Generative AI role in Cardiac MRI Segmentation: A Comprehensive Review

Posted on July , 2025

Halima FOUADI [1,2,*], Mohamed KAS [2], Yassine RUICHEK [2], Youssef EL MERABET [1]

1 [SETIME Laboratory Faculty of Sciences Ibn Tofail University, Kenitra, Morocco](#)

2 [CIAD, UMR 7533, Univ. Bourgogne Franche-Comté, UTBM, Montbéliard, France](#)

* firstname.lastname@utbm.fr

Abstract: The integration of artificial intelligence (AI) has led to notable advancements in medical imaging. However, this progress is limited by the lack of expert-annotated data, particularly for rare pathologies, which hampers scientific research and the training of machine learning models. Generative artificial intelligence meets this need by synthesizing realistic images that correspond to ground-truth data, thereby increasing the data available for training and evaluating

Algorithms. In this review, we focus on the cardiac domain, and more specifically on magnetic resonance imaging (MRI), which is an important tool in the diagnosis of cardiac pathologies (the world's leading cause of death). We examine various generative AI methodologies, such as generative adversarial networks (GANs), variational autoencoders (VAEs), and diffusion models, applied to cardiac MRI data. Furthermore, we discuss the implications of these techniques in generating synthetic datasets, augmenting rare pathological cases, and improving segmentation accuracy and diagnostic outcomes. Finally, we highlight the challenges, limitations, and future directions of integrating generative AI into cardiac MRI workflows, aiming to guide further research and clinical translation.

2. Introduction

Cardiac magnetic resonance imaging is an important modality in the diagnosis, intervention and management of cardiovascular diseases which is one of the most common causes of death in the world according to the world health organization (Jafari et al., 2023) . By providing high-resolution images of the heart, accurate identification of anatomy, functions and tissues characterization, it can identify cardiac pathologies such as myocardial infarction, ischemic heart disease, cardiomyopathy, and congenital heart defects. The application of cardiac MRI is based on the segmentation of heart structures and regions of interest for analysis. However, manual segmentation is time-consuming and labor-intensive, making it

susceptible to inter-observer variability. This highlights the strong need for automated segmentation methods (Kanakatte et al., 2022).

The integration of artificial intelligence (AI) into the automation of cardiac segmentation has experienced significant advancements in recent years. In particular, the adoption of deep learning techniques such as recurrent neural networks (RNNs) and convolutional neural networks (CNNs), has revolutionized the field by achieving exceptional accuracy and efficiency. These methods excelled over traditional algorithms in effectively delineating complex anatomical structures and variations in cardiac magnetic resonance imaging (MRI) data (Taraboulsi et al., 2023). Nevertheless, these models face significant challenges mainly in terms of relying on a database with expert-annotation. Considering these limitations, generative artificial intelligence can meet the need by generating realistic synthetic data, thus increasing the diversity and quantity of available databases, while preserving pathology-related characteristics. By leveraging techniques such as (GANs), (VAEs), and diffusion models, generative AI can synthesize realistic cardiac MRI images with corresponding ground-truth annotations. These synthetic datasets have the potential to augment existing ones and enhance the training of segmentation algorithms. and improving diagnostic accuracy by enabling more robust and diverse model development (Al Khalil et al., 2023).

In this review, we aim to provide a comprehensive analysis of the role of generative AI in cardiac MRI segmentation. We begin by outlining the different techniques of generative AI followed by an exploration of state-of-the-art databases devoted to cardiac MRI semantic segmentation. Next, we discuss the existing works in the literature and their contributions to synthetic data. Finally, we examine the challenges and limitations of these approaches and propose future directions to guide the integration of generative AI into cardiac MRI workflows.

2. Background

Generative AI is an Artificial intelligence field that can generate realistic images, text and sounds by using deep learning algorithms that are trained on large amounts of data. Generative AI has seen tremendous growth in recent years and has been applied to a wide range of practical and creative fields, from art and entertainment to healthcare and engineering. We briefly introduce the different generative AI paradigms in the following.

