

# Motion correction and super-resolution for multi-slice cardiac magnetic resonance imaging via an end-to-end deep learning approach

Zhenhong Chen, Hui Ren, Quanzheng Li, Xiang Li<sup>\*</sup>

Center for Advanced Medical Computing and Analysis, Massachusetts General Hospital and Harvard Medical School, Boston, USA

## ARTICLE INFO

### Keywords:

Cardiac magnetic resonance  
Motion correction  
Super-resolution

## ABSTRACT

Accurate reconstruction of a high-resolution 3D volume of the heart is critical for comprehensive cardiac assessments. However, cardiac magnetic resonance (CMR) data is usually acquired as a stack of 2D short-axis (SAX) slices, which suffers from the inter-slice misalignment due to cardiac motion and data sparsity from large gaps between SAX slices. Therefore, we aim to propose an end-to-end deep learning (DL) model to address these two challenges simultaneously, employing specific model components for each challenge. The objective is to reconstruct a high-resolution 3D volume of the heart ( $V_{HR}$ ) from acquired CMR SAX slices ( $V_{LR}$ ). We define the transformation from  $V_{LR}$  to  $V_{HR}$  as a sequential process of motion correction and super-resolution. Accordingly, our DL model incorporates two distinct components. The first component conducts motion correction by predicting displacement vectors to re-position each SAX slice accurately. The second component takes the motion-corrected SAX slices from the first component and performs the super-resolution to fill the data gaps. These two components operate in a sequential way, and the entire model is trained end-to-end. Our model significantly reduced inter-slice misalignment from originally  $3.33 \pm 0.74$  mm to  $1.36 \pm 0.63$  mm and generated accurate high resolution 3D volumes with Dice of  $0.974 \pm 0.010$  for left ventricle (LV) and  $0.938 \pm 0.017$  for myocardium in a simulation dataset. When compared to the LAX contours in a real-world dataset, our model achieved Dice of  $0.945 \pm 0.023$  for LV and  $0.786 \pm 0.060$  for myocardium. In both datasets, our model with specific components for motion correction and super-resolution significantly enhance the performance compared to the model without such design considerations. The codes for our model are available at [https://github.com/zhenhongchen/CMR\\_MC\\_SR\\_End2End](https://github.com/zhenhongchen/CMR_MC_SR_End2End).

## 1. Introduction

In cardiac imaging, accurate reconstruction of a high-resolution three-dimensional (3D) volume of the left ventricle (LV) and myocardium is crucial for comprehensive assessments of both their anatomy and function. Cardiac magnetic resonance (CMR) imaging is considered as the gold standard imaging modality (Garg and Friedrich, 2018; von Knobelsdorff-Brenkenhoff and Schulz-Menger, 2016; von Knobelsdorff-Brenkenhoff et al., 2017) to provide accurate and reproducible measurement of the LV and myocardial anatomy and function. For LV, the LV morphology, volumes, ejection fraction (EF) derived from CMR are critical parameters in the diagnosis, therapeutic management, prognostication and follow-up of numerous cardiovascular diseases, underscored by clinical guidelines (Heidenreich et al., 2022; Zeppenfeld et al., 2022; Otto et al., 2021). Moreover, the accurate detection of

myocardial dysfunction from CMR can not only indicate previous infarcts, cardiomyopathies, or inflammatory disease (Zeppenfeld et al., 2022), but also, through quantitative myocardial strain measurement, can identify subtle changes in cardiac function while LVEF is still preserved and have greater prognostic values after acute myocardial infarction than LVEF (Reindl et al., 2019; Eitel et al., 2018; Xu et al., 2022).

While obtaining 3D CMR data is attainable in research environments, its integration into clinical practice is hindered by the protocol mandating extended breath-holding (an impractical requirement for patients with severe cardiovascular diseases) and by the low temporal resolution. Consequently, in clinical settings, CMR data is usually acquired as a stack of two-dimensional (2D) short-axis (SAX) slices from multiple subsequent breath holds. The cardiac contours (e.g., endocardial and epicardial contours) can be segmented from each SAX slice to

<sup>\*</sup> Correspondence to: Center for Advanced Medical Computing and Analysis, Massachusetts General Hospital and Harvard Medical School, 100 Cambridge Street, Boston, MA 02114, USA.

E-mail address: [xli60@mg.harvard.edu](mailto:xli60@mg.harvard.edu) (X. Li).

<https://doi.org/10.1016/j.compmedimag.2024.102389>

Received 16 November 2023; Received in revised form 8 March 2024; Accepted 19 April 2024

Available online 29 April 2024

0895-6111/© 2024 Elsevier Ltd. All rights reserved.

delineate the boundary of each cardiac structure. Unfortunately, these 2D slices suffer from two main limitations: the presence of inter-slice misalignment and the data sparsity. Inter-slice misalignment refers to the change of the heart position in each SAX slice caused by cardiac and respiratory motion. This misalignment of the cardiac geometry across all SAX slices can cause errors in the 3D reconstruction of heart. On the other hand, data sparsity refers to the low resolution in the inter-slice axis (z-axis), leading to a notable data gap (typically around 10 mm) between adjacent slices. Therefore, two functions, including motion correction and super-resolution, are required to address these two limitations respectively, thereby reconstructing a more accurate 3D volume of the heart from 2D SAX contours.

Traditional motion correction leverages motion-free high-resolution long-axis (LAX) views (Villard et al., 2017; Sinclair et al., 2017; Yang et al., 2017). The key assumption is the SAX will align perfectly with the LAX when their line profiles (i.e., contours) along the intersection plane match. However, this approach only relies on limited information at the intersection. Besides, the LAX planes are not always available. The data sparsity is traditionally tackled by registering a high-resolution 3D mesh, made either by roughly connecting the contours (Joyce et al., 2022; de Marvao et al., 2014) or by shape priors (Yang et al., 2017; Bai et al., 2015; Villard et al., 2018; Banerjee et al., 2021), towards the 2D contours. However, the registration is computationally expensive, and its accuracy may be comprised by the over-smoothing and the bias of shape priors when dealing with the abnormal shapes (Albà et al., 2016).

Recently, researchers have leveraged the power of deep learning (DL) to address both the motion correction and super-resolution in CMR. Regarding motion correction, Dangi et al (Dangi et al., 2018). employed a convolutional neural network (CNN) to determine the correct center of left ventricle (LV) in each slice. However, this method handles each slice individually and thus fails to consider the morphology consistency over the entire LV. Chang et al (Chang et al., 2022a). proposed an unsupervised 3D recurrent neural network with the objective function to minimize the distance between misaligned SAX contours and LAX contours. Regarding super-resolution, Wang et al (Wang et al., 2021). implemented a 3D version of the enhanced deep residual network (EDSR) (Lim et al., 2017) which is one of the state-of-the-art super-resolution CNN. Some other researchers approached the super-resolution task in different problem settings. For example, Beetz et al (Beetz et al., 2021). framed it as a point completion task by converting 2D contours into a sparse point cloud and using a point completion network to fill in the missing points between adjacent slices. Furthermore, Xu et al (Xu et al., 2019). conceptualized it as an image inpainting task, inpainting the missing parts between adjacent slices. However, both studies (Beetz et al., 2021; Xu et al., 2019) lack thorough validation on the real-world dataset. Additionally, Beetz et al (Beetz et al., 2023). addressed super-resolution as a mesh deformation problem, transforming 2D contours into surface meshes and employing a U-Net to enhance the z-resolution of meshes. Meanwhile, Biffi et al (Biffi et al., 2019). formulated it as an image generation problem by employing a conditional variational autoencoder (VAE) to generate high-resolution 3D volume based on both SAX and LAX as the conditioning inputs for the VAE. Chang et al (Chang et al., 2022b). also utilized the interpolation in the latent space to generate a better spatial interpolation to fill the data gap between two adjacent SAX slices.

It is noteworthy that some of these super-resolution studies (Beetz et al., 2021; Jiang et al., 2021; Wang et al.; Zhu et al., 2021; Beetz et al., 2023; Xu et al., 2019) also address the motion correction concurrently. However, their DL models does not incorporate specific architectural designs for motion correction and super-resolution separately. Instead, their models are tailored for super-resolution, while the motion correction capability is achieved by preparing specific training data which includes pairs of the simulated motion-misaligned 2D contours and the ground truth motion-free 3D volume. We hypothesized that models without specific designs for each function may compromise the accuracy of the reconstructed 3D volumes. Recently, Wang et al (Wang

et al., 2021). proposed a DL model with joint ability to do super-resolution concurrently, which consist of specific designs for each function individually. Concretely, a latent space to generate plausible high-resolution 3D volumes and a motion layer to predict plausible inter-slice alignment are both optimized to minimize the loss defined on 2D contours. However, they did not investigate the effectiveness of each function individually.

In this paper, we have three aims. First, we aim to propose a new DL model comprising of specific model components for motion correction and super-resolution separately. Second, we aim to demonstrate that our model with specific components for these two functions outperforms a model without such design considerations. Third, we seek to investigate the difference in model performance between integrating two functions in an end-to-end model versus separating two functions in two individual models that operate in a sequential fashion.

