A Framework for Simulating Cardiac MR Images With Varying Anatomy and Contrast

Sina Amirrajab, Yasmina Al Khalil, Cristian Lorenz, Jürgen Weese, Josien Pluim, *Fellow, IEEE*, and Marcel Breeuwer

***Abstract*— One of the limiting factors for the development and adoption of novel deep-learning (DL) based medical image analysis methods is the scarcity of labeled medical images. Medical image simulation and synthesis can provide solutions by generating ample training data with corresponding ground truth labels. Despite recent advances, generated images demonstrate limited realism and diversity. In this work, we develop a flexible framework for simulating cardiac magnetic resonance (MR) images with variable anatomical and imaging characteristics for the purpose of creating a diversified virtual population. We advance previous works on both cardiac MR image simulation and anatomical modeling to increase the realism in terms of both image appearance and underlying anatomy. To diversify the generated images, we define parameters: 1) to alter the anatomy, 2) to assign MR tissue properties to various tissue types, and 3) to manipulate the image contrast via acquisition parameters. The proposed framework is optimized to generate a substantial number of cardiac MR images with ground truth labels suitable for downstream supervised tasks. A database of virtual subjects is simulated and its usefulness for aiding a DL segmentation method is evaluated. Our experiments show that training completely with simulated images can perform comparable with a model trained with real images for heart cavity segmentation in mid-ventricular slices. Moreover, such data can be used in addition to classical augmentation for boosting the performance when training data is limited, particularly by increasing the contrast and anatomical variation, leading to better regularization and generalization. The database is publicly available at `https://osf.io/bkzhm/` and the simulation code will be available at `https://github.com/sinaamirrajab/CMRI_Simulation`.**

***Index Terms*— Cardiac MRI, image simulation, image synthesis, database generation, physics-based, deep-learning.**

Manuscript received 20 June 2022; revised 1 October 2022; accepted 14 October 2022. Date of publication 19 October 2022; date of current version 2 March 2023. This work was supported in part by the OpenGTN Project, European Union in the Marie Curie Innovative Training Networks (ITN) Fellowship Program, under Project 764465 (website: opengtn.eu). *(Corresponding author: Sina Amirrajab.)*

Sina Amirrajab, Yasmina Al Khalil, and Josien Pluim are with the Department of Biomedical Engineering, Eindhoven University of Technology, 5600 MB Eindhoven, The Netherlands (e-mail: s.amirrajab@tue.nl; y.al.khalil@tue.nl; j.pluim@tue.nl).

Cristian Lorenz and Jürgen Weese are with the Philips Research Laboratories, 22335 Hamburg, Germany (e-mail: cristian.lorenz@philips.com; juergen.weese@philips.com).

Marcel Breeuwer is with the Department of Biomedical Engineering, Eindhoven University of Technology, 5600 MB Eindhoven, The Netherlands, and also with Philips Healthcare, MR Research and Development-Clinical Science, 5684 PC Best, The Netherlands (e-mail: marcel.breeuwer@philips.com).

Digital Object Identifier 10.1109/TMI.2022.3215798

I. Introduction

SIMULATION and synthesis have recently received great recognition in the medical imaging community. This has been achieved thanks to the solutions that image generation approaches can provide to medical data challenges such as data scarcity, privacy, expert dependency, and expensive collection procedure. The synergies between different approaches, various applications, challenges, and opportunities have been highlighted in the recent editorial by Frangi et al. [1]. With developments in machine learning methods, there is an ever-growing demand for a large heterogeneous medical database that represents enough variability in both anatomical representation and image appearances. Such a diverse database can pave the way for developing, validating, and benchmarking accurate and robust medical image analysis methods that can be employed in routine clinical practice.

In attempts towards generating realistic Cardiac Magnetic Resonance (CMR) images, there has been recent progress in three categories: i) physics-driven image simulation; ii) data-driven image synthesis; and iii) hybrid image generation. In the physics-driven simulation approach, the underlying physics for MR contrast formation, governed by Bloch equations, is implemented and a computerized anatomical model is used to resemble a virtual patient. MR relevant tissue properties are assigned to each label class in the anatomical model to be fed to an image simulator to produce image contrast. The second category uses existing cardiac MR image data to train a generative model that learns the appearance of images and in turn synthesizes similar-looking images. In the final category, the hybrid approach combines patient image data with a biophysical model of the heart to generate images with altered geometry that is informed by mechanical motion parameters. For the purpose of generating diversified images with variable contrast and anatomy, each approach has its distinct advantages and disadvantages.

A. Physics-Driven Image Simulation

Image simulation is performed by combining a spatio-temporal model representing anatomy of interest and a simulator that encompasses the physics of image formation given a set of controllable parameters. In this category, the Virtual Imaging Platform provides an integrated open-access platform for sharing object models and multi-modality medical image simulation pipelines [2]. However, it consists

of only one MR sequence with pre-defined scan parameters on one anatomical model with a simplistic heart geometry and a limited number of surrounding anatomical structures. Among one of the first attempts to generate CMR image data, Tobon-Gomez et al. [3] investigate physics-based image simulation using the MRISIM simulator developed by Kwan et al. in [4] and anatomical models generated from eXtendec CArdiac and Torso (XCAT) phantoms [5]. The authors put effort into making the simulated images more realistic by modeling the left ventricular papillary muscles and trabeculation. They use real patient data to create a simplified mathematical model of cylindrical objects for the papillary muscles and small discs at random regions adjacent to the ventricular wall for heart trabeculation. More recently, based on the original XCAT anatomical phantom, Wissmann et al. in [6] developed a numerical simulation framework suitable for optimizing CMR sampling trajectory, post-processing methods, and reconstruction strategies in the presence of beating and breathing motions, which is referred to as MRXCAT. Despite the modifications to the anatomical model and the simulation approach, the resulting simulated images for both approaches are still far from reality in terms of the detailed structures of the used heart model, the number of neighboring organs visible in the field of view, and the realism of the image contrast and resolution.

B. Data-Driven Image Synthesis

With the advent of generative modelling and the emergence of various techniques for synthesizing images using available clinical data, new methods have been adopted by researchers in the medical imaging community [23], [24]. In particular, Generative Adversarial Networks (GANs) [22] are at the center of most recently proposed models. To provide labeled data for training a segmentation model, some works proposed multi-modal style transfer method for transferring the image appearance (known as style) of the real images to the anatomical information (known as content) of another imaging modality where the content of the synthetic data comes from the heart annotations of cine CMR images in [28] and [29], or CT cardiac images in [30]. While these methods are able to generate realistic-looking images, they allow limited control over the image synthesis procedure, meaning that neither the underlying anatomy content, nor the local tissue intensity and global image contrast of the generated images can be controlled. Disentangling the anatomy factors from the modality features, Joyce et al. in [31] designed separate variational auto encoder (VAE) models [32] for simultaneously learning multi-tissue anatomical model, a deformation model, and an image intensity rendering model. Factorizing the information in the data in this way can provide partial control over generating variable anatomy and overall image appearance. However, the synthetic data may not necessarily represent an accurate anatomy that is required for the downstream supervised tasks. Furthermore, the contrast is not controllable locally and the generated tissue intensity is not based on the physics of MR signal evolution.

C. Hybrid Image Generation

Combining a biophysical model of the heart with a set of real clinical images in a hybrid approach, Prakosa et al. in [7] propose to use a registration algorithm for generating realistic-looking images with controllable cardiac motion. After fusing a surface model of the heart into real cardiac images, the heart geometry is adapted according to a predefined motion model. Along the same direction, in recent work by Duchateau et al. [8], an optimized pipeline is proposed for reducing the registration errors and improving the model-to-image adaptation, therefore generating more realistic images. Similarly, Zhou et al. in [9] introduce a multi-modality pipeline to generate cine CMR images, tagged CMR, and echocardiography sequences from the same virtual patient. With the aim of augmenting data, similar work was done by Acero et al. in [27], who use a heart statistical model of deformation to generate similar looking examples of images from a cardiac MRI database through altering the anatomy. While the generated images using the above-mentioned approaches are realistic in terms of the underlying anatomy, they depend on the availability of real cardiac images. Combining the controllable anatomical model with the imaging features has gained more attention in recent years. Abbasi-Sureshjani et al. in [26] and Amirrajab et al. in [25] propose to integrate the anatomical information of the XCAT phantoms [5] with modality-specific appearance of real data to synthesize data for creating a virtual database of realistic CMR images with ground truth labels. Although the anatomical variability can be created using the heart model, new image appearances can not be generated. Furthermore, the gray values of the images are not governed by the underlying physics of MRI and controlling image contrast at the tissue level is not yet feasible.

D. Motivation and Contribution

In this work, we develop a flexible framework suitable for generating a database of heterogeneous cardiac MR images that present variations in acquisition parameters, tissue properties, image contrasts, and anatomical representation. The proposed framework is tailored towards generating a plethora of realistic-looking images with corresponding ground truth labels. We build upon and advance previous works in both areas of anatomical modeling and cardiac MR image simulation. The simulation pipeline consists of three main elements: i) a parameterized anatomical model based on an improved version of the XCAT phantoms; ii) a set of controllable tissue parameters for more than 10 tissue types within various organs; iii) an optimized CMR simulation model to generate images with variable contrast. We save the output of the simulation together with multi-tissue ground truth labels in the NIfTi file format with proper metadata of acquisition parameters.

