Theory and Applications of
Mathematics & Computer Science

(ISSN 2067-2764)
http://www.uav.ro/applications/se/journal/index.php/tamcs

Theory and Applications of Mathematics & Computer Science 11(1) (2026) 39–50

Convolutional Neural Networks for Cardiac MRI Analysis: An Overview

Marcela-Florina Florea

Politehnica University of Timişoara, Romania

"Aurel Vlaicu" University of Arad, Romania

Abstract

Cardiac magnetic resonance imaging (cardiac MRI) provides detailed information about cardiac anatomy and function, but analysing a large number of images can be time-consuming when performed manually. Deep learning methods, particularly convolutional neural networks (CNNs), are widely used for automated medical image analysis. This paper provides an overview of CNNs and their applications in cardiac MRI analysis, with a focus on cardiac image segmentation. The paper first presents the basic principles of CNNs, including convolution, pooling and activation functions, followed by a discussion of U-Net and its use in biomedical image segmentation. Applications of CNNs in cardiac MRI, including cardiac structure segmentation, disease classification and quantitative assessment, are also discussed. Recent developments such as 2D and 3D CNNs, attention mechanisms and hybrid CNN-Transformer architectures are presented. The paper also discusses current challenges, including limited annotated datasets, differences between imaging settings, generalization, interpretability and clinical validation. Overall, CNNs remain important tools for cardiac MRI analysis, while recent research is exploring more robust models that can be applied across different clinical settings.

Keywords: Cardiac magnetic resonance imaging, Convolutional neural networks, Deep learning, Image segmentation, U-Net, Cardiac MRI, Medical image analysis, Computer vision
2020 MSC: 68T07, 68T45, 92C55

1. Introduction

Cardiac magnetic resonance imaging (cardiac MRI) is a medical imaging technique that can be used to evaluate the anatomy and function of the heart. Different MRI sequences can provide information about cardiac structures, motion and function. Cine MRI is particularly useful for assessing cardiac motion and ventricular function (Tao *et al.*, 2019; Chen *et al.*, 2020).

Analysing cardiac MRI images manually can be time-consuming and depends on the experience of the specialist. This can become difficult when a large number of images and cardiac structures need to be analysed. Manual segmentation can be used to obtain measurements such as ventricular volumes and ejection fraction, but the process requires considerable time and effort. Therefore, there is a need for automated methods (Chen *et al.*, 2020; Bernard *et al.*, 2018).

E-mail address: `marcela-florina.florea@student.upt.ro` (Marcela-Florina Florea)

Computer vision can be used to analyse medical images and identify relevant structures or patterns. Deep learning allows the model to learn relevant features directly from the images instead of relying on manually selected features. Convolutional neural networks (CNNs) are particularly suitable for image analysis because they can learn spatial features through convolutional operations. For this reason, CNNs have been widely used in medical imaging, including cardiac MRI analysis (LeCun *et al.*, 2015; Litjens *et al.*, 2019).

In cardiac MRI, CNN-based methods have been applied to several tasks, including cardiac structure segmentation, disease classification and quantitative assessment. Segmentation is particularly important because it can provide the anatomical information needed for further analysis, such as the calculation of ventricular volumes and cardiac function (Chen *et al.*, 2020; Bernard *et al.*, 2018).

This paper provides an overview of CNNs and their applications in cardiac MRI, with a focus on image segmentation and recent developments. The paper first presents the basic concepts of CNNs and their main components, followed by an overview of their applications in cardiac MRI. Recent developments, including U-Net variants, attention mechanisms and CNN-Transformer hybrid architectures, are also discussed. Finally, the paper briefly addresses current challenges related to generalization, explainability and clinical use.

2. Cardiac MRI

2.1. Cardiac Magnetic Resonance Imaging

Cardiac magnetic resonance imaging (cardiac MRI/CMR) is a non-invasive imaging technique that provides detailed images of cardiac anatomy and function. Different MRI sequences can provide different types of information. Cine MRI is particularly important for assessing cardiac motion and ventricular function (Tao *et al.*, 2019). Cardiac MRI images can be analysed manually, but this can be time-consuming, especially when many images need to be examined. These limitations have created an important motivation for the development of automated methods based on deep learning (Chen *et al.*, 2020; Bernard *et al.*, 2018).

2.2. Main Cardiac Structures

The heart has four chambers: the left atrium, right atrium, left ventricle (LV), and right ventricle (RV). The myocardium represents the muscular wall of the heart. For cardiac MRI segmentation, the ventricles and myocardium are particularly important, although other structures can also be analysed depending on the specific task (Chen *et al.*, 2020).