## 2. Methods

### 2.1. Problem formulation

We focus on the analysis of LV and myocardium in this study, as their importance has been illustrated in the introduction. We assume that the 2D contours of endocardium and epicardium have been segmented in a pre-processing step. In our study, the segmentation was prepared manually. We represent a stack of acquired 2D CMR SAX slices as  $V_{LR} = \{v_k\}_{k \in N}$ , where  $LR$  denotes low-resolution in the z-axis, and  $v_k \in \{LV, myocardium\}$  denotes the segmentation at the  $k^{th}$  voxel.  $N$  stands for the total number of voxels, equal to  $x \times y \times z_{LR}$ , with  $x$  and  $y$  being the x- and y-dimensions and  $z_{LR}$  being the number of SAX slices. The corresponding high-resolution 3D volume of LV and myocardium is denoted by  $V_{HR} = \{v_k\}_{k \in x \times y \times z_{HR}}$  where  $HR$  denotes high-resolution in the z-axis. The image dimension in x-, y- and z-axis is  $x$ ,  $y$  and  $z_{HR}$  respectively. Note that the dimensionality solely differs in the z-axis (i.e., the number of SAX slices) between  $V_{LR}$  and  $V_{HR}$ .

The transformation from a high-resolution 3D volume  $V_{HR}$  into a stack of SAX slices  $V_{LR}$  can be formulated as:

$$V_{LR} = M_t(S_t(V_{HR})) \quad (1)$$

where  $S_t$  denotes the down-sampling factor equal to  $z_{LR}/z_{HR}$  in z-axis.  $M_t = [(t_{x,i}, t_{y,i})]_{i \in \{1,2,3, \dots, z_{LR}\}}$  represents the shift of the heart in each SAX slice caused by cardiac and respiratory motion. Here,  $(t_{x,i}, t_{y,i})$  denote the shift along x- and y-axes in the  $i^{th}$  slice. Eq. 1 demonstrates that during a CMR scan, a high-resolution 3D volume of heart is sampled typically every 10 mm (down-sampling) into a stack of SAX slices with the position of heart in each SAX slice being misaligned due to the motion.

In contrast, the reconstruction of a high-resolution 3D volume  $V_{HR}$  from its corresponding acquired SAX slices  $V_{LR}$  can be formulated as:

$$V_{HR} = S_t(-M_t(V_{LR})) \quad (2)$$

where  $S_t$  denotes the upsampling factor equal to  $z_{HR}/z_{LR}$  in the z-axis, while  $-M_t(V_{LR})$  denotes the motion-corrected SAX slices yielded by repositioning each SAX slice using the inverse of heart shift in that slice,  $-M_t$ . We call  $-M_t$  as “displacement vectors” in this paper. This Eq. 2 demonstrates that the reconstruction should occur by performing motion correction followed by super-resolution in the z-axis in a sequential fashion.

### 2.2. Deep learning model design

Following the concept in Eq. 2, we have designed our DL model with two distinct components: (1) the first component conducts motion correction by predicting displacement vectors  $-M_t = [-t_{x,i}, -t_{y,i}]_{i \in \{1,2,3, \dots, z_{LR}\}}$  and subsequently applying  $-M_t$  onto the

original SAX slices  $V_{LR}$ , thus producing motion-corrected SAX slices,  $-M_t(V_{LR})$ ; (2) the second component performs super-resolution  $S_1$  on the output of the first component, thus producing  $V_{HR}$ . These two components operate sequentially. The overall design of our DL model is illustrated in Fig. 1.

### 2.2.1. The motion correction component

The architecture of the motion correction component is shown in Figs. 1 and 2A. This component takes a stack of low-resolution and misaligned SAX slices,  $V_{LR}$  as the model input. The motion correction component employs a 3D residual neural network to predict  $z_{LR}$  pairs of displacement vectors  $[-t_{x,i}, -t_{y,i}]_{i \in \{1,2,3, \dots, z_{LR}\}}$ . In this study, we consistently zero-pad  $V_{LR}$  to maintain the number of slices  $z_{LR} = 12$ . Additionally, this component also incorporates a spatial transformer module to reposition each slice in  $V_{LR}$  by the predicted displacement vectors, generating the motion-corrected slices  $-M_t(V_{LR})$  as defined in Eq. 2. The loss function for this component equals to the summation of the loss associated with the predicted displacement vectors (measured by mean-absolute-error, MAE) and the loss of motion-corrected slices (measured by Dice coefficient), as shown in Eq. 3:

$$\mathcal{L}_{\text{motion-correction}} = \beta_1 \mathcal{L}_{-M_t} + \beta_2 \mathcal{L}_{-M_t(V_{LR})} \quad (3)$$

where  $\beta_1$  and  $\beta_2$  represent two loss weights.

### 2.2.2. The super-resolution component

Following the motion correction component, we design a super-resolution component. This component uses a 3D version of the EDSR (Lim et al., 2017) which is one of the state-of-the-art super-resolution CNN. Wang et al (Wang et al., 2021). have shown the effectiveness of EDSR in achieving super-resolution,  $S_1$ , along the z-axis of CMR

contours. As depicted in Figs. 1 and 2B, the EDSR takes the output of the first component, which is the motion-corrected SAX slices  $-M_t(V_{LR})$ , as its input. It then predicts  $V_{HR}$  by upsampling the z-resolution of the input by a factor of 5. Consequently,  $V_{HR}$  comprises  $z_{HR} = z_{LR} \times 5 = 60$  slices, with a reduced slice thickness from the original 10 mm to 2 mm. The loss function for this component quantifies the discrepancy between the predicted  $V_{HR}$  and the ground truth  $V_{HR}$  in terms of Dice coefficient.

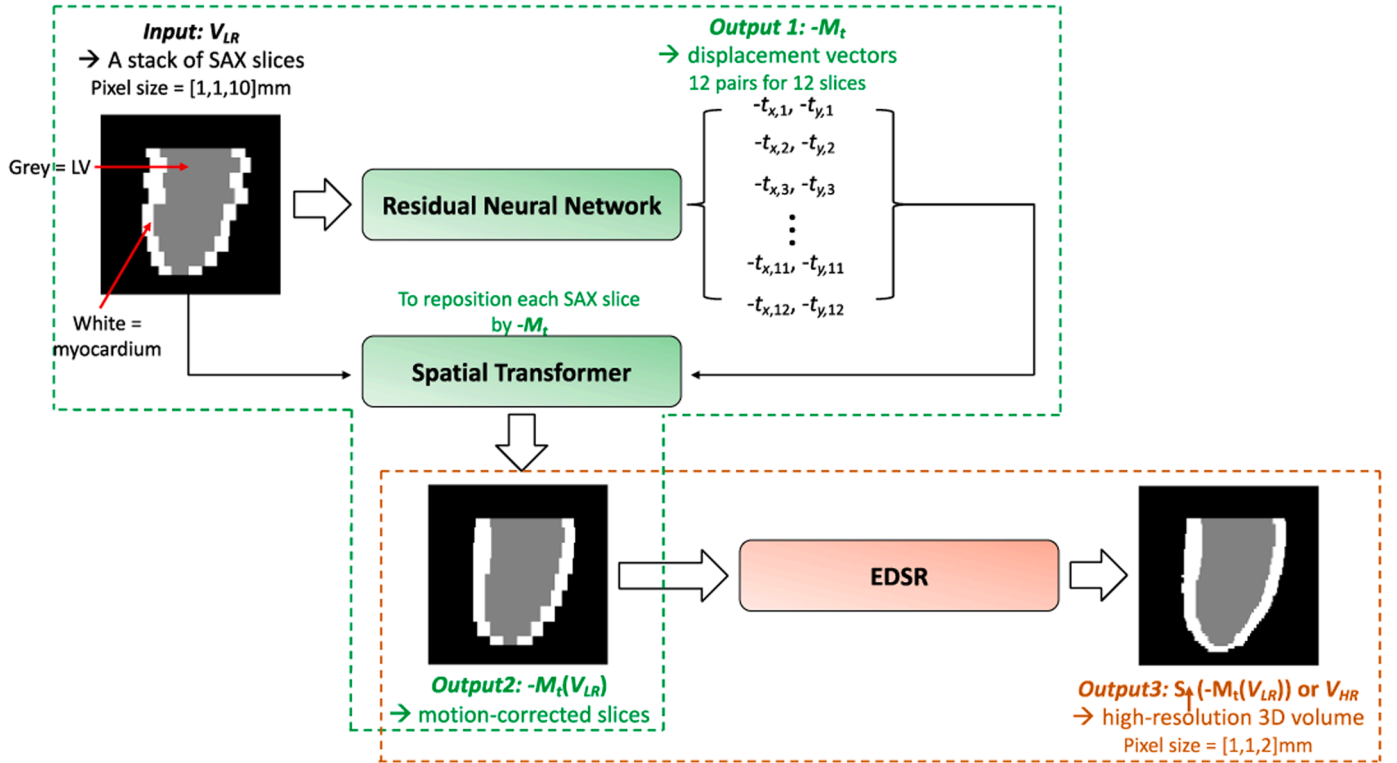
In summary, our DL model takes the original SAX slices as model input. It first predicts the displacement vectors that are subsequently applied to each SAX slice for motion correction. Following this, it performs super-resolution on the motion-corrected slices to fill the data gaps between slices and thereby produce a high-resolution 3D volume of LV and myocardium. In total, this model has three outputs: (1) the displacement vectors  $-M_t$ , (2) the motion-corrected SAX slices,  $-M_t(V_{LR})$  and (3) the high-resolution 3D volume  $V_{HR}$ . The overall loss function encompasses the summation of the loss in the motion correction component (the first two terms on the right-hand side of Eq. 4) and the loss in super-resolution component (the last term):

$$\mathcal{L}_{DL} = \beta_1 \mathcal{L}_{-M_t} + \beta_2 \mathcal{L}_{-M_t(V_{LR})} + \beta_3 \mathcal{L}_{V_{HR}} \quad (4)$$

