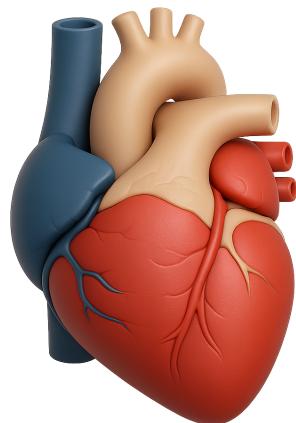


Master Project in Robotics

Enhancing Diagnostic Insights with AI-Assisted Segmentation of Whole-Heart MR Images

Gabriel Paffi



under the supervision of :

Direction :

Prof. Dimitrios Karampinos
Associate Professor
EPFL

Supervisor:

Dr. Ruud VAN HEESWIJK,
PhD, Senior Lecturer
CHUV-UNIL

External expert:

Dr. Augustin OGIER,
PhD Postdoctoral Fellow
CHUV-UNIL



École Polytechnique Fédérale de Lausanne (EPFL) 1015 Lausanne, Switzerland

Acknowledgements

I am especially grateful to Senior Lecturer Ruud van Heeswijk for his trust in assigning me this project and for his valuable advice throughout the semester. I would also like to sincerely thank Professor Dimitrios Karampinos for his attentive supervision, and Professor Matthias Stuber for warmly welcoming me into the lab. I am deeply thankful to Dr. Augustin Ogier for his supervision, guidance, feedback, and insightful comments during the project, as well as for patiently answering my questions, which allowed me to discover part of the medical MR world. I further wish to acknowledge Jean-Baptiste Ledoux for his technical assistance and support at the scanner, Jérôme Yerli for his explanations on k-space, and Kevin Borsos for his help and commitment in the scanner deployment. Finally, I would like to thank the entire laboratory team for sharing their knowledge and for the many enjoyable moments together.

Contents

1	Introduction	5
1.1	Cardiac Anatomy	5
1.2	Medical Imaging Modalities for Cardiac Assessment	6
1.3	Principles and Limitations of Conventional Cardiac MRI	6
1.4	Free-Running	7
1.5	Cardiac MRI Segmentation	8
1.6	Aim	8
2	State of the art	9
2.1	Cardiac MRI Segmentation Techniques	9
2.2	Deep Learning Approaches in Medical Imaging	9
2.3	Applications to High-Dimensional (4D/5D) Cardiac Data	10
2.4	Deep Model Robustness to Image Quality and Reconstruction Variants	10
2.5	Toward Real-Time Clinical Deployment on MRI Systems	10
3	Methods	11
3.1	Study Population and Acquisition Protocol	11
3.2	5D Image Reconstruction Techniques	13
3.3	Ground Truth Segmentations	15
3.4	Deep Learning-Based Segmentation Model	16
3.5	Evaluation Criteria for Segmentation	16
4	Results	18
4.1	Qualitative Segmentation Results	18
4.2	Geometric metrics	20
4.3	Clinical metrics	23
4.4	Physiological Consistency metrics	26
5	Discussion	28
6	On scanner deployment	29
6.1	Communication Protocol and Configurations	30
6.2	Image Processing Workflow	32
6.3	Image Transmission to Console	32
6.4	Modular Package for On-Scanner Deployment	34
7	Conclusion	36
8	Additional contributions	37
8.1	Density Compensation	37
8.2	Extraction of 4CH and SAX Views from Segmented MRI Reconstructions	40
8.3	Principal Component Analysis for Cardiac View Orientation	40
8.4	Conference and Retreat Presentations	41

Abstract

This project demonstrates that deep learning-based segmentation applied to fast gridded reconstructions on 5D free-running cardiac MRI achieves geometric performance comparable to compressed sensing reconstructions (DSC: 0.93 ± 0.01 for Gridded-FR_A vs. 0.94 ± 0.01 for CS-FR_A in the left ventricle) and provides physiologically consistent results for the myocardium (volume mismatch: 4.96% for Gridded-FR_A, 3.96% for CS-FR_A, compared to 5.57% for manual reference). Additionally, the reconstruction and segmentation pipeline can be deployed directly on the scanner, delivering results within 5 minutes after data acquisition, thereby enabling near real-time integration into clinical workflows.

Abbreviations

2CH	Two-Chamber View
3CH	Three-Chamber View
4CH	Four-Chamber View
AI	Artificial Intelligence
bSSFP	Balanced Steady-State Free Precession
CHUV	Centre Hospitalier Universitaire Vaudois
CMRI	Cardiac Magnetic Resonance Imaging
CNN	Convolutional Neural Network
CT	Computed Tomography
CVD	Cardiovascular Disease
DSC	Dice Similarity Coefficient
ECG	Electrocardiogram
EDV	End-Diastolic Volume
EF	Ejection Fraction
ESV	End-Systolic Volume
FOV	Field of View
FR	Free-Running
fNav	Self-Navigation with Respiratory Motion Correction
GT	Ground Truth
HD	Hausdorff Distance
HFpEF	Heart Failure with Preserved Ejection Fraction
HFrEF	Heart Failure with Reduced Ejection Fraction
ICC	Intraclass Correlation Coefficient
ISMRMRD	International Society for Magnetic Resonance in Medicine Raw Data format
k-space	Spatial Frequency Domain in MRI
LV	Left Ventricle
LVB	Left Ventricle Blood Pool
LVM	Left Ventricle Myocardium
MRI	Magnetic Resonance Imaging
MVC	Myocardial Volume Consistency
NUFFT	Non-Uniform Fast Fourier Transform
PCA	Principal Component Analysis
RF	Radiofrequency
ROI	Region of Interest
RV	Right Ventricle
RVB	Right Ventricle Blood Pool
RVD	Relative Volume Difference
SAX	Short-Axis View
SEM	Standard Error of Measurement
SV	Stroke Volume
TCP/IP	Transmission Control Protocol / Internet Protocol
TE	Echo Time
TR	Repetition Time
UNIL	Université de Lausanne
XML	Extensible Markup Language

1 Introduction

Among the wide range of diseases affecting humans, those involving the heart and circulatory system are particularly concerning. Cardiovascular diseases (CVD) are currently the leading cause of death worldwide, with the notable exception of Africa. In 2019, they accounted for approximately 17.9 million deaths, representing 32.1% of global mortality, up from 12.3 million in 1990 [1].

The two main causes are coronary artery disease and stroke, which together are responsible for about 80% of CVD-related deaths in men and 75% in women. While these conditions primarily affect older adults, early detection remains crucial to improve care and reduce their overall impact.

1.1 Cardiac Anatomy

Understanding the impact of cardiovascular diseases requires a clear overview of the heart's anatomical structure.

The human heart is a muscular organ responsible for pumping blood throughout the body via the circulatory system. It is composed of four chambers: two upper chambers called atria (left and right atrium) and two lower chambers called ventricles (left and right ventricle). The right side of the heart receives deoxygenated blood from the body and pumps it into the lungs through the pulmonary artery, where gas exchange occurs. Oxygenated blood returns to the left side of the heart and is then pumped to the rest of the body through the aorta.

The walls of the heart are primarily made of cardiac muscle tissue known as the myocardium, which is responsible for the contractile force needed to pump blood effectively. The myocardium is particularly thick in the left ventricle, reflecting the higher pressure required to distribute blood throughout the systemic circulation.

The chambers are separated by valves that ensure unidirectional blood flow: the tricuspid valve between the right atrium and ventricle, the pulmonary valve between the right ventricle and pulmonary artery, the mitral valve between the left atrium and ventricle, and the aortic valve between the left ventricle and the aorta. This complex structure allows for the precise coordination necessary to maintain efficient circulation (Figure 1).

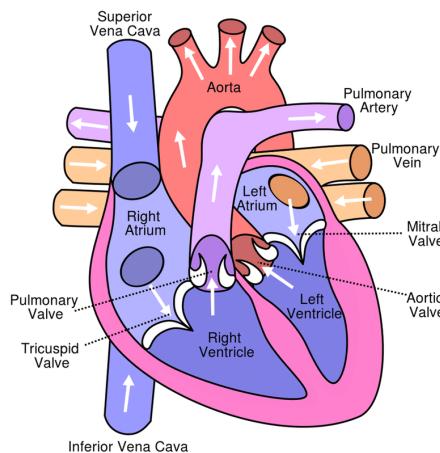


Figure 1: Illustrated heart's structure

1.2 Medical Imaging Modalities for Cardiac Assessment

Given the complexity of the heart's anatomical and functional structure, accurate visualization is essential for diagnosis and treatment planning. Non or minimal-invasive medical imaging techniques such as echocardiography, magnetic resonance imaging (MRI), computed tomography (CT), and nuclear medicine imaging are widely used to assess cardiac morphology and function.

Echocardiography, which uses ultrasound waves, is one of the most common and accessible tools for real-time cardiac evaluation. It provides dynamic information about heart wall motion, valve function, and blood flow.

Computed tomography (CT), particularly with contrast enhancement, offers high-resolution images of the coronary arteries and surrounding vasculature, making it valuable in detecting coronary artery disease. However, it involves exposure to ionizing radiation.

Cardiac magnetic resonance imaging (MRI) is considered the gold standard for soft tissue characterization and functional analysis. It enables high-resolution 3D imaging of cardiac structures, myocardial tissue properties, and blood flow, without radiation.

In this work, we focus specifically on cardiac MRI, as it offers superior image quality and the potential for quantitative analysis of both structure and function. While echocardiography remains the most widely used imaging modality in clinical practice due to its portability, low cost, and rapid acquisition, MRI provides a more comprehensive and detailed view of the heart. Our objective is to explore advanced MRI techniques that can further enhance image quality and extract meaningful quantitative metrics, thereby contributing to improved diagnosis and patient care.

1.3 Principles and Limitations of Conventional Cardiac MRI

Magnetic Resonance Imaging (MRI) is a non-invasive imaging technique that relies on the quantum properties of hydrogen atoms, which are abundant in the human body. More specifically, MRI exploits a quantum characteristic called spin, an intrinsic form of angular momentum carried by subatomic particles such as protons. Due to their spin, hydrogen nuclei behave like tiny bar magnets [2].

When placed in a strong external magnetic field generated by the MRI scanner, these magnetic moments align along the direction of the field. A radiofrequency (RF) pulse is then applied at a specific resonance frequency, temporarily disturbing this alignment. As the protons return to their equilibrium state, they emit RF signals that are detected by receiver coils positioned around the body [3].

The raw signals captured are not directly interpretable as images. Instead, they are stored in what is called k-space, a spatial frequency domain that encodes the information required to reconstruct the final image. The transformation from k-space to image space is performed using the inverse Fourier transform, enabling the generation of high-resolution anatomical images.

While MRI provides detailed anatomical images based on fundamental physical principles, conventional cardiac MRI protocols face several practical challenges in clinical application. Standard imaging often involves 2D cine acquisitions, with thick slices (e.g., 6–8 mm) and inter-slice gaps (e.g., 2 mm). These acquisitions are ECG-gated and require multiple breath holds one for each slice to minimize respiratory

motion. This not only lengthens the scan time but also can be challenging for the patient, especially elderly individuals or those with cardiac or respiratory difficulties.

Furthermore, the dependence on ECG synchronization and precise breath-holding introduces complexity in both scan setup and execution, requiring trained personnel and additional equipment. It may also reduce patient comfort and create a more intrusive examination environment, affecting compliance and image quality.

To overcome the limitations related to spatial coverage and inter-slice inconsistency, 3D acquisition methods such as balanced steady-state free precession (bSSFP) have been introduced. These sequences capture the entire cardiac volume in a single scan without inter-slice gaps, improving anatomical coherence. However, they typically acquire only a single cardiac phase, such as end-diastole, limiting their ability to analyze dynamic cardiac motion. Moreover, they still rely on ECG gating and breath-holding, making them sensitive to patient abilities and prone to motion artifacts in clinical situations.

1.4 Free-Running

Free-Running MRI is a self-gated, continuous acquisition technique designed to overcome the major limitations of conventional cardiac MRI. Unlike standard protocols that require ECG synchronization and multiple breath-holds, FR MRI captures both cardiac and respiratory motion simultaneously, without the need for external gating devices or patient cooperation for breath-holding [4].

This method acquires data continuously while the patient breathes freely, and uses intrinsic information from the MRI signal itself to retrospectively extract both cardiac and respiratory motion states. As a result, it enables the reconstruction of motion-resolved images across all phases of both the cardiac and respiratory cycles, effectively producing a high-dimensional dataset often referred to as "5D MRI" (3D spatial + cardiac phase + respiratory phase).

By removing the reliance on ECG and breath-holds, FR MRI significantly improves patient comfort and simplifies scan setup. Additionally, the ability to retrospectively access and separate cardiac and respiratory motion states provides useful physiological information and contributes to improved image quality, which can support more reliable and quantitative cardiac assessments.

Compared to conventional 2D or 3D cardiac imaging, 5D MRI leads to a substantially larger dataset, typically comprising up to 100 3D volumes per scan by combining multiple respiratory (e.g., 4) and cardiac phases (e.g., 25). This increase not only results in longer reconstruction times but also poses challenges for clinical interpretation, as the volume of data to be reviewed and analyzed grows considerably.

1.5 Cardiac MRI Segmentation

Cardiac MRI segmentation involves delineating anatomical structures of the heart such as the left and right ventricles, myocardium, and atria from MR images to enable quantitative analysis. This step is essential for extracting clinically relevant metrics. Traditionally, segmentation has been performed manually or semi-automatically by clinicians, which is time-consuming, subject to inter-observer variability, and increasingly impractical with large datasets such as those produced by FR MRI. As a result, there is growing interest in developing automated segmentation methods that can provide consistent, accurate, and efficient analysis, particularly in the context of high-dimensional cardiac imaging.

In recent years, deep learning has emerged as a powerful approach for automating medical image segmentation, including in cardiac MRI. These methods, particularly convolutional neural networks (CNNs), are capable of learning complex spatial patterns directly from imaging data, enabling accurate and reproducible delineation of cardiac structures. Once trained, deep learning models can process large volumes of data rapidly, making them well suited for high-dimensional imaging such as 5D MRI. Beyond reducing manual workload, automated segmentation can also help standardize analysis across institutions and improve the scalability of quantitative cardiac assessment in both research and clinical settings.

1.6 Aim

The main objective of this project is to develop and evaluate a deep learning segmentation model that is capable of delivering accurate outputs on rapid reconstruction approaches for 5D FR cardiac MRI data. Fast reconstructions can be computed directly on the scanner within acceptable clinical time limits, and hence are vital for enabling automatic analysis integration into the acquisition pipeline. However, the reconstructions are also more vulnerable to artifacts and noise, and therefore segmentation is particularly challenging. In addition to the segmentation task itself, this contribution also strives to craft and implement the engineering platform required to integrate smoothly into the clinic. This involves the development of a modular communication pipe between the MRI scanner and an external server, automated data transfer, reconstruction, segmentation, and quantitative metric extraction. Such a system ensures that results can be processed efficiently and returned to the scanner console with minimal user interaction, paving the way for strong and scalable deployment in actual clinical settings.

2 State of the art

In recent years, significant advances have been made in cardiac MRI acquisition, image reconstruction, and segmentation. This section reviews the current state of the art in cardiac MRI segmentation methods, the application of deep learning techniques, and recent efforts toward handling high dimensional data and real-time clinical deployment. Particular attention is given to approaches addressing variability in reconstruction quality and the integration of automated tools into the clinical workflow.

2.1 Cardiac MRI Segmentation Techniques

Cardiac MRI segmentation is an extremely crucial step for the measurement of ventricular volumes, myocardial mass, and functional parameters such as stroke volume and ejection fraction. Segmentation has traditionally been performed manually by expert readers, slice by slice in the short-axis plane. Despite being considered the reference standard, it is highly time-consuming, requires expert-level expertise, and is prone to intra- and inter-observer variability, especially in large or high-dimensional datasets.

To reduce annotation time, semi-automatic methods were introduced. They include contour initialization with subsequent active contour models or level-set evolution [5], and propagation methods wherein annotations from one or more reference slices are propagated through the stack [6]. Although these methods were more efficient than manual annotation, they would suffer in the presence of imaging artifacts, non-uniform contrast, or complex pathologies.

With the introduction of FR 4D/5D cardiac MRI, the limitations of the conventional methods were even more pronounced. The large number of images over both cardiac and respiratory phases made manual or semi-automatic segmentation infeasible. Furthermore, conventional methods tend to generalize poorly to increasing data dimensionality and are not always scanner and protocol-robust. These challenges pave the way for more adaptable, automated solutions opening the gate to machine learning and, more recently, deep learning-based methods that learn discriminative features from imaging data itself.

2.2 Deep Learning Approaches in Medical Imaging

Encoder-decoder CNNs are currently the gold standard for medical image segmentation. The original U-Net introduced a symmetric contracting/expanding path with skip connections for precise localization from limited data [7]. Its 3D variants 3D U-Net and V-Net applied this paradigm to 3D MRI with dense volumetric prediction and Dice-optimized losses, improving small structure segmentation [8, 9]. On these foundations, nnU-Net established an end-to-end, self-configuring pipeline (preprocessing, architecture, training, post-processing) that fits a new dataset with minimal manual tuning and remains a competitive baseline across modalities and tasks [10]. Reviews consistently report DL segmentation to outperform classical methods and enable scalable workflows in clinical research [11].