2.3. Why Automated Analysis Is Useful

Manual analysis of cardiac MRI images can be time-consuming and requires considerable expertise, especially when multiple images and cardiac structures need to be analysed. Automated methods can reduce the amount of manual work and can provide more consistent measurements when analysing large datasets.

CNN-based methods can automatically identify cardiac structures and generate segmentation maps from MRI images. These segmentations can then be used to obtain quantitative information such as ventricular volumes and ejection fraction. For example, Tao et al. showed that deep learning-based segmentation can be used for the automatic quantification of left ventricular function from cine MRI (Tao *et al.*, 2019).

CNNs can therefore be used for more than image segmentation. They can also support quantitative assessment and, in some cases, disease classification. The ACDC challenge demonstrated the potential of deep learning for both cardiac structure segmentation and pathology classification (Bernard *et al.*, 2018).

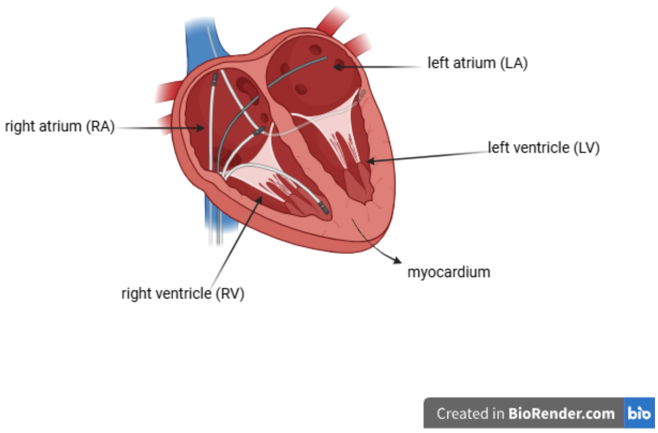

Figure 1. Main cardiac structures in a cardiac MRI image. Source: Created by the authors using BioRender.

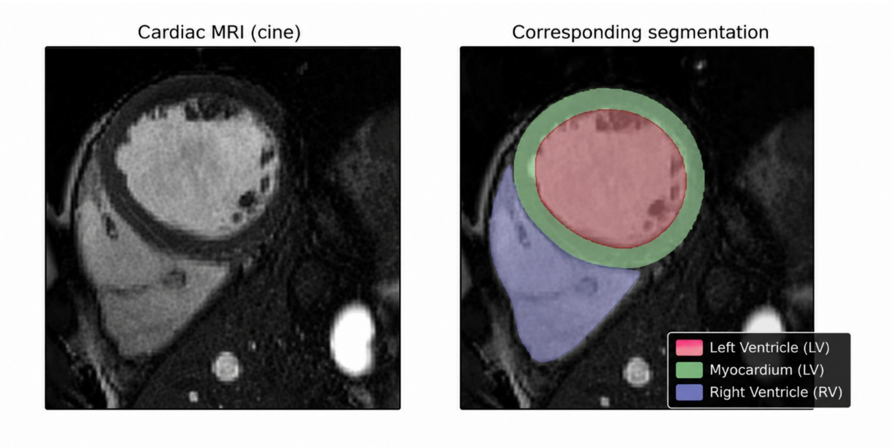

Figure 2: Example of a cardiac magnetic resonance image and corresponding segmentation of the left ventricle, left ventricular myocardium and right ventricle. Adapted and rearranged from Schilling et al. (Schilling *et al.*, 2024), licensed under CC BY 4.0.

3. Convolutional Neural Networks

3.1. Basic CNN Architecture

Deep learning relies on computational models composed of multiple processing layers that learn representations of data at different levels of abstraction. Convolutional neural networks (CNNs) are a class of deep learning architectures that have demonstrated effectiveness in processing spatially structured data, such as images and videos. Unlike traditional approaches that rely heavily on manually designed features, deep learning models can learn relevant representations directly from the input data (LeCun *et al.*, 2015).

CNNs exploit the local structure of images through convolutional operations, allowing the network to learn spatial features. A conventional CNN consists of an input layer, a sequence of convolutional and pooling operations, and an output stage that produces predictions according to the target task (Chen *et al.*, 2020).