- **Generative Adversarial Networks:**

Presented by Goodfellow et al. in 2014 (Goodfellow et al., 2014) ,are a novel class of deep learning techniques (a type of artificial intelligence algorithm). GANs consist of two models: a discriminator D, which is tasked with distinguishing between real and fake images, and a generator G, which learns to create realistic data through training. One type of GANs that is widely used in medical imaging is Pix2Pix GAN that is designed for image-to-image translation tasks. Pixel-to-pixel, indicating that the model operates on a pixel-level mapping between input and output images. The goal is to learn a mapping between an input and a corresponding output image. Pix2pix uses a conditional GAN architecture, where both discriminator and generator are conditioned on the input image. This adversarial training process allows the model to learn to generate high-quality image transformations (Isola et al., 2024).

- **Diffusion models:**

Diffusion model is a class of deep learning models used for generating high-quality images from text descriptions. The name comes from the idea of "diffusion" as a process of gradually transforming noise into a desired output, and "stable" reflects the model's ability to produce consistent high-quality results(Rombach et al., 2022). Thanks to this iterative denoising process, Stable Diffusion models reach higher quality than GANs. However, their application in medical imaging remains limited due to the scarcity of training data with text annotations and their high computational complexity.

- **VAE:**

Is a deep learning model designed to generate data similar to the ground truth by leveraging the principles of autoencoders. It consists of three main components: an encoder, a decoder, and a loss function. The goal of VAE is to learn both an encoder and a decoder that map data x to and from a continuous latent space z. The encoder receives an input image and reduces it to a more compact vector in latent space, capturing the essential features of the data. The decoder then processes this compressed vector to reconstruct it, transforming it back into a format that facilitates prediction of the output image. This process ensures that the data produced is very similar to the original data, while preserving the diversity of the results (Kusner et al., 2025), (Kingma and Welling, 2013).

3. Dataset

Annotated datasets play a crucial role in the training and evaluation of GAI models. In the context of CMRI datasets, it enables models to learn complex patterns and generate realistic, high-quality results. This section (Table 1) presents some of the most popular datasets available from CMRI.

Table 1 : Summary of available datasets of CMRI for semantic segmentation (Annex).

Reference	Dataset	Number of cases	Citations	Years
(Bernard et al., 2018)	ACDC	100 train 50 test	1848	2018
(Campello et al., 2021)	M&Ms	175 train 136 test		2020
(Perry et al., 2009)	SCD	45 Cine	445	2009
(Kadish et al., 2009)	LVSC	100 train 100 test	134	2011
(Petitjean et al., 2015)	RVSC	16 train 32 test	264	2015
(Andreopoulos et al., 2007)	York University	33 Cine	393	2008

4. **Applications of generative AI in cardiac MR imaging**

The state of the art in GAI applications for cardiac MR imaging can be classified into two main approaches: studies that focus exclusively on synthetic scan generation without pixel-wise semantic labels, and those that integrate image generation with segmentation. The following items briefly present the papers corresponding to each of these two approaches, highlighting the main contributions.

• **Unlabeled MRI Scan generation**

(Yoon et al., 2023): The Sequence-Aware Diffusion Model (SADM) was introduced for the generation of longitudinal medical images, such as cardiac and brain MRIs. This model learns to generate medical images from image sequences, considering their temporal order. In this way, it can synthesize the last image of a cardiac cycle from

the first image of that cycle. The model was evaluated on public cardiac MRI data, using the ACDC database.