The main contributions of this paper can be summarized as follows:

- We enhance the heart model by adding patient-specific detailed structures for the trabeculation anatomy of the left and right ventricles. Moreover, we make use of available anatomical parameters in the XCAT phantom to

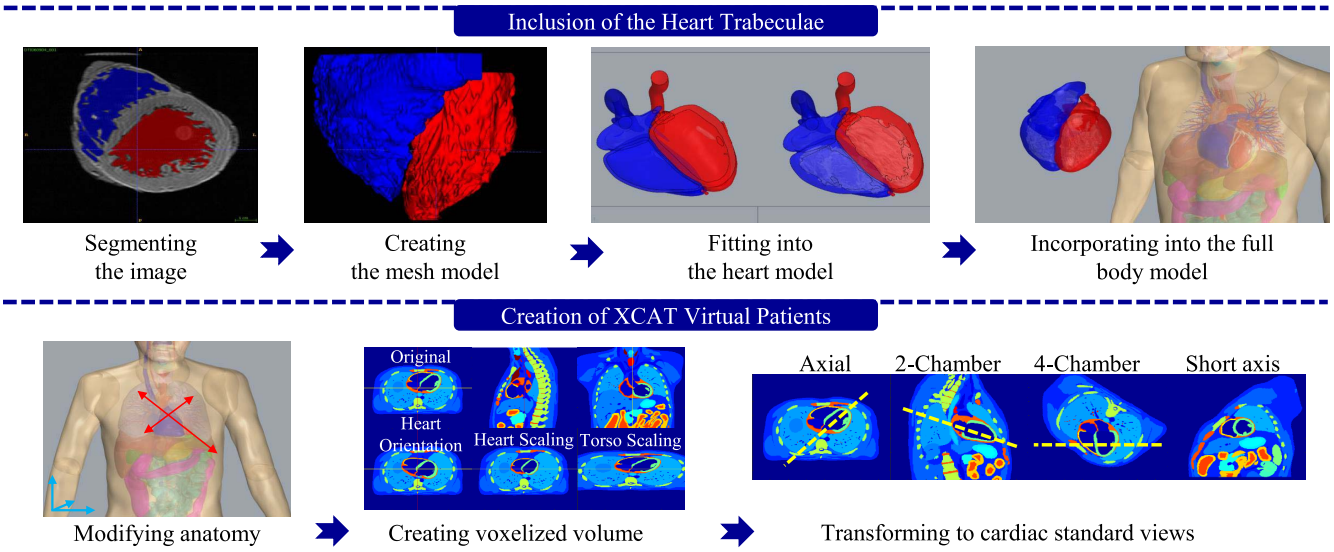

Fig. 1. Inclusion of the trabeculation structures into the XCAT phantom and creation of virtual patients.

create virtual subjects with variable organ size, geometry, volume and location.

- We increase the realism of the CMR image simulation by assigning numerous tissue properties to various organs that are visible in the imaging field of view. We utilize physics-based analytical solutions for fast MR contrast computation with controllable acquisition parameters.
- We create a population of virtual patients (25-30 cardiac phases each) with different anatomical characteristics for cine cardiac MR image simulation for functional analysis and make the database with ground truth labels publicly available to the medical imaging research community upon request via `https://osf.io/bkzhm/`

An initial database of CMR images was simulated on virtual patients using the early version of our framework for the recent work of Al Khalil et al. in [33]. We showed that such a heterogeneous database of images with variation in both anatomy and appearance can be used for pre-training a deep learning cardiac segmentation model that can generalize better to the variability of real cardiac data and experiments showed that similar performance can be achieved when replacing up to 80% of the real data with simulated data. The initial results indicate the usefulness of a simulated database of CMR images for transfer learning in medical imaging. In this work, we made substantial improvements to our simulation framework by 1) optimizing sequence parameters and image contrast, 2) improving the anatomical model by incorporating patient-specific details of the heart trabeculation, 3) improving the image realism by increasing the number of tissue properties used for simulation, and 4) by modeling the partial volume effects using sampling and filtering in the k-space. Here, we present our framework for cardiac MR image simulation and provide detailed explanation of each module. In addition, we evaluate the usefulness of simulated data in the context of data augmentation for training a neural network for cardiac cavity segmentation. We show that with improved image realism we can directly add the newly simulated data to the real image and demonstrate the benefits of data augmentation and data replacement using the simulated images in this study, which was not possible before due to limited image realism.

The structure of the paper is as follows. We give a brief overview of the XCAT anatomical phantom in section II-A.1, the explanation of our approach for incorporating more anatomical details into the XCAT heart model in section II-A.2, the introduction of the anatomical parameters for creating virtual patients in section II-A.3, the description of steps involved for cardiac MR simulation in section II-B, the experimental design for evaluating the generated data in a deep-learning setup in section II-C, and qualitative and quantitative analysis of the results in sections III-A and III-B, followed by discussion and conclusion in IV and V, respectively.

II. Proposed Framework

A. Virtual Subjects

1) The XCAT Phantom: The 4D eXtended CArdiac and Torso phantom (XCAT) [5] is used as the basis for creating virtual patients for image simulation. The anatomies in the XCAT phantom are based on segmented patient data, modelled as Non-Uniform Rational B-Splines (NURBS) surfaces to accurately capture each structure in the body. Defined anatomies using NURBS surfaces provide a realistic representation of a patient with great flexibility to alter and create models with geometrical deformations. For accurately modeling the regional myocardial contracting and twisting motion of the left and right ventricles, the cardiac motion model was improved using the analysis of tagged CMR images [34]. The analysis of the tagged MRI data resulted in a comprehensive motion model that can produce accurate heart geometries at any given time point during the cardiac cycle and represent the complex geometrical deformation of the beating heart. Note that the motion of the XCAT heart includes the shortening along the long axis of the heart that accounts for the through-plane motion during image simulation. These features make the

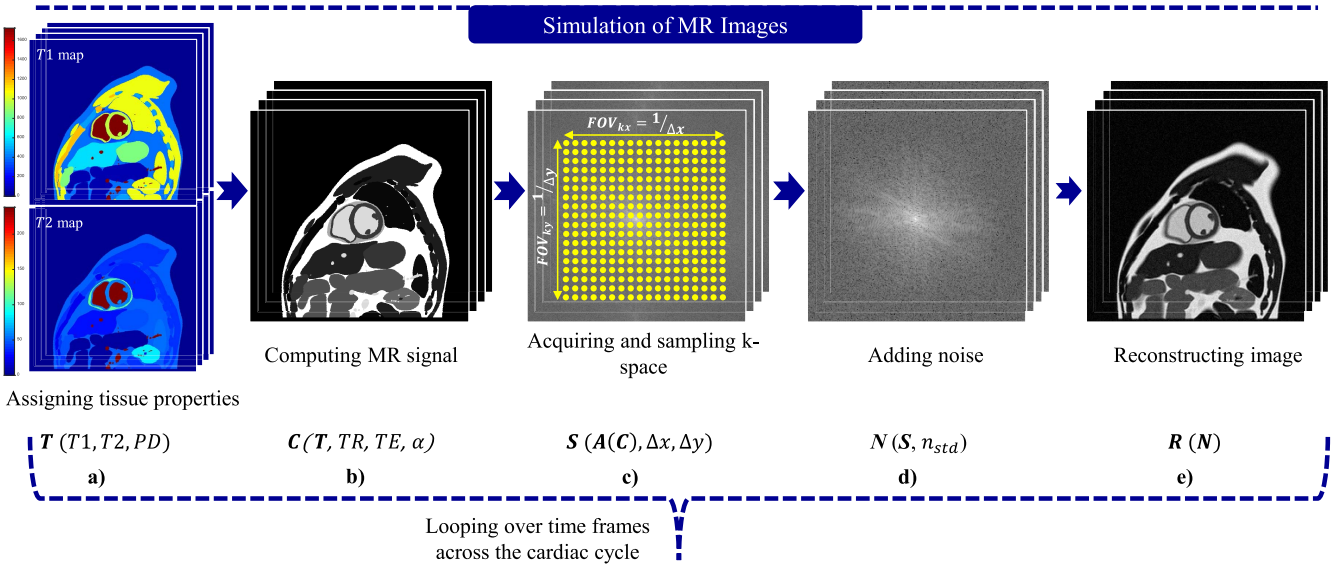

Fig. 2. Simulation pipeline of the cardiac MR images consisting of a) module for assigning MR tissue properties, b) for computing image contrast, c) for data acquisition and sampling k-space, d) for noise addition and e) for image reconstruction.

XCAT a suitable source for creating times series of varying models for simulating the dynamics of the heart. However, the anatomical details of the trabeculae muscles are missing in the current version of the XCAT phantom which hampers both the realism of the image simulation and the performance of image-based assessment of cardiac function.

2) Inclusion of Myocardial Trabeculae in the XCAT Heart: Previous studies have explored the importance of the papillary muscles and trabeculae anatomy in analyzing cardiac images. Their inclusion in the left ventricular cavity or left myocardial volumes can have a considerable impact on the final assessment of the cardiac function [10]. Therefore, from a simulation perspective, the construction of a detailed heart model that comprises the mentioned substructures is crucial for quantitative analysis of the images.

To model the geometry of the trabeculae of the myocardium, we utilized open access ex-vivo high-resolution 3D MRI data of normal human heart with $256 \times 256 \times 136$ matrix size and $0.4297 \times 0.4297 \times 1\ mm^3$ voxel dimensions [11]. As shown in Figure 1, the irregularity of the trabeculae muscular geometry was accurately segmented from the images. The spatial patterns of the tiny structures of the trabeculae anatomy were accurately captured by manually segmenting the right and left ventricle of the heart slice-by-slice using ITK-SNAP software [12]. It was then converted to a 3D polygon mesh model, while preserving the details of the jagged-like structures, and transformed into the inner surfaces of the XCAT heart chambers. The alignment was done for the end-diastolic phase of the heart and the motion model was applied to the trabeculae mesh model to be altered smoothly during the cardiac cycle. These steps can be seen in Figure 1.

3) Anatomical Parameters: To provide large quantities of varying anatomies in CMR image simulation, we use the pre-defined XCAT parameters to create new subjects. We assign a specific set of anatomical parameters for body size and heart geometry and location with respect to neighboring organs. These parameters include body scaling in different dimensions, orientation angles and translation of the heart within the torso, and LV volumes at end diastole and end systole. As shown in Figure 1 the parameters of the anatomy are modified to create a new virtual patient. The XCAT program outputs a three-dimensional volume of binary labels, shown with different colors, for a desired body area that is defined by field of view, matrix size, and the voxel resolution. The voxelized XCAT anatomy is shown for three orthogonal planes (axial, sagittal and coronal), as well as three examples for changing the anatomical parameters of the axial view. We create an isotropic 4D (3D + time) model for each virtual subject. The slices of the volume are presented in the axial view, while the standard views for assessment of cardiac function using MRI can be different. Since it is common to scan in short-axis view of the heart, rotation and re-slicing is performed on the axial slices following the recommendations provided in the CMR pocket guide [39]. The rotation angles are obtained given the spatial location of the heart within the torso provided by the XCAT metadata information. The generated subjects are used as the input for our CMR image simulation pipeline described in the following section.