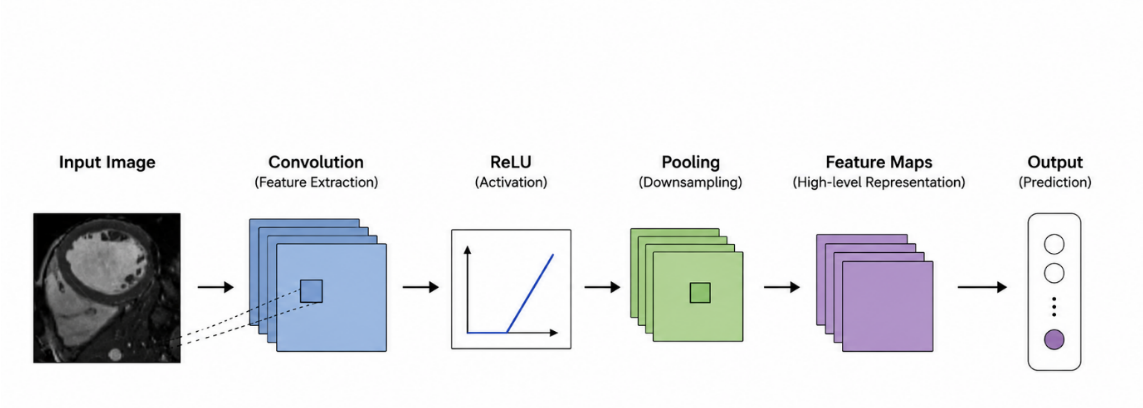

Figure 3: Simplified workflow of a convolutional neural network for image analysis. Source: Created by the authors using AI-assisted image generation.

Through successive convolutional layers, CNNs learn hierarchical representations of the input image. Each convolutional layer applies multiple learnable filters to the input and produces corresponding feature maps (Chen *et al.*, 2020).

3.2. Convolution Operation

The convolution operation is the fundamental operation of a CNN architecture. During this operation, a learnable filter (kernel) is moved across the input image or feature map. The values of the input region are multiplied by the corresponding filter weights and summed, producing a new value at each position in the output feature map (Chen *et al.*, 2020).

$$Y(i, j) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} X(i+m, j+n) K(m, n) \tag{1}$$

where X represents the input image or feature map, K the convolutional kernel of size $M \times N$, and Y the resulting feature map.

Different filters can learn to detect different visual patterns in the input. These patterns may correspond to simple features, such as edges and contours, in earlier layers, whereas filters in deeper convolutional layers can learn increasingly complex structures. The filters are learned automatically by the model during the training process from the available data (LeCun *et al.*, 2015; Chen *et al.*, 2020).

3.3. Pooling and Activation Functions

Pooling operations are used in CNNs to reduce the spatial dimensions of feature maps and retain the most relevant information. This process reduces the computational cost and can also contribute to more robust feature representations. Max pooling is a common pooling operation, where the maximum value within a defined local region is selected.

Activation functions introduce non-linearity into the network, allowing CNNs to learn more complex relationships from the input data. One example is the Rectified Linear Unit (ReLU), which is defined as $f(x) = \max(0, x)$ (Chen *et al.*, 2020).

3.4. CNNs for Image Segmentation

Image segmentation is the process of assigning a class label to individual pixels or voxels. In contrast, image classification assigns a single label to the entire image. This distinction is important in medical imaging, where the precise localization of anatomical structures is often required. CNNs can be adapted for segmentation by replacing the final classification layers with fully convolutional structures that preserve spatial information and produce pixel-wise predictions (Chen *et al.*, 2020).

CNN-based segmentation has been applied extensively in cardiac MRI to automatically delineate cardiac structures, including the left ventricle, right ventricle, and myocardium. These approaches need to combine high-level contextual information with precise spatial details to accurately identify anatomical boundaries. The ACDC challenge demonstrated the potential of deep learning methods for automated multi-structure cardiac MRI segmentation, achieving high agreement with expert-derived measurements (Bernard *et al.*, 2018). This has contributed to the development and adoption of architectures such as U-Net for biomedical image segmentation (Ronneberger *et al.*, 2015).

3.5. U-Net

U-Net is a convolutional neural network designed for biomedical image segmentation and was introduced by Ronneberger et al. The architecture follows an encoder-decoder pattern and is formed by two paths: a contracting path that extracts features from the input image and an expanding path that gradually restores the spatial resolution of the feature maps. The contracting path forms the encoder, and the expanding path forms the decoder. The decoder progressively restores the spatial resolution and produces the final segmentation map (Ronneberger *et al.*, 2015).

One of the key features of the U-Net architecture is the use of skip connections between the layers of the encoder and decoder. These connections transfer feature maps from the encoder to the decoder, helping to preserve spatial details. This process allows the network to combine contextual information with more precise spatial information when generating the final segmentation (Ronneberger *et al.*, 2015).

