited atypical feature profiles, particularly near-zero energy in high-frequency wavelet decompositions (HHH band) and extreme values in low-frequency percentile features, placing them as outliers within the learned feature space of both classifiers. Future work incorporating larger, more diverse training cohorts and histology-stratified validation may help address these edge cases and further improve model robustness.

In terms of potential clinical integration, CardioRadNet is designed to be incorporated into the routine CMR workflow as a decision-support tool applicable even in peripheral imaging centers where standard contrast-free T1-weighted CMR sequences are acquired. In this context, the model could provide an early, automated risk stratification of CMs directly at the point of image acquisition. In practical terms, after routine CMR acquisition in a spoke hospital, the semiautomated segmentation and classification pipeline could assist in identifying patients with a higher likelihood of malignancy, thereby supporting the decision on whether referral to a tertiary hub center is warranted. This could be particularly valuable in improving triage efficiency, ensuring that patients with suspicious findings are appropriately escalated for further diagnostic workup or specialized management, while potentially reducing unnecessary referrals for benign conditions. Importantly, the tool is intended as a decision-support system rather than a replacement for expert interpretation, with the goal of improving diagnostic consistency, optimizing referral pathways, and facilitating earlier identification of clinically relevant cases in a real-world distributed healthcare setting.

Overall, the diagnostic accuracy achieved by our CardioRadNet in characterizing the biological nature of CMs from contrast-free CMR scans is promising. However, some limitations must be acknowledged. First, although reasonable given the rarity of the pathology, the cohort remains relatively limited and monocentric, which constrains model generalizability across diverse patient populations and imaging protocols. It is also worth noting, that the segmentation was performed at the slice level, meaning that each individual slice was treated as a separate input to the deep learning model; consequently, the effective number of training samples was substantially higher than the number of patients alone would suggest, partially compensating for the modest dataset size. Nonetheless, this approach cannot replace genuine inter-patient and inter-center variability. Second, the use of non-contrast, single-sequence imaging, although advantageous for patients with contraindications to contrast agents and for its minimally invasive nature, may limit the amount of complementary information available for feature extraction compared to a multimodal approach, which could be explored in future studies to further enhance model performance. Third, the semiautomated workflow still relies on a user-defined point mask for mass localization, introducing operator dependency and limiting scalability in fully automated clinical settings. Fourth, while the current segmentation performance is reliable and sufficient to support downstream radiomic analysis, its further refinement may benefit from larger and more diverse datasets. Future developments should therefore focus on fully automated segmentation and validation on larger, multicenter and multi-modal cohorts. This would not only increase sample size but also introduce greater variability in patient characteristics and clinical settings, thereby strengthening robustness, minimizing overfitting, and enhancing external validity and applicability in real-world clinical contexts.

5. Conclusions

Our study introduces CardioRadNet, a semiautomated pipeline that integrates a deep learning-based segmentation network for mass delineation with a radiomics-based classification model, enabling objective and reproducible differentiation between benign and malignant CMs. Unlike prior studies, which often focus on specific CMs' subtypes and rely on manual delineation, CardioRadNet considers the full spectrum of CM types and uses semiautomated segmentation for broader clinical applicability. Importantly, our approach relies exclusively on routinely acquired contrast-free T1-weighted CMR sequences, thereby avoiding the risks associated with contrast agent administration, particularly relevant for patients with contraindications to gadolinium, such as those with severe renal impairment or allergies. By leveraging non-contrast imaging and automation, our study addresses the major limitations of prior works and proposes a scalable, clinically applicable tool for accurate CM characterization. Future studies including larger, multicenter and multi-modal cohorts will serve to confirm these promising results and assess the applicability in real-world clinical contexts.

Acknowledgement

Research partially funded by the Italian Ministry of Health, PSC Salute, Traiettoria 2-eHealth, Project RADIOAMICA (CUP: B47G22000220008)

References

[1] F. Angeli *et al.*, «Multimodality Imaging in the Diagnostic Work-Up of Patients With Cardiac Masses: JACC: CardioOncology State-of-the-Art Review», *JACC CardioOncology*, vol. 6, fasc. 6, pp. 847–862, dic. 2024, doi: 10.1016/j.jaccao.2024.09.006.