TABLE I
MR TISSUE PROPERTY STATISTICS AT 1.5 TESLA USED FOR SIMULATION

Tissue type	T1 (Mean ± Std) ms	T2 (Mean ± Std) ms
Myocardium	(977 ± 42)*(1198.7 ± 30.3)**	(55 ± 4)
Blood	(1700 ± 63)	(237 ± 50)
Liver	(581 ± 35)	(48 ± 7)
Kidney	(1080 ± 42)	(86 ± 5)
Spleen	(1057 ± 42)	(79 ± 15)
Body fat	(338 ± 27)	(11 ± 7)
Cartilage	(1168 ± 18)	(27 ± 3)
Skeletal muscle	(1034 ± 87)	(39 ± 5)
Bone	(549 ± 52)	(49 ± 8)
Lung	(1000 ± 82)	(40 ± 8)
Stomach	(765 ± 75)	(58 ± 24)
Intestine	(343 ± 37)	(58 ± 4)
References	[15], [16], [18]–[20], [35]–[38] * inversion-recovery and ** saturation-recovery sequences	

B. Cardiac MR Image Simulation

An overview of the pipeline for simulating CMR images is provided in Figure 2. Firstly, MR properties are assigned to each tissue label in the high-resolution voxelized XCAT anatomy (a). Given these parameters and the imaging settings, image contrast is computed using the analytical MR signal model of the desired pulse sequence (b). In c), the simulated contrast data is transformed to high-resolution k-space data by applying a Fourier transformation followed by a sampling operation (e.g. Cartesian sampling grid) to map the k-space data to a coarser grid, given acquisition resolution. Thereafter in d), complex noise is calculated based on the chosen signal-to-noise ratio to be added to the real and imaginary parts of the complex-valued k-space data. Finally the reconstruction operation is carried out to create the final image in e). For dynamic imaging of the heart motion, the same operations from a-e are performed for each time frame of the voxelized anatomy that includes changes in the heart geometry for one cardiac cycle. The reconstructed 4D image data is saved in the NIfTi file format together with the corresponding ground truth labels. In the following section the individual steps of the pipeline are described in more detail.

1) MR Tissue Properties: The voxelized anatomical models derived from the previous section needs to be complemented with MR tissue properties ($\boldsymbol{T}$) in Figure 2 a). It is important to not only assign tissue properties to the heart tissue and blood, but also to the surrounding organs. By adding tissue properties to the organs in the field of view, we expect the simulated images to be a more realistic representation of real images. The T1 and T2 relaxation times are obtained from literature and summarized in Table I. We found that there is a substantial diversity in the normal range of values reported in the literature. Therefore, we combine the statistics to derive one value for mean and one value for standard deviation for each tissue type and use these values to generate random numbers for T1 and T2 relaxation times for each specific virtual patient.

2) MR Contrast: Image contrast is one of the most important features of any imaging sequence in MRI. It ultimately depends on the selection of the acquisition parameters such as repetition time (TR), echo time (TE), and flip angle (α). The balanced steady-state free precession (bSSFP) sequence has become the most widely used clinical sequence for the cardiac functional assessment using the cine CMR because of producing high contrast between blood and myocardium. We therefore use its analytical solution to, given all parameters, compute the CMR image contrast (Figure 2 b). The bSSFP signal ($\boldsymbol{C}$) exhibits a relatively complicated contrast composed of T1 and T2 relaxations that in the absence of any off-resonance is given as follows [40]:

$$\boldsymbol{C} = \frac{PD_0\sqrt{E_2(1-E_1)}sin\alpha}{1-(E_1-E_2)cos\alpha - E_1E_2} \tag{1}$$

where $E_{1,2} = \exp(-\frac{TR}{T_{1,2}})$ and the echo time in the balanced sequence is set to $TE = \frac{TR}{2}$ that is represented as an extra weighting of the signal ($\sqrt{E_2}$).

Note that the simulated MR signal in Equation 1 contains the magnitude information under the assumption that the transversal magnetization will not dephase between RF pulses in the ideal cases hence there will be zero or negligible phase accumulation. In the presence of magnetic field inhomogeneity, RF coil sensitivity, tissue susceptibility, motion, diffusion and any other sources of off-resonance, the signal formula should be modified to account for simulation of the phase information [49].

3) K-Space Acquisition and Sampling: MR contrast computed on the high-resolution (HR) input model is transformed to HR complex k-space data as follows:

$$\boldsymbol{A} = \mathcal{F}(\boldsymbol{C}) \tag{2}$$

Here $\mathcal{F}(.)$ is the Fast Fourier Transformation operator for transferring the HR simulated contrast to the HR simulated k-space data that has complex values ($\boldsymbol{A}$). This data is then sampled to a desired (lower) resolution matrix using a sampling operator ($\boldsymbol{S}$) given the in-plane $(\Delta x, \Delta y)$ acquisition resolution.

The field of view of the simulation is dictated by the size of the input voxelized model. That causes the space between the sampling points at different spatial frequencies to be fixed, given the equation $\Delta kx, y = 1/FOV_{x,y}$. The maximum spatial frequency is therefore derived based on the acquisition resolution given $FOV_{kx,ky} = 1/\Delta x, y$. This is the cut-off for the spatial frequency defining the extent to which we sample the k-space data (Figure 2 c). In order to avoid ringing artifact due to sharp truncation in the frequency domain, we apply a Tukey window with $\alpha = 0.5$ which results in smoother reconstructed images.

4) Noise Addition: Based on the signal-to-noise (SNR) ratio defined by the imaging parameters, a Gaussian complex noise is generated. We add this noise to the real and imaginary part of the simulated k-space data ($\boldsymbol{S}$). The amplitude of the noise signal in the final image depends on the ratio of the magnitude of the simulated signal and the noise standard deviation. It is evaluated by $SNR = \boldsymbol{C}(ROI)/n_{std}$, where $\boldsymbol{C}(ROI)$ is the mean value of the simulated contrast at the region of interest (around the heart) and n_{std} is the standard deviation of the Gaussian distribution, from which random samples are generated. The contrast $\boldsymbol{C}$ depends on the tissue-specific and sequence-specific parameters and is calculated from noise-free k-space data. Given a desired SNR level per simulation, the noise standard deviation is obtained and used to generate a noise complex data ($Noise(n_{std})$) that is added to k-space in Fourier domain. For calculating SNR with the given equation, we assume single-coil data acquisition. As a result of noise averaging, adding extra receiver coils would lead to images with improved SNR. As shown in Figure 2 d), the noisy k-space data at acquisition resolution ($\boldsymbol{N}$) is computed as:

$$\boldsymbol{N} = \boldsymbol{S} + Noise(n_{std}) \tag{3}$$

5) Reconstruction: The high-resolution input is considered to account for the continuous nature of the underlying anatomy in the real-world scenario. The final image is reconstructed by performing inverse Fourier transformation $\boldsymbol{R} = \mathcal{F}^{-1}(\boldsymbol{N})$ after adding complex noise (Figure 2 e). The simulated images are

re-sampled to a uniform grid (e.g. $256 \times 256 \times 13$) and the intensity value for the whole image is scaled to the interval of $[0 - 4095]$, accounting for a 12-bit digital value of the images. Subsequently, the images are saved together with the proper metadata information, making them easily visualizable using standard image viewing software, such as ITK-SNAP and ImageJ [12], [13]. The corresponding ground truth binary labels for all of the tissue types in the simulation are provided alongside the images.

C. Usefulness of the Simulated Data

We evaluate the usefulness of the simulated CMR images in the context of training a deep convolutional neural network for the task of cardiac cavity segmentation. We investigate three different scenarios of utilizing simulated data, which include: 1) exploring the performance of a segmentation model, trained only using simulated images, on real MR test images, 2) assessing the usability of simulated images as a data augmentation method, and 3) analyzing segmentation performance retention when real data is reduced during training, while the number of simulated data remains the same.

1) Real Data: : We deploy all networks on images acquired from the Automated Cardiac Diagnosis Challenge (ACDC) challenge[1] [42]. The ACDC data-set includes end-systolic (ES) and end-diastolic (ED) images acquired from 100 patients, containing both normal and pathological subjects. Images are acquired by two scanners with different magnetic field strengths and contain expert annotations for left ventricular (LV) blood pool, right ventricular (RV) blood pool, as well as left ventricular myocardium (MYO). Out of the available 100 subjects (200 ED and ES MR images), we reserve 90 (180 MR images) for training and 10 (20 MR images) for testing.

2) Segmentation Network and Training: We adopt a 3D nnU-Net [44] model for a a multi-class segmentation task with several modifications to improve the adaptation from simulated to real MR images during test time. The real cardiac MR images of the ACDC challenge have different in-plane spatial resolution, matrix size, and slice thickness. It is a necessary practice to harmonize that for training a deep-learning segmentation algorithm. All real images used during training and testing are re-sampled to the in-plane resolution of $1.25 \times 1.25\ mm^2$ and cropped around the heart area to the size of 128×128 pixels while keeping the original slice thickness and number of slices. However, to ensure more variation in the FOV and vary the size of the heart in simulated data, we apply a range of different resolutions when resampling the simulated images. We normalize input images at both the training and inference time to an intensity range from [0,1] using a 99th percentile-based approach. During training, we augment the data on-the-fly by utilizing random horizontal and vertical flips ($p = 0.5$), random rotation by integer multiples of $\frac{\pi}{2}$ ($p = 0.5$), scaling (scale factor $s \in [0.85,1.25]$, $p = 0.3$) and random translations ($p = 0.3$), gamma and brightness transformation ($p = 0.6$), as well as elastic transformations ($p = 0.3$).