For example, U-Net can be used to segment the left ventricle in a cardiac MRI image. In the contracting path, the network reduces the spatial dimensions of the image and extracts features related to the ventricle and its surroundings. In the expanding path, the spatial resolution is gradually restored. The skip connections bring higher-resolution information from the earlier layers, which helps the network preserve the ventricular boundaries. The final output is a segmentation map in which the left ventricle is separated from the surrounding tissue.

U-Net has been widely applied to cardiac MRI segmentation, particularly for the segmentation of structures such as the left ventricle, right ventricle, and myocardium. Its relatively simple architecture and ability to produce detailed pixel-wise predictions have made it a common choice for cardiac MRI segmentation approaches (Chen *et al.*, 2020).

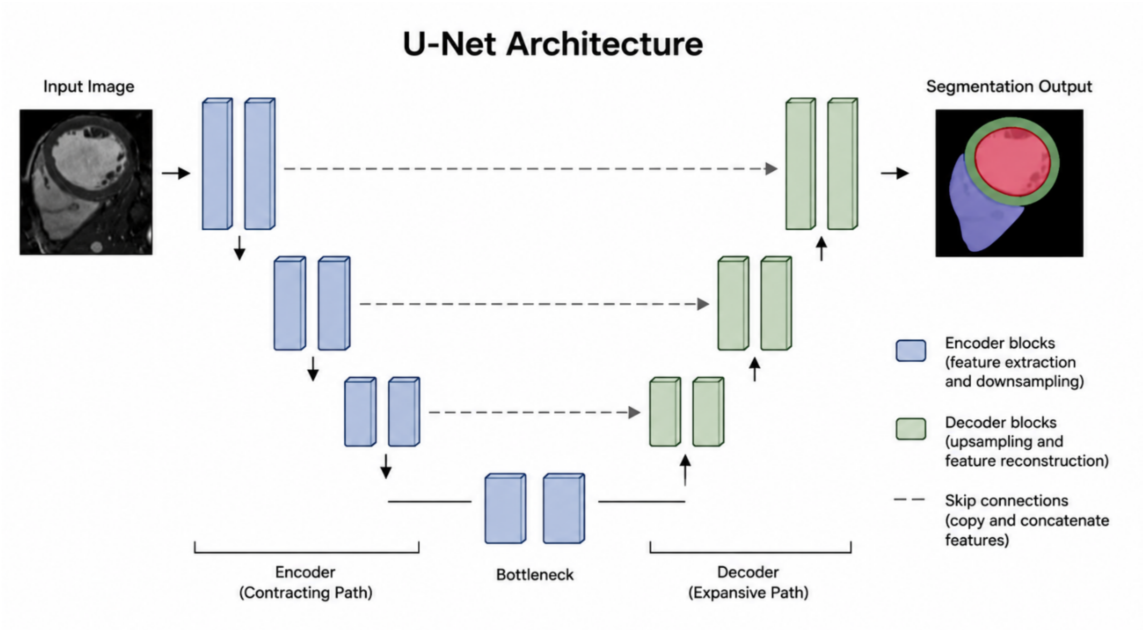

Figure 4: Simplified U-Net architecture showing the encoder, bottleneck, decoder and skip connections. Source: Created by the authors using AI-assisted image generation.

4. Applications of CNNs in Cardiac MRI

4.1. Cardiac Image Segmentation

Cardiac image segmentation is an important first step for many cardiac imaging applications. It is the process of dividing an image into anatomically meaningful regions. In cardiac MRI, the main structures of interest include the left ventricle (LV), right ventricle (RV), and myocardium. The resulting segmentation maps can also be used to extract quantitative measures such as ventricular volumes, myocardial mass, wall thickness, and ejection fraction (Chen *et al.*, 2020).

CNNs can learn the visual features required to distinguish different cardiac structures and generate segmentation maps. Fully convolutional networks (FCNs) and U-Net-based architectures are among the most used approaches for cardiac image segmentation (Chen *et al.*, 2020).

A well-known example is the Automated Cardiac Diagnosis Challenge (ACDC). The ACDC dataset contains 150 multi-equipment cardiac MRI recordings with reference measurements and expert annotations. The challenge focused on the automatic segmentation of the LV cavity, myocardium, and RV, as well as the classification of cardiac pathologies. The ACDC results showed that deep learning methods can support the automatic extraction of clinical indices and cardiac MRI analysis (Bernard *et al.*, 2018).

4.2. Cardiac Disease Classification

Classification is the process of assigning an image or examination to a specific class. CNNs can also be used to classify cardiac MRI images according to predefined pathological categories. CNNs can learn image features associated with different cardiac conditions and use these features to distinguish between different groups.

