Journal of Applied Research and Technology

www.jart.icat.unam.mx

ICAT Instituto de Ciencias Aplicadas y Tecnología

Journal of Applied Research and Technology 23 (2025) 651-658

Original

Cardiac Region Detection Using YOLO-Based Deep Learning Models: A Performance Analysis of YOLO Models

A. Nasuha[1] • A. T. Musaddid[1] • D. Irmawati[*1] • P. T. Aji[1]
M. A. H. Sofyan[1] • S. R. Hakim[1]

[1]*Universitas Negeri Yogyakarta, Indonesia*

Received: 12 21 2023; Accepted: 09 24 2025
Available: 12 31 2025

Abstract: Cardiac region detection is a critical task for the diagnosis and management of cardiovascular diseases (CVDs), where precise identification of key areas, including the left ventricle (LV), right ventricle (RV), and myocardium (MYO), from cardiac magnetic resonance (CMR) images is essential. This study presents a comprehensive performance analysis of various YOLO (You Only Look Once) deep learning models, including YOLOv5, YOLOv8, YOLOv9, and YOLOv10, for the automatic detection of these cardiac regions. The dataset utilized comprises 2D images derived from 3D MRI scans, segmented into training and testing sets. The models were evaluated based on standard metrics such as precision, recall, and mean average precision (mAP). Results indicate that YOLOv5l achieved the highest precision of 99.11%, while YOLOv5s recorded the best recall at 97.99%. The findings demonstrate that smaller models, such as YOLOv5n and YOLOv9t, exhibited superior precision despite their reduced computational requirements. The analysis suggests that model size does not always correlate with performance, highlighting the potential of lighter models for real-time applications in clinical settings.

Keywords: **cardiac region detection, YOLO, deep learning, left ventricle, myocardium, right ventricle.**

*Corresponding author.
E-mail address: dessy.irmawati@uny.ac.id (D. Irmawati)).
Peer Review under the responsibility of Universidad Nacional Autónoma de México.

1. Introduction

Cardiovascular diseases (CVDs) remain a leading cause of morbidity and mortality worldwide, with cardiomyopathies and other cardiac abnormalities playing a major role in their prevalence. (Banerjee & Annamalai, 2024; Yasmin et al., 2021; Zakariah & AlShalfan, 2020). Accurate evaluation of cardiac structures, including the myocardium, left ventricle (LV), and right ventricle (RV), is essential for diagnosing and monitoring cardiovascular conditions such as heart failure, myocardial infarction, and valve abnormalities. (Subha et al., 2024). Cardiac magnetic resonance (CMR) is preferred among the available imaging modalities for its superior image quality and comprehensive cardiac coverage. (Zakariah et al., 2020). Detecting and analyzing these key cardiac regions in CMR images is vital for assessing heart function and guiding clinical decisions.

The image acquisition technique for cardiac MRI may not always result in adequate image quality due to the motion of the patient throughout the scan, the capability of the MR equipment to focus on the right region, or the patient's arrhythmia (Karakamış et al., 2021). Cardiac MRI images pose various obstacles to researchers in identifying CMR regions, including increased noise, intensity inhomogeneity, and partial volume effects. The volume of a complex topological structure, with significant variation between distinct slices (Zakariah et al., 2020). Heart abnormalities can be identified from MRI images of the heart by analyzing important areas or parts of the heart, including the Myocardium (MYO), Left Ventricle (LV), and Right Ventricle (RV). Cine imaging allows precise visualization of global and localized myocardial function, and LV contouring of the myocardium at distinct cardiac phases can provide a detailed estimate of ejection fraction. (Situ et al., 2020).

Advances in artificial intelligence (AI) and computer vision have led to the development of more efficient and accurate techniques for detecting cardiac structures, with deep learning models emerging as powerful tools for this task. One such model, You Only Look Once (YOLO) (Redmon et al., 2016), is widely recognized for its real-time object detection capabilities and high accuracy. A similar study by Kermani et al. (developed a technique dubbed NF-RCNN that employs MRI data to identify anomalies in the heart ventricle's right wall. In the meantime, Chen et al. (2024) employed a variant of YOLOv4-DenseNet to identify ventricular septal defects in ultrasonic images, and Alsanea & Dutta (2024) used DenseNet161 as a classifier and YOLOv7 for feature extraction to identify cardiac anomalies. In this study, various versions of YOLO were employed to detect important parts of the cardiac, such as the myocardium, LV, and RV in CMR images. By comparing the performance of different YOLO versions, this research aims to improve detection accuracy and optimize system performance, providing cardiologists with more precise information for diagnosing and assessing heart conditions.