[1]ACDC data can be found at https://www.creatis.insa-lyon.fr/Challenge/acdc/

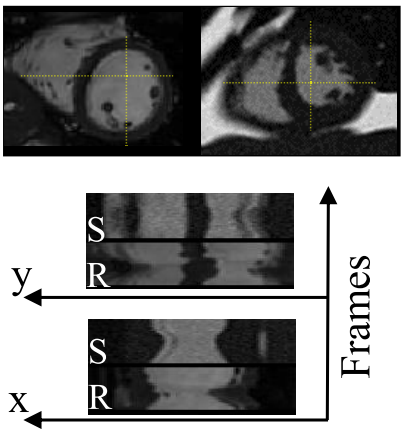

Fig. 3. Dynamic cine simulation for 25 frames across one cardiac cycle with 1 second period. Time profiles along x and y lines for simulated (S) and real (R) images are shown. The Animated gif is available at https://bit.ly/3DcU7q3.

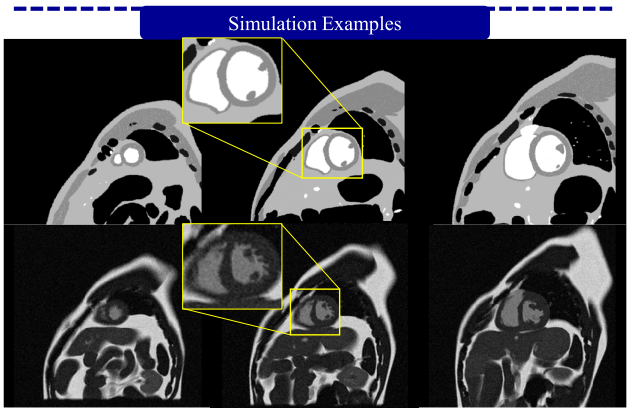

Fig. 4. Examples of simulation with improved image realism for apex, mid, and base locations of the heart. The results for the original MRXCAT approach and our framework is shown in first row and second row, respectively.

At inference time, we apply in-plane resampling, center-cropping and percentile-based normalization on the real images. We additionally apply histogram matching to match the intensity distribution of real images to that of simulated images used during training, using a landmark-based histogram standardization approach described in [43]. Moreover, we apply total variation (TV) denoising on real images, which removes high frequency noise and textural features, but retains sharp edges and outlines of larger tissues. This procedure reduces the bias of the trained network towards tissue texture, which is difficult to realistically simulate. All images are filtered using a strength parameter $\alpha = 15$, which was visually determined to smooth out the texture, while retaining the cavity shape.

3) Experiments: We explore the usability of simulated data for cardiac cavity segmentation by performing three experiments that outline different aspects of simulated images. To compare the performance of each model, we calculate the Dice similarity metric and Hausdorff Distance (HD) on all slices of each subject in the test set.

Experiment 1: Performance of the Simulated model; By employing a training and testing procedure described in Section II-C.2, we first train two segmentation models

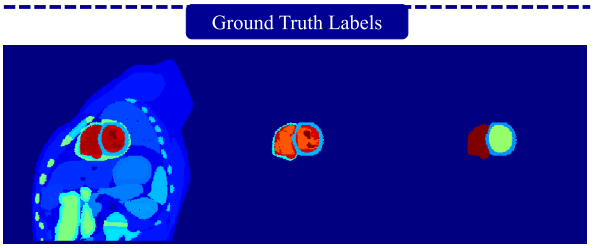

Fig. 5. Ground truth labels of a simulated cardiac MR image. Full label map with 49 separate tissues shown with different colors (left), the heart tissues (middle), and simplified classes for heart cavity (right).

(**Simulated 1 and Simulated 2**) using a total of 200 simulated volumes (ED and ES) each for subjects with normal and abnormally thick myocardium, respectively. We create XCAT subjects with thickened myocardium by modifying the NURBS surfaces for the left ventricle. We evaluate the model on 20 ED and ES volumes extracted from the ACDC test set, where all slices are considered during evaluation. We compare the performance of this model to a model trained using real images from the ACDC training set, which is our baseline. All models are evaluated using the Dice score and Hausdorff distance.

Experiment 2: Augmentation using simulated data;
In this experiment, we evaluate the use of simulated data as an augmentation method, whereby we wish to observe whether adding simulated data to the training set of real images has the potential to improve the model's generalization performance and reduce over-fitting. The baseline model for this experiment is trained with 180 ACDC (ED and ES) volumes, extensively augmented using geometric and intensity transformations, as well as transformations affecting the quality of images by adding (Gaussian) noise. The **Augmented models** Aug 1 and Aug 2 are trained by adding 200 simulated volumes (corresponding to Simulated 1 and Simulated 2 data) to the same set of real images used for training the baseline model. The evaluation of both models is performed on all slices of the test set.

Experiment 3: Real data reduction;
In the last experiment, we wish to evaluate the extent to which the simulated data can compensate for the lack of real MR images during training, in scenarios where annotated MR images are limited. Since limitation of expert-annotated data is often a challenge for supervised deep learning algorithms, having the ability to train well-performing models with the help of simulated data is of the utmost importance during model development. To demonstrate this, we systematically reduce the number of real images available during training, while retaining the number of simulated images (200) and evaluate the models on the same set of 20 test volumes as before (all slices included).

III. Results

A. Qualitative Analysis

Visual comparison of simulated CMR images using the original MRXCAT approach [6] and our proposed framework is shown in Figure 4. Both the underlying anatomy and the image contrast have been improved, resulting in increased image quality and realism. As can be observed from the zoom area around the heart, inclusion of the small structures of the trabeculae anatomy in the myocardium has made the XCAT heart more similar to the anatomy of the real human heart. This modelled trabeculation has improved the realism of myocardium-to-blood borders for left and right cavities compared to the previous simulation. Furthermore, increasing the number of tissue types in the image simulation as well as assigning relevant relaxation properties is beneficial to enhance image quality. Using low-pass filtering for partial volume effects in the original MRXCAT approach gave a rather simplified look to the images. By adding the sampling operation in the pipeline we could better account for the effects of partial volume on the smoothness and blurriness of the organ and tissue boundaries.

Our framework can also take the time series of the XCAT models to perform dynamic simulation for the cine CMR imaging study. As can be observed in the movie version of Figure 3 (available at `https://bit.ly/3DcU7q3`) the papillary and trabeculae structures inside the left and right ventricles are deformed according to the motion model of the beating heart, which replicates the twisting of the myocardium. The simulated heart motion resembles the real one shown next to it and the time profile of simulated (S) and real (R) image along two perpendicular directions (x; top-down and y; left-right) for all frames show good similarity in terms of myocardial displacement.

For each virtual subject we have a number of labels to which we assigned tissue properties for image simulation. These input labels provide accurate information about the underlying anatomy and can be utilized as ground truth labels of the simulated images. Different labels (49 tissue types) are represented by different colors in Figure 5. The heart-only label map with its different components is shown in Figure 5. The combined version representing the heart as 3 simplified classes may be more suitable for training a heart cavity segmentation network.

Generating diversified images with varying parameters can enrich the virtual population. Within our framework, we made numerous parameters available to alter the simulation characteristics. This additional flexibility yields the advantage to perform arbitrary changes in the imaging settings or anatomical features. Two examples for male and female anatomies are shown in Figure 6. It also illustrates some examples of changing image characteristics by modifying the imaging parameters such as flip angle, in-plane resolution, and SNR level. We can observe the alternations on the anatomical features such as location of the heart within torso and scaling of the whole body, as well as changes of the image appearance.

Using the proposed framework, we simulate a population of virtual patients with varying anatomical and contrast characteristics, tissue properties and sequence parameters. The heart model of the virtual male is enhanced by adding an additional layer of the heart trabeculae as explained in section II-A.2, whereas the virtual female only has the papillary muscles. Figure 7 shows the T1 and T2 relaxation times, repetition time (TR), simulated signal-to-noise ratio (SNR) and flip angle

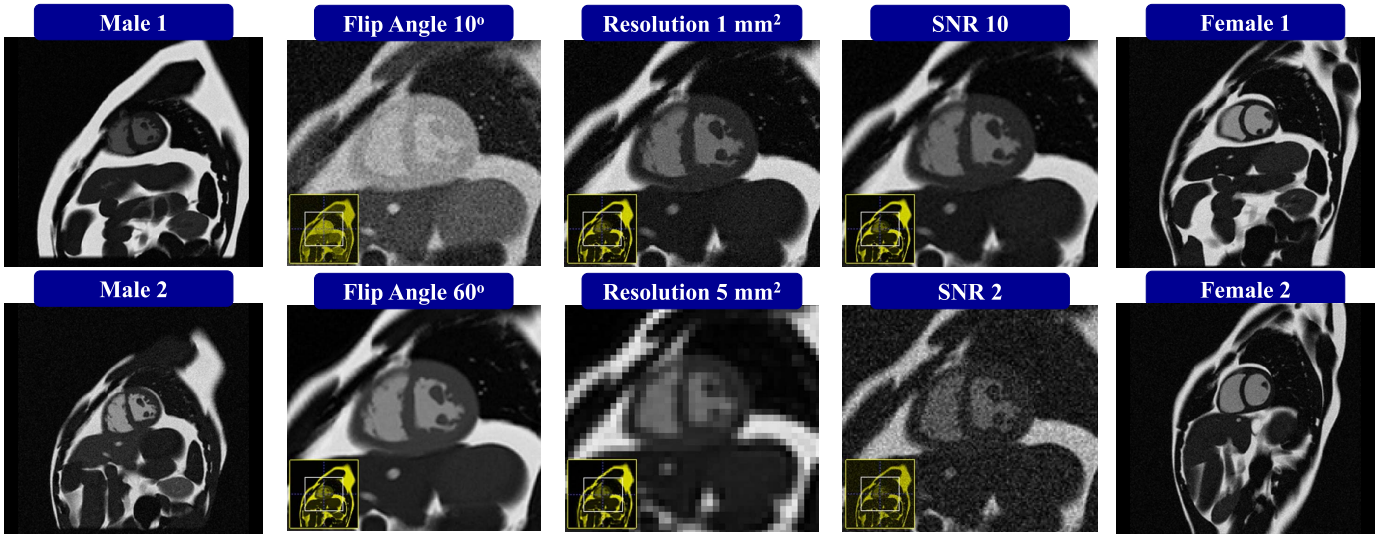

Fig. 6. Two examples of the simulated images for virtual subject male (first column) and virtual subject female (last column), and examples for varying sequence parameters flip angle (degree), in-plane resolution (mm) and simulated SNR levels. Note that the heart trabeculation was only added to the male subjects.