The ACDC challenge also included a classification task in which cardiac MRI examinations were divided into five groups: normal subjects, previous myocardial infarction, dilated cardiomyopathy, hypertrophic cardiomyopathy, and abnormal right ventricle. The results also demonstrated the potential of deep learning for automated cardiac pathology classification (Bernard *et al.*, 2018).

4.3. *Quantitative Assessment*

The segmentation of cardiac structures can also be used to obtain quantitative information about cardiac function. From the segmentation maps, parameters such as left ventricular end-diastolic volume (LVEDV), left ventricular end-systolic volume (LVESV), ejection fraction (EF), and myocardial mass can be calculated. These measurements are commonly used to assess cardiac function and can be obtained automatically from cardiac MRI images using deep learning-based segmentation methods (Chen *et al.*, 2020).

A study by Tao et al. used a CNN-based approach for the automatic quantification of left ventricular function from cine cardiac MRI. The study used data from four medical centres and three MRI vendors and showed that the use of more diverse training data can improve the robustness of automatic LV quantification across different imaging settings (Tao *et al.*, 2019).

Table 1. Representative Dice scores reported on major cardiac MRI segmentation benchmarks.

Dataset	**Structure**	**Dice**	**Evaluation context**
ACDC	LV	0.967 (ED), 0.928 (ES)	Official leaderboard
ACDC	RV	0.946 (ED), 0.904 (ES)	Official leaderboard
ACDC	Myocardium	0.896 (ED), 0.919 (ES)	Official leaderboard
M&Ms	LV + Myocardium + RV	0.8813 overall	Multi-centre, multi-vendor

Reported Dice scores are provided as reported by the respective benchmarks and should not be interpreted as directly comparable across challenges, since the datasets, evaluation protocols, and testing conditions differ.

4.4. *Other Applications*

CNNs and other deep learning approaches have also been applied to other cardiac MRI tasks beyond segmentation and classification. One important application is MRI reconstruction, where deep learning methods can be used to reconstruct images from undersampled data and potentially reduce acquisition time. Other applications include image enhancement, flow imaging, late gadolinium enhancement, and tissue characterization. These applications show that deep learning can be used at different stages of the cardiac MRI workflow, from image reconstruction to the analysis of cardiac structures and tissue (Oscanoa *et al.*, 2023).

5. Current State of Research

5.1. *CNN and U-Net Variants*

As CNNs became more common in image segmentation, research focused on adapting these architectures for more difficult segmentation problems. U-Net has become an important baseline for biomedical image segmentation, but researchers have developed different U-Net-based variants to improve feature extraction and segmentation performance (Chen *et al.*, 2020).

One direction has been to use pretrained CNN backbones, attention mechanisms and multiscale features within the U-Net framework. These modifications aim to improve the representation of relevant structures and retain the encoder-decoder structure of the original U-Net. Despite these developments, U-Net-based architectures remain important in cardiac MRI segmentation and continue to serve as a baseline for newer approaches (Aladwani *et al.*, 2026).

Table 2. Comparison of CNN architectures relevant to cardiac MRI analysis.

Architecture	Application	Advantages	Limitations	Computational cost
CNN	Classification, feature extraction	Effective feature learning	Limited spatial output	Lower than volumetric models; depends on depth and resolution
FCN	Segmentation	Pixel-wise prediction	Less precise spatial boundaries	Depends on feature-map resolution and depth
U-Net	Medical image segmentation	Preserves contextual and spatial information	Requires labelled data	Moderate; skip connections increase memory use
3D U-Net	Volumetric segmentation	Uses 3D spatial context	Higher memory requirements	High memory and computational requirements

5.2. 2D vs. 3D CNNs

CNNs can operate on 2D images or 3D volumes. In cardiac MRI, 2D CNNs process individual image slices, and 3D CNNs process a volume and can use information from neighbouring slices. Three-dimensional models can therefore capture additional spatial context, but they generally require more computational resources and memory.

In cardiac MRI, 3D approaches can provide additional spatial context, but this does not automatically lead to better performance. The characteristics of the acquisition, including slice thickness and spatial resolution, can influence the usefulness of 3D information (Baumgartner *et al.*, 2017). Therefore, a more complex architecture does not necessarily result in better performance. The choice between 2D and 3D approaches depends on the imaging data, available training data and computational resources (Baumgartner *et al.*, 2017; Aladwani *et al.*, 2026).

5.3. Attention Mechanisms

Attention mechanisms allow the network to give more importance to relevant features or regions instead of treating all information equally. In MRI images, boundaries can be difficult to identify because neighbouring structures may have similar appearances. Attention mechanisms can help the network focus on the most relevant information.