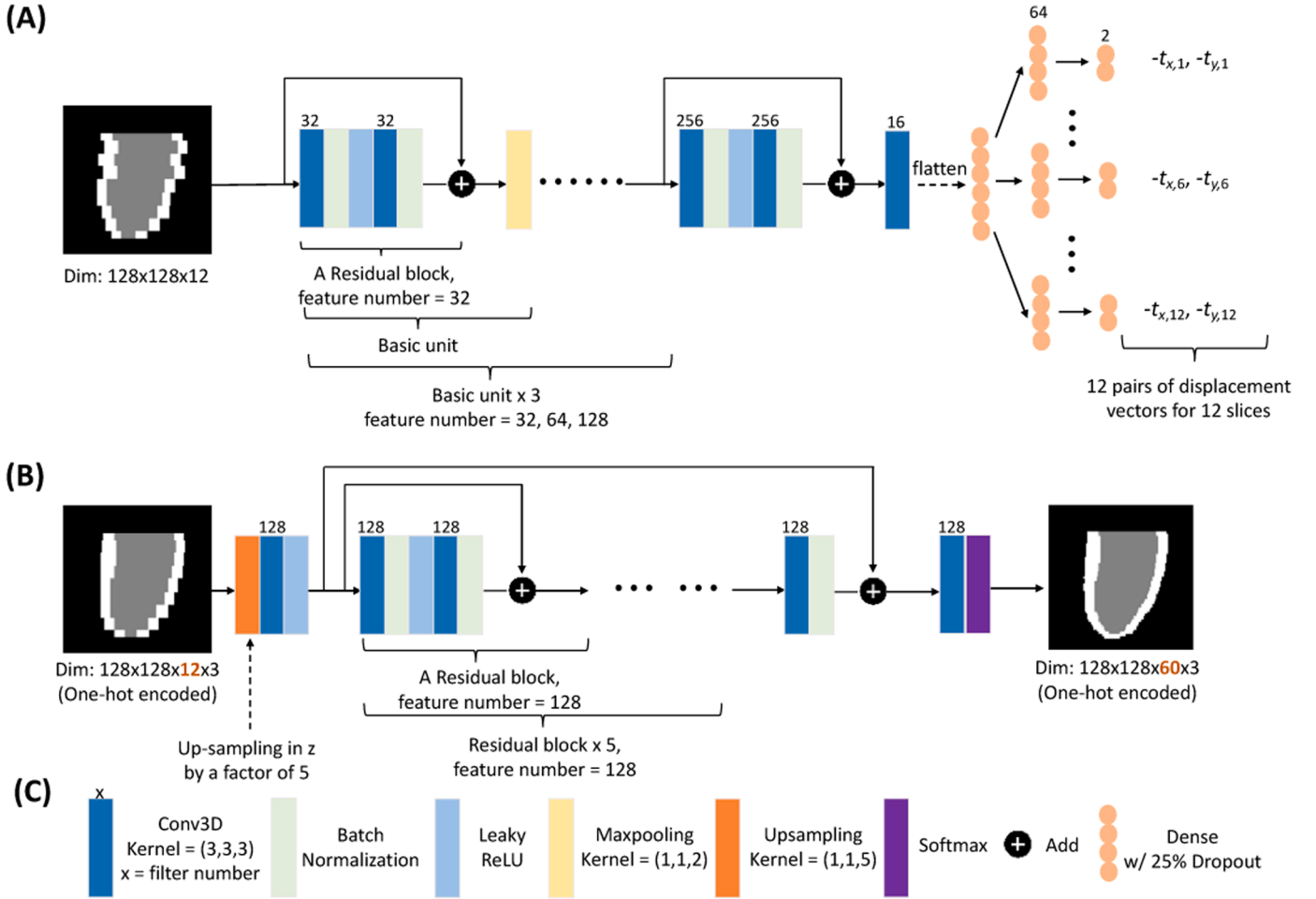
In our study, we maintained uniform loss weights, setting  $\beta_1 = \beta_2 = \beta_3$  without further exploration of their impact. In this paper, we denote this model as DL<sub>end2end</sub> to emphasize this model is trained end-to-end.

### 2.3. Comparison to the model only tailored for super-resolution

We investigate whether our DL model, DL<sub>end2end</sub>, with specific components for motion correction and super-resolution offers superior performance compared to a model lacking such specific design considerations (in the other words, only tailed for super-resolution). Our



**Fig. 1. Model Designs.** Fig. 1 presents an overview of the model. It consists of two different components for two functions. First, it has a residual neural network and spatial transformer module for motion correction function (this component is highlighted by the green dashed line). This component has two related outputs (words in green): Output 1 represents the displacement vectors, and Output 2 represents the motion-corrected SAX slices. Second, it has an EDSR for super resolution function (this component is highlighted by the orange dashed line). This component takes Output 2 as input and yields the third output of the entire model (words in orange): Output 3, which is the high-resolution 3D volume. Two components operate sequentially. CMR contour images denote pixels labeled as LV in grey, pixels as myocardium in white, and pixels as background in black. SAX = short axis. EDSR = enhanced deep residual network. Definitions for  $V_{LR}$ ,  $M_t$  and  $V_{HR}$  can be found in Section 2.1.



**Fig. 2. Component architecture.** Detailed architectures for the residual neural network in the motion correction component and the EDSR in the super-resolution component are illustrated in (A) and (B) respectively. Notably, (B) shows the EDSR has an upsampling layer as the first layer to enhance the z-resolution by a factor of 5. (C) provides the breakdown of each layer in the network.

intuition is that our model architecture reflecting the transformation process from  $V_{LR}$  to  $V_{HR}$  in Eq. 2 coupled with the inclusion of losses from both tasks might have the potential to yield enhanced results.

To ensure a fair comparison, we maintain consistency in the model input and keep the model architecture as similar as possible between two models. Therefore, we selected the 3D version of EDSR (Lim et al., 2017) as the model without specific designs (see Fig. 2B). EDSR was originally crafted solely for super-resolution, and it serves as the super-resolution component in our DL model. Much like the approach in Wang et al (Wang et al., 2021), this EDSR was trained to fulfill both motion correction and super-resolution functions by utilizing training data composed of pairs of  $V_{LR}$  (low-resolution, motion-misaligned) and the corresponding ground truth  $V_{HR}$  (high-resolution, motion-free). In the other words, the model is designed to directly connect the Input and the Output 3 in Fig. 1 by the EDSR.

We evaluated two models' performance (DL<sub>end2end</sub> vs. EDSR) by comparing their predicted  $V_{HR}$  against the ground truth  $V_{HR}$  in terms of Dice coefficient and Hausdorff distance (HD). We used one-tailed paired t-test to assess whether one model outperforms the other and the statistical significance was determined at a threshold of  $p < 0.05$ .

#### 2.4. Comparison between the end-to-end model and the two-step pipeline

We investigate whether training the motion correction and super-resolution components as two individual models and employing them as a two-step pipeline can influence model performance. Our hypothesis

is that the end-to-end model might have superior performance than the two-step pipeline since in the end-to-end model, the loss in the predicted  $V_{HR}$  can backpropagate to the motion correction component, potentially improving the motion correction results and thus providing better input for super-resolution.

To train two models individually, the first motion correction model (comprising a residual neural network and spatial transformer module, highlighted by the green dashed line in Fig. 1) was trained on pairs of original SAX slices ( $V_{LR}$ ) and its corresponding displacement vectors ( $-M_t$ ) along with the ground truth motion-corrected slices ( $-M_t(V_{LR})$ ). The second super-resolution model (comprising EDSR, highlighted by the orange dashed line in Fig. 1) was trained by pairs of motion-corrected slices ( $-M_t(V_{LR})$ ) and its corresponding ground truth high-resolution 3D volume ( $V_{HR}$ ). Once trained, the new  $V_{LR}$  undergoes the motion correction model followed by the super-resolution model as a two-step pipeline to yield  $V_{HR}$ . This approach is denoted as DL<sub>2-step</sub> in this paper.

We evaluated the performance of two models (DL<sub>end2end</sub> vs. DL<sub>2-step</sub>) from two aspects. First, we compared the inter-slice misalignment in their motion-corrected SAX slices (averaged across each slice that contains the heart) against the ground truth motion-free SAX slices in terms of MAE. We used the Mann-Whitney U test to assess whether one model outperforms the other. Second, we compared their predicted  $V_{HR}$  against the ground truth  $V_{HR}$  in terms of Dice coefficient and HD and used one-tailed paired t-test to assess whether one model outperforms the other. Statistical significance was determined at a threshold of  $p < 0.05$ .



## 2.5. Model training and evaluation

### 2.5.1. Simulation dataset

First, we prepared a simulation dataset to train and validate our model with known ground truth high-resolution 3D volumes  $V_{HR}$  and known ground truth heart shifts  $M_t$ . Concretely, a public dataset of high-resolution 3D balanced steady-state free precession cine CMR from 1331 healthy adults (Savioli et al., 2021) called “Cardiac Super-resolution Label Maps” was used. Each CMR study was provided as segmentations of LV and myocardium rather than greyscale images in this dataset. These segmentations serve as the ground truth high-resolution 3D volumes  $V_{HR}$ . The original high-resolution data was resampled to have  $1 \times 1 \times 2$  mm spatial resolution and was cropped to an image dimension of  $128 \times 128 \times 60$  while ensuring the full LV coverage. We only used the end-diastolic (ED) frame in our study.