[2] C. J. T. Karigyo *et al.*, «Cardiac Tumors: Review», *Braz. J. Cardiovasc. Surg.*, vol. 39, fasc. 6, p. e20230405, lug. 2024, doi: 10.21470/1678-9741-2023-0405.

[3] M. J. Leja, D. J. Shah, e M. J. Reardon, «Primary cardiac tumors», *Tex. Heart Inst. J.*, vol. 38, fasc. 3, pp. 261–262, 2011.

[4] N. Prabhu e J. M. DeCara, «Imaging Cardiac Masses in Patients with Cancer», *Isr. Med. Assoc. J. IMAJ*, vol. 24, fasc. 3, pp. 186–190, mar. 2022.

[5] R. Calabretta e M. Hacker, «A PET-derived tumor expansion pattern to differentiate between primary cardiac lymphoma from primary cardiac sarcoma», *J. Nucl. Cardiol. Off. Publ. Am. Soc. Nucl. Cardiol.*, vol. 29, fasc. 6, pp. 2878–2880, dic. 2022, doi: 10.1007/s12350-022-03097-z.

[6] T. J. Poterucha, J. Kochav, D. S. O'Connor, e G. F. Rosner, «Cardiac Tumors: Clinical Presentation, Diagnosis, and Management», *Curr. Treat. Options Oncol.*, vol. 20, fasc. 8, p. 66, giu. 2019, doi: 10.1007/s11864-019-0662-1.

[7] B. Yanagawa *et al.*, «Surgery for Tumors of the Heart», *Semin. Thorac. Cardiovasc. Surg.*, vol. 30, fasc. 4, pp. 385–397, 2018, doi: 10.1053/j.semtcvs.2018.09.001.

[8] D. Kassop *et al.*, «Cardiac Masses on Cardiac CT: A Review», *Curr. Cardiovasc. Imaging Rep.*, vol. 7, fasc. 8, p. 9281, 2014, doi: 10.1007/s12410-014-9281-1.

[9] S. Liddy, C. McQuade, K. P. Walsh, B. Loo, e O. Buckley, «The Assessment of Cardiac Masses by Cardiac CT and CMR Including Pre-op 3D Reconstruction and Planning», *Curr. Cardiol. Rep.*, vol. 21, fasc. 9, p. 103, set. 2019, doi: 10.1007/s11886-019-1196-7.

[10] H. Uehara, K. Ohba, M. Ono, T. Imazuru, e T. Shimokawa, «Cardiac myxoma with high standardized uptake value of FDG-PET-CT in the right ventricular outflow tract», *Gen. Thorac. Cardiovasc. Surg. Cases*, vol. 3, fasc. 1, p. 55, dic. 2024, doi: 10.1186/s44215-024-00179-x.

[11] Y. J. Jeong, H. J. Yoon, e D.-Y. Kang, «Growing Cardiac Hemangioma on Serial F-18 FDG PET/CT», *Nucl. Med. Mol. Imaging*, vol. 46, fasc. 3, pp. 223–226, set. 2012, doi: 10.1007/s13139-012-0155-7.

[12] S. Nóbrega, C. Martins Da Costa, A. F. Amador, S. Justo, e E. Martins, «Cardiovascular Magnetic Resonance Versus Histopathologic Study for Diagnosis of Benign and Malignant Cardiac Tumours: A Systematic Review and Meta-Analysis», *J. Cardiovasc. Imaging*, vol. 31, fasc. 4, p. 159, 2023, doi: 10.4250/jcvi.2023.0028.

[13] J. E. Van Timmeren, D. Cester, S. Tanadini-Lang, H. Alkadhi, e B. Baessler, «Radiomics in medical imaging—"how-to" guide and critical reflection», *Insights Imaging*, vol. 11, fasc. 1, p. 91, dic. 2020, doi: 10.1186/s13244-020-00887-2.