Sun et al. introduced Stack Attention U-Net (SAUN), which extends the U-Net model by introducing attention mechanisms and using information from neighbouring slices. This provides additional spatial context when segmenting cardiac structures. The method was evaluated on the ACDC dataset and showed good performance for left ventricle segmentation (Sun *et al.*, 2021).

Other studies have also combined attention mechanisms with pretrained CNN backbones, such as ResNet and DenseNet, in U-Net-based architectures. These approaches represent further modifications of the standard U-Net architecture aimed at improving feature extraction and segmentation performance. Attention mechanisms have therefore become one of the directions explored to improve CNN-based cardiac MRI segmentation (Das & Das, 2024; Aladwani *et al.*, 2026).

5.4. CNN + Transformer / Hybrid Architectures

CNNs are effective at extracting local spatial features, but their ability to model relationships between distant regions can be limited. Transformers can capture longer-range dependencies through attention mechanisms. Hybrid architectures have therefore been introduced to combine the local feature extraction of CNNs with the broader contextual information provided by Transformers.

An example is TF-Unet, which combines a U-Net-based CNN structure with Transformer components for cardiac MRI segmentation. The method was evaluated on the ACDC dataset and showed the potential of combining CNN and Transformer architectures (Fu *et al.*, 2022).

Research is also moving towards larger pretrained models that can potentially be adapted to more than one cardiac MRI task. Such models may support tasks such as segmentation, cardiac function assessment and disease classification, rather than requiring a separate model for each task. This represents a shift from task-specific CNNs towards more general-purpose models for cardiac MRI analysis (Aladwani *et al.*, 2026; Fu *et al.*, 2026).

5.5. Generalization

One of the biggest challenges in cardiac MRI research is that a model can perform very well on the dataset used for development but perform differently on data from another hospital or scanner. This can be influenced by various factors, such as different MRI vendors, acquisition protocols, imaging conditions, and patient populations (Tao *et al.*, 2019; Campello *et al.*, 2021).

The M&Ms challenge was created specifically to study this problem in cardiac MRI segmentation. The dataset included 375 cardiac MRI datasets acquired from four different scanner vendors, six hospitals, and three countries. The results showed the importance of testing models on data from unseen vendors and imaging protocols and highlighted the need for further work to improve generalization (Campello *et al.*, 2021). Recent research also identifies generalization and external validation as important requirements for the clinical use of deep learning in cardiac MRI (Aladwani *et al.*, 2026).

5.6. Explainability and Clinical Use

Good performance on a benchmark dataset does not automatically mean that a CNN model is ready to be used in clinical practice. The model also needs to be tested on external data and under different clinical conditions. Explainability is another important aspect, as clinicians may need to understand how the model reached a particular result. The model should also be reliable and fit into the existing clinical workflow. Recent studies highlight external validation, explainability and robustness as important aspects that need to be considered before deep learning models can be used in cardiac MRI (Aladwani *et al.*, 2026; van der Zande *et al.*, 2026).

6. Challenges and Future Directions

The availability of annotated data is one of the main challenges in cardiac MRI analysis. Creating large annotated datasets is difficult because annotation requires time and expert knowledge. Another challenge is that a model trained on one dataset may not perform in the same way on data from another centre or scanner. Model performance can also be influenced by factors such as MRI vendors, acquisition protocols, image resolution and patient populations. These issues highlight the importance of using larger and more diverse datasets from different centres (Chen *et al.*, 2020; Skandarani *et al.*, 2021; Campello *et al.*, 2021).

6.1. Class Imbalance and Limited Annotations

Another challenge is class imbalance in cardiac MRI segmentation. The regions of interest, such as the myocardium and cardiac chambers, occupy substantially smaller areas than the background in many images. As a result, a model may achieve good overall performance by correctly predicting the background while performing less accurately on smaller cardiac structures. This issue can affect the training process and may require appropriate loss functions or sampling strategies to improve the segmentation of smaller regions.

6.2. Foundation Models and Self-Supervised Learning

Foundation models and self-supervised learning represent promising directions for reducing the dependence on large manually annotated datasets. Self-supervised approaches can use large collections of unlabeled cardiac MRI images to learn general image representations before being adapted to specific downstream tasks. Foundation models may further support the transfer of learned representations across different cardiac MRI tasks and datasets. However, their effectiveness in cardiac MRI still depends on the availability and diversity of suitable pre-training data, as well as on their ability to generalize across scanners, vendors and acquisition protocols (Liu *et al.*, 2024; Shad *et al.*, 2026).