2.3 Applications to High-Dimensional (4D/5D) Cardiac Data

Motion-resolved FR whole-heart MRI reconstructs a high-dimensional data set (3D + cardiac + respiratory), commonly referred to as 5D. XD-GRASP brought sorting into additional motion-state dimensions with compressed sensing, allowing for free-breathing acquisitions without navigators [12]. Later 5D sparse-MRI frameworks and fully self-gated FR approaches provided automated pipelines and high-quality coronary MRA / whole-heart cine without ECG or breath-holds [13, 4]. Aside from anatomy, FR 5D flow imaged respiratory- and cardiac-resolved hemodynamics within minutes, with the potential for complete physiologic assessment in a single scan . Application papers and reviews have recently addressed push-button workflows and larger patient cohorts [14, 15].

2.4 Deep Model Robustness to Image Quality and Reconstruction Variants

Segmentation models are faced with domain shifts (scanner, protocol, contrast, and reconstruction differences) in practice. Systematic research in cardiac MRI documents substantial performance drops under domain shift and proposes data augmentation and normalization strategies to improve generalization, often within the nnU-Net paradigm [16]. More general overviews and cardiac-dedicated research highlight domain adaptation (unsupervised/semi-supervised), generative augmentation, and few-shot target annotations to recover performance across centers and protocols [17, 18]. These findings validate the potential for using models that remain precise on fast (and noisier) reconstructions when adaptation and training prioritize robustness.

2.5 Toward Real-Time Clinical Deployment on MRI Systems

Scanner-independent pipelines have evolved to support near real-time reconstruction and analysis. ISMRMRD provides a standard raw-data format for streamed acquisitions and interoperable tooling [19]. Gadgetron enables modular, streaming recon pipelines and can be embedded on clinical scanners or clusters for low-latency processing [20]. Open-source end-to-end workflows now extend from sequence to recon and analysis, accelerating translation and reproducibility [21]. Together, these components ensure on-scanner or near-scanner deployment where fast recon + robust segmentation can yield preliminary results during the exam.

3 Methods

3.1 Study Population and Acquisition Protocol

A total of 35 subjects were included in this study, anonymized to ensure patient privacy and ethical guidelines. Imaging was performed on two different clinical MRI scanners: Dataset D1 ($n = 15$) was acquired on a 1.5 Tesla Siemens Sola system and included only healthy volunteers scanned without contrast agent. Dataset D2 ($n = 20$) was acquired on a 3 Tesla Siemens Prisma system and included 10 healthy subjects, 5 patients with heart failure with preserved ejection fraction (HFpEF), and 5 with heart failure with reduced ejection fraction (HFrEF). In D2, all subjects received a gadolinium-based contrast agent to enhance visualization and tissue contrast. HFpEF is defined by typical heart failure symptoms with normal or near-normal ejection fraction, while HFrEF involves a decreased ejection fraction due to impaired systolic function. A summary of the dataset structure is provided in Figure 2.

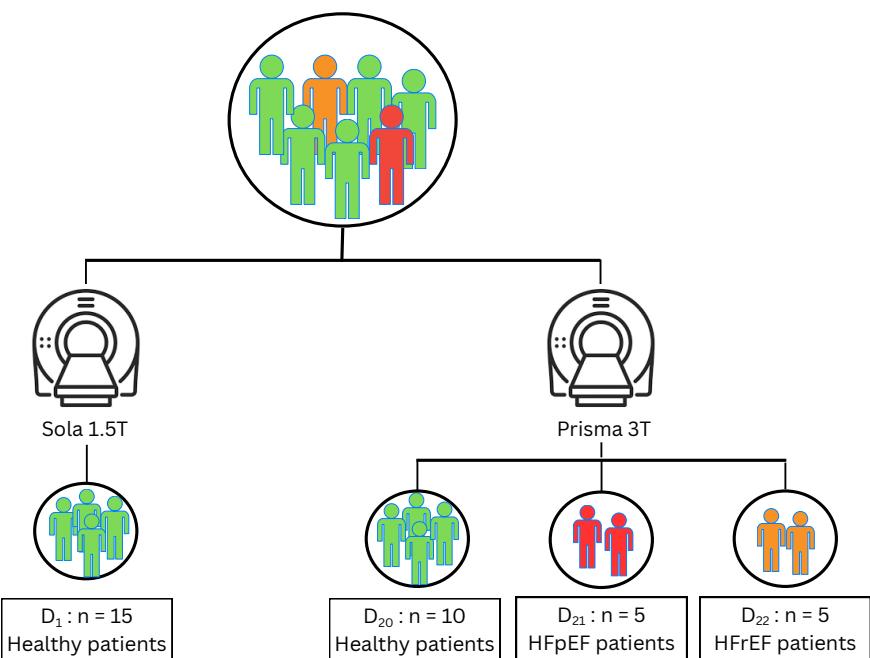


Figure 2: Distribution of the study population across scanners

FR MRI acquisitions were performed using a slab-selective 3D balanced steady-state free precession (bSSFP) sequence with self-gating. The acquisition was conducted without ECG triggering or breath-holding, enabling continuous free-breathing scanning. The imaging volume covered a field of view (FOV) of $220 \times 220 \times 220$ mm 3 , with isotropic spatial resolution of $1.4 \times 1.4 \times 1.4$ mm 3 , using a slab-selective RF excitation. A radial k-space acquisition trajectory was used to enable motion-resolved reconstruction. Imaging parameters included a repetition time (TR) of 3.5 ms and an echo time (TE) of 1.7 ms. Each scan lasted approximately 4 minutes, during which cardiac and respiratory motion signals were intrinsically recorded from the acquired data for retrospective sorting.

The FR acquisition provides a 3D volume of the heart, captured in standard anatomical planes axial, sagittal, and coronal. This volumetric dataset represents

a cube enclosing the entire heart, allowing for flexible reformatting into any desired imaging orientation. On the scanner, specific cardiac views are selected based on their clinical relevance for assessing ventricular morphology and myocardial function. In this study, we focus on three reformatted planes: short-axis (SAX), pseudo-four-chamber (4CH), and pseudo-two-chamber (2CH) views. These orientations are widely used in cardiac imaging due to their ability to visualize the left and right ventricles, myocardium, and overall cardiac function in a standardized and interpretable manner.

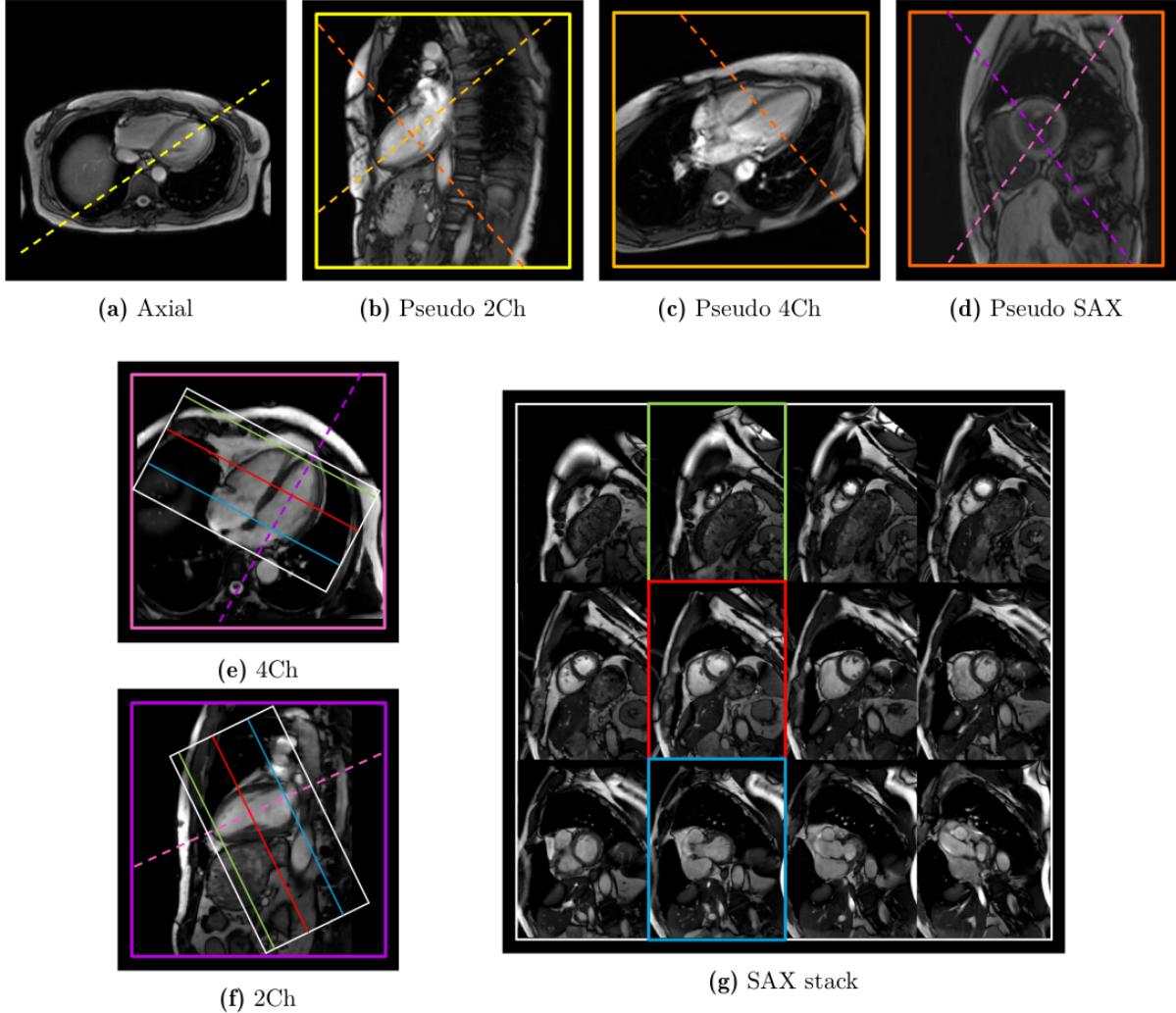


Figure 3: Illustration of views planification for 2D cine acquisition with LV focus. Colors link one orientation to another. In g), the SAX stack represents the apposition of all the acquired slices in SAX orientation spanning over the entire LV, from apex (top-left) towards base (bottom-right). In e) f) and g), green indicates the apical slice which is the start of the LV from the apex of the heart, red indicates a middle slice and blue denotes the basal slice which is the end of the LV towards the base of the heart.

3.2 5D Image Reconstruction Techniques

The FR MRI acquisition used in this study generates a continuous stream of undersampled radial k-space data that must be retrospectively sorted and reconstructed into motion-resolved 5D images. Depending on the reconstruction method, image quality, reconstruction time, and motion fidelity may vary significantly. These differences are critical when training and evaluating segmentation models, as they influence both the visual clarity of anatomical structures and the reliability of extracted metrics.

In this work, we consider three reconstruction techniques:

- **Gridded reconstruction**
- **fNav-based reconstruction**
- **Compressed sensing**

In all reconstruction methods, cardiac and respiratory bins were retrospectively extracted from the k-space data using self-gating. Although the binning process is shared across techniques, each method handles the binned data differently during reconstruction.

Gridded Reconstruction

Gridded reconstruction is the simplest and fastest method used in this study. After cardiac and respiratory binning, only the data corresponding to the end-expiration respiratory phase were retained, while all cardiac phases were preserved. The selected radial k-space data were interpolated onto a Cartesian grid using a non-uniform fast Fourier transform (NUFFT), which approximates the inverse Fourier transform for non-Cartesian sampling trajectories. Mathematically, this operation can be expressed as

$$I(\mathbf{r}) = \sum_k s(\mathbf{k}) \cdot e^{i2\pi\mathbf{k} \cdot \mathbf{r}}$$

where $s(\mathbf{k})$ represents the measured k-space signal and \mathbf{r} the spatial coordinates. Due to undersampling, many regions of Cartesian k-space remain unpopulated after gridding. These gaps are implicitly zero-filled, resulting in noise and artifacts in the reconstructed images. While this approach provides fast reconstruction and minimal computational overhead, the resulting image quality is degraded.

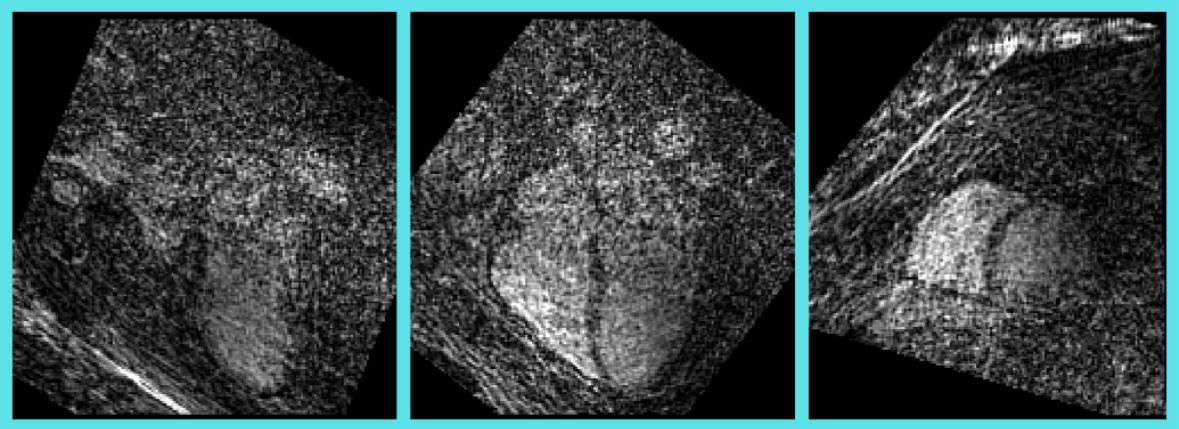


Figure 4: Gridded reconstruction at end-diastolic phase, showing pseudo-2CH (left), pseudo-4CH (center), and SAX (right) views.

fNav Reconstruction

The fNav-based reconstruction also begins with cardiac and respiratory binning of the radial k-space data, similar to the gridded approach. However, instead of discarding all but one respiratory phase, fNav assumes that respiratory motion is approximately linear along the x, y, and z axes. Using motion signals extracted from fNav, a respiratory motion model is estimated and applied to shift and realign the k-space data from all respiratory bins toward a common reference phase typically end-expiration. This recentering allows the inclusion of significantly more k-space samples for each cardiac phase, while minimizing respiratory motion artifacts. The resulting images exhibit reduced blurring and improved anatomical clarity compared to simple gridded reconstructions [22].

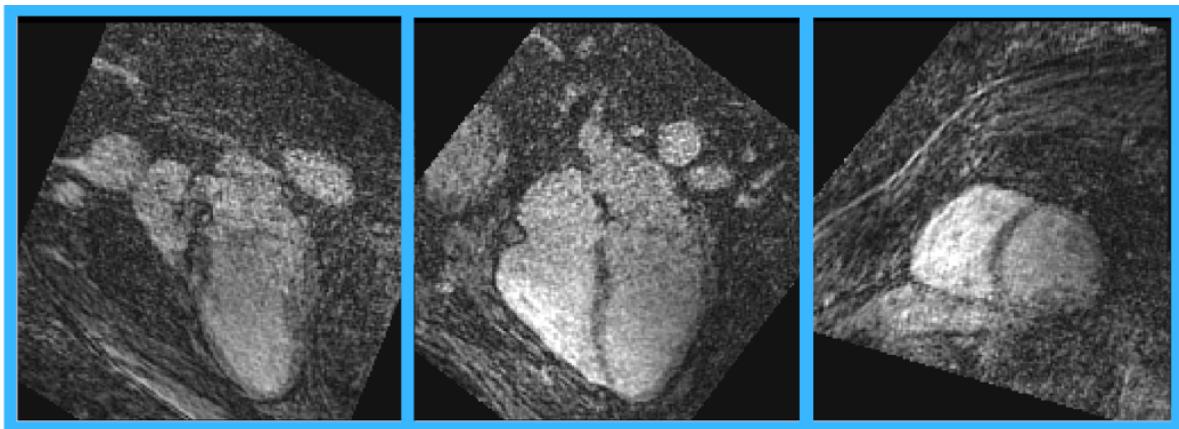


Figure 5: fNav reconstruction at end-diastolic phase, showing pseudo-2CH (left), pseudo-4CH (center), and SAX (right) views.