To simulate the transformation from  $V_{HR}$  to  $V_{LR}$ , we followed the process outlined in Eq. 1. We first sampled  $V_{HR}$  every 5 SAX slices, effectively down-sampling by a factor of 5 ( $S_1$  in Eq. 1). The resulting SAX stack has an image dimension of  $128 \times 128 \times 12$  with the slice thickness increasing from original 2 mm to 10 mm. This low-z-resolution SAX stack was free of motion misalignment and thus served as the ground truth motion-free SAX slices in our study (i.e.,  $-M_t(V_{LR})$ ). Subsequently, we introduced heart shift  $M_t = [(t_{x,i}, t_{y,i})]_{i \in \{1,2,3, \dots, z_{LR}\}}$  in each SAX slice. Each pair of  $t_{x,i}, t_{y,i}$  was assumed to be independent and randomly generated following a Gaussian distribution fitted from the UK Biobank (Tarroni et al., 2020). The resulting misaligned SAX slices formed  $V_{LR}$ , which served as the input of our model. For each of 1331 CMR study, we simulated 15 different motions (i.e., 15 different  $M_t$ ), resulting in a dataset of 19,965 pairs of  $V_{LR}$  and  $V_{HR}$ .

The simulation dataset was split at the patient-level, with 80% for training and validation and the remaining 20% for testing. We trained and tested our end-to-end DL model ( $DL_{end2end}$ ) as well as two models for comparison (EDSR in Section 2.3 and  $DL_{2-step}$  in Section 2.4) using the same training and testing dataset to ensure a fair competition. We only reported the results in the testing dataset in this paper.

### 2.5.2. Real-world Data

The trained models ( $DL_{end2end}$ , EDSR and  $DL_{2-step}$ ) were tested on a real-world dataset, Sunnybrook Cardiac Data (also known as the 2009 Cardiac MR Left Ventricle Segmentation Challenge data) (Radau et al., 2009), without further fine-tuning or re-training. This dataset contains 45 cine CMR studies. Each study includes a SAX stack as well as LAX views including two-chamber (2CH), three-chamber (3CH) and four-chamber (4CH) views. These LAX views serve as high-resolution motion-free ground truth. For evaluation, we re-sliced the predicted  $V_{HR}$  along the LAX views and compared the re-sliced contours with the ground truth LAX contours in terms of Dice coefficient.

## 3. Results

### 3.1. Results on simulation dataset

#### 3.1.1. Performance of our Model, $DL_{end2end}$

To recall, our model  $DL_{end2end}$  provides three outputs: two motion correction-related outputs including (1) the displacement vectors  $-M_t$  and (2) the motion-corrected SAX slices,  $-M_t(V_{LR})$ , and one super-resolution-related output which is (3) the high-resolution 3D volume  $V_{HR}$ .

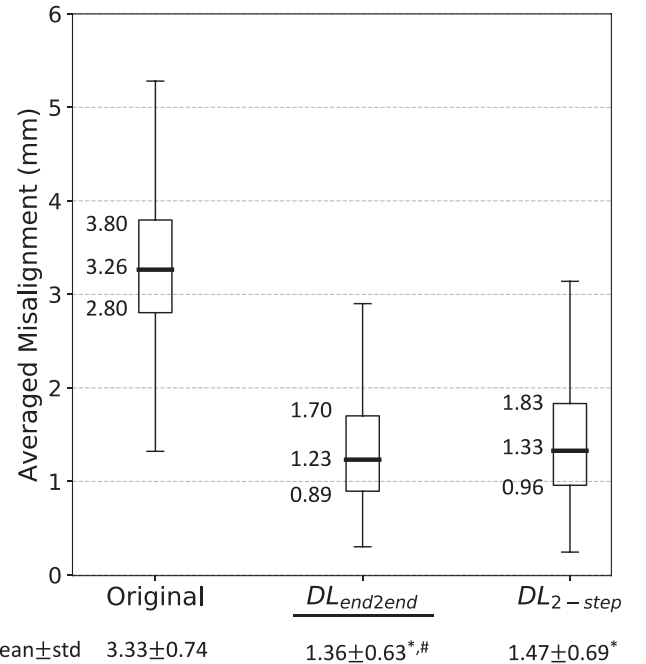
We first assess the performance of motion correction. We compare the motion-corrected SAX slices against the ground truth motion-free SAX slices in terms of inter-slice misalignment averaged across each SAX slice (only SAX slices that contain the LV and myocardium). Concretely, the original misalignment was  $3.33 \pm 0.74$  mm in the simulated testing dataset. After motion correction by our model, the inter-slice misalignment in motion-corrected SAX slices significantly

decreased to  $1.36 \pm 0.63$  mm ( $p < 0.001$  by Mann-Whitney U test). This comparison is shown in the first and second columns in Fig. 3. Next, we evaluate the performance of super-resolution by comparing the predicted  $V_{HR}$  with the ground truth  $V_{HR}$  in terms of Dice coefficient as well as Hausdorff Distance (HD). As shown in Fig. 4, our model’s predictions achieved Dice of  $0.974 \pm 0.010$  for LV and Dice of  $0.938 \pm 0.017$  for myocardium. Additionally, the HD measurement for LV were  $2.08 \pm 0.39$  mm and HD for myocardium were  $2.21 \pm 0.39$  mm.

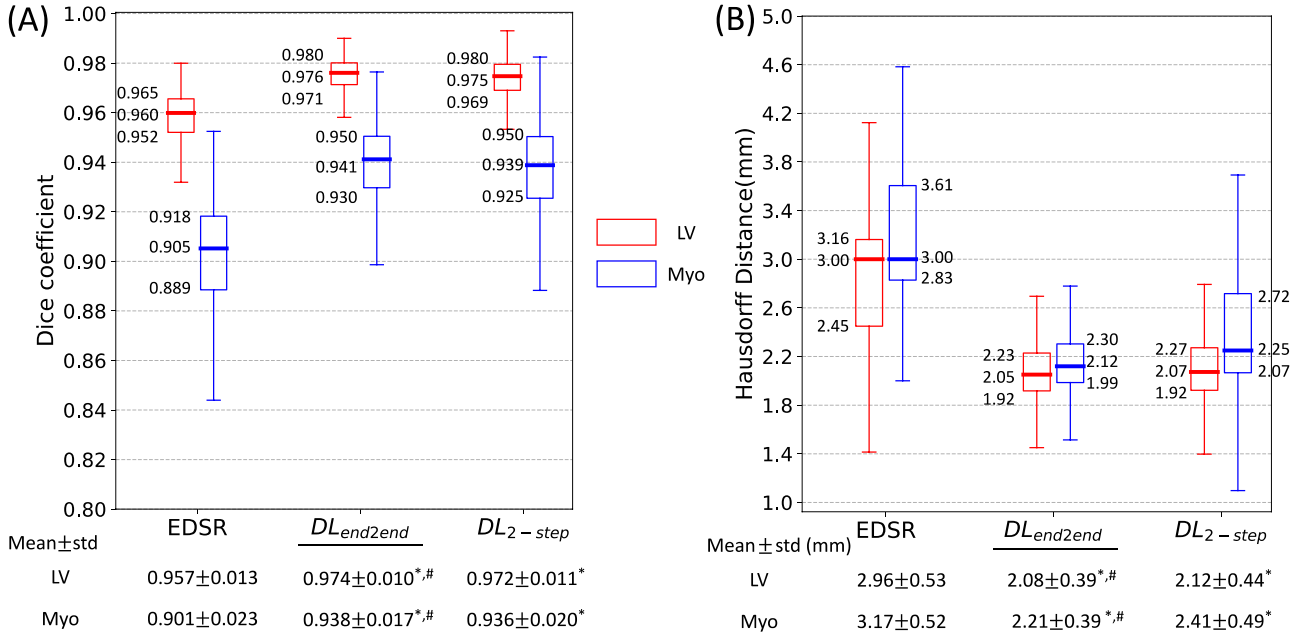
#### 3.1.2. Comparison to performance of EDSR

To recall, EDSR is the model without specific components for motion correction and super-resolution. Since it directly outputs  $V_{HR}$  without predicting intermediate motion-corrected results, we solely report the comparison between the predicted  $V_{HR}$  and the ground truth  $V_{HR}$ .

Comparatively, our model  $DL_{end2end}$  exhibited higher Dice and lower HD measurements for both LV and myocardium compared to EDSR. As shown in Fig. 4 (the 1st and 2nd columns), EDSR achieved a Dice of  $0.957 \pm 0.013$  for LV while our model achieved a significantly higher Dice of  $0.974 \pm 0.010$  ( $p < 0.001$  by one-tailed t-test). Similarly, EDSR obtained a Dice  $0.901 \pm 0.023$  for myocardium while our model had significantly higher Dice of  $0.938 \pm 0.017$  ( $p < 0.001$ ). In terms of HD, EDSR achieved  $2.96 \pm 0.53$  mm for LV and  $3.17 \pm 0.52$  mm for myocardium, while our model achieved significantly smaller HD of  $2.08 \pm 0.39$  mm ( $p < 0.001$ ) for LV and  $2.21 \pm 0.39$  mm ( $p < 0.001$ ) for myocardium. Qualitatively, we prepared two illustrative examples of the  $V_{HR}$  produced by our model and EDSR in Fig. 5 (refer to the second and the fourth column). It shows that our model tends to reconstruct more accurate boundary of the 3D volumes.