(FA) for each generated virtual subjects. For simulations in this study, we use myocardial T1 values of 977 ± 42 ms, which is measured using inversion-recovery techniques [15], [16]. However, saturation-recovery can provide more accurate tissue quantification that can yield higher T1 values as discussed in [17] and these are therefore included in Table I. We observe that selecting slightly higher/lower tissue property values for myocardium has a negligible effect on the simulated image contrast. To visualize the range of anatomical variations, Figure 8 depicts left ventricular (LV) volumes at end-diastolic (ED) and end-systolic (ES) phases of the heart for the simulated subjects.

B. Quantitative Analysis

1) Experiment 1: Performance of the Simulated Models: Figure 9 showcases some of the predictions for mid-ventricular slices for the segmentation models trained on simulated images only with normal (Simulated 1) and thickened myocardium (Simulated 2), compared to the baseline model trained using real images. The models are evaluated on all slices of testing subjects, resulting in the overall Dice scores and HD scores shown in Figure 11. We perform a two-tailed Wilcoxon signed-rank test for the two models with $p < 0.01$, indicating statistical significance between the performance of the two models. By visual observation, we can determine that the performance drop observed in the Simulated 1 is largely due to thicker myocardium appearing in the test set, especially with patients containing pathology. While tissue segmentation in the presence of pathology is generally a challenge when evaluating any segmentation approach, additional challenges stem from the nature of simulation, where most simulated myocardial tissue is thinner due to exclusion of papillary muscles and trabeculae. This also affects the performance of the model when it comes to LV segmentation. To address this limitation, we simulate new subjects with thickened myocardium to train the Simulated 2 model. This newly simulated data help increasing the segmentation performance for thick myocardium present in the testing data, as well as boosting the performance for LV and RV segmentation as can be observed from Figure 11. A drop in segmentation performance in comparison with the baseline can stem from the basal and apical slices and can be linked to the complex anatomy of the heart for pathological cases, which is a significant challenge for a network that exclusively learns from the appearance and shape of simulated hearts. Additionally, the performance of the Simulated models is drastically affected by changes in appearance and texture, despite our attempts to minimize these effects through TV filtering. However, these results suggest a strong potential of utilizing a cost-effective artificial data-set for training neural networks that can perform on par with traditional approaches requiring large annotated sets of real MR data.

2) Experiment 2: Augmentation Using Simulated Data: Figure 11 depicts the performance of the Augmented models (Aug 1 and Aug 2) compared to the baseline trained without simulated images. We can observe that both RV and LV segmentation performance improves with the addition of simulated data, with the accuracy of MYO segmentation slightly reducing for Aug 1 model. This is again the effect of thinner myocardial tissue in simulated images in Aug 1, which hampers model performance in the presence of myocardial thickening. To address this issue and improving the performance of the augmented model on the pathological cases with thick myocardium we simulate subjects with abnormally thick myocardium and add them to the training. We observe improved performance for the segmentation of the myocardium as well as right and left ventricles in terms of Dice score and substantial reduction in the HD score for Aug 2 model where subjects with thickened myocardium are simulated. The obtained Dice and HD scores for segmentation using the Aug 1 and Aug 2 models are statistically significant with $p < 0.05$.

Samples of slices where augmentation with simulated data improved the performance can be observed in the first three columns in Figure 10. The majority of visually observed

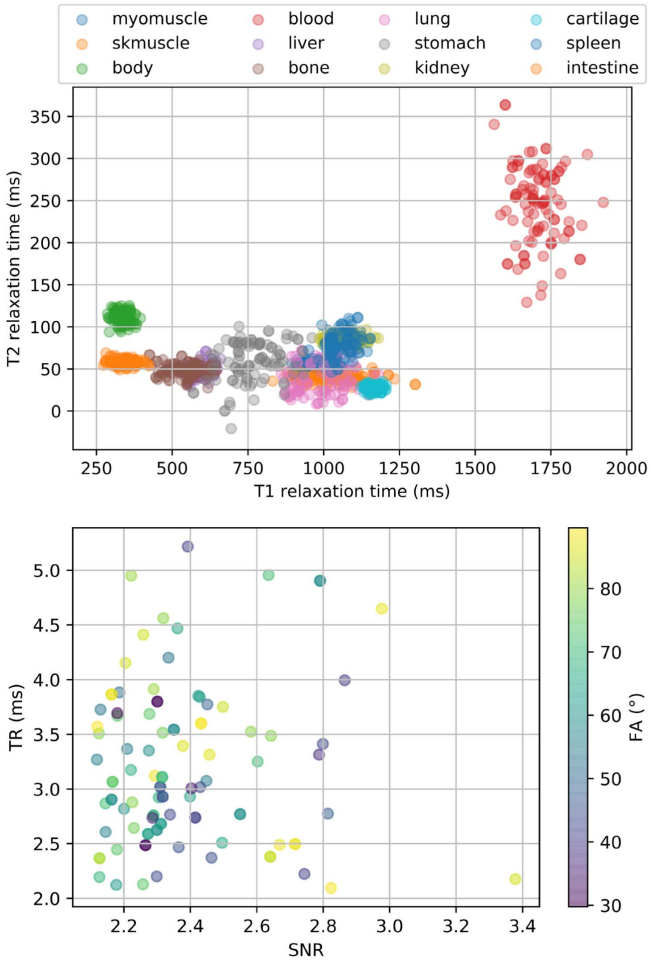

Fig. 7. The distribution of T1 and T2 relaxation times for tissue properties, repetition time (TR), signal-to-noise ratio (SNR) and flip angle (FA) for sequence parameters used for simulating subjects in the heterogeneous population.

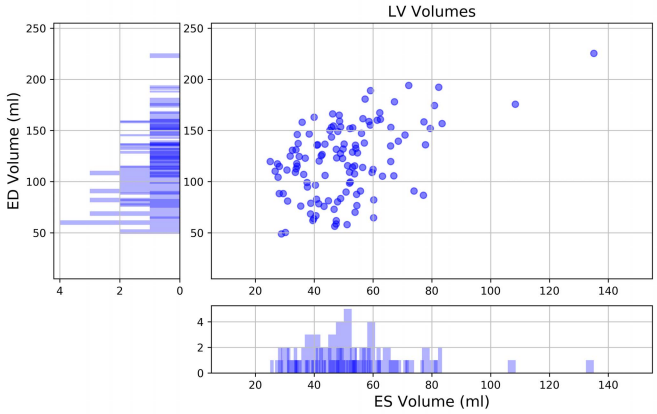

Fig. 8. The distribution of the left ventricular (LV) volumes at end-diastolic (ED) and end-systolic (ES) phases for simulated subjects.

improvement in predictions is typically related to the RV segmentation performance, but it is also notable for LV and myocardium segmentation, especially when simulated images with thickened myocardium is added. We hypothesize that adding simulated data to the model has reduced over-fitting and given more emphasis to cavity shapes during training. Additionally, we find that simulating subjects with thickened

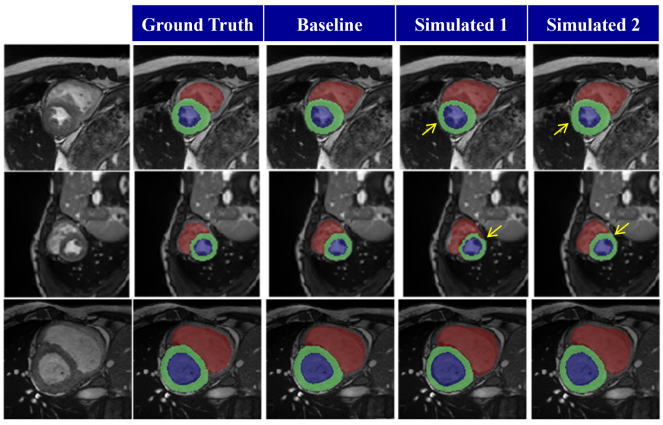

Fig. 9. Comparison between a segmentation model trained on 180 real images (baseline) and identical model trained completely on 200 simulated images (simulated).

myocardium is important to improve the performance of the model on the pathological cases present in the test set. To further demonstrate the contribution of simulated pathological data to the segmentation performance, we train an additional model augmented with a combination of both Simulated 1 and Simulated 2 data. In total, we add 200 simulated volumes, where 50% of images belong to the Simulated 1 data-set, with the other half extracted from Simulated 2 data. To avoid selection bias when choosing the 50% of total simulated data available, we perform a five-fold cross validation experiment and randomly select 100 images per simulated set for each fold. A model trained under such a setup performs better than the Aug 1 model, whereby we observe a reduction in the number of outliers and improvement in HD scores across all tissues, particularly the LV myocardium. We attribute this to the presence of the Simulated 2 data, which we already show can tackle the appearance of thickened myocardium (Aug 2 model). However, while better than Aug 1, this model under-performs compared with Aug 2, indicating that Simulated 1 images may not introduce enough variability beneficial for improving the segmentation performance of ACDC data. However, this may not be the case for other data-sets (such as those containing examples of healthy subjects only).