Compressed Sensing Reconstruction

Compressed Sensing reconstruction uses the same cardiac and respiratory binning strategy as in the previous methods but retains all bins across both motion dimensions. Instead of selecting or merging respiratory phases, the full 5D dataset is reconstructed jointly using an iterative optimization algorithm that exploits temporal and spatial sparsity. The reconstruction solves an inverse problem with regularization terms that promote sparsity across both cardiac and respiratory dimensions [23]. The objective can be written as:

$$\min_I \|F_u I - y\|_2^2 + \lambda \cdot \mathcal{R}(I),$$

where I is the image to reconstruct, F_u is the under-sampled Fourier operator (including the NUFFT for radial trajectories), y is the measured k-space data, and $\mathcal{R}(I)$ is a regularization term (e.g., total variation or temporal finite differences). This method enables high-fidelity reconstruction of motion-resolved images, preserving both respiratory and cardiac motion while significantly reducing undersampling artifacts. However, due to its computational complexity, it requires substantially longer reconstruction times.

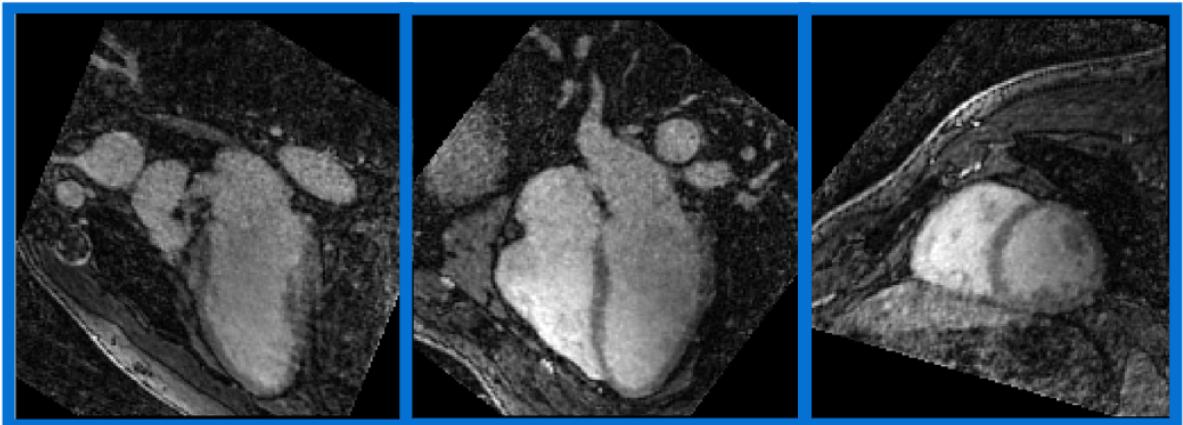


Figure 6: CS reconstruction at end-diastolic phase, showing pseudo-2CH (left), pseudo-4CH (center), and SAX (right) views.

3.3 Ground Truth Segmentations

Manual segmentation was performed on compressed sensing reconstructed images using three standard cardiac views: SAX, pseudo-2CH, and pseudo-4CH. The segmentation was primarily carried out in the SAX view, which provides optimal visibility of ventricular structures and myocardial thickness. The two long-axis views were used as complementary references to assist in accurately delineating the myocardium, left ventricle (LV), and right ventricle (RV). The segmentation process was conducted slice by slice in the SAX orientation (120-140 slices) and focused exclusively on two key cardiac phases: end-systole and end-diastole. To reduce the high time and labor demands of manual annotation, a semi-automatic propagation technique was employed, allowing segmentation of only a subset of slices instead of the entire image

set [6]. Segmentations were generated and reviewed by a team composed of clinicians, postdoctoral researchers, and master’s students. Due to the high time and labor demands associated with manual annotation, full cardiac cycle segmentation was not performed.

3.4 Deep Learning-Based Segmentation Model

For the segmentation of cardiac structures, we employed the nnU-Net framework [10], which has become a state-of-the-art tool in biomedical image segmentation due to its self-configuring nature. nnU-Net automatically adapts its preprocessing, network architecture, training, and post-processing pipelines to the dataset at hand, removing the need for extensive manual tuning. This flexibility makes it particularly suitable for heterogeneous medical imaging datasets, such as those arising from different reconstruction techniques (CS, fNav, and Gridded).

In this project, the model was implemented with a residual encoder architecture, which enhances feature extraction by improving gradient flow and enabling deeper representations without compromising training stability. Residual connections allow the network to capture both low-level and high-level spatial information, which is critical for accurately delineating thin myocardial walls and complex ventricular structures in undersampled cardiac MR images.

The choice of nnU-Net with a residual encoder was motivated by two main factors: (1) its demonstrated efficiency across a broad range of medical imaging challenges, and (2) its ease of deployment in real-world clinical environments. Once trained, the network can be integrated into the scanner-side workflow with minimal adjustments, ensuring rapid inference on new datasets and making it an attractive solution for on-scanner applications.

Building on recent work [24], we adapted the nnU-Net pipeline to 5D FR cardiac MRI data on different reconstruction techniques. The trained model produced accurate segmentations of the left ventricular blood pool, right ventricular blood pool, and myocardium, supporting the extraction of clinically relevant volumetric and functional parameters in near real-time.

3.5 Evaluation Criteria for Segmentation

To comprehensively evaluate our segmentation pipeline, we considered both qualitative and quantitative results. In addition to visual inspection of the segmentation outputs, we focused on quantitative metrics including reconstruction time, geometric accuracy, clinical relevance, and physiological consistency across the cardiac cycle. These criteria aim to identify the reconstruction technique that offers the best trade-off between computational efficiency and segmentation performance, with the goal of enabling reliable, real-time clinical deployment directly on the scanner.

Geometric segmentation accuracy was assessed using well-established evaluation metrics commonly employed in deep learning-based medical image analysis. These metrics provide quantitative measures of spatial overlap and boundary agreement between the predicted and reference segmentations (ground truth segmentation).

Specifically, we computed the Dice Similarity Coefficient (DSC), a volume overlap metric used to evaluate differences between two contours (eq 1).

$$DSC = \frac{2 \times TP}{2 \times TP + FP + FN} = \frac{2 \times |X \cap Y|}{|X| + |Y|} \quad (1)$$

We also evaluate the Hausdorff distance (HD), the maximum distance between the surface points of the predicted segmentation and the reference, reflecting the worst-case boundary error between the two (eq 2).

$$HD(X, Y) = \max \left\{ \sup_{x \in X} \inf_{y \in Y} \|x - y\|, \sup_{y \in Y} \inf_{x \in X} \|y - x\| \right\} \quad (2)$$

We then compute the relative volume difference (RVD) for each region of interest of the heart, which quantifies the percentage difference in volume between the predicted and reference segmentation. This metric helps assess the impact of segmentation errors on clinically relevant volumetric measurements (eq 3).

$$RVD = 100 \times \frac{V_{ref} - V_{pred}}{V_{ref}} \quad (3)$$

Clinical and Physiological metrics were derived from the predicted segmentations at end-diastolic and end-systolic phases, including end-diastolic volume (EDV), end-systolic volume (ESV), stroke volume (SV), ejection fraction (EF), and myocardial volume. These values were computed from the 3D segmentations of the left and right ventricles and the myocardium and reflect key indicators used in routine clinical assessment of cardiac function.

Stroke volume (SV) and ejection fraction (EF) were computed from the end-diastolic volume (EDV) and end-systolic volume (ESV) of the left ventricle using the following equations:

$$SV = EDV - ESV \quad (4)$$

$$EF = \frac{EDV - ESV}{EDV} \times 100 = \frac{SV}{EDV} \times 100 \quad (5)$$

Ejection fraction is a key parameter for stratifying heart failure patients. Heart failure with reduced ejection fraction (HFrEF) is characterized by an EF below 40%, indicating impaired systolic function. In contrast, heart failure with preserved ejection fraction (HFpEF), defined by an EF of 50% or higher, is associated with diastolic dysfunction despite a normal or near-normal EF. These two subtypes differ in pathophysiology, prognosis, and treatment strategies, making EF an essential clinical marker in the diagnostic workflow.

4 Results

In this section, we report the results of the evaluation of a nnU-Net model applied to cardiac MR images reconstructed using three distinct techniques: CS, fNav, and Gridded. Manual segmentation performed on CS-reconstructed images serve as the reference. The analysis focuses on three key dimensions: (1) geometric performance, assessed through overlap and distance-based metrics; (2) clinical validity, evaluated via volumetric measurements and derived functional indices such as end systole, diastole volume and ejection fraction; and (3) physiological consistency, based on the temporal stability of segmentation across cardiac phases.

In addition to segmentation accuracy, we firstly report the reconstruction time associated with each method (shown on Table 1), as it is a key factor for clinical feasibility and real-time pipeline deployment. This multi-level assessment aims to determine the extent to which nnU-Net maintains manual-level performance when applied to varying reconstruction inputs, while also considering the trade-off between quality and computational efficiency.

Reconstruction Method	Reconstruction Time (min)
Gridded	5
fNav	10
Compressed Sensing	300

Table 1: Average reconstruction time per subject for each reconstruction method.

As shown in Table 1, Gridded and fNav-based reconstructions remain reasonably fast, making them compatible with the goal of deploying an end-to-end pipeline directly on the scanner for clinical use. In contrast, compressed sensing requires significantly more time, which limits its feasibility for direct integration during patient acquisition. While it provides higher image fidelity, we investigate whether this gain in quality translates into meaningful improvements in segmentation accuracy by comparing geometric metrics across all reconstruction methods.

4.1 Qualitative Segmentation Results

Qualitative inspection of the segmentation results across the three reconstruction methods gridded, fNav, and compressed sensing as well as the manual reference, reveals a high degree of visual similarity (Figure 7). For the selected cardiac phase and imaging views (SAX, pseudo-4CH, and pseudo-2CH), no major differences are observed at first glance between the automatic segmentations. All methods capture the anatomical boundaries of the ventricles and myocardium consistently. Notably, the manual segmentations exhibit slightly less smooth contours, which is expected due to the inherent variability and limited spatial precision associated with human annotation, particularly when drawn slice-by-slice. This qualitative consistency suggests that even lower-fidelity reconstructions can yield visually accurate segmentations under favorable conditions. However, to validate this observation, we next assess segmentation performance using quantitative geometric metrics across the entire dataset.

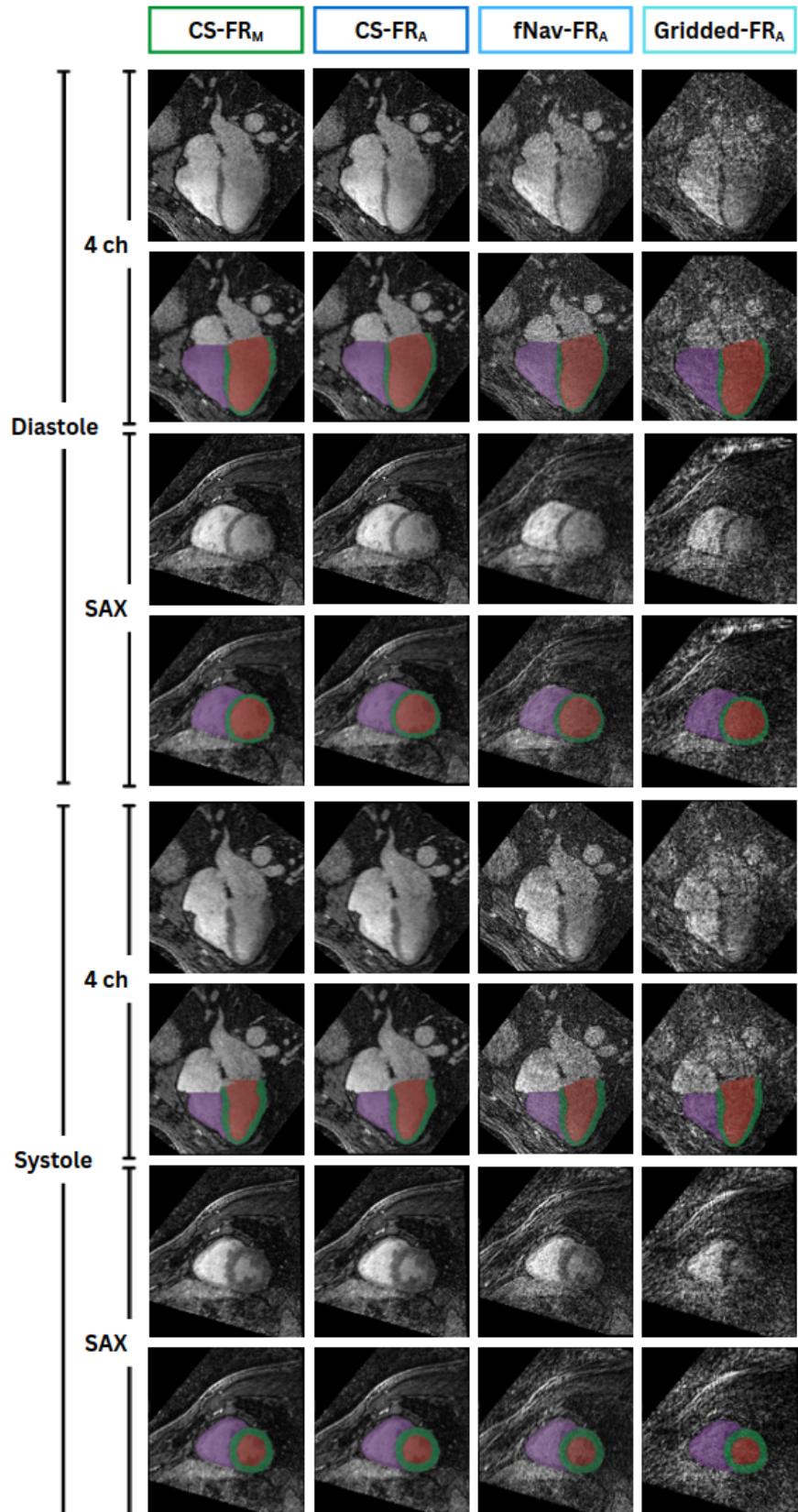


Figure 7: Segmentation results on a healthy volunteer across the four methods: FR_M, CS-FR_A, fNav-FR_A and Gridded-FR_A. Each method is shown with two views: pseudo-4CH and SAX.

4.2 Geometric metrics

We begin by evaluating geometric metrics such as Dice Similarity Coefficient, Hausdorff Distance, and Relative Volume Differences, which quantify the spatial agreement between predicted and reference segmentation.

Dice Similarity Coefficient

DSC scores were closely matched between FR_A and FR_M across all reconstruction techniques. The LVM region showed slightly lower agreement, regardless of the technique used. This reduced similarity is likely due to the lower number of annotated voxels in the myocardium, which can increase variability in segmentation metrics.

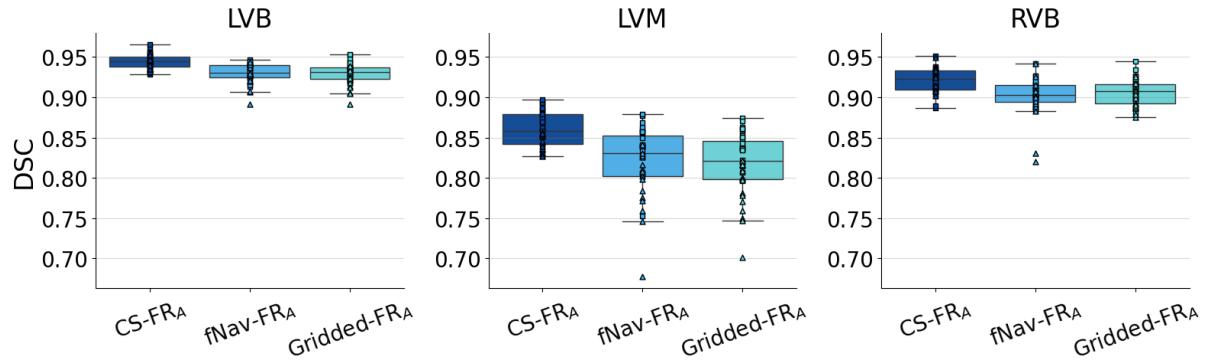


Figure 8: DSC for left ventricular blood pool (LVB), left ventricular myocardium (LVM), and right ventricular blood pool (RVB), with boxplots comparing the three reconstruction models in each anatomical region.

Hausdorff Distances

HD were comparable between FR_A and FR_M across all reconstruction techniques. However, the RVB region exhibited consistently higher HD values, reflecting greater boundary variability. This may be attributed to the complex anatomical structure and segmentation challenges associated with the right ventricular blood pool. In contrast, the LVB and LVM regions demonstrated lower HD values, indicating more consistent and accurate boundary alignment. Subject responsible for the HD Outlier of RV segmentation is shown on Figure 10.