**Fig. 3. Motion correction performance.** These box plots show the inter-slice misalignment (y-axis) averaged across all SAX slices that contain the heart. The three columns, from left to right, represent the original data in the simulation dataset, the motion-corrected SAX slices generated by our model  $DL_{end2end}$  (underlined), and the motion-corrected SAX slices generated by the two-step pipeline  $DL_{2-step}$ , respectively. For each box, the three numbers besides denote the upper quartile, median (indicated by the bold line in the middle of the box), and lower quartile from top to bottom respectively. The mean and standard deviation (std) are also listed below each box plot. An asterisk (\*) represents “significantly lower than original”, and a hash (#) represents “significantly lower than  $DL_{2-step}$ ”. All measurements are in millimeters. Std = standard deviation.



**Fig. 4.** Predicted  $V_{HR}$  compared with ground truth  $V_{HR}$  in the simulation dataset. (A) illustrates Dice coefficient and (B) illustrates Hausdorff distance. The three columns, from left to right, represent  $V_{HR}$  generated by EDSR,  $V_{HR}$  generated by our model  $DL_{end2end}$  (underlined), and  $V_{HR}$  generated by the two-step pipeline  $DL_{2-step}$ , respectively. The red box represents LV while the blue box represents myocardium. An asterisk (\*) represents “significantly higher than EDSR” in (A) and “significantly lower than EDSR” in (B), and a hash (#) represents “significantly higher than  $DL_{2-step}$ ” in (A) and “significantly lower than  $DL_{2-step}$ ” in (B). Myo = myocardium.

### 3.1.3. Comparison to performance of the two-step pipeline, $DL_{2-step}$

To recall, in  $DL_{2-step}$  we trained the motion correction and super-resolution components in our model as two individual models and employing these two as a two-step pipeline.

We first compare the motion correction performance in  $DL_{end2end}$  and  $DL_{2-step}$  in terms of inter-slice misalignment.  $DL_{2-step}$  achieved inter-slice misalignment as  $1.47 \pm 0.69$  mm in its output of motion correction step, while our model  $DL_{end2end}$  had significantly lower misalignment as  $1.36 \pm 0.63$  mm ( $p < 0.001$  by Mann-Whitney U test). This comparison is shown in the second and third columns in Fig. 3. We then compare the super-resolution performance. As shown in Fig. 4 (the 2nd and 3rd columns), we found out  $DL_{end2end}$  achieved better performance in terms of Dice and HD compared to  $DL_{2-step}$ . Concretely,  $DL_{2-step}$  achieved a Dice of  $0.972 \pm 0.011$  for LV and  $0.936 \pm 0.020$  for myocardium while our model achieved a significantly higher Dice of  $0.974 \pm 0.010$  ( $p < 0.001$  by one-tailed t-test) and  $0.938 \pm 0.017$  ( $p < 0.001$ ) for LV and myocardium, respectively. Similarly,  $DL_{2-step}$  achieved HD of  $2.12 \pm 0.44$  mm for LV and  $2.41 \pm 0.49$  mm for myocardium, while our model achieved significantly smaller HD of  $2.08 \pm 0.39$  mm ( $p < 0.001$ ) for LV and  $2.21 \pm 0.39$  mm ( $p < 0.001$ ) for myocardium. Qualitatively, Fig. 5 shows that the  $V_{HR}$  produced by these two methods are very similar to each other (refer to the second and the third column).

Note that although  $DL_{2-step}$  had quantitatively lower performance than  $DL_{end2end}$ , it still had significantly higher Dice and lower HD compared to EDSR ( $p < 0.001$  for all comparisons, refer to the first and third columns in Fig. 4).

### 3.2. Results on real-world data

Once we obtained the high-resolution 3D volumes of LV and myocardium ( $V_{HR}$ ) from the DL models, we re-sliced the  $V_{HR}$  along the LAX views and compared to the ground truth LAX contours. The Dice coefficient was averaged across all LAX views and reported for each of three trained approaches ( $DL_{end2end}$ , EDSR and  $DL_{2-step}$ ) in Fig. 6. Concretely, regarding Dice for LV,  $DL_{2-step}$  had the highest value as

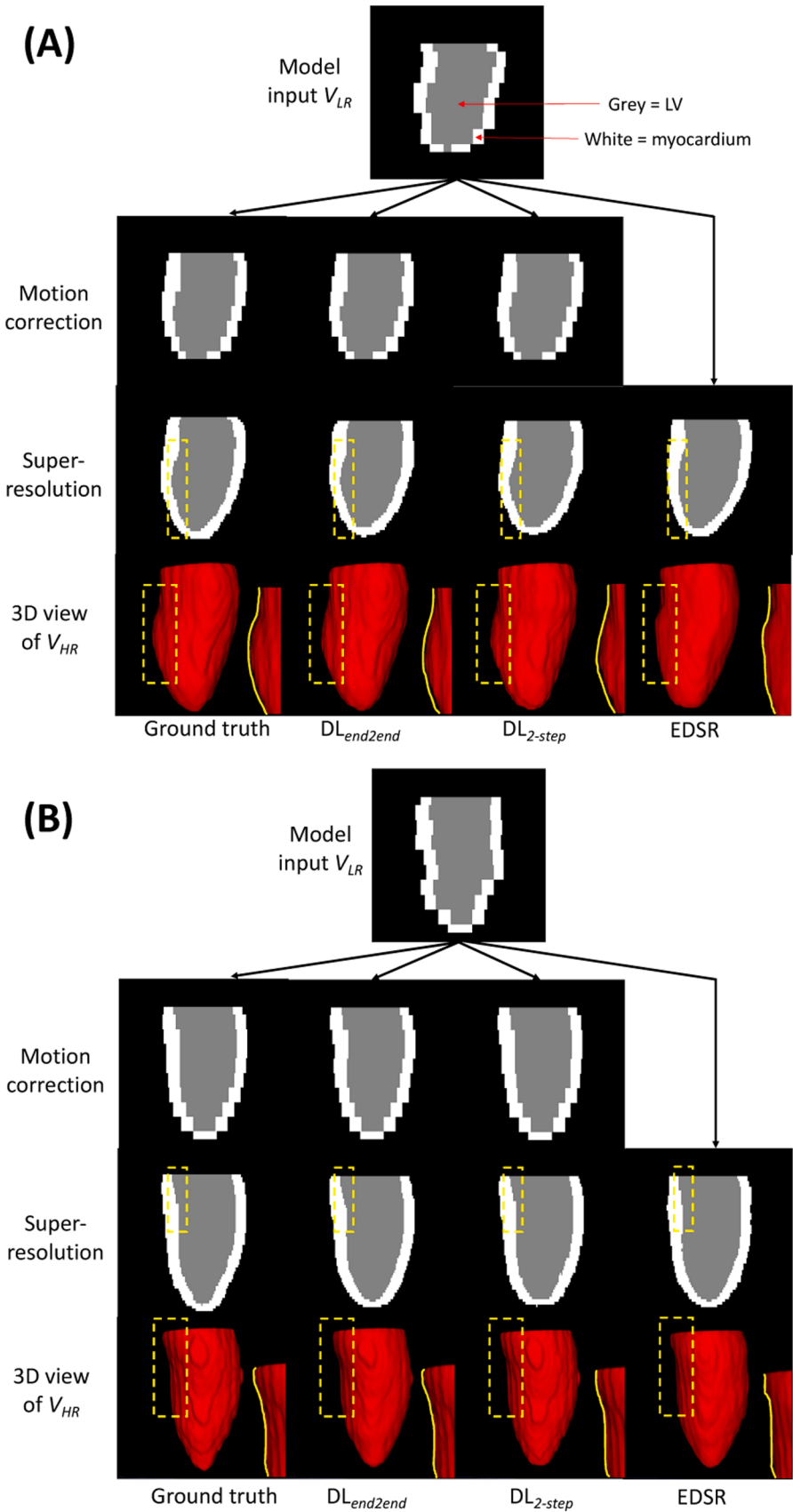
$0.947 \pm 0.023$ , while  $DL_{end2end}$  had the second highest as  $0.945 \pm 0.023$  and EDSR had the lowest as  $0.922 \pm 0.058$ . There is no significant difference between  $DL_{end2end}$  and  $DL_{2-step}$  ( $p = 0.917$ ) while both are significantly higher than EDSR ( $p = 0.002$  for both). Regarding Dice for myocardium,  $DL_{end2end}$  had the highest value as  $0.786 \pm 0.060$ , while  $DL_{2-step}$  had the second highest as  $0.783 \pm 0.070$  and EDSR had the lowest as  $0.743 \pm 0.075$ . There is no significant difference between  $DL_{end2end}$  and  $DL_{2-step}$  ( $p = 0.362$ ) while both are significantly higher than EDSR ( $p < 0.001$  for both). Two visual examples of the results of our model are presented in Fig. 7.

## 4. Discussion

In this paper, we addressed two critical challenges in reconstructing a high-resolution 3D volume of the heart ( $V_{HR}$ ) from CMR SAX slices ( $V_{LR}$ ): motion correction and super-resolution. We proposed a new end-to-end model comprising two dedicated components for motion correction and super-resolution respectively. Two components operate sequentially to transform  $V_{LR}$  to  $V_{HR}$ . Our model was trained on a simulation dataset with pairs of ground truth high-resolution 3D volumes  $V_{HR}$  and corresponding  $V_{LR}$  with simulated inter-slice misalignment and z-axis down-sampling, enabling the supervised learning. In the simulated testing dataset, our model demonstrated excellent performance in both motion correction (reduced inter-slice misalignment from originally  $3.33 \pm 0.74$  mm to  $1.36 \pm 0.63$  mm) and super-resolution (Dice of  $0.974 \pm 0.010$  for LV and Dice of  $0.938 \pm 0.017$  for myocardium). Furthermore, we evaluated our model in a real-world dataset characterized by naturally occurring low-resolution SAX slices and potential inter-slice misalignment. Without further fine-tuning, our model achieved Dice of  $0.945 \pm 0.023$  for LV and  $0.786 \pm 0.060$  for myocardium when compared to the ground truth LAX contours in this real-world dataset.