6.3. Federated Learning

Federated learning is another promising direction for cardiac MRI analysis, particularly when data are distributed across multiple hospitals. Instead of transferring patient images to a central location, models can be trained locally and their learned parameters or updates can be shared between participating institutions. This approach could facilitate the use of larger and more diverse datasets while reducing the need to directly exchange patient data. Nevertheless, differences between institutions, scanners and acquisition protocols remain important challenges for federated cardiac MRI analysis (Puyol-Anton *et al.*, 2022; He *et al.*, 2025).

Another important challenge is interpretability and clinical validation. Models can provide good results, but it can be difficult to understand why a model produces a particular prediction. This is important when the results may influence clinical decisions.

Moreover, good performance in a controlled environment does not automatically mean that a model is ready for clinical use. Models require external validation, robustness testing and integration into the clinical workflow (Aladwani *et al.*, 2026; van der Zande *et al.*, 2026).

Future research is likely to focus on larger and more diverse datasets, hybrid CNN-Transformer architectures and larger pretrained or foundation models. Improved cross-centre validation and explainable AI may further support the transition from research to clinical use. Recent large-scale models show the possibility of using one pretrained model for several cardiac MRI tasks, but these approaches still require further validation (Shad *et al.*, 2026; Fu *et al.*, 2026). The main goal is therefore not only to improve performance on benchmark datasets, but also to develop models that are reliable and useful in different clinical settings.

7. Conclusions

CNNs have become important tools for cardiac MRI analysis, especially for cardiac image segmentation. They can be used to automatically identify cardiac structures and provide information that can be further used for quantitative assessment and disease classification.

U-Net and its variants remain important for cardiac MRI segmentation and continue to provide a starting point for newer approaches.

Recent research has focused on improving these methods through 3D architectures, attention mechanisms and Transformer-based or hybrid architectures.

However, several challenges remain. The availability of annotated data, differences between datasets and MRI acquisition settings, generalization to new centres and scanners, and clinical validation are still important problems.

Overall, CNNs remain an important part of cardiac MRI analysis, while current research is moving towards models that are more robust and can be used across different clinical settings.

References

Aladwani, F., A. Perelli, I. Mordi and F. Khan (2026). Deep learning for cardiac mri: Performance evidence and barriers to clinical integration. a systematic review and meta-analysis. *European Heart Journal – Imaging Methods and Practice* **4**(1), qyag045.

Baumgartner, C. F., L. M. Koch, M. Pollefeys and E. Konukoglu (2017). An exploration of 2d and 3d deep learning techniques for cardiac mr image segmentation. In: *Statistical Atlases and Computational Models of the Heart – STACOM 2017*. Springer. pp. 111–119.

Bernard, O., A. Lalande, C. Zotti, F. Cervenansky, X. Yang, P.-A. Heng, I. Cetin, K. Lekadir, O. Camara, M. A. Gonzalez Ballester, G. Sanroma, S. Napel, S. Petersen, G. Tziritas, E. Grinias, M. Khened, V. A. Kollerathu, G. Krishnamurthi, M.-M. Rohe, X. Pennec, M. Sermesant, F. Isensee, P. Jager, K. Maier-Hein, P. M. Full, I. Wolf, S. Engelhardt, C. F. Baumgartner, L. M. Koch, J. M. Wolterink, I. Isgum, Y. Jang, Y. Hong, S. Patravali, S. Jain, O. Humbert and P.-M. Jodoin (2018). Deep learning techniques for automatic mri cardiac multi-structures segmentation and diagnosis: Is the problem solved?. *IEEE Transactions on Medical Imaging* **37**(11), 2514–2525.

Campello, V. M., P. Gkontra, C. Izquierdo, C. Martin-Isla, A. Sojoudi, P. M. Full, K. Maier-Hein, Y. Zhang, Z. He, J. Ma, M. Parreno, A. Albiol, F. Kong, S. C. Shadden, J. Corral Acero, V. Sundaresan, M. Saber, M. Elattar, H. Li, B. Menze, F. Khader, C. Haarburger, C. M. Scannell, M. Veta, A. Carscadden, K. Punithakumar, X. Liu, S. A. Tsaftaris, X. Huang, X. Yang, L. Li, X. Zhuang, D. Vilades, M. L. Descalzo, A. Guala, L. La Mura, M. G. Friedrich, R. Garg, J. Lebel, F. Henriques, M. Karakas, E. Cavus, S. E. Petersen, S. Escalera, S. Segui, J. F. Rodriguez-Palomares and K. Lekadir (2021). Multi-centre, multi-vendor and multi-disease cardiac segmentation: The m&ms challenge. *IEEE Transactions on Medical Imaging* **40**(12), 3543–3554.