2. Related Work

The development of deep learning technology, such as object detection, has gained significant interest across multiple computing domains as a key task of computational vision.

Diagnosis of myocarditis using cardiac magnetic resonance imaging (CMRI) is challenging due to the low contrast of the images. A deep learning-based computer-aided diagnostic system (CADS) utilizing transformer models offers significant improvements in accuracy. The proposed CADS follows a series of steps, including data selection, preprocessing, feature extraction, classification, and post-processing. The preprocessing phase involves denoising, resizing, and applying data augmentation techniques such as CutMix and MixUp. This approach achieved a 99.73% accuracy in detecting myocarditis. (Wamil et al., 2022). In this study, only one part of the cardiac system, which is the myocardium, is studied to diagnose myocarditis.

This study investigates the use of a 3D convolutional neural network (CNN) combined with a salient region detection model to enhance the accuracy of cardiovascular magnetic resonance (CMR) imaging, specifically for ensuring complete cardiac coverage, which is critical for accurate volumetric and functional assessments. The salient region detection model applies a three-step algorithm to extract key regions from short-axis cine CMR stacks, effectively reducing false negatives. Tested on a dataset of over 6,200 participants from the UK Biobank, the model outperformed previous approaches, improving the accuracy of detecting basal and apical slices. The baseline model's accuracy in identifying these slices increased from 96.25% to 96.88% for basal slices and from 94.51% to 95.72% for apical slices after integrating the salient region detection model. This improvement highlights the model's potential to enhance image-quality control in large-scale population datasets and to enable real-time post-imaging assessments. (Nabavi et al., 2022). The increase in slice identification still needs further improvement, because both the basal (upper part near the

atria) and apical (lower tip) slices are critical for accurate volumetric and functional measurements, which are essential for diagnosing various cardiac conditions. Misidentification or incomplete coverage of these slices can result in inaccurate assessment of the heart's overall performance, especially for calculating ejection fraction, cardiac output, and chamber volumes.

(Ahmad et al., 2023) created a fully automated pipeline for segmenting the right ventricle (RV), myocardium (MYO), and left ventricle (LV) using short-axis CMRI sequence images. The approach introduces a dilated residual network (DRN) that captures features at full resolution in the UNet's bottleneck layer, enhancing both spatial and temporal information while preserving localization accuracy. To mitigate overfitting and address class imbalance, it employed data augmentation techniques. Additionally, pixel-wise outputs from each expanding path are combined to enhance the training response. The proposed method for the automatic cardiac diagnostic challenge. The test set comprises 50 patient records. The model constructed attained an overall Dice Similarity Coefficient (DSC) of 0.924 ± 0.03, 0.907 ± 0.01, and 0.949 ± 0.05 for RV, MYO, and LV, respectively. We found Hausdorff distance (HD) scores of 10.09 ± 0.01, 7.25 ± 0.05, and 6.86 ± 0.02 mm for RV, MYO, and LV, respectively. The results demonstrate outstanding performance, outperforming cutting-edge approaches in terms of accuracy and achieving expert-level segmentation. As a result, the total DSC and HD scores improved by 1.0% and 1.5%, respectively.

Alsanea & Dutta (2024) introduced a deep learning approach that combines YOLOv7 with DenseNet161 to detect cardiovascular diseases in radiographic images. Their model achieved an accuracy of 98.5% and a sensitivity of 97.8%, demonstrating its robustness in identifying cardiovascular abnormalities. This work integrates YOLOv7's real-time detection capabilities with DenseNet161's feature-extraction efficiency, offering a solution suitable for clinical applications that require both precision and speed.