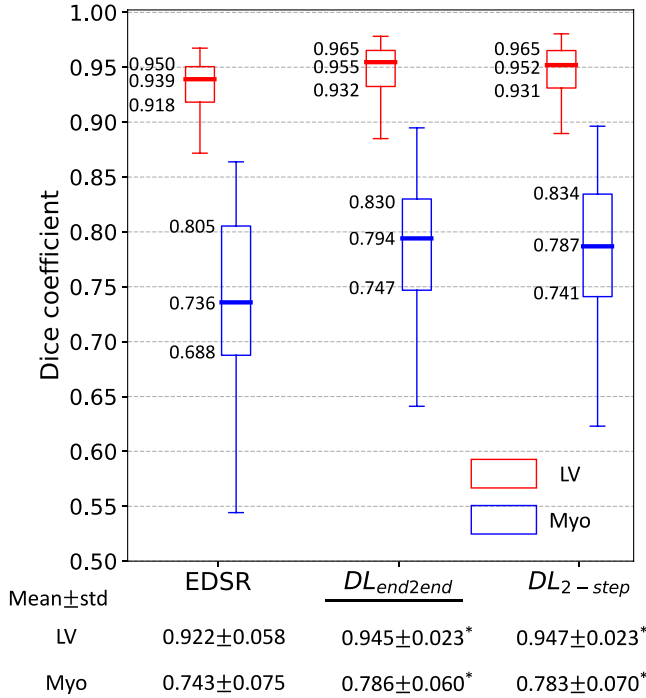
### 4.1. Comparison to Performance of EDSR

In this paper, an EDSR directly trained on pairs of  $V_{LR}$  and  $V_{HR}$  served



(caption on next page)

**Fig. 5. Examples of predicted high-resolution 3D volumes,  $V_{HR}$  in the simulation dataset.** (A) and (B) present two examples of transforming a stack of SAX slices ( $V_{LR}$ ) into a high-resolution 3D volume ( $V_{HR}$ ) using different DL approaches. Each of the four columns, from left to right, represent the ground truth, and the results of  $DL_{end2end}$ ,  $DL_{2-step}$ , and EDSR, respectively. The top first row displays the  $V_{LR}$  used as the model input. The second row displays the motion-corrected SAX slices. Notably, no image is shown in the second row for EDSR since this approach does not produce the intermediate motion correction results. The third row showcases the predicted  $V_{HR}$ , and the fourth row displays the 3D visualization of the endocardium (red) within  $V_{HR}$ . The yellow boxes highlight the region where different approaches yield different reconstruction. A zoomed-in view of the yellow box is presented on the right side of the fourth row, with the endocardial boundary delineated by a yellow line. In both (A) and (B), our model  $DL_{end2end}$  yields  $V_{HR}$  that closely resembles the ground truth, particularly in the endocardial boundary in the yellow box region.



**Fig. 6. Predicted  $V_{HR}$  compared with ground truth LAX contours in the real-world dataset.** The Dice coefficient comparisons are shown. The three columns, from left to right, represent  $V_{HR}$  generated by EDSR,  $V_{HR}$  generated by our model  $DL_{end2end}$  (underlined), and  $V_{HR}$  generated by the two-step pipeline  $DL_{2-step}$ , respectively. The red box represents LV while the blue box represents myocardium. An asterisk (\*) represents “significantly higher than EDSR”. Myo = myocardium.

as an example of DL models that lack specific model components for motion correction and super-resolution. We show that our model  $DL_{end2end}$  exhibited significantly higher performance both in the simulation dataset (as depicted in the first and second columns in Fig. 4) and the real-world dataset (as depicted in the first and second columns in Fig. 6). A potential explanation is that both model has an EDSR architecture (in  $DL_{end2end}$  we have super-resolution component as an EDSR architecture). However, in our model the super-resolution component operates on motion-corrected input, whereas in the case of EDSR, it directly operates on uncorrected data. In our model, the task is more manageable for the EDSR component as it can focus exclusively on super-resolution aspect. Empirically, we observed that the  $V_{HR}$  produced by a single EDSR tended to produce less accurate boundary in the reconstructed volumes compared to our DL model (refer to Fig. 5 for visual examples).

#### 4.2. Comparison to performance of the two-step pipeline, $DL_{2-step}$

We demonstrated that our  $DL_{end2end}$ , which incorporates both motion correction and super-resolution as an end-to-end process, outperforms a single motion correction model (the first step in the two-step pipeline) in reducing inter-slice misalignment. Specifically,  $DL_{end2end}$  reduced the misalignment to  $1.36 \pm 0.63$  mm, while  $DL_{2-step}$  achieved a larger value

of  $1.47 \pm 0.69$  mm. Regarding the final predicted  $V_{HR}$ ,  $DL_{end2end}$  had slightly better performance in LV and myocardium volume reconstruction in the simulation dataset, while both models showed no significant difference in performance in the real-world dataset. Overall, the distinction between an end-to-end model and a two-step pipeline did not yield a substantial difference in performance.

#### 4.3. Comparison to other task-specific models

Our model integrates two task-specific components: motion correction followed by super-resolution component. In motion correction, we draw inspiration from Chang et al (Chang et al., 2022a), that takes misaligned SAX slices,  $V_{HR}$ , as model input and outputs the correct displacement vector,  $-M_t$ , for each slice. The difference is that Chang et al. uses LAX views as reference to optimize the prediction in an unsupervised fashion while we utilize simulation data to enable the supervised learning. For super-resolution, prevalent models in published studies adopt an autoencoder architecture to encode low-resolution image feature into a latent space and then either directly generate high-resolution volumes in the decoder (Wang et al., 2021; Biffi et al., 2019) or utilize latent space interpolation to generate missing information between neighboring slices (Chang et al., 2022b). We, however, leverages a 3D version of EDSR which contains an upsampling layer to expand the z-axis dimension first followed by a long series of residual blocks to refine the high-resolution features. EDSR has been shown to outperform the simple ResNet or U-Net (Lim et al., July 10, 2017) in the natural image super-resolution.

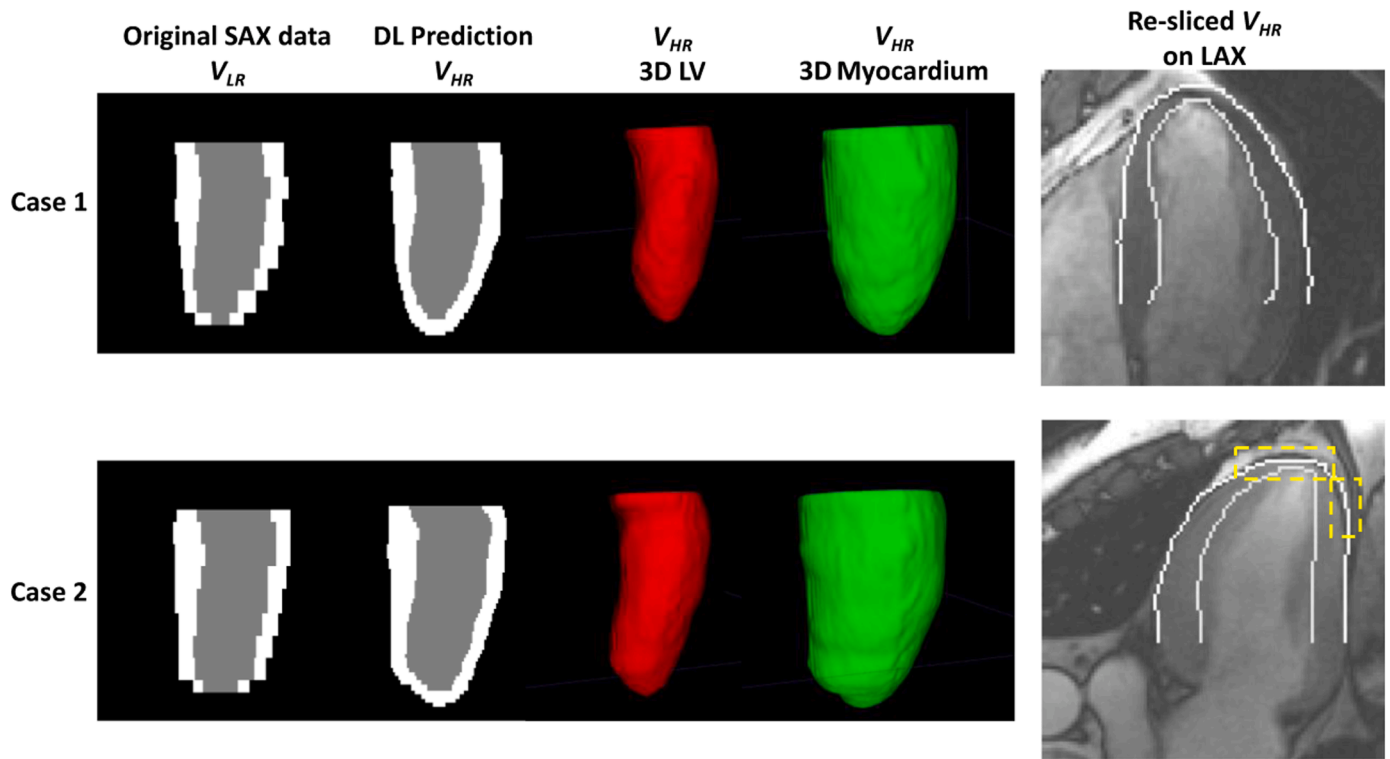
It is imperative to highlight that the innovation of our method lies in its end-to-end architecture that processes both tasks concurrently. This end-to-end design not only streamlines the workflow by employing a singular model, but also enhance final results. Our findings demonstrate that two tasks mutually benefit each other during the end-to-end training, as evident by the superior performance of  $DL_{end2end}$  over both standalone EDSR and  $DL_{2-step}$ .