(Kim and Ye, 2022): This study proposed a model for generating 4D cardiac cycle images, enabling the visualization of continuous anatomical changes. This model is particularly suited for generating 4D images of the cardiac cycle, allowing for continuous and progressive visualization of anatomical deformations throughout the cardiac cycle. This model relies on a structure similar to 3D UNet, with skip connections to preserve essential spatial information. This architecture helps generate high-quality volumetric images. It includes a Deformation Module based on VoxelMorph-1that generates deformation fields in 3D images. This module enables smooth deformation between the different phases of the cardiac cycle. Scan to scan without segmentation.

(Campello et al., 2022): This study presented a Conditional GAN (cGAN) for synthesizing heart scans of different ages using only cross-sectional data. The used cGAN architecture is based on a U-Net architecture with residual blocks and attention mechanisms. The model is conditioned by age and body mass index (BMI) to adjust images according to these covariates. A Wasserstein-GAN algorithm with gradient penalty (WGAN-GP) is used to stabilize training.

- **Labled**

(Ossenberg-Engels and Grau, 2020) : The authors proposed a Conditional Generative Adversarial Network to predict cardiac deformation between end-diastolic (ED) and end-systolic (ES) frames. Using the UK Biobank dataset, their model learned a deterministic mapping between ED and ES short-axis frames, enabling the modeling of cardiac sequences and the functional behavior of the heart. This learning helped to increase the data by transforming the scans from each phase to the other one respecting their corresponding semantic labels.

(Al Khalil et al., 2022): This framework, trained on the M&Ms dataset, focuses on right ventricle segmentation and integrates three key components: Detection of the region of interest (ROI) by cropping the image to center the heart within the field of view (FOV), image synthesis through the application of a mask-conditional GAN that learns the mapping from segmentation labels to corresponding realistic images. The application of random elastic deformation, morphological dilation, and erosion to the labels to generate anatomical variations of the heart, including pathological cases. Finally, a modified U-Net network was proposed to enhance cardiac

segmentation through the integration of both real and synthetic images.

(Al Khalil et al., 2023): The authors proposed a conditional synthesis approach using GANs to generate realistic cardiac MRI images in the short-axis view. This study is based on three main steps: image synthesis, a conditional synthesis approach based on GANs is used to generate realistic cardiac MRI images in short-axis view. The quality of these images is enhanced using labels of different tissues surrounding the heart, generated by a multi-tissue segmentation network trained on simulated XCAT-based images. This strategy helped the GAN to generate coherent MRI scans of the heart and its surroundings. The next steps of their framework consist of region of interest (ROI) detection and heart chamber segmentation using the generated images to train a convolutional neural network (CNN), based on a U-Net architecture, for heart chamber segmentation (right ventricle, left ventricle and myocardium).

(Diller et al., 2020): This study utilized cardiac MRIs from patients with Tetralogy of Fallot to develop and compare segmentation models. Progressive GANs (PG- GANs) are trained, on the collected data sourced from 14 German centers, to generate synthetic MRI frames. The synthetized frames were manually segmented to create training data for a U-Net based segmentation model. To evaluate the quality of synthetic data a random selection of 200 PG-GAN-generated images and 200 original MRI images was submitted to human investigators who had to identify the PG-GAN-generated image which reflected their realism.

(Amirrajab et al., 2022): This study introduced a two-module framework for generating high-fidelity cardiac MR images. The first module utilizes a U-Net model for multi-tissue segmentation of cardiac MR images. The output of this module is a segmentation mask that labels various tissues, including the myocardium (MYO), right ventricle (RV), and left ventricle (LV). These segmentation masks serve as input labels for generating new images using a cGAN trained on M&Ms dataset, which produces realistic cardiac MR images based on the anatomical structures encoded in the segmentation masks. The simulated anatomies of virtual subjects are derived from the 4D XCAT phantoms, and the images are simulated through a physics-based simulation tool that implements the Bloch equations for cine studies.

(Kim and Ye, 2022): DiffuseMorph is an unsupervised model for deformable image registration using diffusion models. Image registration aims to align multiple images taken from different angles or at different time points by deforming them to match a reference image or atlas. DiffuseMorph achieves deformable image registration in

an unsupervised manner by utilizing a diffusion model. The training is based on ACDC benchmarks. These data were resampled, normalized, and cropped to fit the model.