(Chen et al., 2021) developed a modified YOLOv4-DenseNet algorithm for detecting ventricular septal defects in ultrasound images. This model achieved a mean Average Precision (mAP) of 92.3% and an F1 Score of 0.88, outperforming existing methods on this task. The integration of DenseNet with YOLOv4 enabled enhanced feature extraction and object detection, making it a valuable tool for accurate diagnosis of ventricular defects in clinical settings.

In their study, Farea Shaaf et al. (2023) utilized Faster R-CNN to detect the left ventricular cavity in cardiac MRI images. The model achieved high performance, with a precision of 94%, a recall of 95%, and an F1 Score of 0.95, demonstrating its ability to localize and identify complex cardiac structures. This research highlights the effectiveness of region-based convolutional neural networks for medical imaging segmentation tasks.

(Kermani et al., 2020) proposed the NF-RCNN model for localizing the heart and detecting right ventricle wall abnormalities in cardiac MRI scans. Their approach yielded a Dice Similarity Coefficient (DSC) of 0.93, demonstrating its ability to segment and identify abnormalities in challenging MRI datasets accurately. This work emphasizes the potential of region-based neural networks for precise segmentation of cardiac structures and detection of abnormalities.

(Lee et al., 2022) applied YOLOv5 with data preprocessing techniques for cardiac detection tasks, focusing on identifying critical cardiac areas from MRI images. Their model demonstrated a precision of 90.2% and a recall of 88.7%, highlighting its capability for real-time detection of cardiac abnormalities. This study underscores the utility of YOLOv5's real-time object detection framework in medical imaging applications, particularly when preprocessing is applied to enhance detection accuracy.

3. Method

3.1 Cardiac Region Detection and Dataset

This study focuses on detecting three critical regions of the heart, namely the Left Ventricle (LV), the Right Ventricle (RV), and the Myocardium (MYO) from Magnetic Resonance Imaging (MRI) scans. The detection task was performed using the state-of-the-art YOLOv5 model. (Jocher et al., 2022), YOLOv8 (Jocher et al., 2023), YOLOv9 (C.-Y. Wang & Liao, 2024) and the latest YOLOv10 (A. Wang et al., 2024).

The dataset used in this study was derived from 3D MRI images obtained from the Automated Cardiac Diagnosis Challenge (ACDC), which were subsequently randomly sampled into 2D images. The final dataset comprises 2,107 images. To ensure robustness, Table 1 shows that the dataset was randomly shuffled and then divided into training and testing subsets, with the training dataset consisting of 1,648 images and the testing dataset containing 459 images.

Table 1. Dataset.

Dataset	Total	Annotation		
		LV	RV	MYO
Training	1,648	1,384	1,640	1619
Testing	459	429	454	452

3.2 YOLO Algorithm

The You Only Look Once (YOLO) algorithm, introduced by Redmon et al. (2016), is a fast and efficient method for real-time object detection. Instead of analyzing different parts of an image in isolation, YOLO processes the entire image at once using a convolutional neural network (CNN). It divides the image into a grid, where each cell predicts bounding boxes and confidence scores for objects. The output includes the box coordinates, confidence scores, and class probabilities. This approach allows YOLO to process up to 45 frames per second, making it highly effective for applications such as autonomous driving and surveillance while maintaining high accuracy.

Since its inception, YOLO has undergone several iterations, including YOLOv5, YOLOv8, YOLOv9, and the latest YOLOv10, each version introducing enhancements in speed, accuracy, and robustness. These improvements have focused on better feature extraction, improved small-object detection, adaptive anchor boxes, and advanced data augmentation techniques, solidifying YOLO's position as a leading choice in object detection.

YOLOv5 is known for balancing detection accuracy and speed. It uses a Convolutional Neural Network (CNN) for feature extraction and an anchor-based detection approach. In YOLOv5, anchor boxes are used to detect objects at different scales. The model optimizes the Intersection over Union (IoU) for localization accuracy, where A and B are the areas of the predicted and ground-truth bounding boxes in Equation (1).

$$\mathrm{IoU} = \frac{\mathrm{A} \cap \mathrm{B}}{\mathrm{A} \cup \mathrm{B}} \tag{1}$$