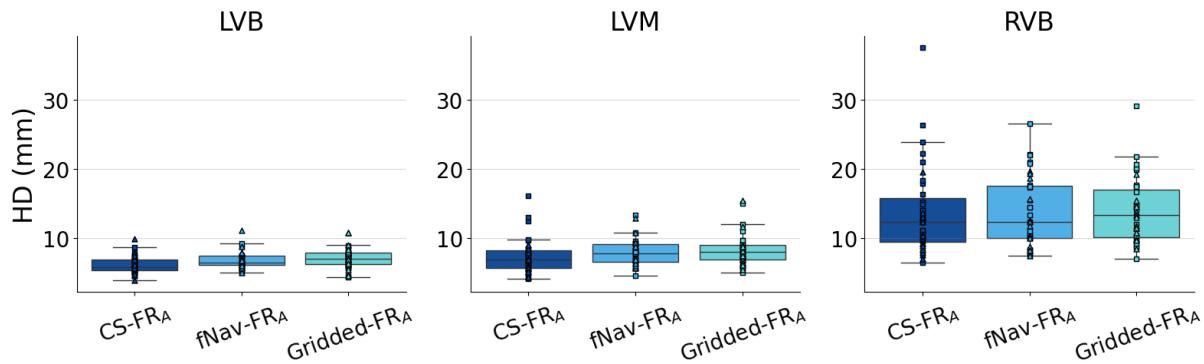


Figure 9: HD for left ventricular blood pool (LVB), left ventricular myocardium (LVM), and right ventricular blood pool (RVB), with boxplots comparing the three reconstruction models in each anatomical region.

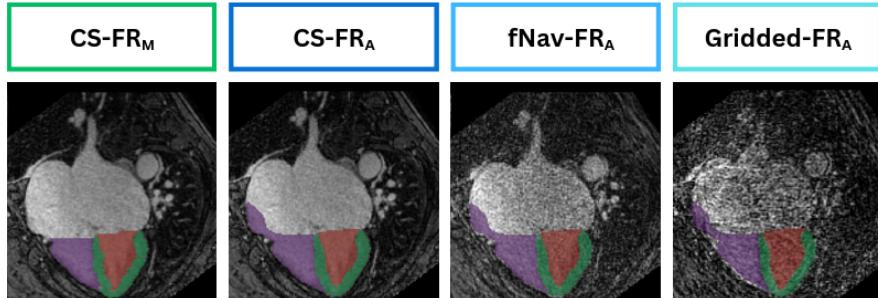


Figure 10: Comparison between manual and automatic segmentation in a subject with a dilated atrium, shown in a systolic phase image in the 4CH view.

Relative Volume Differences

RVD remained relatively low across all reconstruction techniques, particularly for the LVB region, indicating good volumetric agreement. The LVM and RVB regions showed slightly higher RVD values, which may reflect greater anatomical complexity and segmentation variability. Among the techniques, CS generally yielded the lowest RVD values, suggesting slightly better volume consistency across regions.

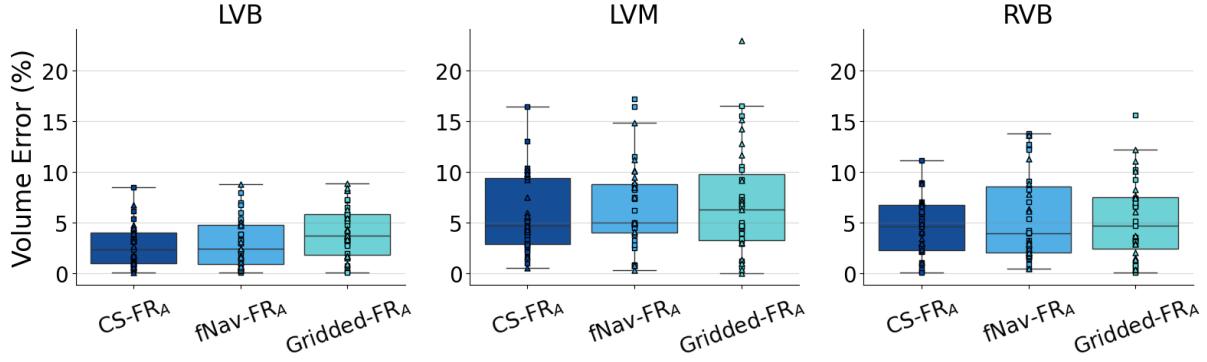


Figure 11: RVD for left ventricular blood pool (LVB), left ventricular myocardium (LVM), and right ventricular blood pool (RVB), with boxplots comparing the three reconstruction models in each anatomical region.

Metric	Region	Model		
		CS-FRA	fNav-FRA	Gridded-FRA
DSC (-)	LVB	0.94 ± 0.01	0.93 ± 0.02	0.93 ± 0.01
	LVM	0.86 ± 0.02	0.82 ± 0.04	0.82 ± 0.04
	RVB	0.92 ± 0.02	0.90 ± 0.02	0.91 ± 0.02
HD (mm)	LVB	6.18 ± 1.27	6.77 ± 1.27	7.11 ± 1.48
	LVM	7.38 ± 2.55	7.93 ± 1.98	8.31 ± 2.41
	RVB	13.83 ± 6.47	13.79 ± 5.14	13.91 ± 4.76
RVD (%)	LVB	2.37 ± 2.06	3.06 ± 2.39	4.08 ± 2.54
	LVM	5.76 ± 3.74	6.58 ± 4.32	7.02 ± 5.32
	RVB	4.51 ± 2.81	5.57 ± 4.36	5.24 ± 3.83

Table 2: Segmentation metrics (DSC, HD, RVD) by anatomical region (LVB, LVM, RVB) and reconstruction model (CS-FRA, fNav-FRA, Gridded-FRA). Values are reported as mean \pm standard deviation.

Overall, the geometric metrics indicate that all reconstruction methods closely approximated the manual segmentation (Table 2). Among them, automatic segmentation on CS consistently achieved slightly better performance compared to fNav and Gridded techniques, suggesting improved spatial accuracy.

4.3 Clinical metrics

Next, we compute clinical metrics including end-diastolic volume, end-systolic volume, stroke volume, ejection fraction, and myocardial volume, which are routinely used by clinicians for cardiac function assessment.

Plot Bland-Altman between Manual and Automatic segmentation

The nnU-Net model demonstrated strong agreement with manual segmentations across all reconstruction techniques for key clinical metrics. Intraclass correlation coefficients (ICCs) were excellent for LV volumes (≥ 0.99) and slightly lower, yet still strong, for RV (≥ 0.96). Ejection fraction (EF) showed more variability, particularly for the RV (ICC 0.91–0.93), but remained within acceptable limits. Standard Error of Measurement (SEM) values were low overall, especially for LV metrics (as low as 1.97% for EF), while RV metrics showed moderately higher SEM, notably with fNav. Among reconstruction methods, CS consistently yielded the highest ICCs and lowest SEMs, suggesting more stable performance.

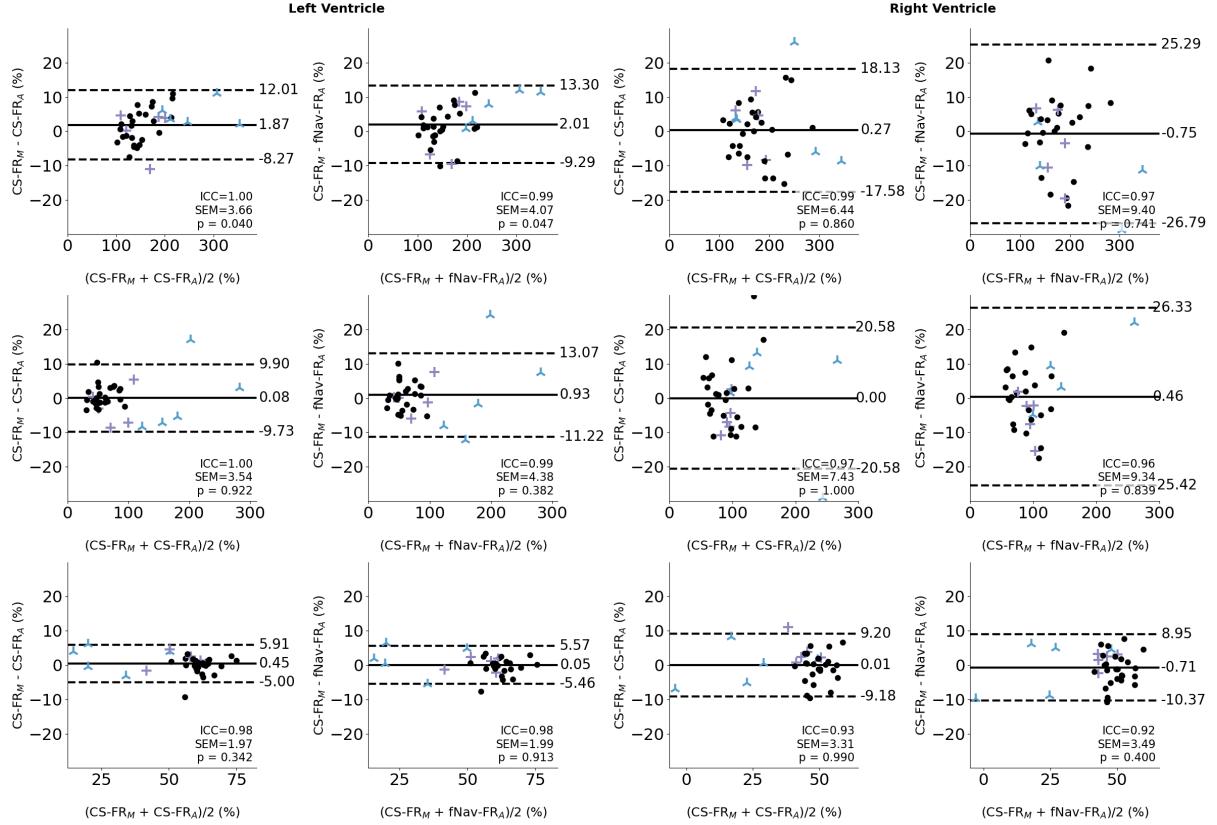


Figure 12: Bland–Altman analysis of ventricular function: top, end-diastolic volumes (EDV); middle, end-systolic volumes (ESV); and bottom, ejection fraction (EF) for both left and right ventricles. Semi-manual segmentation on compressed-sensing images (CS-FRM) is compared with DL-based自动 segmentations on compressed-sensing images (CS-FRA) and respiratory motion–corrected images (fNav-FRA). Black points represent healthy volunteers, while purple and blue crosses denote HFP EF and HFrEF patients, respectively.

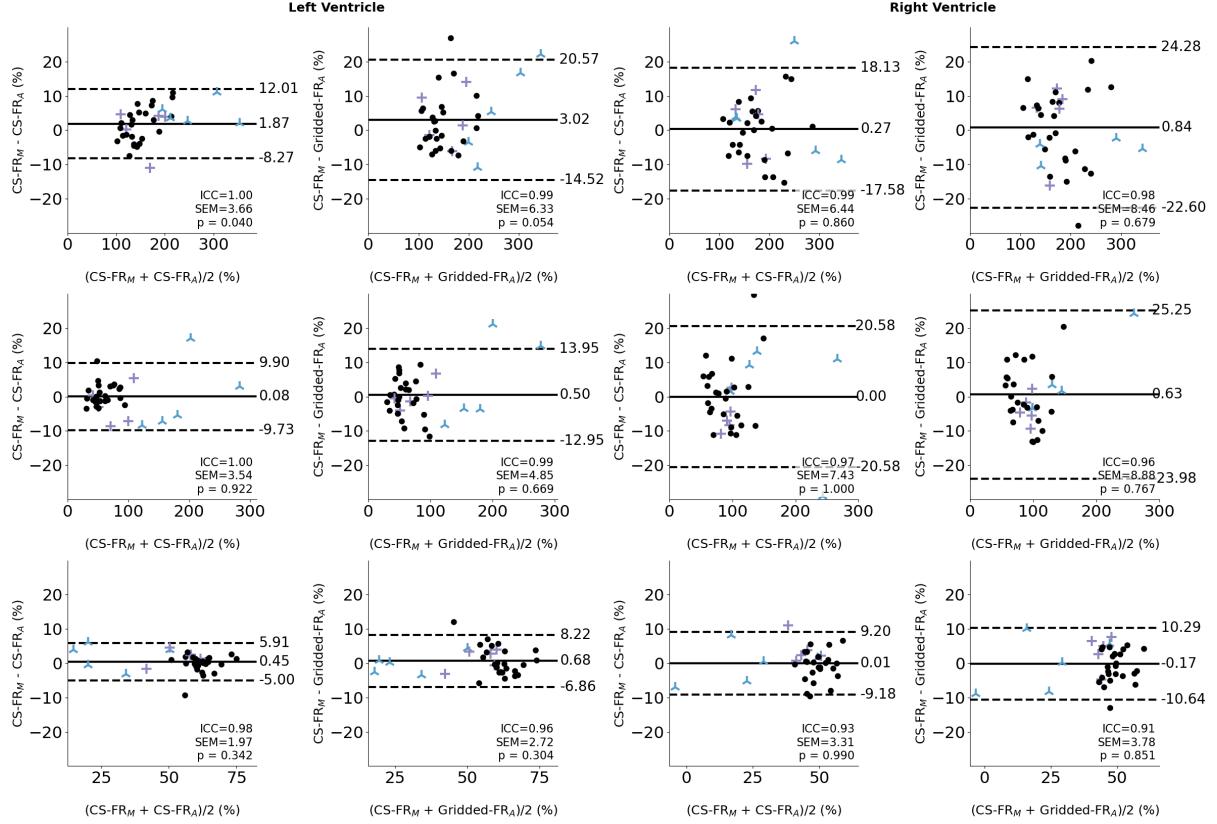


Figure 13: Bland–Altman analysis of ventricular function: top, end-diastolic volumes (EDV); middle, end-systolic volumes (ESV); and bottom, ejection fraction (EF) for both left and right ventricles. Semi-manual segmentation on compressed-sensing images (CS-FR_M) is compared with DL-based automatic segmentations on compressed-sensing images (CS-FR_A) and gridded images (Gridded- FR_A). Black points represent healthy volunteers, while purple and blue crosses denote HFrEF and HFpEF patients, respectively.

Plot Bland-Altman between CS and fNav/Gridded segmentation

Goal: Show robustness of nnUnet and adaptability to reconstruction techniques

Comparing fNav and Gridded segmentations to CS-based predictions demonstrates the robustness of nnU-Net across varying image qualities. ICC values remained consistently high across all clinical metrics (≥ 0.97), indicating excellent agreement regardless of reconstruction type. Ejection fraction showed slightly reduced agreement for LV and RV with Gridded (ICC 0.97), but remained strong. SEM values were low overall, particularly for EF (as low as 1.18% for LV with fNav), further confirming the model’s adaptability. These results highlight the resilience of nnU-Net to image noise and reconstruction differences, with CS serving as a stable reference.

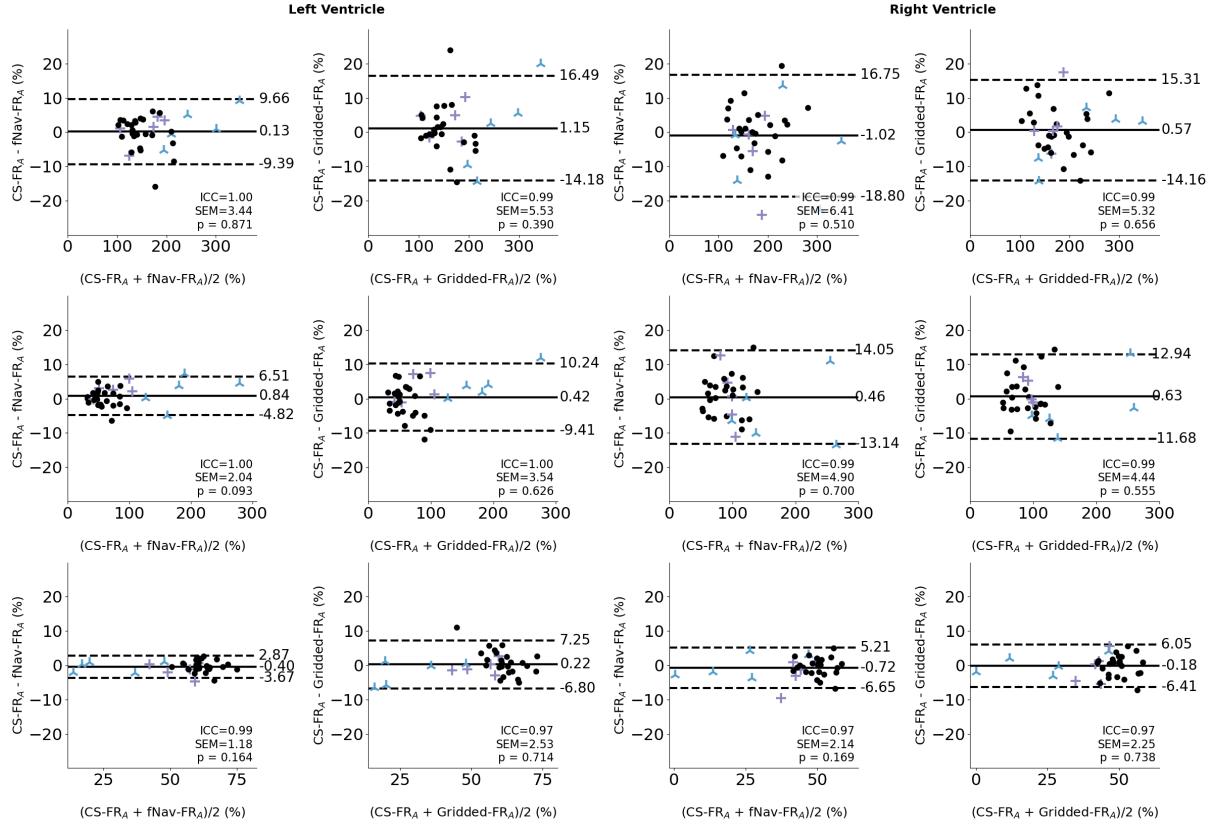


Figure 14: Bland–Altman analysis of ventricular function: top, end-diastolic volumes (EDV); middle, end-systolic volumes (ESV); and bottom, ejection fraction (EF) for both left and right ventricles. DL-based automatic segmentations on compressed-sensing(CS-FR_A) is compared with respiratory motion–corrected images (fNav-FR_A) and gridded images (Gridded-FR_A). Black points represent healthy volunteers, while purple and blue crosses denote HFrEF and HFpEF patients, respectively.