[14] M. K. Halle *et al.*, «Radiomic profiles improve prognostication and reveal targets for therapy in cervical cancer», *Sci. Rep.*, vol. 14, fasc. 1, p. 11339, mag. 2024, doi: 10.1038/s41598-024-61271-4.

[15] C. You *et al.*, «Multicenter radio-multiomic analysis for predicting breast cancer outcome and unravelling imaging-biological connection», *Npj Precis. Oncol.*, vol. 8, fasc. 1, p. 193, set. 2024, doi: 10.1038/s41698-024-00666-y.

[16] M. Kolossváry *et al.*, «Radiomic Features Are Superior to Conventional Quantitative Computed Tomographic Metrics to Identify Coronary Plaques With Napkin-Ring Sign», *Circ. Cardiovasc. Imaging*, vol. 10, fasc. 12, p. e006843, dic. 2017, doi: 10.1161/CIRCIMAGING.117.006843.

[17] F. Lo Iacono, R. Maragna, G. Pontone, e V. D. A. Corino, «A robust radiomic-based machine learning approach to detect cardiac amyloidosis using cardiac computed tomography», *Front. Radiol.*, vol. 3, p. 1193046, giu. 2023, doi: 10.3389/fradi.2023.1193046.

[18] A. Corti *et al.*, «Enhancing cardiovascular risk stratification: Radiomics of coronary plaque and perivascular adipose tissue – Current insights and future perspectives», *Trends Cardiovasc. Med.*, vol. 35, fasc. 1, pp. 47–59, gen. 2025, doi: 10.1016/j.tcm.2024.06.003.

[19] J. Son *et al.*, «Radiomics Feature Analysis Using Native T1 Mapping for Discriminating Between Cardiac Tumors and Thrombi», *Acad. Radiol.*, vol. 29, pp. S1–S8, apr. 2022, doi: 10.1016/j.acra.2020.12.009.

[20] M. Kassi *et al.*, «Differentiating benign from malignant cardiac tumors with cardiac magnetic resonance imaging», *J. Thorac. Cardiovasc. Surg.*, vol. 157, fasc. 5, pp. 1912-1922.e2, mag. 2019, doi: 10.1016/j.jtcvs.2018.09.057.

[21] X. Hu, P. Yang, D. Pan, e P. Wang, «18F-FDG PET/CT metabolic parameters can semi-quantitatively evaluate the nature of the heart and pericardial masses: a retrospective study», *Sci. Rep.*, vol. 14, fasc. 1, lug. 2024, doi: 10.1038/s41598-024-67336-8.

[22] P. Paolisso *et al.*, «Cardiac Magnetic Resonance to Predict Cardiac Mass Malignancy: The CMR Mass Score», *Circ. Cardiovasc. Imaging*, vol. 17, fasc. 3, mar. 2024, doi: 10.1161/circimaging.123.016115.

[23] P. Yue *et al.*, «Differential and prognostic value of cardiovascular magnetic resonance derived scoring algorithm in cardiac tumors», *Int. J. Cardiol.*, vol. 331, pp. 281–288, mag. 2021, doi: 10.1016/j.ijcard.2021.01.068.

[24] P. Pazos-López *et al.*, «Value of CMR for the Differential Diagnosis of Cardiac Masses», *JACC Cardiovasc. Imaging*, vol. 7, fasc. 9, pp. 896–905, set. 2014, doi: 10.1016/j.jcmg.2014.05.009.

[25] J. W. Lee *et al.*, «CT-based radiomics signature for differentiation between cardiac tumors and thrombi: a retrospective, multicenter study», *Sci. Rep.*, vol. 12, fasc. 1, mag. 2022, doi: 10.1038/s41598-022-12229-x.