In Equation (2), the total loss L in YOLOv5 includes objectness, classification, and localization losses where $\mathrm{L}_{\mathrm{obj}}$, $\mathrm{L}_{\mathrm{cls}}$ and $\mathrm{L}_{\mathrm{loc}}$ represent the objectness, classification, and localization losses.

$$\mathrm{L} = \delta_{\mathrm{obj}}\mathrm{L}_{\mathrm{obj}} + \delta_{\mathrm{cls}}\mathrm{L}_{\mathrm{cls}} + \delta_{\mathrm{loc}}\mathrm{L}_{\mathrm{loc}} \tag{2}$$

YOLOv8 enhances YOLOv5 by introducing a Cross-Stage Partial Network (CSPNet) to improve gradient flow and reduce computational cost during feature extraction. CSPNet improves efficiency by partitioning feature maps into two paths, allowing gradients to flow more efficiently through the network, reducing computational load. YOLOv8 dynamically adjusts its loss components in (3) where α and β are adjusted during training for better optimization.

$$\mathrm{L}_{\mathrm{adaptive}} = \alpha\mathrm{L}_{\mathrm{cls}} + \beta\mathrm{L}_{\mathrm{loc}} \tag{3}$$

YOLOv9 introduces the Generalized Efficient Layer Aggregation Network (GELAN), a novel way to combine spatial information at multiple scales. This architecture enables multi-scale spatial pooling and efficient layer aggregation, improving feature learning for object detection. The total loss in YOLOv9 includes a new term L_{gelan} that accounts for the layer aggregation as computed in (4). This term enhances learning across different feature map scales.

$$\mathrm{L} = \mathrm{L}_{\mathrm{cls}} + \mathrm{L}_{\mathrm{loc}} + \mathrm{L}_{\mathrm{gelan}} \tag{4}$$

YOLOv10 eliminates the need for Non-Maximum Suppression (NMS), reducing inference time while maintaining accuracy. This is achieved through dual label assignments, where the model performs one-to-one matching during inference and one-to-many matching during training. This method improves detection efficiency without the computational burden of traditional post-processing as calculated in (5).

$$\mathrm{L} = \mathrm{L}_{\mathrm{one-to-one}} + \mathrm{L}_{\mathrm{one-to-many}} \tag{5}$$

In this study, we focused on newer YOLO models (YOLOv5 to YOLOv10) for detecting cardiac regions, given their advancements in computational efficiency and accuracy. Earlier models, such as YOLOv2, YOLOv3, and YOLOv4, though foundational in object detection, were not included in the comparison. These earlier models often featured larger sizes and higher computational demands than some newer variants, making them less suitable for real-time applications.

3.3 Metrics

Each model (YOLOv5, YOLOv8, YOLOv9, YOLOv10) was trained and evaluated on MRI scans for cardiac region detection, where the following metrics were used in Equations (6)-(9):

$$\mathrm{Precision} = \frac{\mathrm{TP}}{\mathrm{TP+FP}} \tag{6}$$

where TP is the number of true positives and FP is the number of false positives.

$$\mathrm{Recall} = \frac{\mathrm{TP}}{\mathrm{TP+FP}} \tag{7}$$

where FN is the number of false negatives.

Precision (P) - Recall (R) Curve is the precision and recall values, which are computed at different IoU (1) thresholds. For a given class *c*, the AP (Average Precision) is the area under the precision-recall curve. Where $P_c(r)$ is the precision as a function of recall *r*.

$$\mathrm{AP_c} = \int_0^1 \mathrm{P_c(r)dr} \tag{8}$$

Mean Average Precision (mAP): The overall mAP is calculated as the mean of the AP scores over all object classes .

$$\mathrm{mAP} = \frac{1}{\mathrm{C}} \sum_{\mathrm{c=1}}^{\mathrm{C}} \mathrm{AP_c} \tag{9}$$

mAP-50 metric evaluates the mean Average Precision (mAP) at an Intersection over Union (IoU) threshold of 0.50, which is a standard benchmark for object detection tasks. On the other hand, mAP50-90 is an extension that computes the mean mAP across IoU thresholds from 0.50 to 0.90 in 0.05 increments. This provides a more comprehensive evaluation by considering detection performance across a range of localization-precision levels, which is particularly useful for scenarios requiring fine-grained analysis.