4.4 Physiological Consistency metrics

Finally, we assess physiological consistency by measuring the temporal stability of the segmentations across the cardiac cycle, reflecting the reliability of the model under dynamic conditions.

Myocardium Volume Consistency

MVC (%)	Model			
	CS-FR _M	CS-FR _A	fNav-FR _A	Gridded-FR _A
	5.57 ± 3.54	3.96 ± 2.36	4.15 ± 2.89	4.96 ± 4.32

Table 3: Myocardium Volume Consistency per reconstruction model. Values are reported as mean ± standard deviation.

All automatic segmentations outperformed manual segmentation in terms of myocardial volume consistency, as indicated by lower mean errors. The CS reconstruction yielded the best results with the lowest mean (3.96) and standard deviation (2.36), suggesting higher reliability. While fNav also showed improved consistency over manual, the Gridded reconstruction exhibited higher variability, with the largest standard deviation (4.32), indicating less stable performance across cases.

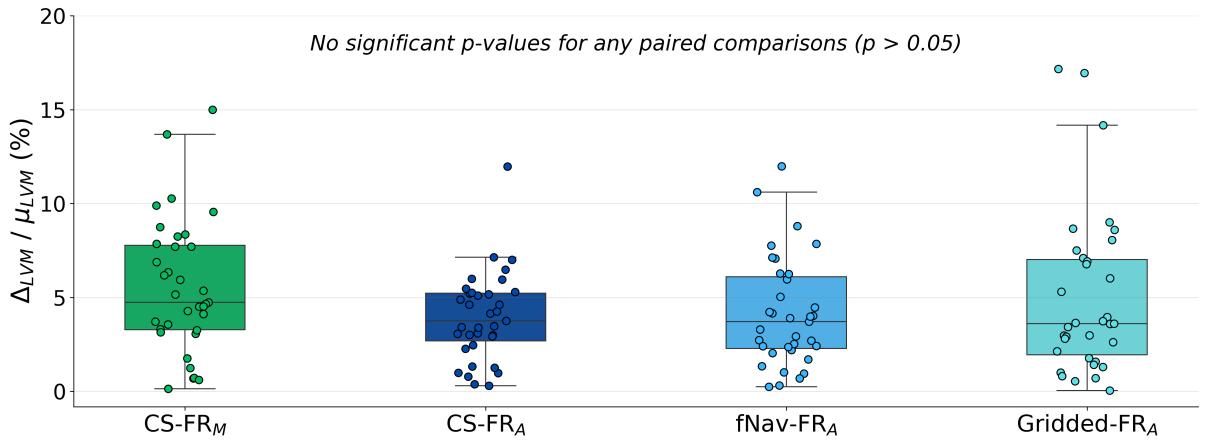


Figure 15: Systolic-diastolic left ventricular myocardial (LVM) volume mismatch computed for each segmentation as the difference between the end-diastole and end-systole volumes (Δ_{LVM}) divided by the mean volume (μ_{LVM}).

Stroke Volume Consistency

In all models, a systematic bias is observed between left and right ventricular stroke volumes, as shown in the Bland–Altman plots. The mean of the differences (LVSV - RVSV) is consistently greater than zero, indicating that the left ventricular stroke volume tends to be overestimated compared to the right across all reconstruction methods.

Moreover, the limits of agreement are wider for the Compressed Sensing (CS-FR_A) and Gridded (Gridded-FR_A) methods, suggesting greater variability in stroke volume estimation with these techniques. In contrast, the manual segmentation and fNav method (fNav-FR_A) show narrower agreement, indicating more consistent estimates between LVSV and RVSV.

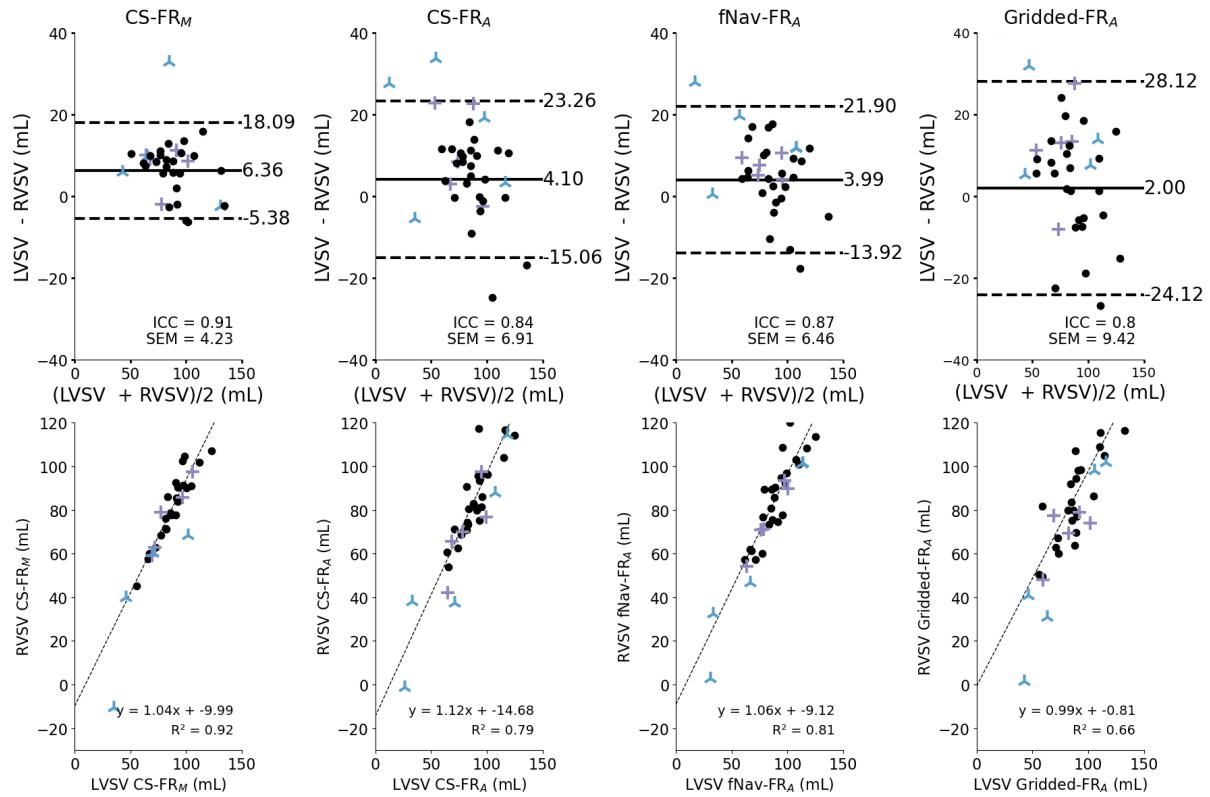


Figure 16: Comparison of left-ventricular stroke volume (LVSV) and right-ventricular stroke volume (RVSV) across semi-manual segmentation and DL-based automatic segmentations. Top row: Bland–Altman plots of LVSV–RVSV agreement. Bottom row: direct correlation plots of RVSV (y-axis) versus LVSV (x-axis).

5 Discussion

The results of this research demonstrate that deep learning segmentation consistently performs well across reconstruction techniques. Automatic segmentations were highly comparable to manual reference, supporting the robustness of the nnU-Net model even when applied to rapidly reconstructed datasets such as gridded and fNav images. Performance was comparable to the compressed sensing high-quality reference, demonstrating that reconstruction speed does not necessarily decrease segmentation accuracy.

Yet, some limitations were observed. In particular, certain special cases of unusual cardiac morphology or image characteristics resulted in less precise automatic segmentation, most notably in the right ventricle (Figure 10). This shows that, while the model generalizes well overall, some structural variations remain challenging.

Quantitative analysis also illustrated that gridded reconstructions exhibited wider limits of agreement than compressed sensing, which is a reflection of increased dispersion of results. However, overall physiological consistency of the segmentation was preserved. Importantly, gridded and fNav reconstructions can be accomplished within a few minutes, which is a clear benefit for on-scanner application and integration in the clinical workflow.

It should be noted that the datasets used in this research were acquired on high-resolution scanners and, in many cases, enhanced with contrast agents, which can have facilitated both reconstruction and segmentation. Future research should extend the evaluation to lower-field scanners, lower-resolution acquisitions, and patient groups with more heterogeneous cardiac anatomies or pathologies, to confirm the model’s robustness for more general clinical practice.

Further experiments can also test the interaction of reconstruction techniques and segmentation models. For example, testing a model trained on gridded images on compressed sensing reconstructions (or vice versa) would reveal the effect of reconstruction-specific features on segmentation accuracy. Similarly, one might examine the impact of imaging plane on training: while manual segmentation was performed on SAX, 2CH, and 4CH views, training the model on axial, sagittal, and coronal planes but still using conventional cardiac planes as ground truth may reveal additional strategies for improving generalization.

Overall, the findings confirm that gridded and fNav reconstructions, although faster and less computationally demanding, are viable options for on-scanner segmentation. The compromise between their speed and accuracy renders them highly suitable for clinical implementation, with comparable performance to compressed sensing reconstructions and close to manual segmentation.

6 On scanner deployment

To enable seamless integration of automated segmentation within the clinical workflow, we developed a communication pipeline between the MRI scanner and a remote processing server, as illustrated in Figure 17. In this system, the scanner acts as a client, transmitting raw FR acquisitions to the server using the ISMRMRD protocol via a dedicated Python client interface. On the server side, incoming data streams are received by a listener script (`server.py`), which triggers the execution of the main reconstruction and analysis routine (`freerunning.py`). This routine first performs image reconstruction using the gridded approach, followed by segmentation using a pre-trained nnU-Net model. After segmenting the left and right ventricular blood pools and myocardium, the pipeline computes key volumetric metrics such as end-diastolic volume (EDV), end-systolic volume (ESV), stroke volume (SV), and myocardial volume. Upon completion, both the reconstructed images with overlaid segmentations and a diagnostic report summarizing all computed metrics are transmitted back to the scanner console for immediate visualization and interpretation.

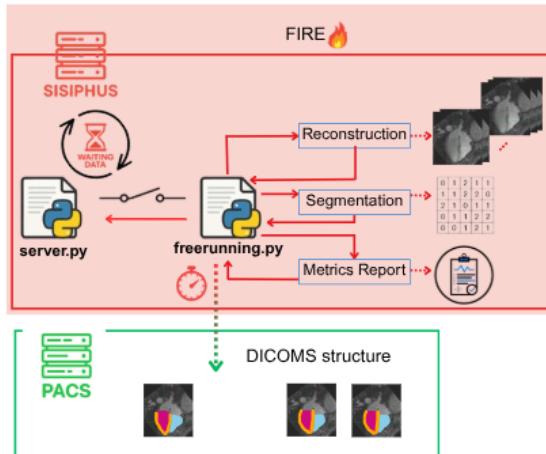
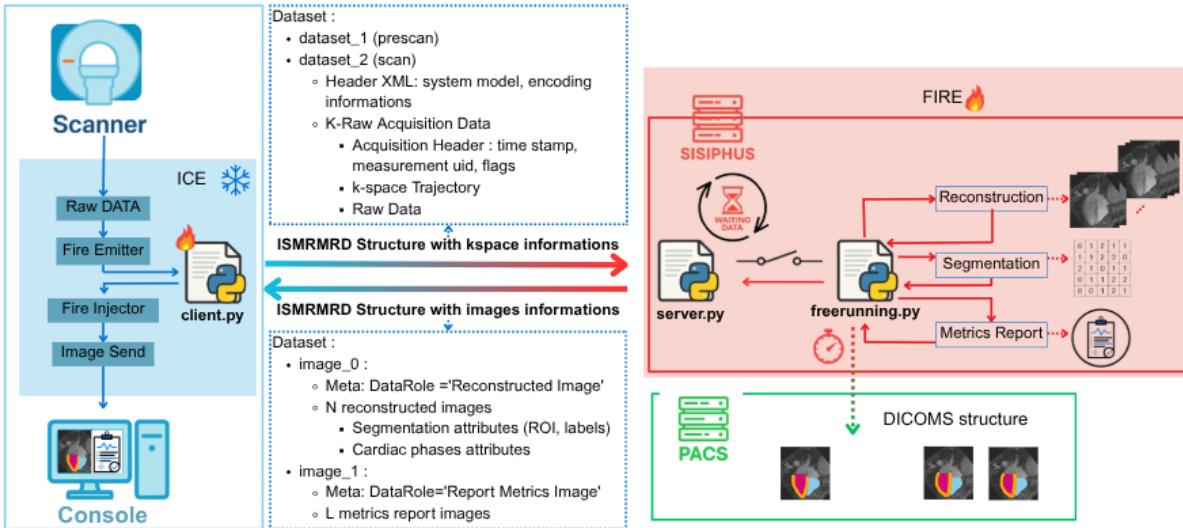


Figure 17: Communication pipeline between the MRI scanner and the remote server

6.1 Communication Protocol and Configurations

ISMRMRD Protocol

To enable communication between the MRI scanner and the processing server, we employed the ISMRMRD (International Society for Magnetic Resonance in Medicine Raw Data) format an open and standardized data format designed to store raw MR acquisition data in a structured and vendor-neutral way. The ISMRMRD framework provides a flexible XML-based header and binary storage for raw k-space readouts, physiological signals, and optional k-space trajectory information, making it particularly suitable for non-Cartesian and FR sequences. The communication pipeline leverages this format to stream data from the scanner to the server, ensuring compatibility with custom reconstruction workflows.

Each acquisition session begins with the transmission of a comprehensive XML header, which describes the essential parameters of the scan. This header follows the ISMRMRD schema definition (`ismrmrd.xsd`) and includes multiple sections such as the system information, sequence parameters, and, most importantly, the encoding block. The `encoding` section defines both the encoded and target reconstructed spaces, including matrix sizes, field-of-view dimensions, oversampling factors, and partial Fourier settings. Additionally, the `encodingLimits` field specifies the bounds of k-space sampling across all dimensions (readout, phase encoding, and slice). These values are critical for the reconstruction server to correctly interpret and process the incoming raw data. For instance, in Cartesian acquisitions with partial Fourier and oversampling, the header allows the reconstruction to adapt the image grid accordingly.

Following the XML header, the scanner sends the actual raw k-space data as a stream of individual Acquisition objects. Each acquisition contains a fixed `AcquisitionHeader`, the complex-valued raw data array for all receiver channels, and in non-Cartesian cases an associated trajectory coordinate set. These acquisitions are sequentially transmitted as binary blocks conforming to the ISMRMRD specification, enabling real-time ingestion and processing on the server side. This format supports both 2D and 3D acquisitions, as well as dynamic imaging frameworks such as FR or motion-resolved sequences.

In addition to k-space data, the ISMRMRD protocol also allows the transmission of physiological waveform data, including ECG, respiratory traces, or external triggering signals. These are encapsulated as Waveform objects, each consisting of a fixed `WaveformHeader` and a raw waveform array.

Scanner Configuration

Beyond the raw data format, the integration of the scanner into the reconstruction pipeline relies on a Python client script (`client.py`) that is triggered during a specific phase of the MRI sequence. Depending on the desired workflow, the client file can be launched either after raw k-space acquisition or once reconstructed images are available. In our implementation, the client is triggered immediately after the acquisition of raw k-space data, enabling direct interaction with the reconstruction server. However, this approach remains flexible, and the same communication framework can be configured to operate with image-level data instead of raw acquisitions, depending on the needs of the sequence and clinical use case.

The communication between the scanner and server is established using the Python socket package, based on a direct TCP/IP connection. The client establishes a secure SSH channel to the reconstruction server using a predefined IP address and port, ensuring secure and consistent data transfer. Once the connection is established, the client script transmits the ISMRMRD data stream and simultaneously specifies which server-side reconstruction script should be launched, in our case, freerunning.py. This server script is responsible for receiving, reconstructing, and optionally segmenting the data. The reconstruction logic is initiated by server.py, which dynamically interprets the received file instruction and launches the corresponding processing script.