(Amirrajab et al., 2020): This study proposed a method with two different configurations one using only the ground truth annotations available for the heart and another increasing the number of labels into 8 classes encompassing the organs surrounding the heart when training the XCAT-GAN model. Their pipeline is composed of three cascaded models: (1) a modified version of UNet that predicts multi-tissue segmentation maps from real images, used only in 8-class image synthesis. (2) a conditional GAN architecture trained on pairs of real images and label maps (4 or 8 classes) to generate synthetic images based on XCAT labels. (3) an adapted version of U-Net in 2D, used to evaluate new synthetic images and their corresponding labels in various experimental scenarios including only valid ones in the augmented training dataset.

(Abbasi-Sureshjani et al., 2020): The authors proposed a GAN-based approach for synthesizing 4D (3D+t) cardiac MR images, using the 4D XCAT model as ground truth. To preserve the spatial and semantic information of the reference anatomy, they used the SPADE (Semantic Image Synthesis with Spatially Adaptive Normalization) model originally proposed for semantic controlled generation. For training, they used images from the ACDC database with their corresponding segmentation masks. During inference, they replace the segmentation masks with voxelated 4D labels from the XCAT to generate new 4D MRI images.

(Lustermans et al., 2022): This work aimed to improve cardiac MRI scan segmentation with late contrast (LGE), particularly in contexts with limited data sets. The first approach involves dividing the segmentation task into simpler sub-problems, and the second relies on the use of synthetic data to increase the amount of data available. A cascade pipeline method has been proposed, comprising three deep-based blocks. The first identifies the left ventricle, the second delineates the left ventricular myocardium, and the third segments the regions of myocardial infarction. The segmentation-conditioned synthetic data generator (using a GAN) was used to augment the training data. The study also showed that augmentation by synthetic data improves scar segmentation, particularly in challenging datasets with noise and artifacts.

(Skandarani et al., 2020): This paper proposed a model to produce highly realistic MRI images (100k) with pixel-accurate ground truth for cardiac segmentation in

cine-MR combining Variational Autoencoder (VAE) with SPADE-GAN. VAE network is trained to learn the latent representations of cardiac shapes, enabling the model to capture the variations in heart shapes across individuals. On the other hand, SPADE-GAN generates realistic MR images based on an anatomical map input. The GAN learns to generate images whose cardiac structures align with the shapes generated by the VAE.

Figure 1 illustrates a mapping of the reviewed literature works based on their ability to generate labelled scans and also depending on the required inputs (Scan, Scan+Label, Label).
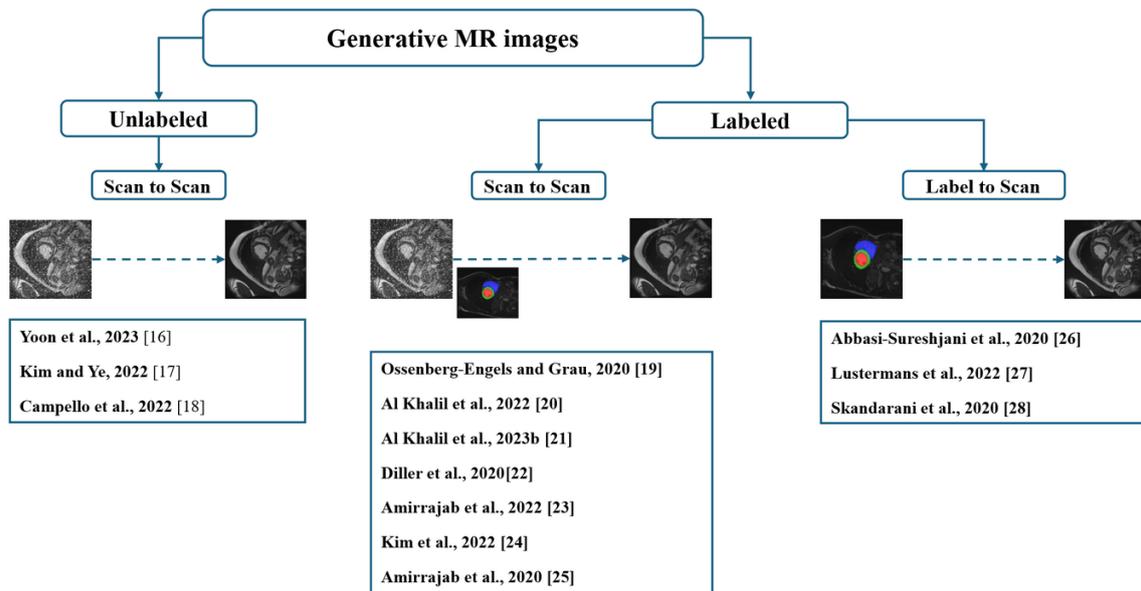
Generative AI in cardiac MRI