4. Implementation Result

4.1 Experimental Setup

This study was conducted on a computer equipped with an Intel Core i5-12400F CPU running at 2.5GHz and a 12GB NVIDIA GeForce RTX 3060 GPU for both training and testing purposes. The system was set up with Pytorch with CUDA 11.8, and Python version 3.8.19 was used as the programming language for the research's implementation. Each network model was configured with a batch size of 16, trained in 100 epochs, and started with an initial learning rate of 0.001, which was gradually reduced after several epochs.

4.2 Training

The models were trained for this study on a total of 16 YOLO variants, as in Figure 1, encompassing YOLOv5, YOLOv8, YOLOv9, and YOLOv10. As presented in Table 2, these models range from the smallest to the largest versions within each YOLO version in 100 epochs. However, experiments were not conducted on all the largest variants of these versions.

In this training phase, model performance exhibits variability and is not strictly proportional to model size. Notably, smaller models such as YOLOv5n and YOLOv9t demonstrate superior precision compared to certain larger models, indicating that increased model size does not consistently correlate with improved performance. However, overall, YOLOv5l achieves the highest precision at 99.11%, while YOLOv5s records the best recall at 97.99%.

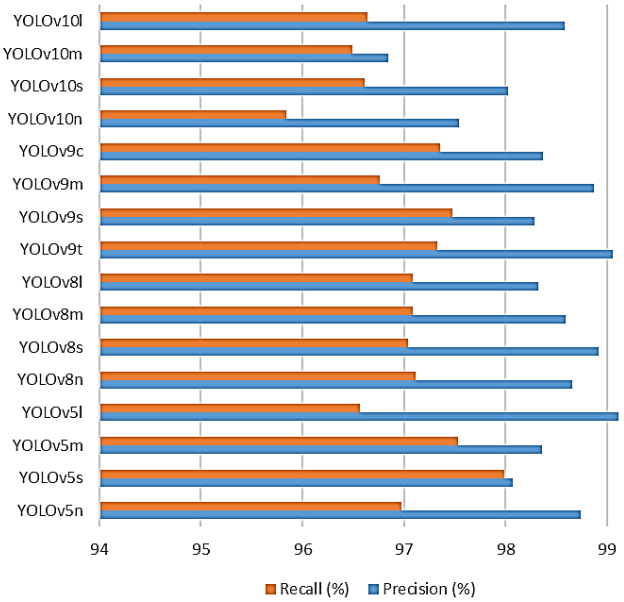

Figure 1. Training results.

Table 2. Training results.

Model	Precision (%)	Recall (%)
YOLOv5n	98.74	96.98
YOLOv5s	98.07	97.99
YOLOv5m	98.36	97.53
YOLOv5l	99.11	96.57
YOLOv8n	98.66	97.12
YOLOv8s	98.92	97.04
YOLOv8m	98.59	97.09
YOLOv8l	98.32	97.09
YOLOv9t	99.06	97.33
YOLOv9s	98.29	97.48
YOLOv9m	98.87	96.76
YOLOv9c	98.37	97.36
YOLOv10n	97.54	95.84
YOLOv10s	98.03	96.61
YOLOv10m	96.85	96.49
YOLOv10l	98.58	96.64

4.3 Evaluation

In the performance evaluation of the YOLO models for cardiac region detection (Table 3) and Figure 2 compare the mAP performance of YOLOv5, YOLOv8, YOLOv9, and YOLOv10 models, showing YOLOv5 achieving the highest mAP as the number of parameters increases, while YOLOv9 excels at lower parameter counts. YOLOv5 models, ranging from nano (YOLOv5n) to large (YOLOv5l), demonstrate robust precision and recall metrics, with mAP-50 consistently at 99.1%. YOLOv5n, characterized by its minimal parameter count and computational overhead, exhibits strong detection capabilities, making it a suitable candidate for real-time applications. As model complexity increases from YOLOv5m to YOLOv5l, there is a slight performance improvement, particularly in the mAP50-90 metric, although this comes with a substantial increase in computational requirements.