In addition to the raw data, the scanner can transmit a JSON configuration file alongside the acquisition. This optional configuration file allows for dynamic control over the processing pipeline on the server side. In our case, parameters defined in the JSON file determine whether segmentation should be performed, whether a diagnostic report should be generated, and whether specific FR analysis metrics (e.g., model of the nnUNet) should be included. This modular structure enhances the flexibility of the server, allowing different tasks to be toggled on or off depending on the context of the scan. An example of such a configuration is provided in the following subsection.

```

1  {
2      "parameters": {
3          "config": "free_running",
4          "segmentation": false,
5          "report": false,
6          "reformat": false
7      },
8      "reconstruction_parameters": {
9          "type": "static",
10         "flag": {
11             "fire_mode": true,
12             "no_motion": false,
13             "coil_sens": false,
14             "reduced_card_bins": true
15         },
16         "p": {
17             "dc_nyq_factor": 2
18         }
19     },
20     "segmentation_parameters": {
21         "model": 762
22     }
23 }
```

Figure 18: Example of a JSON configuration file for on-scanner deployment

6.2 Image Processing Workflow

Once all acquisitions have been received on the server side, the reconstruction process is initiated within the `freerunning.py` script. This script launches the FreeRunningPy package developed by Augustin Ogier, which is specifically designed for the reconstruction of 5D FR MRI data. The package takes as input the raw acquisition data along with several processing flags. These flags are dynamically configured based on the JSON configuration file sent by the scanner, allowing flexible control over each processing step.

The reconstruction workflow begins with motion extraction, where respiratory and cardiac surrogate signals are derived directly from the raw k-space data. These signals are then used to perform signal-based binning, allowing the sorting of data into motion-resolved respiratory and cardiac phases. Following this, the 5D image volumes are reconstructed using a non-uniform fast Fourier transform (NUFFT) combined with density compensation. The resulting multi-dimensional image dataset (3D spatial + respiratory + cardiac dimensions) is then returned to the `freerunning.py` script for subsequent segmentation and metric extraction. In our implementation, only the end-respiratory phase is retained from the 5D volume, resulting in a 4D dataset (3D spatial + cardiac) used for segmentation and subsequent transmission back to the scanner.

The resulting 4D image (3D spatial + cardiac phases) is saved in NIfTI format and passed to the nnU-Net segmentation pipeline. The model used is specified in the JSON configuration file; in our case, model 762 corresponds to the network trained specifically on gridded reconstructions. The nnU-Net model produces a 4D segmentation mask delineating the left and right ventricular blood pools and myocardium across all cardiac phases. From this segmentation, key volumetric metrics are computed, including end-diastolic volume (EDV), end-systolic volume (ESV), stroke volume (SV), ejection fraction (EF), and myocardial volume.

6.3 Image Transmission to Console

At the end of the processing pipeline, the output consists of a 4D anatomical image volume, a corresponding 4D segmentation mask, and a dictionary of computed metrics. Before transmission back to the scanner, the data are reorganized to follow a standardized temporal ordering in which the first cardiac phase corresponds to end-diastole. For each cardiac phase and for each slice along the axial-axis, an image is sent individually to the scanner, accompanied by a configuration header that includes essential metadata such as spatial position, cardiac phase index, image number, and series index.

To preserve a clean separation between anatomical image content and segmentation information, the associated mask for each image slice is transmitted as part of the ISMRMRD metadata field, rather than being embedded directly into the image data. This approach maintains image fidelity while enabling proper overlay of the segmentation on the scanner console. Additionally, the cardiac function metrics (e.g., EDV, ESV, EF) are compiled into a tabular summary using Matplotlib. This summary is rendered as an image and sent back to the scanner in compliance with the ISMRMRD image message format, allowing it to be viewed directly on the console alongside the reconstructed images (Figure 19,20).

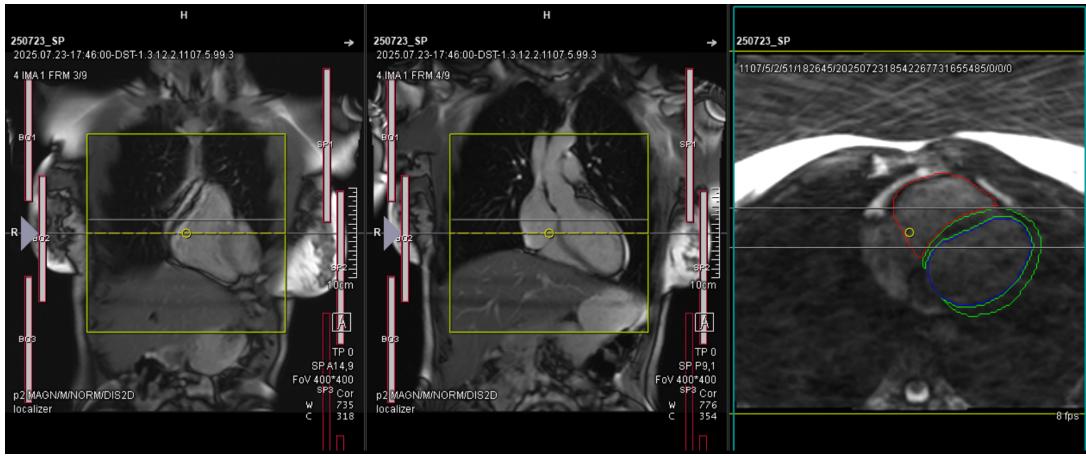


Figure 19: On-scanner reconstruction and segmentation images. From left to right: (1),(2) localizer used for planning, (3) gridded raw gridded reconstruction.

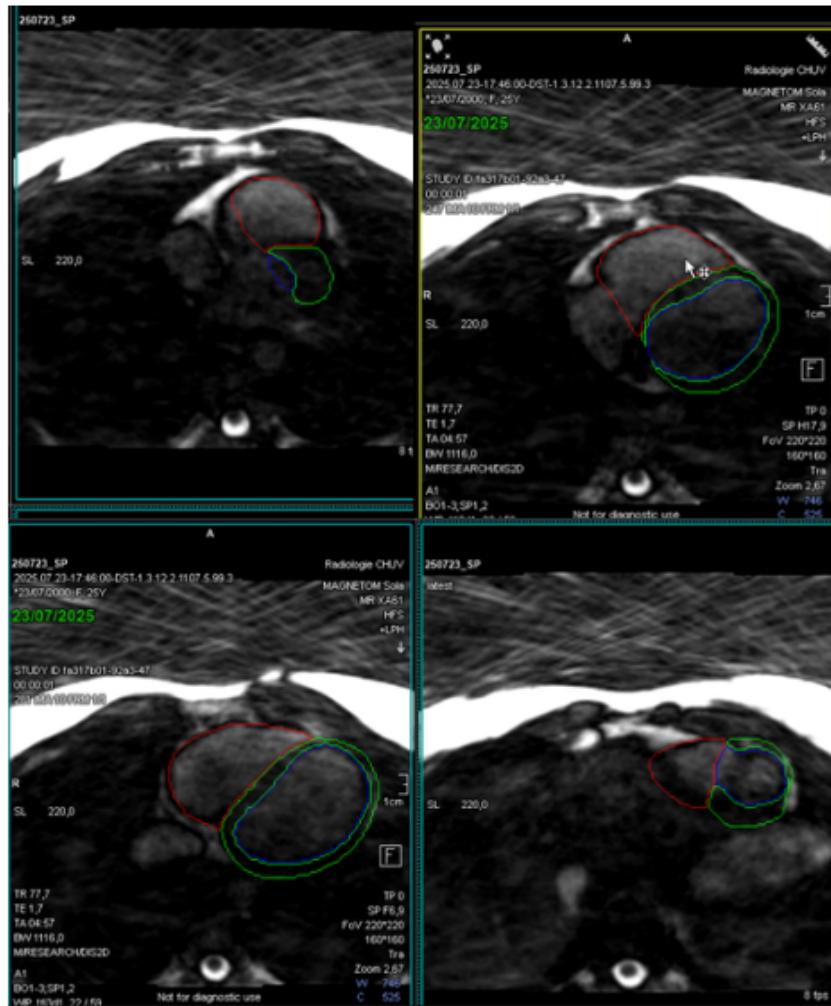


Figure 20: On-scanner reconstruction and segmentation results. Example of raw gridded reconstructions at different axial sections.

6.4 Modular Package for On-Scanner Deployment

To promote flexibility and future extension, we developed a modular software package named wildFFS, designed to streamline scanner-side deployment of FR cardiac MRI processing. The package is organized into distinct modules for communication, reconstruction, and segmentation, each of which can be executed independently depending on the desired workflow. This modular architecture enables easy integration of new components and methods, whether for research or clinical adaptation.

To demonstrate this flexibility, we incorporated external modules developed by collaborators. Notably, a deep learning-based denoising model developed by Kevin Boros (ReconAI-CardiacFR) was integrated into the pipeline to operate directly on reconstructed gridded images, acting as a pseudo compressed sensing approach to improve image quality while maintaining computational efficiency. Additionally, ongoing collaborations with Pauline Calarnou and Amaury George have focused on the integration of T1 and T2 mapping workflows into the same framework, further extending the utility of wildFFS beyond volumetric segmentation. This design philosophy allows wildFFS to evolve as a unified yet extensible tool for real-time processing of advanced cardiac MRI applications [25].

ReconAI-CardiacFR

To accelerate the reconstruction of high-dimensional FR cardiac MRI, Kevin Borsos developed a deep learning-based framework that replaces traditional compressed sensing algorithms. By leveraging a residual neural network trained on motion-resolved data, the model enables efficient reconstruction directly from gridded images. This approach is particularly well-suited for non-Cartesian and highly undersampled acquisitions, such as those used in 5D Ferumoxytol-enhanced imaging. Integrated within the wildFFS pipeline, the model functions as a learnable denoiser, offering a flexible and computationally efficient alternative to iterative reconstructions, and demonstrating the potential of modular AI-based tools for advanced cardiac MRI workflows [26]. Figure 21 shows an example of a gridded reconstruction enhanced on-scanner using Kevin’s ReconAI model integrated within the wildFFS package, with segmentation results obtained by applying the nnU-Net model trained on compressed sensing reconstructions.

Deep Learning Module for Quantitative T1 and T2 Mapping

We also integrated a deep learning-based T1 and T2 mapping module developed by Amaury George and Pauline Calarnou into the communication pipeline. In this workflow, the server receives raw MR acquisitions and, upon detection of a T1 or T2 mapping request, launches a MATLAB-based compressed sensing reconstruction tailored for rapid processing of the small number of 2D images typically acquired (e.g., 25 images). Once reconstruction is complete, the resulting image series is passed to the trained deep learning model, which estimates pixel-wise T1 or T2 values across the myocardium. The final quantitative maps are then transmitted back to the scanner console, enabling real-time visualization of myocardial tissue characteristics directly within the clinical interface (Figure 22).

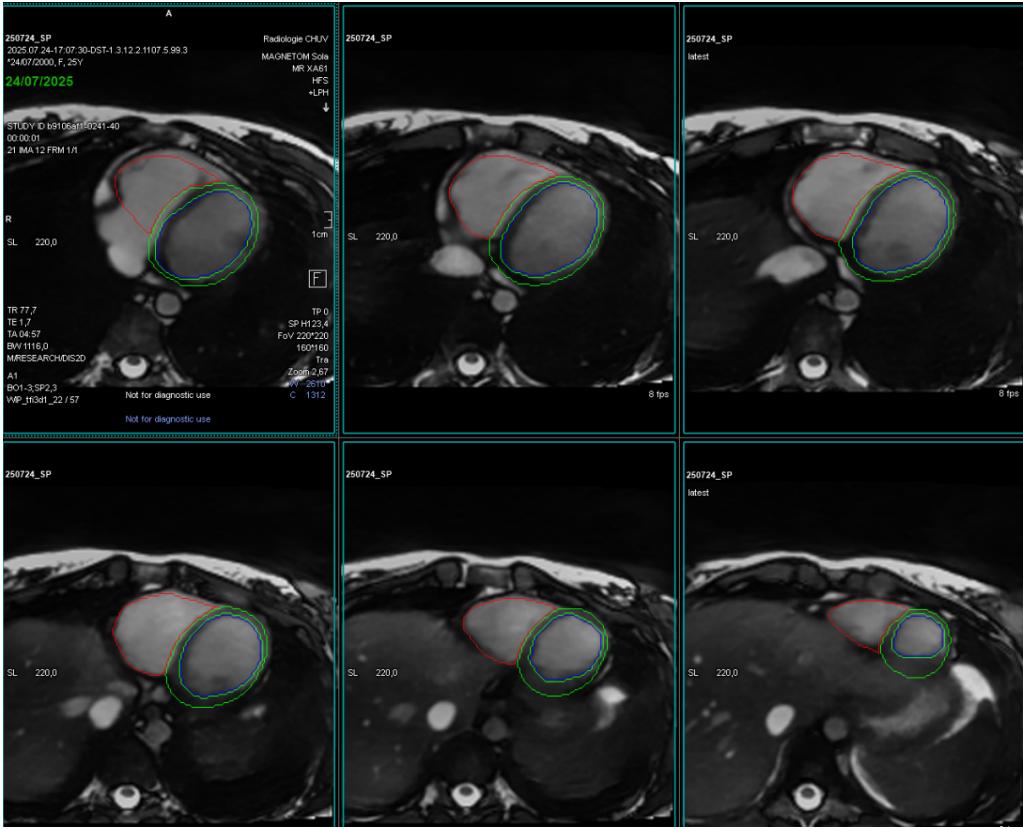


Figure 21: On-scanner reconstruction and segmentation results. Example of raw gridded reconstructions enhanced ReconAI model at different axial sections.

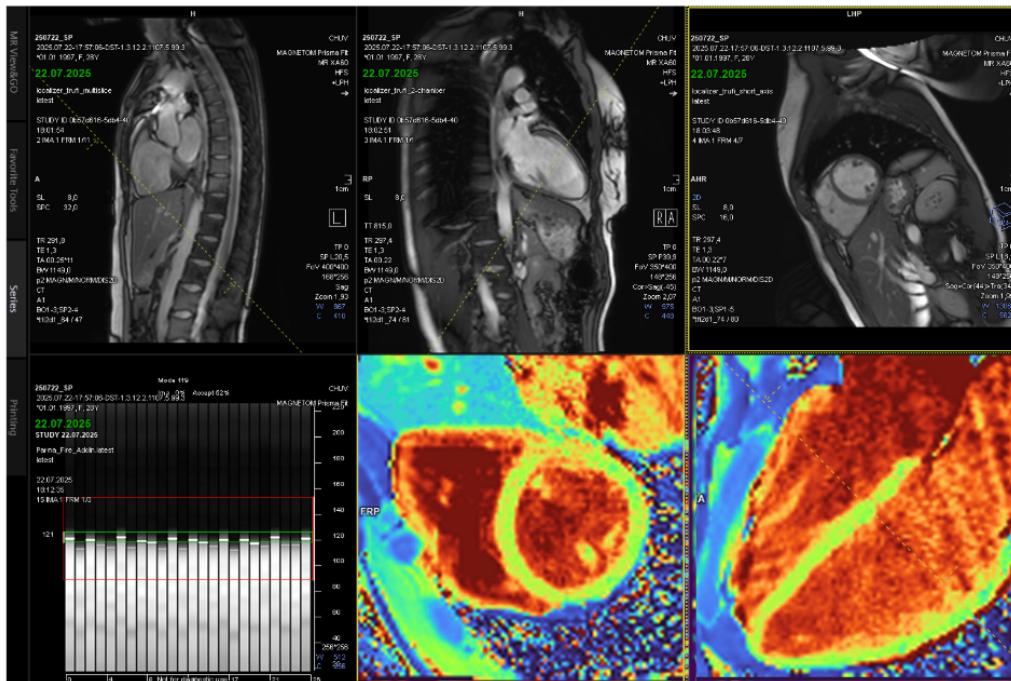


Figure 22: On-scanner reconstruction maps for T1 and T2 characterization (bottom middle and right).

7 Conclusion

This work demonstrates that deep learning-based segmentation is possible on rapidly reconstructed 5D FR cardiac MRI with precision similar to more computationally intensive methods and physiologically concordant with manual references. These findings confirm that rapid reconstruction techniques are sufficient for reliable quantitative analysis, and as such, they are suitable for clinical adoption. Along with the segmentation performance, an engineering framework was developed to integrate reconstruction, segmentation, and quantitative metric extraction into an optimized pipeline directly connected to the MRI scanner. The system enables automatic delivery of clinically relevant measurements minutes after acquisition, bridging the gap between state-of-the-art image reconstruction methods and real-time clinical applicability. Together, the scientific and engineering innovations allow for practical on-scanner implementation of AI-enhanced cardiac examination at scale, with the potential to bring increased efficiency and accessibility to clinical cardiovascular imaging.

8 Additional contributions

8.1 Density Compensation

One of my contributions to the reconstruction pipeline was the implementation and evaluation of a new density compensation strategy for radial k-space reconstruction. Previously, the reconstruction relied on a cell-count based density compensation approach, which estimates sampling density by counting the number of spokes passing through each region of k-space. In radial acquisitions, the sampling density in k-space is inherently non-uniform denser near the center and sparser toward the periphery. Without correction, this imbalance leads to intensity variations and reconstruction artifacts. Density compensation addresses this issue by applying a weight to each k-space sample, effectively normalizing the sampling distribution to approximate uniform Cartesian coverage.