3) Experiment 3: Real Data Reduction: The effect of reducing real data during training, where we try to imitate the scenario where acquiring annotated real MR data is challenging, can be observed in Figures 10 and 11. We reduce the number of real MR images available during training to 100, 50 and 25 volumes only, while retaining 200 Simulated 1 images. We compare the performance of such models to models trained with only a limited amount of real MR volumes, without the addition of simulated data. As expected, the performance of such models tends to drop for all tissues, with a significant reduction in performance when only 50 and 25 real MR images are used during training. However, if such limited data-sets are aided with simulated images, the performance drop is less significant and in some cases the performance of the augmented model is retained compared to the model trained with the maximum amount of real MR volumes available (compare 180R with 100R + 200S in Figure 11). In Figure 10, we showcase the most extreme improvements, which occur for

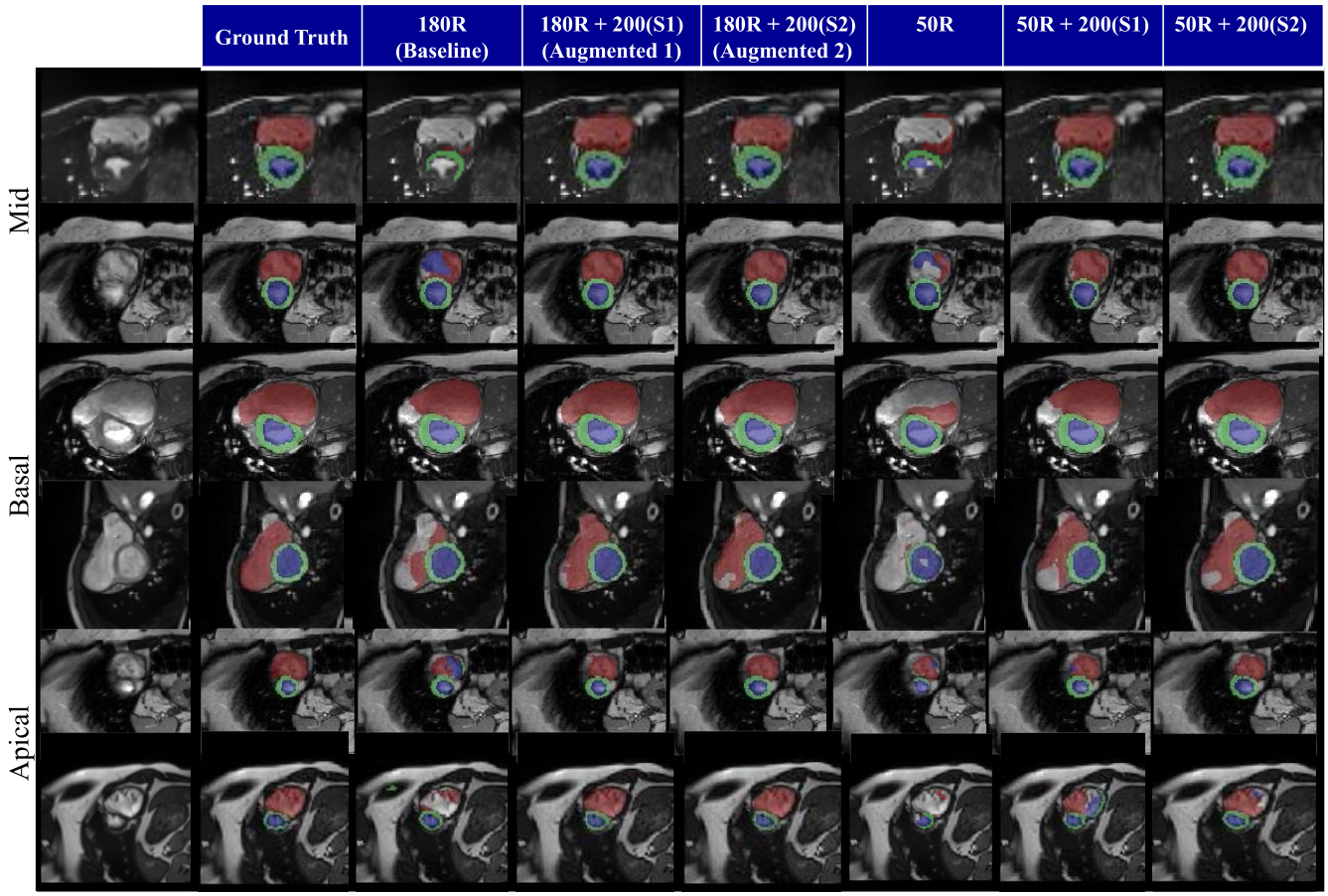

Fig. 10. Comparison between the baseline segmentation model trained on 180 real images (180R), the augmented 1 model with 200 simulated 1 images (180R + 200(S1)), the augmented 2 model with 200 simulated 2 images (180R + 200(S2)), reduced model trained with 50 real images (50R) and augmented version of that with simulated 1 and 2 images (50R + 200(S1)) and (50R + 200(S2)). Results for mid-ventricular, basal and apical slices of the heart are shown.

models trained with 50 real images only, where we observe the 50R model struggling with the segmentation of the RV and producing false positive predictions for the LV. These are successfully compensated by the addition of simulated images during training.

IV. Discussion

In this paper we propose a flexible framework for physics-based CMR image simulation with the purpose of generating a heterogeneous population comprising diversified virtual subjects. We improved the quality of the simulated images by enhancing and optimizing the three main components: i) computerized human anatomical model, ii) magnetic tissue properties, and iii) physics of MR image formation.

Our proposed image simulation pipeline has separated modules that could easily be adapted and replaced to account for more comprehensive experiments. For instance, in the sampling module c, for the fast generation of CMR images, uniform Cartesian sampling for the k-space data can be replaced by more advanced acquisition trajectories. It should be noted that the reconstruction module needs to be modified accordingly, and also the coil sensitivity map may be required for parallel imaging. The analytical description of the MR signal was used for fast generation of the imaging contrast. Alternatively, the signal model could be replaced with the extended phase graph formulation [21] or full numerical simulation of the MR signal using the JEMRIS simulator [41], at the cost of substantially increasing the complexity and computation time of the simulation. The JEMRIS simulation approaches, based on numerically solving the Bloch equations, is considered more suitable for investigating the effects of various pulse sequences design on the spin system, while in the context of generating ample images for deep-learning application, the effects of the parameters on the global contrast of the final images is more relevant.

Similar to the MRXCAT approach [6], the contrast is governed by the analytical solution of the Bloch equations at the steady state of the magnetization for the bSSFP sequence. It was shown that bSSFP sequence is less sensitive to field inhomogeneities when a short TR is used [40]. We use short TR and therefore ignore the presence of any off-resonance effects, which in real world scenario might be due to RF inhomogeneity, imperfect shimming, magnetic susceptibility, and T2* effects. Moreover, the RF slice profile is assumed to be perfectly rectangular, and no spin dephasing is present during the acquisition. Simulating imperfections in MR scanner and designing different RF pulses require numerically solving Bloch equations using software packages such as JEMRIS [41] and MRiLab [45] which found to be extremely time consuming and unsuitable for image database generation.

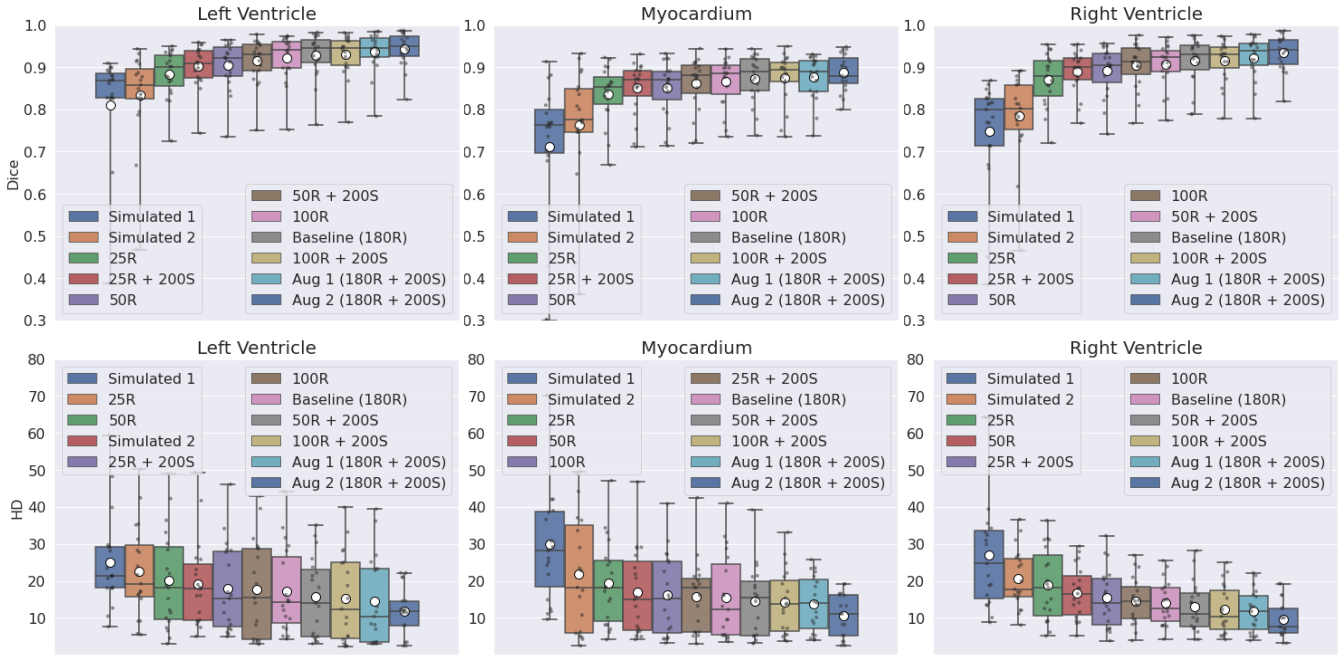

Fig. 11. Performance evaluation based on Dice score (higher better) and Housdorff Distance (HD) (lower better) for the baseline segmentation model trained on 180 real (180R) images, the augmented models (Aug 1 and Aug 2) with 200 simulated (180R + 200S) images, reduced models trained with 100 real images (100R), 50R, 25R images and augmented versions 100R + 200S, 50R + 200S, and 25R + 200S images. The Simulated 1 and Simulated 2 models are trained using simulated images only with normal and thickened myocardium, respectively.