Chen, C., C. Qin, H. Qiu, G. Tarroni, J. Duan, W. Bai and D. Rueckert (2020). Deep learning for cardiac image segmentation: A review. *Frontiers in Cardiovascular Medicine* **7**, 25.

Das, N. and S. Das (2024). Attention-unet architectures with pretrained backbones for multi-class cardiac mr image segmentation. *Current Problems in Cardiology* **49**(1), 102129.

Fu, Y., W. Bai, W. Yi, C. Manisty, A. N. Bhuva, T. A. Treibel, J. C. Moon, M. J. Clarkson, R. H. Davies and Y. Hu (2026). Development and validation of a versatile foundation model for cine cardiac magnetic resonance image analysis. *Communications Medicine* **6**, 408.

Fu, Z., J. Zhang, R. Luo, Y. Sun, D. Deng and L. Xia (2022). Tf-unet: An automatic cardiac mri image segmentation method. *Mathematical Biosciences and Engineering* **19**(5), 5207–5222.

He, Xiaoxiao, Haizhou Shi, Ligong Han, Chaowei Tan, Bo Liu, Zihao Xu, Meng Ye, Leon Axel, Kang Li and Dimitris N. Metaxas (2025). Rate-my-lora: Efficient and adaptive federated model tuning for cardiac mri segmentation. In: *2025 IEEE 22nd International Symposium on Biomedical Imaging (ISBI)*.

LeCun, Y., Y. Bengio and G. Hinton (2015). Deep learning. *Nature* **521**, 436–444.

Litjens, G., F. Ciompi, J. M. Wolterink, B. D. de Vos, T. Leiner, J. Teuwen and I. Isgum (2019). State-of-the-art deep learning in cardiovascular image analysis. *JACC: Cardiovascular Imaging* **12**(8), 1549–1565.

Liu, Zelong, Komal Kainth, Alexander Zhou, Timothy W. Deyer, Zahi A. Fayad, Hayit Greenspan and Xueyan Mei (2024). A review of self-supervised, generative, and few-shot deep learning methods for data-limited magnetic resonance imaging segmentation. *NMR in Biomedicine* **37**(8), e5143.

Oscanoa, J. A., M. J. Middione, C. Alkan, M. Yurt, M. Loecher, S. S. Vasanawala and D. B. Ennis (2023). Deep learning-based reconstruction for cardiac mri: A review. *Bioengineering* **10**(3), 334.

Puyol-Anton, Esther, Bram Ruijsink, Maria Mariscal et al. (2022). Federated learning for multi-center imaging diagnostics: a simulation study in cardiovascular disease. *Scientific Reports*.

Ronneberger, O., P. Fischer and T. Brox (2015). U-net: Convolutional networks for biomedical image segmentation. In: *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*. Springer. pp. 234–241.

Schilling, M., C. Unterberg-Buchwald, J. Lotz and M. Uecker (2024). Assessment of deep learning segmentation for real-time free-breathing cardiac magnetic resonance imaging at rest and under exercise stress. *Scientific Reports* **14**, 3754.

Shad, R., C. Zakka, D. Kaur, M. Mathur, R. Fong, J. Cho, R. W. Filice, J. Mongan, K. Kallianos, D. Eng, M. Leipzig, W. R. Witschey, A. de Feria, V. A. Ferrari, E. A. Ashley, M. A. Acker, C. Langlotz and W. Hiesinger (2026). A generalizable deep learning system for cardiac mri. *Nature Biomedical Engineering*.

Skandarani, Y., P.-M. Jodoin and A. Lalande (2021). Deep learning based cardiac mri segmentation: Do we need experts?. *Algorithms* **14**(7), 212.

Sun, X., P. Garg, S. Plein and R. J. van der Geest (2021). Saun: Stack attention u-net for left ventricle segmentation from cardiac cine magnetic resonance imaging. *Medical Physics* **48**(4), 1750–1763.

Tao, Q., W. Yan, Y. Wang, E. H. M. Paiman, D. P. Shamonin, P. Garg, S. Plein, L. Huang, L. Xia, M. Sramko, J. Tintera, A. de Roos, H. J. Lamb and R. J. van der Geest (2019). Deep learning-based method for fully automatic quantification of left ventricle function from cine mr images: A multivendor, multicenter study. *Radiology* **290**(1), 81–88.

van der Zande, J. L., L. Alvarez-Florez, R. H. J. A. Volleberg, C. Bras, D. Karkalousos, R. Nijveldt, N. van Royen, T. Leiner, N. Khalili, G. Litjens, J. Thannhauser and I. Isgum (2026). Deep learning for cardiac image analysis: Unveiling advances in deep learning architectures. *JACC: Cardiovascular Imaging*.