Taxonomy of existing works in cardiac MRI segmentation

5.Opening

Generative AI has enabled significant advances in the improvement of medical image databases, particularly for rare or difficult-to-annotate cases, such as cardiac MRI. However, beyond the generation of realistic and diverse images to train

segmentation models, many other applications are possible in this field, notably decision-making through image classification, as well as temporal image synthesis to study the evolution of cardiac pathologies over time, an important area in the monitoring of patients with chronic cardiovascular diseases.

6.Conclusion:

The integration of generative AI into cardiac imaging, particularly for MRI segmentation, represents an essential lever for cardiovascular management. Techniques such as GANs, VAEs and diffusion models have demonstrated their potential to generate realistic images and increase the diversity of training data, while improving the accuracy of segmentation models. However, several challenges remain, particularly regarding the quality of the images generated, their generalizability to varied clinical populations and the need for high-quality annotated data. As technology continues to advance, these models could not only enrich available databases but also improve diagnostic and clinical outcomes. The future of GAI in cardiac MRI lies in better clinical integration, with particular attention to model validation and adaptation to specific patient needs.

Acknowledgements

References :

10. Andreopoulos and J. K. Tsotsos, "Efficient and generalizable statistical models of shape and appearance for analysis of cardiac MRI," Med Image Anal, vol. 12, no. 3, pp. 335–357, Jun. 2008, doi: 10.1016/j.media.2007.12.003.
11. H. Kadish et al., "Rationale and Design for the Defibrillators To Reduce Risk By Magnetic Resonance Imaging Evaluation (DETERMINE) Trial," J Cardiovasc Electrophysiol, vol. 20, no. 9, pp. 982–987, Sep. 2009, doi: 10.1111/j.1540-8167.2009.01503.x.
12. Kanakatte, D. Bhatia, and A. Ghose, "3D Cardiac Substructures Segmentation from CMRI using Generative Adversarial Network (GAN)," in 2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), Jul. 2022, pp. 1698–1701. doi: 10.1109/EMBC48229.2022.9871950.
13. Kim and J. C. Ye, "Diffusion Deformable Model for 4D Temporal Medical Image

Generation," in Medical Image Computing and Computer Assisted Intervention – MICCAI 2022, L. Wang, Q. Dou, P. T. Fletcher, S. Speidel, and S. Li, Eds., Cham: Springer Nature Switzerland, 2022, pp. 539–548. doi: 10.1007/978-3-031-16431-6_51.

14. Kim, I. Han, and J. C. Ye, "DiffuseMorph: Unsupervised Deformable Image Registration Using Diffusion Model," Sep. 29, 2022, arXiv: arXiv:2112.05149. doi: 10.48550/arXiv.2112.05149.