To support the understanding of the density compensation process, we illustrate the trajectory of k-space acquisitions used in 3D radial MRI acquisitions and the corresponding projection on a 2D plane (Figure 23).

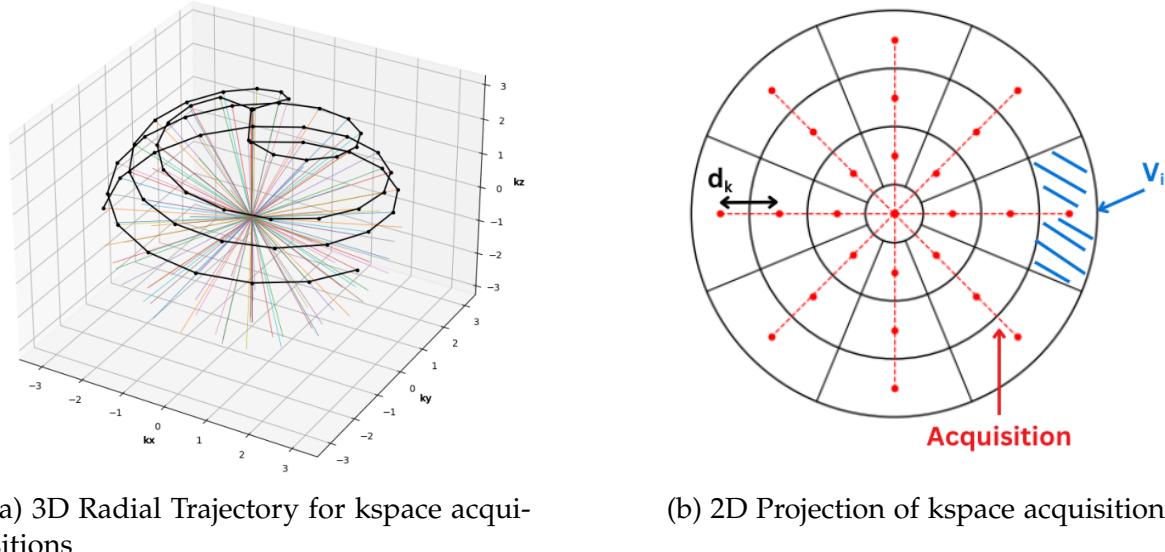


Figure 23: Representation of kspace trajectory for radial trajectory

For this density compensation analysis, we assume a uniform distribution of acquisitions in 3D space. The key idea behind this analytical approach is that density compensation weights are determined point-by-point along the acquisition trajectory and depend on the distance from the center. The farther a sampling point is from the center, the larger the surrounding volume it represents (denoted as V_i) and thus, the greater its influence on the image reconstruction (through the weight W_i).

To compute the density compensation for each sampling point, we use the following parameters:

- d_k : the distance between two neighboring points along a single readout (i.e., along a radial line)
- N_p : the number of sampling points per readout

- N_{acq} : the total number of radial readouts in the 3D acquisition

The density compensation weight for a sample point i , located at a distance $d_k * i$ from the center, is computed based on the volume of the spherical shell (or ring) it belongs to. This shell corresponds to the region between two neighboring sample radii. The weight W_i is proportional to the volume of this shell divided by the number of sampling points lying at the same radial distance. We have the following formula :

$$W_i = \frac{V_i}{N_i} \quad (6)$$

where :

- V_i is the volume of the spherical shell around radius r_i
- $N_i = 2 \times N_{acq}$ is the number of samples lying within that shell.

To compute the volume of the shell V_i , we compute the difference in volume between two spheres S_i :

$$V_i = \begin{cases} S_0 & \text{if } i = 0 \\ S_i - S_{i-1} & \text{if } i > 0 \end{cases}$$

With the volume of the sphere i :

$$S_i = \frac{4}{3}\pi((\frac{1}{2} + i)d_k)^3$$

We then obtain :

$$V_0 = \frac{\pi}{6}d_k^3$$

$$V_i = \frac{4}{3}\pi((\frac{1}{2} + i)d_k)^3 - \frac{4}{3}\pi((\frac{1}{2} + (i-1)d_k)^3 = \frac{4}{3}\pi d_k^3(\frac{1}{4} + 3i^2)$$

Finally we obtain the density compensation :

$$W_i = \begin{cases} \frac{\pi}{6}d_k^3 & \text{if } i = 0 \\ \frac{2\pi}{3N_{acq}}(\frac{1}{4} + 3i^2) & \text{if } i > 0 \end{cases}$$

To prevent overweighting of the outermost samples which may be less reliable or contribute disproportionately we introduce a Nyquist-based cutoff in the density compensation (Figures 24,25) . This cutoff is based on a scaled Nyquist radius, which approximates the boundary of fully sampled k-space.

We define the cutoff radius r_{cut} as:

$$r_{cut} = \alpha \cdot r_{Nyq}$$

where:

- $\alpha \in \mathbb{R}^+$ is a user-defined *Nyquist scaling factor*,
- r_{Nyq} is the theoretical Nyquist radius, approximated by:

$$r_{Nyq} = \sqrt{\frac{2N_{acq}}{\pi}}$$

In practice, we apply the cutoff to the density compensation weights W_i as follows:

$$W_i^{\text{cut}} = \begin{cases} W_i & \text{if } r_i \leq r_{\text{cut}} \\ W_{\text{clip}} & \text{if } r_i > r_{\text{cut}} \end{cases}$$

where $W_{\text{clip}} = \max\{W_i \mid r_i \leq r_{\text{cut}}\}$ is the maximum weight within the Nyquist region. This ensures the weights remain bounded and smooth while suppressing excessive influence from outer k-space samples.

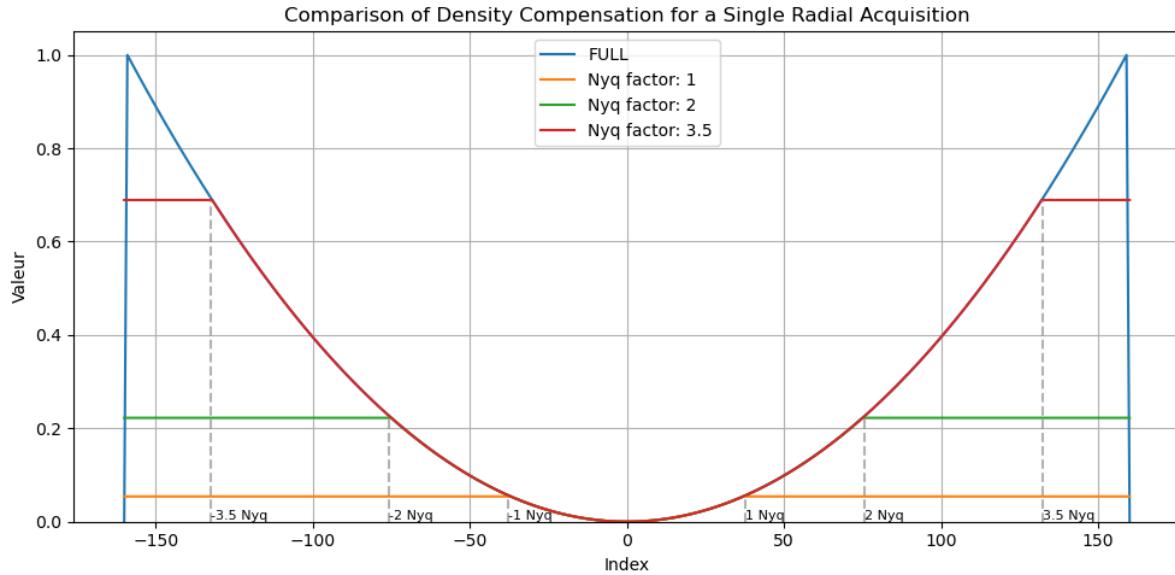


Figure 24: Density-compensation weights (DCF) as a function of normalized k-space radius for multiple Nyquist factors (1,2,3.5 and no cut-off). The specific cut-off radius is highlighted, showing how the Nyquist setting determines the retained high-frequency support.

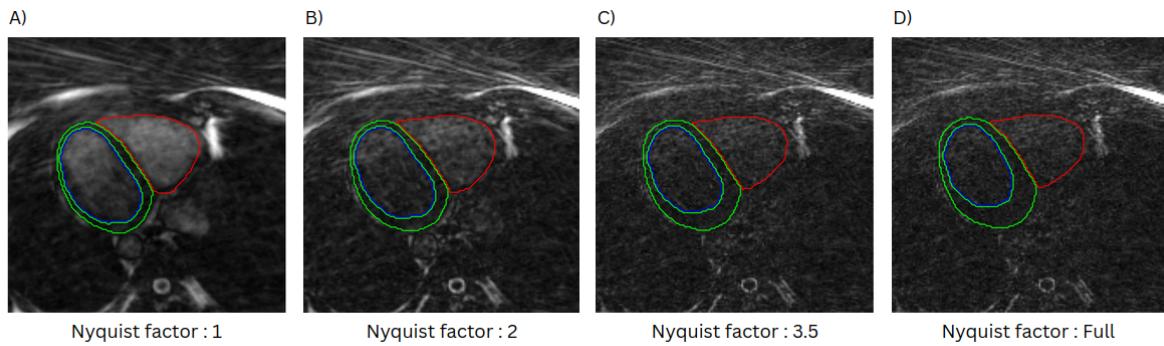


Figure 25: Effect of Nyquist factor on density compensation in gridded reconstruction. End-diastolic heart images are shown with Nyquist factors of A) 1, B) 2, C) 3.5, and D) no cutoff.

8.2 Extraction of 4CH and SAX Views from Segmented MRI Reconstructions

Following segmentation of the reconstructed 4D MRI volumes, a nearly finalized post-processing workflow was implemented to reformat the data into clinically relevant cardiac views. Using the segmented masks of the left and right ventricles, principal component analysis (PCA) was applied to estimate the heart’s principal axes and derive its orientation. This orientation was then used to compute affine transformation matrices to resample the original 4D volume into SAX and 4CH views (Figure 26). The current implementation preserves spatial resolution and temporal dynamics, enabling motion-resolved cardiac planes to be reconstructed directly from the original axial, sagittal, and coronal images. This step facilitates standardized visualization and quantitative analysis in orientations familiar to clinicians, without requiring additional acquisitions.

8.3 Principal Component Analysis for Cardiac View Orientation

To derive cardiac-oriented planes from the segmentation, we applied Principal Component Analysis (PCA) to the coordinates of the left ventricular (LV) mask . Let $\mathbf{x}_i \in \mathbb{R}^3$ denote the coordinates of all LV voxels. We first compute the centroid:

$$\bar{\mathbf{x}} = \frac{1}{N} \sum_{i=1}^N \mathbf{x}_i$$

The coordinates are then centered:

$$\mathbf{X}_c = \{\mathbf{x}_i - \bar{\mathbf{x}}\}$$

and the covariance matrix is computed as:

$$\mathbf{C} = \frac{1}{N-1} \mathbf{X}_c^\top \mathbf{X}_c$$

Eigen-decomposition of \mathbf{C} yields the principal axes $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ and variances $\lambda_1 \geq \lambda_2 \geq \lambda_3$:

$$\mathbf{C}\mathbf{v}_k = \lambda_k \mathbf{v}_k$$

The first principal axis \mathbf{v}_1 corresponds to the **long axis** of the LV, which in theory should pass through the apex and the center of the mitral valve plane at the base. The second principal axis \mathbf{v}_2 lies within the **SAX plane**, orthogonal to the long axis, and the third principal axis \mathbf{v}_3 is perpendicular to both.

From these axes, a new affine transformation is defined to resample the 4D volume into:

- **SAX view:** stack of slices perpendicular to \mathbf{v}_1 , covering the LV from base to apex.
- **4CH view:** plane containing \mathbf{v}_1 and \mathbf{v}_3 , intersecting both ventricles and atria.

This PCA-based approach ensures consistent and reproducible alignment across subjects.

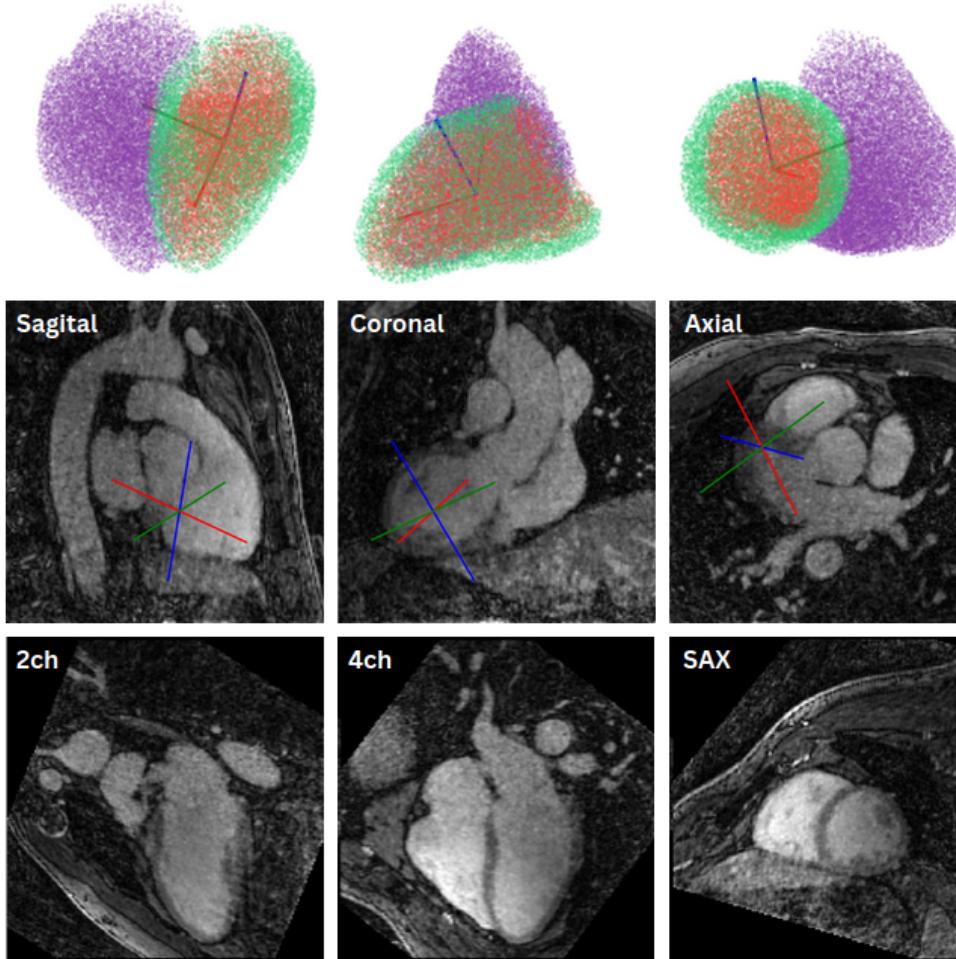


Figure 26: From segmentation to cardiac views at diastole. (Top row) Multi-view renderings of the cardiac segmentation. (Middle row) Axial, sagittal, and coronal slices annotated with the PCA-derived axes from the diastolic segmentation mask (long axis direction—red, short axis direction—blue, orthogonal direction—green). (Bottom row) Standard cardiac planes (2CH, 4CH, SAX) produced by rotating the diastolic volume into the PCA-defined cardiac coordinate system.

8.4 Conference and Retreat Presentations

In addition to this thesis work, I contributed as second author to an abstract accepted for presentation at the SCMR 2026 Conference in Rio de Janeiro, Brazil (*Deep-Learning-Based Segmentation of Gridded Reconstructions from Undersampled Free-Running 4D Data for On-Scanner Deployment*). I also prepared and presented a scientific poster at the 2025 CIBM–CHUV–MR Retreat in Bourgogne (Figure 30).