#### 4.4. Application scenarios, limitations and potential directions

**Application scenarios:** 3D and 4D (3D + time) cardiac shape analysis of LV and myocardium using CMR is crucial to detect abnormal myocardial regions (Sliman et al., 2014; Bao et al., 2019) and provides diagnostic and prognostic values in cardiovascular diseases (Corral et al., 2022; Gilbert et al., 2018a, 2018b; Medrano-Gracia et al., 2014; Zhang et al., 2010). However, conventional CMR suffers from inter-slice misalignment due to motion as well as the missing information due to large slice thickness, which all comprise the fidelity of a 2D SAX slice stack as a representation of the actual 3D cardiac volume (Petersen et al., 2016). Therefore, our model can serve as an important image pre-processing step to address these two issues and reconstruct motion-free and high-resolution 3D cardiac volumes in a research or clinical cardiac analysis workflow.

**Limitation and potential directions:** Our study has several limitations. Methodology-wise, the following points are noteworthy. First, several prior studies have shown that leveraging the intersection between LAX and SAX can enhance both motion correction (Villard et al., 2017; Sinclair et al., 2017; Yang et al., 2017; Chang et al., 2022a) and super-resolution (Biffi et al., 2019). This paper focuses on developing an end-to-end model tailored to simultaneously address motion correction





**Fig. 7.** Examples of predicted high-resolution 3D volumes,  $V_{HR}$  in the real-world dataset. Each row represents a different case from the real-world dataset. The first column displays the original SAX slices  $V_{LR}$ , which is used as the model input. The second column displays the predicted  $V_{HR}$  produced by our model, while the third and fourth columns show the 3D visualization of the endocardium (red) and the epicardium (green) within the  $V_{HR}$ , respectively. The figures on the far right overlay the re-sliced  $V_{HR}$  (along the corresponding LAX view) on the LAX view. The inner grey contour represents the endocardial boundary, while the outer white contour represents the epicardial boundary within the predicted  $V_{HR}$ . It is observed that in the first case the predicted boundaries resemble the true boundaries, whereas in the second case the epicardial boundary in the apex does not align with the true boundary accurately (see yellow boxes, resulting in slightly lower myocardium Dices). Note we define the most basal slice as the slice below the Left Ventricular Outflow Tract (LVOT) plane rather than the slice on the mitral valve plane.

and super-resolution with specific model component designs for each function. We did not explore the utilization of LAX in this context, and the public dataset (Savioli et al., 2021) we used to simulation and model training did not include the LAX view either. In the future, we will investigate the incorporation of LAX into our end-to-end model to further enhance the reconstruction performance. Second, the attention mechanism has demonstrated effectiveness in emphasizing both local and global features to enhance MR super-resolution (Jiang et al., 2021). We therefore intend to incorporate attention mechanism into the future model development. Third, our model is restricted to upscaling the z-resolution by a factor of 5. We aim to extend our model's flexibility to enable multi-scale super-resolution using models with sequential outputs of different scales (Zhu et al., 2019). Fourth, our method operates on the segmentation contours, meaning that the accuracy of the segmentation in the pre-processing step may impact the outcomes of our method. However, conversely, our method that only relies on segmentations rather than images has the potential to be adapted to other imaging modalities.

Evaluation-wise, there are areas that require further exploration. First, our model was evaluated on a relatively small set of real-world CMR data. Its performance in a more extensive real-world CMR dataset such as the UK Biobank (Bai et al., 2015) needs further investigation. Moreover, we observed a degradation in the accuracy of the predicted  $V_{HR}$  in the real-world dataset compared to the simulation dataset, particularly in myocardium (Dice = 0.938 in the simulation dataset versus 0.786 in the real-world dataset). We plan to address this domain shift challenge in our future work. Another potential factor contributing to the lower myocardium Dice may be the inherent intra-reader variability in manual contour segmentation in the SAX slices (used as the

model input) and the LAX views (used as the comparison reference).

## 5. Conclusion

In this study, we propose an end-to-end DL model comprising two specific components for motion correction and super-resolution individually. This model demonstrates high performance in simultaneously addressing these two challenges, resulting in the accurate reconstruction of high-resolution 3D volumes of the heart from acquired CMR SAX slices.

## Funding sources

This work was supported by NHLBI [grant number 1R01HL159183-01A1].

## CRediT authorship contribution statement

**Xiang Li:** Writing – review & editing, Visualization, Validation, Supervision, Methodology, Conceptualization, Investigation. **Quanzheng Li:** Conceptualization, Funding acquisition, Project administration, Resources. **Hui Ren:** Writing – review & editing, Validation, Project administration, Methodology, Investigation. **Zhennong Chen:** Writing – original draft, Visualization, Validation, Methodology, Investigation, Data curation, Conceptualization.

## Declaration of Competing Interest

The authors declare that they have no known competing financial

interests or personal relationships that could have appeared to influence the work reported in this paper.

## Data Availability

Only public datasets are used in this manuscript.

## Acknowledgement

None.