[26] Y. Bao *et al.*, «Development and validation of a novel echocardiography-based nomogram for the streamlined classification of cardiac tumors in cancer patients», *Quant. Imaging Med. Surg.*, vol. 15, fasc. 3, pp. 1873–1887, mar. 2025, doi: 10.21037/qims-24-1096.

[27] J. J. Maleszewski *et al.*, «The 2021 WHO Classification of Tumors of the Heart», *J. Thorac. Oncol.*, vol. 17, fasc. 4, pp. 510–518, apr. 2022, doi: 10.1016/j.jtho.2021.10.021.

[28] A. Fedorov *et al.*, «3D Slicer as an image computing platform for the Quantitative Imaging Network», *Magn. Reson. Imaging*, vol. 30, fasc. 9, pp. 1323–1341, nov. 2012, doi: 10.1016/j.mri.2012.05.001.

[29] A. Zwanenburg *et al.*, «The Image Biomarker Standardization Initiative: Standardized Quantitative Radiomics for High-Throughput Image-based Phenotyping», *Radiology*, vol. 295, fasc. 2, pp. 328–338, mag. 2020, doi: 10.1148/radiol.2020191145.

[30] S. Gitto *et al.*, «Deep learning for automated segmentation of central cartilage tumors on MRI», *Eur. Radiol. Exp.*, vol. 9, fasc. 1, p. 91, set. 2025, doi: 10.1186/s41747-025-00633-7.

[31] J. J. M. Van Griethuysen *et al.*, «Computational Radiomics System to Decode the Radiographic Phenotype», *Cancer Res.*, vol. 77, fasc. 21, pp. e104–e107, nov. 2017, doi: 10.1158/0008-5472.CAN-17-0339.

[32] «Assessing Left Ventricle Radiomic Features Robustness by Segmentation Perturbations», in *IFMBE Proceedings*, Cham: Springer Nature Switzerland, 2024, pp. 356–362. doi: 10.1007/978-3-031-49068-2_36.

[33] A. Zwanenburg *et al.*, «Assessing robustness of radiomic features by image perturbation», *Sci. Rep.*, vol. 9, fasc. 1, p. 614, gen. 2019, doi: 10.1038/s41598-018-36938-4.

[34] T. K. Koo e M. Y. Li, «A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research», *J. Chiropr. Med.*, vol. 15, fasc. 2, pp. 155–163, giu. 2016, doi: 10.1016/j.jcm.2016.02.012.