15. Petitjean et al., "Right ventricle segmentation from cardiac MRI: A collation study," Medical Image Analysis, vol. 19, no. 1, pp. 187–202, Jan. 2015, doi: 10.1016/j.media.2014.10.004.

16. R. P. R. M. Lustermans, S. Amirrajab, M. Veta, M. Breeuwer, and C. M. Scannell, "Optimized automated cardiac MR scar quantification with GAN-based data augmentation," Computer Methods and Programs in Biomedicine, vol. 226, p. 107116, Nov. 2022, doi: 10.1016/j.cmpb.2022.107116.

G.-P. Diller et al., "Utility of deep learning networks for the generation of artificial cardiac magnetic resonance images in congenital heart disease," BMC Med Imaging, vol. 20, p. 113, Oct. 2020, doi: 10.1186/s12880-020-00511-1.

2014. Goodfellow et al., "Generative Adversarial Nets," in Advances in Neural Information Processing Systems, Curran Associates, Inc., 2014. Accessed: Dec. 23, 2024. [Online]. Available: https://proceedings.neurips.cc/paper_files/paper/2014/hash/5ca3e9b122f61f8f06494c97b1afccf3-Abstract.html

2015. El-Taraboulsi, C. P. Cabrera, C. Roney, and N. Aung, "Deep neural network architectures for cardiac image segmentation," Artificial Intelligence in the Life Sciences, vol. 4, p. 100083, Dec. 2023, doi: 10.1016/j.ailsci.2023.100083.

2016. Ossenberg-Engels and V. Grau, "Conditional Generative Adversarial Networks for the Prediction of Cardiac Contraction from Individual Frames," in Statistical Atlases and Computational Models of the Heart. Multi-Sequence CMR Segmentation, CRT-EPiggy and LV Full Quantification Challenges, M. Pop, M. Sermesant, O. Camara, X. Zhuang, S. Li, A. Young, T. Mansi, and A. Suinesiaputra, Eds., Cham: Springer International Publishing, 2020, pp. 109–118. doi: 10.1007/978-3-030-39074-7_12.

2017. S. Yoon, C. Zhang, H.-I. Suk, J. Guo, and X. Li, "SADM: Sequence-Aware Diffusion Model for Longitudinal Medical Image Generation," vol. 13939, 2023, pp. 388–400. doi: 10.1007/978-3-031-34048-2_30.

Kingma, D. P., & Welling, M. (2013). Auto-encoding variational bayes. arXiv preprint

arXiv:1312.6114.

1954. J. Kusner, B. Paige, and J. M. Hernández-Lobato, "Grammar Variational Autoencoder," in Proceedings of the 34th International Conference on Machine Learning, PMLR, Jul. 2017, pp. 1945–1954. Accessed: Jan. 02, 2025. [Online]. Available: https://proceedings.mlr.press/v70/kusner17a.html

1955. Jafari et al., "Automated diagnosis of cardiovascular diseases from cardiac magnetic resonance imaging using deep learning models: A review," Computers in Biology and Medicine, vol. 160, p. 106998, Jun. 2023, doi: 10.1016/j.compbiomed.2023.106998.

1956. Bernard et al., "Deep Learning Techniques for Automatic MRI Cardiac Multi-Structures Segmentation and Diagnosis: Is the Problem Solved?," IEEE Transactions on Medical Imaging, vol. 37, no. 11, pp. 2514–2525, Nov. 2018, doi: 10.1109/TMI.2018.2837502.

1957. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, "Image-To-Image Translation With Conditional Adversarial Networks," presented at the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 1125–1134. Accessed: Dec. 24, 2024. [Online]. Available: https://openaccess.thecvf.com/content_cvpr_2017/html/Isola_Image-To-Image_Translation_With_CVPR_2017_paper.html

1958. Perry, L. Yingli, C. Kim, P. Gideon, A. J. Dick, and G. A. Wright, "Evaluation Framework for Algorithms Segmenting Short Axis Cardiac MRI.," The MIDAS Journal, Jul. 2009, doi: 10.54294/g80ruo.