## References

- Albà, X., Pereañez, M., Hoogendoorn, C., et al., 2016. An algorithm for the segmentation of highly abnormal hearts using a generic statistical shape model. *IEEE Trans. Med. Imaging* 35 (3), 845–859. <https://doi.org/10.1109/TMI.2015.2497906>.
- Bai, W., Shi, W., de Marvao, A., et al., 2015. A bi-ventricular cardiac atlas built from 1000+ high resolution MR images of healthy subjects and an analysis of shape and motion. *Med Image Anal.* 26 (1), 133–145. <https://doi.org/10.1016/j.media.2015.08.009>.
- Banerjee, A., Camps, J., Zacur, E., et al., 2021. A completely automated pipeline for 3D reconstruction of human heart from 2D cine magnetic resonance slices. *Philos. Trans. A Math. Phys. Eng. Sci.* 379 (2212), 20200257 <https://doi.org/10.1098/rsta.2020.0257>.
- Bao, H., Ren, H., Zhou, Z., Li, X., Guo, N., Li, Q., 2019. 3D Regional shape analysis of left ventricle using mr images: abnormal myocardium detection and classification. 2019 IEEE 16th Int. Symp. . Biomed. Imaging (ISBI 2019) 789–792. <https://doi.org/10.1109/ISBI.2019.8759451>.
- Beetz, M., Banerjee, A., Grau, V., 2021. Biventricular surface reconstruction from cine mri contours using point completion networks. 2021 IEEE 18th Int. Symp. . Biomed. Imaging (ISBI) 105–109. <https://doi.org/10.1109/ISBI48211.2021.9434040>.
- Zhu, J., Tan, C., Yang, J., Yang, G., Lio, P., 2021. Arbitrary scale super-resolution for medical images. *International Journal of Neural Systems* 31 (10), 2150037.
- Beetz M., Banerjee A., Grau V. Reconstructing 3D Cardiac Anatomies from Misaligned Multi-View Magnetic Resonance Images with Mesh Deformation U-Nets. In: Proceedings of the First International Workshop on Geometric Deep Learning in Medical Image Analysis. PMLR; 2022:3-14. Accessed February 12, 2023. <https://proceedings.mlr.press/v194/beetz22a.html>.
- Biffi, C., Cerrolaza, J.J., Tarroni, G., et al., 2019. 3D High-resolution cardiac segmentation reconstruction from 2d views using conditional variational autoencoders. 2019 IEEE 16th Int. Symp. . Biomed. Imaging (ISBI 2019) 1643–1646. <https://doi.org/10.1109/ISBI.2019.8759328>.
- Chang, Q., Yan, Z., Zhou, M., et al., 2022b. DeepRecon: Joint 2D Cardiac Segmentation and 3D Volume Reconstruction via a Structure-Specific Generative Method. In: Wang, L., Dou, Q., Fletcher, P.T., Speidel, S., Li, S. (Eds.), *Medical Image Computing and Computer Assisted Intervention – MICCAI 2022*. Lecture Notes in Computer Science. Springer Nature Switzerland, pp. 567–577. [https://doi.org/10.1007/978-3-031-16440-8\\_54](https://doi.org/10.1007/978-3-031-16440-8_54).
- Chang Q., Yan Z., Ye M., et al. An Unsupervised 3D Recurrent Neural Network for Slice Misalignment Correction in Cardiac MR Imaging. In: Puyol Antón E, Pop M, Martín-Isla C, et al., eds. Statistical Atlases and Computational Models of the Heart. Multi-Disease, Multi-View, and Multi-Center Right Ventricular Segmentation in Cardiac MRI Challenge. Lecture Notes in Computer Science. Springer International Publishing; 2022a:141-150. doi:10.1007/978-3-030-93722-5\_16.
- Corral, A.J., Schuster, A., Zacur, E., et al., 2022. Understanding and improving risk assessment after myocardial infarction using automated left ventricular shape analysis. *JACC: Cardiovasc. Imaging* 15 (9), 1563–1574. <https://doi.org/10.1016/j.jcmg.2021.11.027>.
- Dangi S., Linte C.A., Yaniv Z. Cine Cardiac MRI Slice Misalignment Correction Towards Full 3D Left Ventricle Segmentation. *Proc SPIE Int Soc Opt Eng.* 2018;10576: 1057607. doi:10.1117/12.2294936.
- Eitel, I., Stiermaier, T., Lange, T., et al., 2018. Cardiac magnetic resonance myocardial feature tracking for optimized prediction of cardiovascular events following myocardial infarction. *JACC: Cardiovasc. Imaging* 11 (10), 1433–1444. <https://doi.org/10.1016/j.jcmg.2017.11.034>.
- Garg R., Friedrich M.G. Cardiomyopathies. In: Camm AJ, Lüscher TF, Maurer G, Serruys PW, Pennell D, eds. The ESC Textbook of Cardiovascular Medicine. Oxford University Press; 2018:0. doi:10.1093/med/9780198784906.003.0106\_update\_001.
- Gilbert, K., Forsch, N., Hegde, S., et al., 2018a. Atlas-Based Computational Analysis of Heart Shape and Function in Congenital Heart Disease. *J. Cardiovasc. Transl. Res* 11 (2), 123–132. <https://doi.org/10.1007/s12265-017-9778-5>.
- Gilbert, K., Pontre, B., Occlshaw, C.J., Cowan, B.R., Suinesiaputra, A., Young, A.A., 2018b. 4D modelling for rapid assessment of biventricular function in congenital heart disease. *Int J. Cardiovasc. Imaging* 34 (3), 407–417. <https://doi.org/10.1007/s10554-017-1236-6>.
- Heidenreich, P.A., Bozkurt, B., Aguilar, D., et al., 2022. 2022 AHA/ACC/HFSA Guideline for the Management of Heart Failure. *J. Am. Coll. Cardiol.* 79 (17), e263–e421. <https://doi.org/10.1016/j.jacc.2021.12.012>.
- Jiang, M., Zhi, M., Wei, L., et al., 2021. FA-GAN: Fused attentive generative adversarial networks for MRI image super-resolution. *Comput. Med. Imaging Graph.* 92, 101969 <https://doi.org/10.1016/j.compmedimag.2021.101969>.
- Joyce, T., Buoso, S., Stoeck, C.T., Kozerke, S., 2022. Rapid inference of personalised left-ventricular meshes by deformation-based differentiable mesh voxelization. *Med. Image Anal.* 79, 102445 <https://doi.org/10.1016/j.media.2022.102445>.
- Lim, B., Son, S., Kim, H., Nah, S., Lee, K.M., July 10, 2017. Enhanced deep residual networks for single image super-resolution. *Publ. Online.* <https://doi.org/10.48550/arXiv.1707.02921>.
- de Marvao, A., Dawes, T.J., Shi, W., et al., 2014. Population-based studies of myocardial hypertrophy: high resolution cardiovascular magnetic resonance atlases improve statistical power. *J. Cardiovasc. Magn. Reson.* 16 (1), 16 <https://doi.org/10.1186/1532-429X-16-16>.
- Medrano-Gracia, P., Cowan, B.R., Ambale-Venkatesh, B., et al., 2014. Left ventricular shape variation in asymptomatic populations: the multi-ethnic study of atherosclerosis. *J. Cardiovasc. Magn. Reson.* 16 (1), 56 <https://doi.org/10.1186/s12968-014-0056-2>.
- Otto, C.M., Nishimura, R.A., Bonow, R.O., et al., 2021. 2020 ACC/AHA Guideline for the Management of Patients With Valvular Heart Disease. *J. Am. Coll. Cardiol.* 77 (4), e25–e197. <https://doi.org/10.1016/j.jacc.2020.11.018>.
- Petersen, S.E., Matthews, P.M., Francis, J.M., et al., 2016. UK Biobank's cardiovascular magnetic resonance protocol. *J. Cardiovasc. Magn. Reson.* 18 (1), 8 <https://doi.org/10.1186/s12968-016-0227-4>.
- Radau, P., Lu, Y., Connelly, K., Paul, G., Dick, A.J., Wright, G.A., 2009. Evaluation framework for algorithms segmenting short axis cardiac MRI. *Published online July 9, MIDAS J.* <https://doi.org/10.54294/g80ruo>.
- Reindl, M., Tiller, C., Holzknecht, M., et al., 2019. Prognostic implications of global longitudinal strain by feature-tracking cardiac magnetic resonance in st-elevation myocardial infarction. *Circ.: Cardiovasc. Imaging* 12 (11), e009404. <https://doi.org/10.1161/CIRCIMAGING.119.009404>.
- Savioli N., de Marvao A., O'Regan D. Cardiac super-resolution label maps. 2021;1. doi:10.17632/pw87p286yx.1.
- Sinclair M., Bai W., Puyol-Antón E., Oktay O., Rueckert D., King A.P. Fully Automated Segmentation-Based Respiratory Motion Correction of Multiplanar Cardiac Magnetic Resonance Images for Large-Scale Datasets. In: Descoteaux M, Maier-Hein L, Franz A, Jannin P, Collins DL, Duchesne S, eds. Medical Image Computing and Computer-Assisted Intervention – MICCAI 2017. Lecture Notes in Computer Science. Springer International Publishing; 2017:332-340. doi:10.1007/978-3-319-66185-8\_38.
- Sliman, H., Elnakib, A., Beache, G.M., et al., 2014. A novel 4D PDE-based approach for accurate assessment of myocardium function using cine cardiac magnetic resonance images. 2014 IEEE Int. Conf. Image Process. (ICIP) 3537–3541. <https://doi.org/10.1109/ICIP.2014.7025718>.
- Tarroni, G., Bai, W., Oktay, O., et al., 2020. Large-scale quality control of cardiac imaging in population studies: application to UK biobank. *Sci. Rep.* 10 (1), 2408. <https://doi.org/10.1038/s41598-020-58212-2>.
- Villard B., Zacur E., Dall'Armellina E., Grau V. Correction of Slice Misalignment in Multi-breath-hold Cardiac MRI Scans. In: Mansi T, McLeod K, Pop M, Rhode K, Sermesant M, Young A, eds. Statistical Atlases and Computational Models of the Heart. Imaging and Modelling Challenges. Lecture Notes in Computer Science. Springer International Publishing; 2017:30-38. doi:10.1007/978-3-319-52718-5\_4.
- Villard, B., Grau, V., Zacur, E., 2018. Surface Mesh Reconstruction from cardiac MRI Contours. *J. Imaging* 4 (1), 16. <https://doi.org/10.3390/jimaging4010016>.
- von Knobelsdorff-Brenkenhoff, F., Schulz-Menger, J., 2016. Role of cardiovascular magnetic resonance in the guidelines of the European Society of Cardiology. *J. Cardiovasc. Magn. Reson.* 18, 6. <https://doi.org/10.1186/s12968-016-0225-6>.
- von Knobelsdorff-Brenkenhoff, F., Pilz, G., Schulz-Menger, J., 2017. Representation of cardiovascular magnetic resonance in the AHA / ACC guidelines. *J. Cardiovasc. Magn. Reson.* 19 (1), 70 <https://doi.org/10.1186/s12968-017-0385-z>.
- Wang, S., Qin, C., Savioli, N., et al., July 8, 2021. Joint motion correction and super resolution for cardiac segmentation via latent optimisation. *Publ. Online.* <https://doi.org/10.48550/arXiv.2107.03887>.
- Xu H., Zacur E., Schneider J.E., Grau V. Ventricle Surface Reconstruction from Cardiac MR Slices Using Deep Learning. In: Coudière Y, Ozenne V, Vigmond E, Zemzemi N, eds. Functional Imaging and Modeling of the Heart. Lecture Notes in Computer Science. Springer International Publishing; 2019:342-351. doi:10.1007/978-3-030-21949-3\_37.
- Xu, J., Yang, W., Zhao, S., Lu, M., 2022. State-of-the-art myocardial strain by CMR feature tracking: clinical applications and future perspectives. *Eur. Radio.* 32 (8), 5424–5435. <https://doi.org/10.1007/s00330-022-08629-2>.
- Yang, D., Wu, P., Tan, C., Pohl, K.M., Axel, L., Metaxas, D., 2017. 3D Motion modeling and reconstruction of left ventricle wall in cardiac MRI. *Funct. Imaging Model Heart* 10263, 481–492. [https://doi.org/10.1007/978-3-319-59448-4\\_46](https://doi.org/10.1007/978-3-319-59448-4_46).
- Zeppenfeld, K., Tfelt-Hansen, J., de Riva, M., et al., 2022. 2022 ESC Guidelines for the management of patients with ventricular arrhythmias and the prevention of sudden cardiac death: developed by the task force for the management of patients with ventricular arrhythmias and the prevention of sudden cardiac death of the European Society of Cardiology (ESC) Endorsed by the Association for European Paediatric and Congenital Cardiology (AEPCC). *Eur. Heart J.* 43 (40), 3997–4126. <https://doi.org/10.1093/eurheartj/ehac262>.
- Zhang, H., Wahle, A., Johnson, R.K., Scholz, T.D., Sonka, M., 2010. 4-D cardiac MR image analysis: left and right ventricular morphology and function. *IEEE Trans. Med. Imaging* 29 (2), 350–364. <https://doi.org/10.1109/TMI.2009.2030799>.
- Zhu J., Yang G., Lio P. How Can We Make Gan Perform Better in Single Medical Image Super-Resolution? A Lesion Focused Multi-Scale Approach. 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019). Published online April 2019:1669-1673. doi:10.1109/ISBI.2019.8759517.