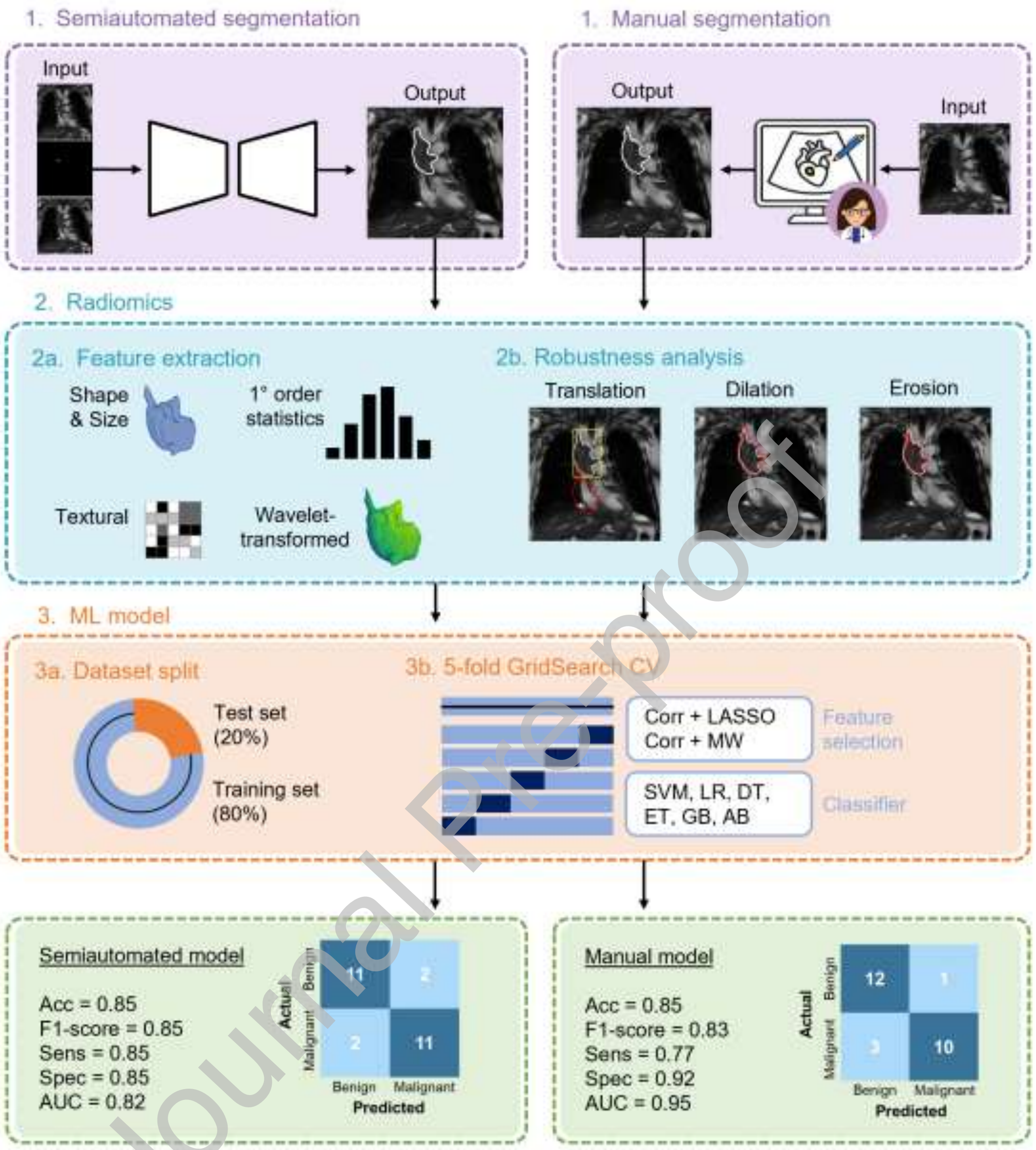
[35] N. V. J. Vanami, L. R. Chintalapati, Y. Challagundla, e S. N. Mohanty, «Feature Selection Using Data Mining Techniques for Prognostication of Cardiovascular Diseases», in *Multimedia Technology and Enhanced Learning*, vol. 532, B. Wang, Z. Hu, X. Jiang, e Y.-D. Zhang, A c. di, in Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, vol. 532., Cham: Springer Nature Switzerland, 2024, pp. 344–353. doi: 10.1007/978-3-031-50571-3_24.

[36] B. Kocak *et al.*, «METhodological RadiomICs Score (METRICS): a quality scoring tool for radiomics research endorsed by EuSoMII», *Insights Imaging*, vol. 15, fasc. 1, p. 8, gen. 2024, doi: 10.1186/s13244-023-01572-w.

[37] W. Xu e X. Shi, «Integrating radiomic texture analysis and deep learning for automated myocardial infarction detection in cine-MRI», *Sci. Rep.*, vol. 15, fasc. 1, p. 24365, lug. 2025, doi: 10.1038/s41598-025-08127-7.

[38] S. Wang *et al.*, «Advancing musculoskeletal tumor diagnosis: Automated segmentation and predictive classification using deep learning and radiomics», *Comput. Biol. Med.*, vol. 175, p. 108502, giu. 2024, doi: 10.1016/j.compbiomed.2024.108502.

Graphical abstract



Ethics Statement

The study was conducted according to the guidelines of the Declaration of Helsinki. The study protocol was approved by the local Ethics Committee (Registration No. 102/2017/Oss/AOUBo) and all the patients signed the informed consent.

Declaration of interests

☐The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☒The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Meri Ferretti reports financial support was provided by Italian Ministry of Health. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.