1959. Rombach, A. Blattmann, D. Lorenz, P. Esser, and B. Ommer, "High-Resolution Image Synthesis With Latent Diffusion Models," presented at the Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022, pp. 10684–10695. Accessed: Dec. 24, 2024. [Online]. Available: https://openaccess.thecvf.com/content/CVPR2022/html/Rombach_High-Resolution_Image_Synthesis_With_Latent_Diffusion_Models_CVPR_2022_paper.html

1960. Abbasi-Sureshjani, S. Amirrajab, C. Lorenz, J. Weese, J. Pluim, and M. Breeuwer, "4D Semantic Cardiac Magnetic Resonance Image Synthesis on XCAT Anatomical Model," May 20, 2020, arXiv: arXiv:2002.07089. doi: 10.48550/arXiv.2002.07089.

1961. Amirrajab et al., "XCAT-GAN for Synthesizing 3D Consistent Labeled Cardiac MR Images on Anatomically Variable XCAT Phantoms," Jul. 31, 2020, arXiv: arXiv:2007.13408. doi: 10.48550/arXiv.2007.13408.

1962. Amirrajab, Y. Al Khalil, C. Lorenz, J. Weese, J. Pluim, and M. Breeuwer, "Label-informed cardiac magnetic resonance image synthesis through conditional

generative adversarial networks," Computerized Medical Imaging and Graphics, vol. 101, p. 102123, Oct. 2022, doi: 10.1016/j.compmedimag.2022.102123.

1963. M. Campello et al., "Cardiac aging synthesis from cross-sectional data with conditional generative adversarial networks," Front. Cardiovasc. Med., vol. 9, Sep. 2022, doi: 10.3389/fcvm.2022.983091.

1964. M. Campello et al., "Multi-Centre, Multi-Vendor and Multi-Disease Cardiac Segmentation: The M&Ms Challenge," IEEE Transactions on Medical Imaging, vol. 40, no. 12, pp. 3543–3554, Dec. 2021, doi: 10.1109/TMI.2021.3090082.

1965. Al Khalil, S. Amirrajab, C. Lorenz, J. Weese, J. Pluim, and M. Breeuwer, "On the usability of synthetic data for improving the robustness of deep learning-based segmentation of cardiac magnetic resonance images," Medical Image Analysis, vol. 84, p. 102688, Feb. 2023, doi: 10.1016/j.media.2022.102688.

1966. Al Khalil, S. Amirrajab, C. Lorenz, J. Weese, J. Pluim, and M. Breeuwer, "On the usability of synthetic data for improving the robustness of deep learning-based segmentation of cardiac magnetic resonance images," Medical Image Analysis, vol. 84, p. 102688, Feb. 2023, doi: 10.1016/j.media.2022.102688.

1967. Al Khalil, S. Amirrajab, J. Pluim, and M. Breeuwer, "Late Fusion U-Net with GAN-Based Augmentation for Generalizable Cardiac MRI Segmentation," in Statistical Atlases and Computational Models of the Heart. Multi-Disease, Multi-View, and Multi-Center Right Ventricular Segmentation in Cardiac MRI Challenge, vol. 13131, E. Puyol Antón, M. Pop, C. Martín-Isla, M. Sermesant, A. Suinesiaputra, O. Camara, K. Lekadir, and A. Young, Eds., in Lecture Notes in Computer Science, vol. 13131. , Cham: Springer International Publishing, 2022, pp. 360–373. doi: 10.1007/978-3-030-93722-5_39.

1968. Skandarani, N. Painchaud, P.-M. Jodoin, and A. Lalande, "On the effectiveness of GAN generated cardiac MRIs for segmentation," May 22, 2020, arXiv: arXiv:2005.09026. doi: 10.48550/arXiv.2005.09026.