We consider variations in the sequence parameters (shown in Figure 7) to be an import feature of our simulation pipeline that results in producing images with different contrasts and appearances. We observe that different scanner vendors and imaging centers use different sequence parameters to optimize the bSSFP contrast hence they produce varying image appearances. This varying contrasts and imaging features hamper the performance of DL segmentation network as discussed in the M&Ms challenge [51]. We believe by altering sequence parameters we can generate images with diverse appearances to help development of generalizable DL method that can robustly perform on heterogeneous data from different clinical centers, imaging conditions and scanner vendors.

Images with variations in the noise level are also generated in this study and the range of simulated SNR is experimentally identified by inspecting the images. Note that complex noise is added to the complex k-space data to resemble a real scenario where the noise is already present in the MR signal during the acquisition. We found that applying filter to the k-space data and resampling to the grid after reconstruction will change the final SNR. In agreement with what discussed in [50], we found that SNR measurements using two region approaches (background noise and region of interest) do not completely agree with the actual simulated SNR. Our measurements of SNR based on two ROIs from the simulated images resulted in overestimating the simulated SNR level. However, the absolute target SNR value is not crucial here, rather variation in the SNR and the ability to change the noise level is considered an important feature of our framework for generating diversified images for the purpose of DL training.

We created a database of CMR images of virtual patients using the proposed framework to aid the development of data-hungry deep-learning medical image analysis methods. One application using the preliminary version of the framework was demonstrated in [33]. The benefit of using the simulated images for training a cardiac cavity segmentation model was investigated and it was shown that pre-training a deep-learning based segmentation model generalizes better to the inherent variability of real cardiac images.

In addition, we explored the application of the virtual CMR database in the task of cardiac cavity segmentation using supervised deep learning. Our experiments demonstrated that models trained with images simulated in this study solely can already perform comparable cavity tissue (LV, RV and MYO) segmentation in mid-ventricular slices to models trained with real MR images annotated by medical experts. We further show that the use of simulated data can be considered an addition to classical augmentation methods, which are typically limited in producing tissue shapes and appearance different to existing data. We demonstrate that simulated data can compensate for the lack of real data during training and be of significant help in settings where data acquisition is challenging. This indicates that artificial data generated through the proposed framework has the potential to extend the variability of anatomical and contrast features available in training data and consequently, help the network generalize and adapt better to unseen data. Of course, simulating a more realistic appearance and contrast, with significant variation in shape and quality is of the utmost importance for achieving this. The added benefit is the fact that through this framework we can simulate and design a number of pulse sequences, generating simulated data representative of the data available in test sets, which is typically costly and infeasible for typical MR acquisition.

Limitations and Future Work

We simulate single-coil acquisition scenario for fast generation of the images. However, three-dimensional coil sensitivity maps can be simulated using the Biot-Savart law as previously discussed in the MRXCAT [6]. The normalized sensitivity map can be multiplied with the simulated contrast resulting in one image per coil. Similarly, one could use the same multi-coil simulation setup inside our framework. Note that the multi-coil acquisition scenario is more suited for reconstruction purposes and optimization for parallel imaging. We noticed that the addition of multiple coils will substantially increase the simulation time and complexity, making our framework less suitable for generating substantial number of images.

Flowing blood, also known as in-flow effects, can change the blood-to-myocardium contrast and has historically been a source of error in cardiac MR images. The simulations in this work do not account for this effect due to the complexity of computational modeling for the blood flow and time-consuming simulation procedure, requiring careful considerations of RF pulse profile, velocity distributions, and slice thickness. It is, however, possible to extend our framework with such model for the purpose of flow quantification using phase contrast image simulation.

Major cardiac MR imaging artifacts such as respiratory motion and ECG-mistriggering artifacts are not simulated in this study. These artifacts are one of the primary sources of failure of deep-learning (DL) based segmentation models. Recently the authors in [46] and [47] proposed k-space models of motion artifacts for the purpose of making DL algorithms more robust to such artifacts. Particularly in [47] authors proposed a data augmentation strategy to simulate cardiac ECG-mistriggering and breathing artifacts based on k-space data corruption. The same approach can be used to simulate these artifacts on our database. Another concurrent work in [48] showed that our framework can be extended for the application of late gadolinium enhancement (LGE) simulation on a virtual subject with myocardial infarction and included respiratory artifact in the simulation procedure. Precisely, the LGE simulation is performed at various time points across one respiratory cycle from the XCAT model and the data is combined in such a way to resemble the slice-misalignment artifact. The authors investigate the effects of such artifact on the electrophysiology modeling of the heart. Modeling relevant artifacts in cardiac MR imaging and simulating subjects with respiratory motion artifacts, ECG-mistriggering, and ghosting artifacts, as well as investigating their impact on the segmentation remains to be explored in future work.

In attempts to simulate pathological cases in this study, we generate subjects with thickened myocardium for hypertrophic cardiomyopathy by modifying the NURBS surfaces of the XCAT heart model. We demonstrate that the addition of these subjects with abnormally thick myocardium would substantially improve the segmentation performance on pathological cases, achieving best overall score for the Aug 2 model in Figure 11. However, the current model lacks subjects with myocardial infarction (potentially thinned heart), congenital heart, and other cardiac diseases. Adding virtual subjects with such diseases would be of great importance and interest for the research community. However, due to the complex nature of the heart modeling for such patients, we believe this could be address as future work. For instance, by anatomical modeling of congenital heart disease in the XCAT phantom, one could use our framework to simulate corresponding images. We showed that the inclusion of the trabeculation anatomy in the male XCAT subject increased the visual realism of the simulated images. However, the impact of this addition on the segmentation results was not directly quantified in the presented paper. A mix of female subjects without and male subjects with trabeculation were simulated for the Simulated 1 data, whereas for the Simulated 2 data with thickened myocardium we only generated male subjects without the addition of the trabeculation. The results of these experiments suggested that even without the added trabeculation anatomy, the simulated images improved the segmentation performance, and whether or not the addition would bring extra benefits to the segmentation remained to be explored in future studies. The presented framework was designed to provide simulation of realistic CMR images for the cine study. Extension to other CMR modalities such as late gadolinium enhancement and first-pass perfusion are considered as future research. These additions to the framework could be of great interest for multi-modality studies, especially for expanding the database to increase its applicability for disease classification. We believe this work is a step towards our aim to establish a unified framework for personalized multi-modal cardiac magnetic resonance image simulation.

V. Conclusion

In this paper, we proposed a flexible framework for realistic simulation of cardiac magnetic resonance images with controllable anatomical features, MR tissue properties, acquisition parameters, and image appearance. We generated a virtual population of CMR images with ground truth labels using our proposed framework and made it publicly available to aid development of deep-learning cardiac image analysis algorithms. Furthermore, our usability experiments suggested that augmentation with the simulated population can boost the segmentation performance, and even retain the baseline performance when just 45% of the real data is available.

References

[1] A. F. Frangi, S. A. Tsaftaris, and J. L. Prince, "Simulation and synthesis in medical imaging," *IEEE Trans. Med. Imag.*, vol. 37, no. 3, pp. 673–679, Mar. 2018, doi: 10.1109/TMI.2018.2800298.

[2] T. Glatard et al., "A virtual imaging platform for multi-modality medical image simulation," *IEEE Trans. Med. Imag.*, vol. 32, no. 1, pp. 110–118, Jan. 2013, doi: 10.1109/TMI.2012.2220154.

[3] C. Tobon-Gomez, F. M. Sukno, B. H. Bijnens, M. Huguet, and A. F. Frangi, "Realistic simulation of cardiac magnetic resonance studies modeling anatomical variability, trabeculae, and papillary muscles," *Magn. Reson. Med.*, vol. 65, no. 1, pp. 280–288, Jan. 2011, doi: 10.1002/mrm.22621.

[4] R. K.-S. Kwan, A. C. Evans, and G. B. Pike, "MRI simulation-based evaluation of image-processing and classification methods," *IEEE Trans. Med. Imag.*, vol. 18, no. 11, pp. 1085–1097, Nov. 1999.

[5] W. P. Segars, G. Sturgeon, S. Mendonca, J. Grimes, and B. M. W. Tsui, "4D XCAT phantom for multimodality imaging research," *Med. Phys.*, vol. 37, no. 9, pp. 4902–4915, Aug. 2010.

[6] L. Wissmann, C. Santelli, W. P. Segars, and S. Kozerke, "MRXCAT: Realistic numerical phantoms for cardiovascular magnetic resonance," *J. Cardiovascular Magn. Reson.*, vol. 16, no. 1, p. 63, Dec. 2014.

[7] A. Prakosa et al., "Generation of synthetic but visually realistic time series of cardiac images combining a biophysical model and clinical images," *IEEE Trans. Med. Imag.*, vol. 32, no. 1, pp. 99–109, Jan. 2013, doi: 10.1109/TMI.2012.2220375.

[8] N. Duchateau, M. Sermesant, H. Delingette, and N. Ayache, "Model-based generation of large databases of cardiac images: Synthesis of pathological cine MR sequences from real healthy cases," *IEEE Trans. Med. Imag.*, vol. 37, no. 3, pp. 755–766, Mar. 2018, doi: 10.1109/TMI.2017.2714343.

[9] Y. Zhou et al., "A framework for the generation of realistic synthetic cardiac ultrasound and magnetic resonance imaging sequences from the same virtual patients," *IEEE Trans. Med. Imag.*, vol. 37, no. 3, pp. 741–754, Mar. 2018, doi: 10.1109/TMI.2017.2708159.

[10] J. W. Weinsaft et al., "Left ventricular papillary muscles and trabeculae are significant determinants of cardiac MRI volumetric measurements: Effects on clinical standards in patients with advanced systolic dysfunction," *Int. J. Cardiol.*, vol. 126, no. 3, pp. 359–365, 2008.

[11] P. A. Helm, H.-J. Tseng, L. Younes, E. R. McVeigh, and R. L. Winslow, "Ex vivo 3D diffusion tensor imaging and quantification of cardiac laminar structure," *Magn. Reson. Med.*, vol. 54, no. 4, pp. 850–859, Oct. 2005.