In comparison, YOLOv8n, despite being a newer version, presents performance metrics closely aligned with YOLOv5n, particularly in precision (98.7%) and recall, with a marginally lower mAP50-90 value of 90.1%. This indicates minimal performance gains between the two versions, with YOLOv8n offering efficiency improvements rather than significant enhancements in detection accuracy, especially at higher intersection over union (IoU) thresholds.

The latest YOLOv10 models show a different balance between computational efficiency and detection performance. YOLOv10n, with 2.69M parameters and 8.2 GFLOPs, reports a precision of 97.8% and an mAP50-90 of 89.1%, slightly lower than its predecessors. The larger models, such as YOLOv10m and YOLOv10l, exhibit incremental improvements in accuracy metrics, though overall performance remains comparable to that of YOLOv5 and YOLOv8, suggesting that YOLOv10 prioritizes computational efficiency and processing speed over substantial accuracy gains.

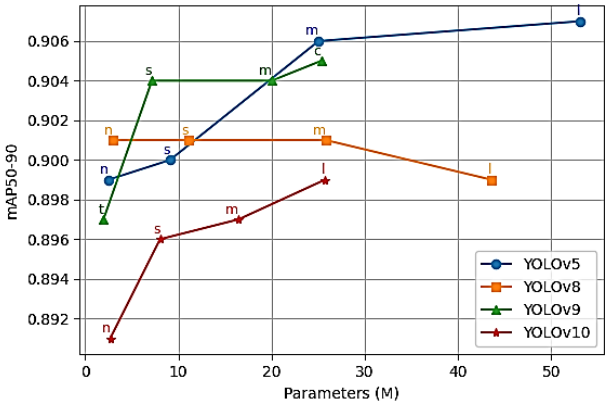

Figure 2. Model mAP50-90 comparison.

Table 3. Testing results.

Model	Params	GFLOPs	Infer. Time/img (ms)	P (%)	R (%)	mAP-50	mAP50-90
YOLOv5n	2503529	7.10	3.10	98.70	97.00	99.10	89.90
YOLOv5s	9112697	23.80	5.50	98.00	98.00	99.10	90.00
YOLOv5m	25046953	64.00	12.60	98.40	97.50	99.10	90.60
YOLOv5l	53133721	134.70	20.40	99.10	96.60	99.10	90.70
YOLOv8n	3006233	8.10	3.10	98.70	97.00	99.00	90.10
YOLOv8s	11126745	28.40	6.30	98.90	97.00	98.90	90.10
YOLOv8m	25841497	78.70	14.90	98.60	97.10	98.90	90.10
YOLOv8l	43608921	164.80	22.80	98.30	97.10	98.90	89.90
YOLOv9t	1971369	7.60	4.30	99.20	97.10	98.90	89.70
YOLOv9s	7168249	26.70	8.60	98.30	97.50	98.90	90.40
YOLOv9m	20015161	76.50	18.30	98.90	96.80	98.80	90.40
YOLOv9c	25321561	102.30	22.30	98.40	97.40	99.00	90.50
YOLOv10n	2695586	8.20	4.20	97.80	95.80	98.90	89.10
YOLOv10s	8037282	24.50	7.90	98.00	96.60	98.80	89.60
YOLOv10m	16453858	63.40	16.80	96.80	96.60	98.70	89.70
YOLOv10l	25720994	126.30	26.50	98.50	96.60	98.90	89.90

To evaluate the performance of cardiac region detection, we selected YOLOv5l due to its superior mAP50-90 score. Figure 3 illustrates the detection results across various cardiac MRI images, while Table 4 presents the confusion matrix for detections in each class. The confusion matrix shows detections, including the misclassification of 16 background areas as the left ventricle (LV). This issue may stem from visual similarities, overfitting, or class imbalance. Potential solutions include improving the training dataset, applying data augmentation techniques, and adjusting model hyperparameters.

Increasing the number of parameters tends to improve model performance (precision, recall, mAP50) as shown in Figure 2, but the improvements diminish as model size grows. There is a trade-off between model size, inference time, and performance. Larger models show better performance but come at the cost of longer inference times, while smaller models offer nearly comparable performance with much lower computational requirements.