References

- [1] World Health Organization, “Cardiovascular diseases (CVDs)”, 2021. Available: [https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-\(cvds\)](https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds)).
- [2] Real Engineering, *L'incroyable Ingénierie Des IRM*. Available: <https://www.youtube.com/watch?v=N1YXqRG7lus&t=828s>.
- [3] Danish Research Centre for Magnetic Resonance (DRCMR), “What is MRI?”, 2009. Available: <https://www.drcmr.dk/MR>.
- [4] Di Sopra, L., Piccini, D., Coppo, S., Stuber, M., & Yerly, J. (2019). An automated approach to fully self-gated free-running cardiac and respiratory motion-resolved 5D whole-heart MRI. *Magnetic Resonance in Medicine*, 82(6), 2118–2132. <https://doi.org/10.1002/mrm.27898>
- [5] Pednekar A.S., Kurkure U., Muthupillai R., Flamm S.D., Kakadiaris I.A. Automatic active contour segmentation of cardiac MRI. In: Proceedings of SPIE Medical Imaging 2006: Image Processing. Vol 6144. 2006:61444L. doi:10.1117/12.653429
- [6] Ogier A, Sdika M, Foure A, Le Troter A, Bendahan D. Individual muscle segmentation in MR images: A 3D propagation through 2D non-linear registration approaches. *Annu Int Conf IEEE Eng Med Biol Soc*. 2017 Jul;2017:317-320. doi: 10.1109/EMBC.2017.8036826. PMID: 29059874.
- [7] Ronneberger O., Fischer P., Brox T. U -Net: Convolutional Networks for Biomedical Image Segmentation. In: MICCAI 2015. Springer; 2015:234–241.
- [8] Çiçek Ö., Abdulkadir A., Lienkamp S.S., Brox T., Ronneberger O. 3D U -Net: Learning Dense Volumetric Segmentation from Sparse Annotation. In: MICCAI 2016. Springer; 2016:424–432.
- [9] Milletari F., Navab N., Ahmadi S.-A. V -Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation. In: 2016 4th Int. Conf. 3D Vision (3DV); 2016:565–571.
- [10] Isensee F., Jaeger P.F., Kohl S.A.A., Petersen J., Maier-Hein K.H. nnU-Net: a self-adapting framework for deep learning-based biomedical image segmentation. *Nat Methods*. 2021;18(2):203–211. doi:10.1038/s41592-020-01008-z
- [11] Chen C., Qin C., Qiu H., et al. Deep Learning for Cardiac Image Segmentation: A Review. *Front Cardiovasc Med*. 2020;7:25. doi:10.3389/fcvm.2020.00025
- [12] Feng L., Axel L., Chandarana H., et al. XD -GRASP: Golden-angle radial MRI with reconstruction of extra motion-state dimensions using compressed sensing. *Magn Reson Med*. 2016;75(2):775–788. doi:10.1002/mrm.25818
- [13] Feng L., Coppo S., Piccini D., et al. Five-Dimensional Whole-Heart Sparse MRI. *Magn Reson Med*. 2017;78(2):589–603. doi:10.1002/mrm.26476

- [14] Roy C.W., Cheng J.Y., Lustig M., et al. Free-running cardiac and respiratory motion-resolved 5D whole-heart imaging: framework and applications. *J Cardiovasc Magn Reson.* 2022;24(1):71. doi:10.1186/s12968-022-00871-3
- [15] Holtackers R.J., et al. Free-Running Cardiac and Respiratory Motion-Resolved 5D Whole-Heart Imaging: A Review. *Diagnostics.* 2024;14(17):1946. doi:10.3390/diagnostics14171946
- [16] Full P.M., Isensee F., Jäger P.F., Maier-Hein K.H. Studying Robustness of Semantic Segmentation under Domain Shift in Cardiac MRI. *arXiv:2011.07592*, 2020.
- [17] Guan H., Liu M. Domain Adaptation for Medical Image Analysis: A Survey. *IEEE Trans Biomed Eng.* 2022;69(3):1173–1185. doi:10.1109/TBME.2021.3122293
- [18] Lu C., Zheng S., Gupta G. Unsupervised Domain Adaptation for Cardiac Segmentation: Toward Structure Mutual Information Maximization. In: CVPR Workshops; 2022:1841–1850.
- [19] Inati S.J., Naegele J.D., Zwart N.R., et al. ISMRM Raw Data Format: A Proposed Standard for MRI Raw Datasets. *Magn Reson Med.* 2017;77(1):411–421. doi:10.1002/mrm.26089
- [20] Hansen M.S., Sørensen T.S. Gadgetron: An Open Source Framework for Medical Image Reconstruction. *Magn Reson Med.* 2013;69(6):1768–1776. doi:10.1002/mrm.24389
- [21] Veldmann M.H., Willig-Onwuachi J., et al. Open-Source MR Imaging and Reconstruction Workflow. *Magn Reson Mater Phy.* 2022;35(2):275–289. doi:10.1007/s10334-021-00949-9
- [22] Roy C.W., Vo H.Q., Hann E., Lustig M., Cheng J.Y. Respiratory motion compensation for free-running 5D whole-heart MRI using an extra motion dimension in XD-GRASP. *J Cardiovasc Magn Reson.* 2021;23(1):33. doi:10.1186/s12968-021-00729-5
- [23] Feng L., Axel L., Chandarana H., Block K.T., Sodickson D.K., Otazo R. XD-GRASP: Golden-angle radial MRI with reconstruction of extra motion-state dimensions using compressed sensing. *Magn Reson Med.* 2018;79(2):2992–3004. doi:10.1002/mrm.26818
- [24] Ogier A.C., Boros K., Roy C.W., Romanin L., Prsa M., Stuber M., Yerly J. Deep learning-based respiratory motion-resolved 5D whole-heart MRI: Towards efficient on-scanner deployment. *medRxiv.* 2025. doi:10.1101/2025.02.07.24302932
- [25] WildFFS: A modular framework for free-running cardiac MRI reconstruction, segmentation, and communication. Available at: <https://gitlab.com/wildffs> (accessed 2025).
- [26] Borsos KB, Ogier AC, Roy CW, Romanin L, Stuber M, Prsa M, Kustner T, Yerly J. Efficient deep-learning-based reconstruction of Ferumoxytol-enhanced whole-heart 5D cardiac MRI. *Proc Int Soc Magn Reson Med.* 2025; 1187

Appendix

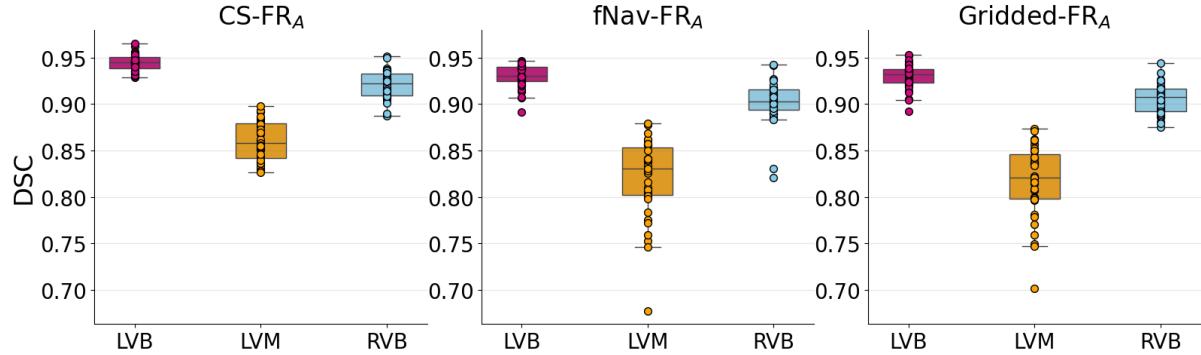


Figure 27: DSC for left ventricular blood pool (LVB), left ventricular myocardium (LVM), and right ventricular blood pool (RVB), with boxplots comparing anatomical region in each reconstruction models.

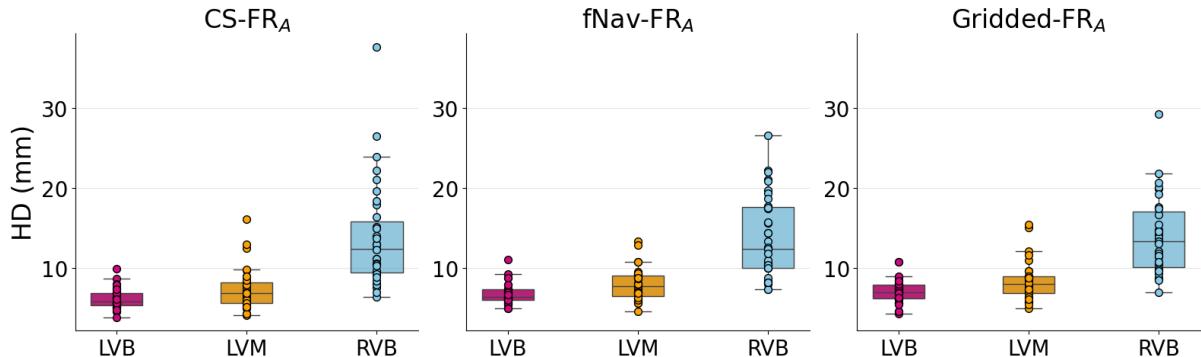


Figure 28: HD for left ventricular blood pool (LVB), left ventricular myocardium (LVM), and right ventricular blood pool (RVB), with boxplots comparing anatomical region in each reconstruction models.

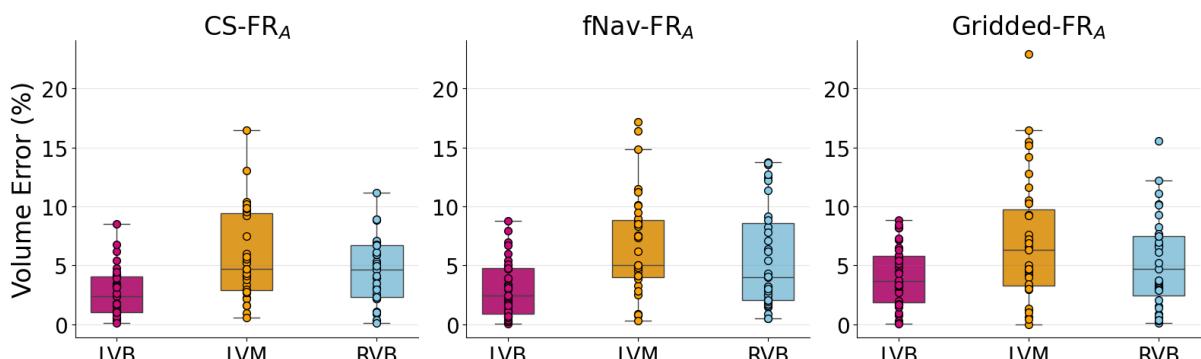


Figure 29: RVD for left ventricular blood pool (LVB), left ventricular myocardium (LVM), and right ventricular blood pool (RVB), with boxplots comparing anatomical region in each reconstruction models.

Deep-Learning-Based Segmentation of Gridded Reconstructions from Undersampled Free-running 4D Data for On-Scanner Deployment

Gabriel Paffi¹, Augustin C. Ogier¹, Jérôme Yerly^{1,2}, Angela Rocca³, Isabel Montón Quesada⁴, Jean-Baptiste Ledoux^{1,3}, Robert J. Holtackers^{1,4}, Matthias Stuber^{1,2}, Roger Hullin³, David Rotzinger¹, Ruud B. van Heeswijk¹

¹ Department of Radiology, Lausanne University Hospital (CHUV) and University of Lausanne (UNIL), Lausanne, Switzerland; ² CIBM Center for BioMedical Imaging, Lausanne, Switzerland; ³ Cardiology Service, Cardiovascular Department, Lausanne University Hospital (CHUV) and University of Lausanne (UNIL), Lausanne, Switzerland

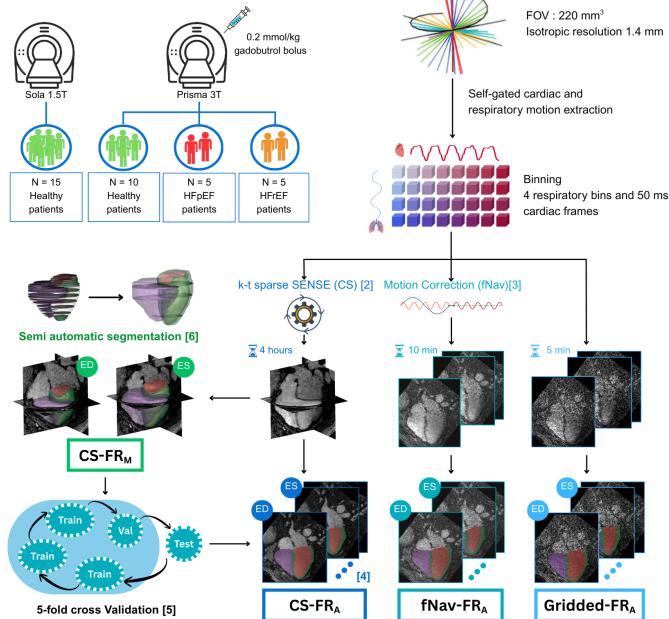
Background

- Free-running MRI enables self-gated, free-breathing cardiac imaging, more comfortable and accessible for patients [1].
- It generates **large 5D datasets** (3D + cardiac + respiratory), which are impractical to segment manually.
- No current solution performs both reconstruction and segmentation directly **on the scanner**.

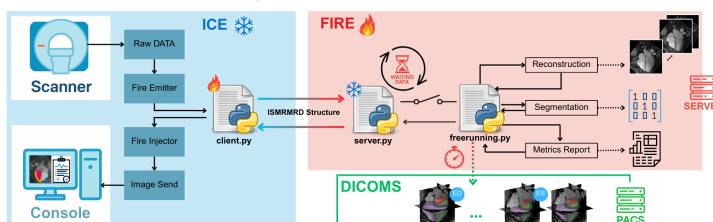
Objective

Deploy a **fully integrated pipeline** for cardiac image reconstruction and segmentation directly **on the MRI scanner**.

Methods



On-scanner deployment



Conclusion

We demonstrated that our deep learning segmentation models deliver **clinically equivalent performance on rapidly reconstructed gridded and fNav images** compared to the high-quality CS reference. This shows that CS is not necessary for accurate whole-heart delineation, enabling immediate post-acquisition analysis. These models have been implemented in an **on-scanner reconstruction framework for on-scanner deployment**. Future prospective studies should quantify their impact on workflow efficiency and diagnostic accuracy to **pave the way for 5D cardiac MRI in clinical routine**.

References : [1] Di Sopra L. et al. MagnResonMed, 2019, 82(6); [2] Feng L. et al. MagnResonMed. 2018, 79(2); [3] Christopher W. Roy et al. Cardiovascular Magnetic Resonance, 2021, 23(33); [4] Ogier A. C. et al. Medrxiv, 2025; [5] Isensee F. et al. NatMeth, 2021, 18(2); [6] Ogier A. C. et al. IEEE Eng Med Biol Soc. 2017;317-320

Results

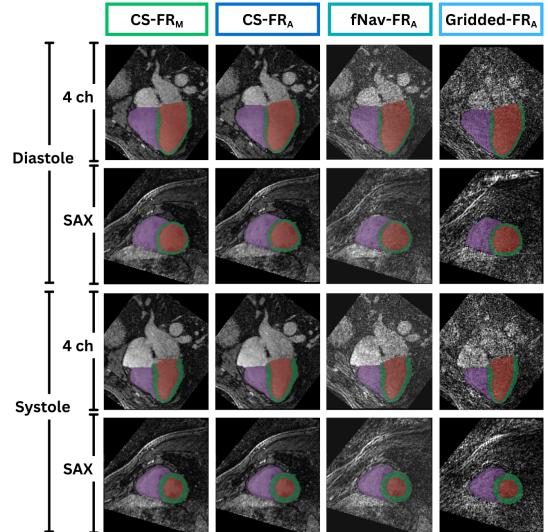


Figure 1: **Qualitative segmentation images.** Visual comparison of manual and DL-based automatic segmentations with short-axis and pseudo four-chamber views shown for each phase.

	Region	Model	Model
DSC (-)	LVB	CS-FRA	Gridded-FRA
	LVM	0.94 ± 0.01	0.89 ± 0.01
RVD (%)	RVB	0.86 ± 0.02	0.82 ± 0.04
	LVB	0.92 ± 0.02	0.90 ± 0.02
	LVM	2.37 ± 2.06	3.06 ± 2.39
	RVB	5.76 ± 3.74	4.08 ± 2.54
	LVB	6.58 ± 4.32	7.02 ± 5.32
	RVB	4.51 ± 2.81	5.57 ± 4.36

Table 1: **Quantitative segmentation performance** across reconstruction models. Dice similarity coefficient and relative volume difference are reported for the LVB, LVM and RVB.

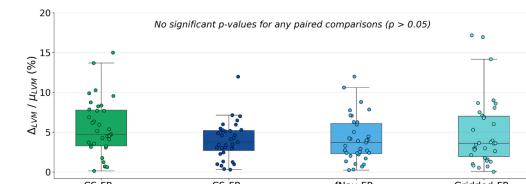


Figure 2: **Physiology consistency metrics.** Myocardium volume consistency error of LV between the end-diastole and end-systole volume for manual and automatic segmentations.

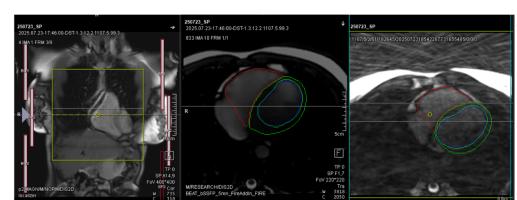


Figure 3: **On-scanner reconstruction and segmentation images.** From left to right: (1) localizer used for planning, (2) gridded reconstruction enhanced with DL-based denoising, and (3) raw gridded reconstruction.

Figure 30: Poster for the 11th CIBM-CHUV-MR Retreat in Bourgogne