[12] P. A. Yushkevich et al., "User-guided 3D active contour segmentation of anatomical structures: Significantly improved efficiency and reliability," *Neuroimage*, vol. 31, no. 3, pp. 1116–1128, Jul. 2006.

[13] C. A. Schneider, W. S. Rasband, and K. W. Eliceiri, "NIH Image to Imagej: 25 years of image analysis," *Nature Methods*, vol. 9, no. 7, pp. 671–675, 2012, doi: 10.1038/nmeth.2089.

[14] *Rhinoceros 3D*. Accessed: 2020. [Online]. Available: https://www.rhino3d.com

[15] D. Dabir et al., "Reference values for healthy human myocardium using a T1 mapping methodology: Results from the international T1 multicenter cardiovascular magnetic resonance study," *J. Cardiovascular Magn. Reson.*, vol. 16, no. 1, p. 69, Dec. 2014.

[16] M. Granitz, "Comparison of native myocardial T1 and T2 mapping at 1.5 T and 3T in healthy volunteers," *Wiener Klinische Wochenschrift*, vol. 131, no. 7, pp. 143–155, 2019.

[17] P. Kellman and M. S. Hansen, "T1-mapping in the heart: Accuracy and precision," *J. Cardiovascular Magn. Reson.*, vol. 16, no. 1, pp. 1–20, Dec. 2014.

[18] K. Chow, J. A. Flewitt, J. D. Green, J. J. Pagano, M. G. Friedrich, and R. B. Thompson, "Saturation recovery single-shot acquisition (SASHA) for myocardial T 1 mapping," *Magn. Reson. Med.*, vol. 71, no. 6, pp. 2082–2095, Jun. 2014.

[19] S. Matsumoto, "Myocardial T1 values in healthy volunteers measured with saturation method using adaptive recovery times for T1 mapping (SMART1Map) at 1.5 T and 3 T," *Heart Vessels*, vol. 34, no. 11, pp. 1889–1894, 2019.

[20] G. J. Stanisz et al., "T1, T2 relaxation and magnetization transfer in tissue at 3T," *Magn. Reson. Med.*, vol. 54, no. 3, pp. 507–512, Sep. 2005.

[21] M. Weigel, "Extended phase graphs: Dephasing, RF pulses, and echoes–pure and simple," *J. Magn. Reson. Imag.*, vol. 41, no. 2, pp. 266–295, Feb. 2015.

[22] I. Goodfellow et al., "Generative adversarial networks," *Commun. ACM*, vol. 63, no. 11, pp. 139–144, 2020.

[23] X. Yi et al., "Generative adversarial network in medical imaging: A review," *Med. Image Anal.*, vol. 58, Dec. 2019, Art. no. 101552.

[24] S. Kazeminia et al., "GANs for medical image analysis," *Artif. Intell. Med.*, vol. 109, Sep. 2020, Art. no. 101938.

[25] S. Amirrajab et al., "XCAT-GAN for synthesizing 3D consistent labeled cardiac MR images anatomically variable XCAT phantoms," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.* Cham, Switzerland: Springer, 2020, pp. 128–137.

[26] S. Abbasi-Sureshjani, S. Amirrajab, C. Lorenz, J. Weese, J. Pluim, and M. Breeuwer, "4D semantic cardiac magnetic resonance image synthesis on XCAT anatomical model," 2020, *arXiv:2002.07089*.

[27] J. C. Acero et al., "SMOD-data augmentation based on statistical models of deformation to enhance segmentation in 2D cine cardiac MRI," in *Proc. Int. Conf. Funct. Imag. Model. Heart*. Cham, Switzerland: Springer, 2019, pp. 361–369.

[28] C. Chen et al., "Unsupervised multi-modal style transfer for cardiac MR segmentation," in *Proc. Int. Workshop STACOM*. Cham, Switzerland: Springer, 2019, pp. 209–219.

[29] C. Ma, "Neural style transfer improves 3D cardiovascular MR image segmentation on inconsistent data," in *Proc. Int. Conf. MICCAI*. Cham, Switzerland: Springer, 2019, pp. 128–136.

[30] A. Chartsias et al., "Adversarial image synthesis for unpaired multimodal cardiac data," in *Proc. Int. Workshop SASHIMI*. Cham, Switzerland: Springer, 2017, pp. 3–13.

[31] T. Joyce et al., "3D medical image synthesis by factorised representation and deformable model learning," in *Proc. Int. Workshop SASHIMI*. Cham, Switzerland: Springer, 2019, pp. 110–119.

[32] D. P. Kingma and M. Welling, "Auto-encoding variational Bayes," 2013, *arXiv:1312.6114*.

[33] Y. Al Khalil et al., "Heterogeneous virtual population simulated CMR images for improving generalization cardiac segmentation algorithms," in *Proc. Int. Workshop SASHIMI*. Cham, Switzerland: Springer, 2020, pp. 68–79.

[34] W. P. Segars, D. S. Lalush, E. C. Frey, D. Manocha, M. A. King, and B. M. W. Tsui, "Improved dynamic cardiac phantom based on 4D NURBS and tagged MRI," *IEEE Trans. Nucl. Sci.*, vol. 56, no. 5, pp. 2728–2738, Oct. 2009, doi: 10.1109/TNS.2009.2016196.

[35] C. M. J. de Bazelaire, G. D. Duhamel, N. M. Rofsky, and D. C. Alsop, "MR imaging relaxation times of abdominal and pelvic tissues measured in vivo at 3.0 T: Preliminary results," *Radiology*, vol. 230, no. 3, pp. 652–659, Mar. 2004.

[36] M. Barth et al., "Proton NMR relaxation times of human blood samples at 1.5 T and implications for functional MRI," *Cellular Mol. Biol.*, vol. 43, no. 5, pp. 91–783, 1997.

[37] J. A. Luetkens et al., "Incremental value of quantitative CMR including parametric mapping for the diagnosis of acute myocarditis," *Eur. Heart J., Cardiovascular Imag.*, vol. 17, no. 2, pp. 154–161, Feb. 2016.

[38] B. Baeßler, F. Schaarschmidt, C. Stehning, B. Schnackenburg, D. Maintz, and A. C. Bunck, "A systematic evaluation of three different cardiac T2-mapping sequences at 1.5 and 3T in healthy volunteers," *Eur. J. Radiol.*, vol. 84, no. 11, pp. 2161–2170, Nov. 2015.

[39] B. Herzog, J. P. Greenwood, and S. Plein, *CMR Pocket Guide*. Accessed: 2013. [Online]. Available: http://www.cmr-guide.com/Page001.html

[40] K. Scheffler and S. Lehnhardt, "Principles and applications of balanced SSFP techniques," *Eur. Radiol.*, vol. 13, no. 11, pp. 2409–2418, 2003.

[41] T. Stöcker, K. Vahedipour, D. Pflugfelder, and N. J. Shah, "High-performance computing MRI simulations," *Magn. Reson. Med.*, vol. 64, no. 1, pp. 186–193, Jul. 2010.

[42] O. Bernard et al., "Deep learning techniques for automatic MRI cardiac multi-structures segmentation and diagnosis: Is the problem solved?" *IEEE Trans. Med. Imag.*, vol. 37, no. 11, pp. 2514–2525, Nov. 2018, doi: 10.1109/TMI.2018.2837502.

[43] A. Madabhushi and J. K. Udupa, "New methods of MR image intensity standardization via generalized scale," *Med. Phys.*, vol. 33, no. 9, pp. 3426–3434, Aug. 2006, doi: 10.1118/1.2335487.

[44] F. Isensee, P. F. Jaeger, S. A. A. Kohl, J. Petersen, and K. H. Maier-Hein, "NnU-Net: A self-configuring method for deep learning-based biomedical image segmentation," *Nature Methods*, vol. 18, no. 2, pp. 203–211, Feb. 2021, doi: 10.1038/s41592-020-01008-z.

[45] F. Liu, J. V. Velikina, W. F. Block, R. Kijowski, and A. A. Samsonov, "Fast realistic MRI simulations based on generalized multi-pool exchange tissue model," *IEEE Trans. Med. Imag.*, vol. 36, no. 2, pp. 527–537, Feb. 2017, doi: 10.1109/TMI.2016.2620961.

[46] R. Shaw, C. H. Sudre, T. Varsavsky, S. Ourselin, and M. J. Cardoso, "A k-space model of movement artefacts: Application to segmentation augmentation and artefact removal," *IEEE Trans. Med. Imag.*, vol. 39, no. 9, pp. 2881–2892, Sep. 2020, doi: 10.1109/TMI.2020.2972547.

[47] I. Oksuz et al., "Automatic CNN-based detection of cardiac MR motion artefacts using K-space data augmentation and curriculum learning," *Med. Image Anal.*, vol. 55, pp. 136–147, Jul. 2019, doi: 10.1016/j.media.2019.04.009.

[48] E. Kruithof, S. Amirrajab, M. J. M. Cluitmans, K. D. Lau, and M. Breeuwer, "Influence of image artifacts on image-based computer simulations of the cardiac electrophysiology," *Comput. Biol. Med.*, vol. 137, Oct. 2021, Art. no. 104773, doi: 10.1016/j.compbiomed.2021.104773.

[49] K. Scheffler, "A pictorial description of steady-states in rapid magnetic resonance imaging," *Concepts Magn. Reson.*, vol. 11, pp. 291–304, Jul. 1999, doi: 10.1002/(SICI)1099-0534(1999)11:5<291::AID-CMR2>3.0.CO;2-J.

[50] O. Dietrich, J. G. Raya, S. B. Reeder, M. F. Reiser, and S. O. Schoenberg, "Measurement of signal-to-noise ratios in MR images: Influence of multichannel coils, parallel imaging, and reconstruction filters," *J. Magn. Reson. Imag.*, vol. 26, no. 2, pp. 375–385, Aug. 2007, doi: 10.1002/jmri.20969.

[51] V. M. Campello et al., "Multi-centre, multi-vendor and multi-disease cardiac segmentation: The M&Ms challenge," *IEEE Trans. Med. Imag.*, vol. 40, no. 12, pp. 3543–3554, Dec. 2021, doi: 10.1109/TMI.2021.3090082.