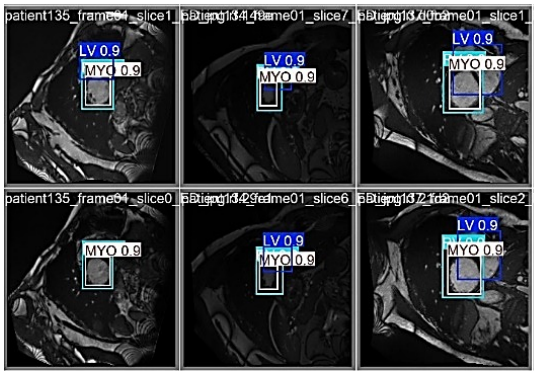

Figure 3. Sample of cardiac region detection results on YOLOv5l.

Table 4. Confusion matrix (YOLOv5l).

Predicted		True			
		LV	RV	MYO	Background
	LV	409	0	0	16
	RV	1	449	0	1
	MYO	2	0	445	4
	Background	17	5	7	0

Conclusions

In this study, the performance of various YOLO-based deep learning models (YOLOv5, YOLOv8, YOLOv9, and YOLOv10) was evaluated for cardiac region detection from CMR images. The results demonstrate that YOLOv5l provides the highest precision (99.11%) and YOLOv5s offers the best recall (97.99%). Interestingly, smaller models such as YOLOv5n and YOLOv9t showed superior precision, indicating that increased model size does not necessarily lead to better performance. This suggests that smaller YOLO models could be highly effective for real-time clinical applications without compromising detection accuracy. Overall, the findings highlight the robustness of YOLO models in detecting key cardiac structures, offering promising prospects for enhancing automated cardiovascular diagnostics.

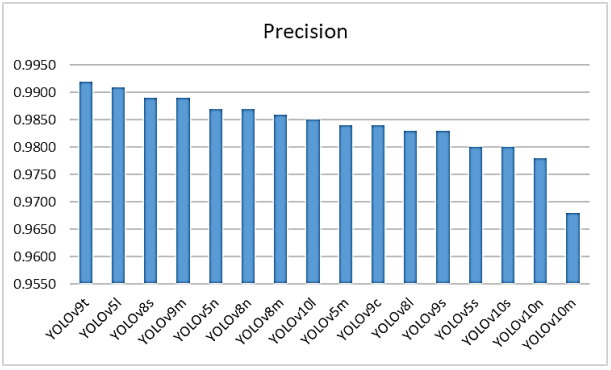

Figure 4. Performance rank by precision.

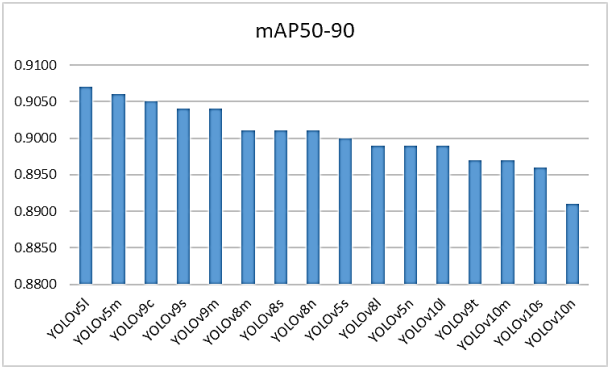

Figure 5. Performance rank by mAP50-90.

The exclusion of older models like YOLOv2, YOLOv3, and YOLOv4 was deliberate, given their larger size and higher computational cost compared to some newer variants, which limits their applicability in real-time clinical scenarios. By focusing on newer, more efficient YOLO versions, this study provides insights into the potential of lightweight deep learning models to improve automated cardiovascular diagnostics.

Conflict of interest

The authors have no conflict of interest to declare.

Funding

This work was supported by Universitas Negeri Yogyakarta.

References

Ahmad, F., Hou, W., Xiong, J., & Xia, Z. (2022). Fully automated cardiac MRI segmentation using dilated residual network. *Medical Physics*, *50*(4), 2162–2175. https://doi.org/10.1002/mp.16108

Alsanea, M., & Dutta, A. K. (2024). Detecting cardiovascular diseases from radiographic images using deep learning techniques. *Expert Systems*, *41*(7). https://doi.org/10.1111/exsy.13542

Banerjee, S., & Annamalai, R. (2024). A Deep Learning Ensemble-Based Approach for Cine MRI Segmentation and Classification Pipeline. *2024 5th International Conference on Image Processing and Capsule Networks (ICIPCN)*, 101–107. https://doi.org/10.1109/ICIPCN63822.2024.00025

Chen, Z., Chen, X., Liu, Y., Chen, E. Z., Chen, T., & Sun, S. (2023). Enhancing cardiac MRI segmentation via Classifier-Guided Two-Stage network and All-Slice Information Fusion transformer. In *Lecture notes in computer science* (pp. 145–154). https://doi.org/10.1007/978-3-031-47076-9_15

Jocher, G., Chaurasia, A., & Qiu, J. (2023). Ultralytics YOLOv8. https://github.com/ultralytics/ultralytics (consulted in October 2024).

Karakamış, K., Özer, C., & Öksüz, İ. (2021). Artifact Detection in Cardiac MRI Data by Deep Learning Methods. *2021 29th Signal Processing and Communications Applications Conference (SIU)*, 1–4. https://doi.org/10.1109/SIU53274.2021.9477844

Kermani, S., Oghli, M. G., Mohammadzadeh, A., & Kafieh, R. (2020). NF-RCNN: Heart localization and right ventricle wall motion abnormality detection in cardiac MRI. *Physica Medica*, *70*, 65–74. https://doi.org/10.1016/j.ejmp.2020.01.011

Nabavi, S., Hashemi, M., Moghaddam, M. E., Abin, A. A., & Frangi, A. F. (2022). Fully Automated Assessment of Cardiac Coverage in Cine Cardiovascular Magnetic Resonance Images using an Explainable Deep Visual Salient Region Detection Model. https://doi.org/10.48550/arXiv.2206.06844

Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You Only Look Once: Unified, Real-Time Object Detection. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 779–788. https://doi.org/10.1109/CVPR.2016.91

Situ, Y., Birch, S. C. M., Moreyra, C., & Holloway, C. J. (2020). Cardiovascular magnetic resonance imaging for structural heart disease. Cardiovascular Diagnosis and *Therapy*, 10(2), 361–375. https://doi.org/10.21037/cdt.2019.06.02

Subha, V., Gomathi, G., & Manivanna Boopathi, A. (2024). An exploration of ventricle regions segmentation and multiclass disease detection using cardiac MRI. *International Journal of Imaging Systems and Technology*, 34(1). https://doi.org/10.1002/ima.22938

Ultralytics/yolov5: v7.0 - YOLOv5 SOTA Realtime Instance Segmentation. (2022). Zenodo. https://zenodo.org/records/7347926

Wamil, M., Goncalves, M., Rutherford, A., Borlotti, A., & Pellikka, P. A. (2022). Multi-modality cardiac imaging in the management of diabetic heart disease. *Frontiers in Cardiovascular Medicine*, 9. https://doi.org/10.3389/fcvm.2022.1043711

Wang, A., Chen, H., Liu, L., Chen, K., Lin, Z., Han, J., & Ding, G. (2024). YOLOv10: Real-Time End-to-End Object Detection. ArXiv.Org. https://doi.org/10.48550/arxiv.2405.14458

Wang, C.-Y., & Liao, H.-Y. M. (2024). YOLOv9: Learning What You Want to Learn Using Programmable Gradient Information. ArXiv Preprint ArXiv:2402.13616.

Yasmin, F., Shah, S. M. I., Naeem, A., Shujauddin, S. M., Jabeen, A., Kazmi, S., Siddiqui, S. A., Kumar, P., Salman, S., Hassan, S. A., Dasari, C., Choudhry, A. S., Mustafa, A., Chawla, S., & Lak, H. M. (2021). Artificial intelligence in the diagnosis and detection of heart failure: the past, present, and future. *Reviews in Cardiovascular Medicine*, 22(4), 1095. https://doi.org/10.31083/j.rcm2204121

Zakariah, M., & AlShalfan, K. (2020). Cardiovascular Disease Detection Using MRI Data with Deep Learning Approach. *International Journal of Computer and Electrical Engineering*, 12(2), 72–82. https://doi.org/10.17706/IJCEE.2020.12.2.72-82