



Missing Slice Imputation in Population CMR Imaging via Conditional Generative Adversarial Nets

Le Zhang^{1,2(✉)}, Marco Pereañez³, Christopher Bowles³, Stefan Piechnik⁴, Stefan Neubauer⁴, Steffen Petersen⁵, and Alejandro Frangi³

- ¹ Centre for Computational Imaging and Simulation Technologies in Biomedicine, Department of Electronic and Electrical Engineering, University of Sheffield, Sheffield, UK
le.zhang@sheffield.ac.uk
- ² Queen Square Institute of Neurology, University College London, London, UK
- ³ Centre for Computational Imaging and Simulation Technologies in Biomedicine, School of Computing and School of Medicine, University of Leeds, Leeds, UK
- ⁴ Oxford Center for Clinical Magnetic Resonance Research (OCMR), Division of Cardiovascular Medicine, John Radcliffe Hospital, University of Oxford, Oxford, UK
- ⁵ William Harvey Research Institute, Barts Heart Centre, Barts Health NHS Trust, Queen Mary University of London, London, UK

Abstract. Accurate ventricular volume measurements depend on complete heart coverage in cardiac magnetic resonance (CMR) from where most immediate indicators of normal/abnormal cardiac function are available non-invasively. However, incomplete coverage, especially missing basal or apical slices in CMR sequences is insufficiently addressed in population imaging and current clinical research studies yet has important impact on volume calculation accuracy. In this work, we propose a new deep architecture, coined Missing Slice Imputation Generative Adversarial Network (MSIGAN), to learn key features of cardiac short-axis (SAX) slices across different positions, and use them as conditional variables to effectively infer missing slices in the query volumes. In MSIGAN, the slices are first mapped to latent vectors with position features through a regression net. The latent vector corresponding to the desired position is then projected onto the slice manifold conditional on slice intensity through a generator net. The latent vector along with the slice features (*i.e.*, intensity) and desired position control the generation vs. regression. Two adversarial networks are imposed on the regressor and generator, encouraging more realistic slices. Experimental results show that our method outperforms the previous state-of-the-art in missing slice imputation for cardiac MRI.

Keywords: Deep learning · Data imputation · Generative adversarial net · Ventricular volume · MRI · Population imaging

1 Introduction

Cardiac Magnetic Resonance Imaging (CMRI) can not only provide anatomical information from the heart, but also yields physiological information associated with cardiovascular diseases. Ejection fraction (EF) and cardiac output (CO) of both ventricles directly relate to the volume of these chambers, whose extents in turn defined between their basal and apical slices, are the most commonly used clinical diagnostic parameters for cardiac function. Most published studies addressing computation of EF and CO assume a complete data set in the sense all imaging planes are available for all samples. In practice, in population imaging, clinical trials, or clinical routine, this assumption may not hold as certain datasets may have missing or corrupted information due to imaging artifacts and acquisition/storage errors [10]. Additionally, many discriminative classifiers, such as SVMs, require training and testing data where a full set of features is available for every sample.

A common strategy to account for unavailable features is the removal of incomplete samples from the study cohort [9]. However, excluding data reduces not only statistical power and cause bias, but is also of ethical and financial concern as partially acquired subject data remains unused, and limits the application of such methods to similarly complete datasets. Some data imputation based methods have been proposed to deal with this problem, such as using the mean of the data or model-based missing data estimation [3]. If the missing mechanism is random, the missing variable can be imputed by the marginal distribution of the observed data using maximum likelihood estimation (MLE) [2]. The stochastic regression imputation method can better use the information provided by the data to solve the collinearity problem caused by the high correlation of predicted variables [7]. When the data missingness is non-random, the missing variable cannot be predicted from the available variables in the database alone, and there is no general method of handling missing data properly [3]. The performance of imputation approaches is ideally assessed by both the feature error and the classification accuracy on the imputed features.

In this paper, we adapt the developed generative adversarial net (GAN) to generate missing CMRI slices after applying quality control (QC). We propose a missing slice imputation based generative adversarial network (MSIGAN) model to infer missing slice features from multi-position input images. After inference, the features of the desired position and slice intensity are concatenated to generate real images in a certain position with the correct appearance. The main contributions of the MSIGAN are highlighted:

- (1) A novel deep MSIGAN architecture is proposed for generating missing SAX slices for CMRI across different positions. First, a regression net learns the intrinsic features of the input slices. Conditioned on these features, and a pre-computed feature of the expected position, a generator and discriminator aim to generate the realistic image in the expected position.
- (2) Given the slice features and expected position, we design a conditional generative network to infer an image matching the missing slice of the input

cardiac volume. The adversarial training mechanism and auxiliary slice position regressor is combined to achieve effective feature generation.

- (3) This is the first paper to exploit deep learning methods, especially GANs, for missing slice imputation in cardiac MRI, which is an important step after QC and before quantitative medical image analysis. It can be learned once and then applied to synthesize missing slices in cases of incomplete heart coverage with no further training.

2 Methodology

2.1 Problem Formulation

The overall target of missing slice imputation (MSI) for cardiac MRI is similar to the missing data imputation problem in data mining [4]. Given a query cardiac MR volume, a regression list of slice positions in the data set is desired, followed by image synthesis to impute the query volume's missing slices. For each input 3D cardiac image \mathbf{X} , we aim to map it to a feature representation \mathbf{f} and synthesize the missing slice $\hat{\mathbf{x}}$ using the following function:

$$\hat{\mathbf{x}} = \Gamma(\text{series}(\{\mathbf{f}_n\}_{n=1}^N)) = \Gamma(\text{series}(\mathcal{R}(\mathbf{X}) \cdot \{\mathcal{Y}_n\}_{n=1}^N)). \quad (1)$$

where the operator $\mathcal{R}(\cdot)$ is to extract the features (*i.e.*, intensity) of the input image \mathbf{X} . $\{\mathcal{Y}_n\}_{n=1}^N$ is obtained by the regression model to identify the slice position information, like the distance to basal/or apical slice, from the cardiac volume. N is the number of slices. The operator $\Gamma(\cdot)$ denotes the transformation from the concatenated features to the inferred slice features in the cardiac volume. Finally, we can synthesize the missing slice from a certain position. Therefore, the most significant factor to achieve effective synthesis is how to design and optimize $\mathcal{R}(\cdot)$, \mathcal{Y} and $\Gamma(\cdot)$.

We formulate the image synthesis component for the MSI problem with these three steps: first, given a 3D cardiac volume $\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$ and the corresponding slice position label $\mathbf{y} = \{y_1, y_2, \dots, y_n\}$, a regression net $\mathcal{R}(\cdot)$ aims to learn the cardiac intensity feature \mathbf{f}_{int} and the slice position maps \mathcal{Y} . Then, using the feature maps for different slice positions and intensities as conditions, we aim to generate the desired slice features via $\Gamma(\cdot)$ with an adversarial training architecture. A generative net takes the intrinsic slice features (*i.e.*, intensity), a random vector and the position feature of the desired slice as input to synthesize a cardiac cine MRI in the corresponding position within the same input volume. Finally, a discriminative net distinguishes the generated samples from the real images, and simultaneously tries to match the inferred slices with correct features and positions. The network architecture is illustrated in Fig. 1. The synthesized slices can be directly adopted for imputing the missing slice in the target CMR volume.

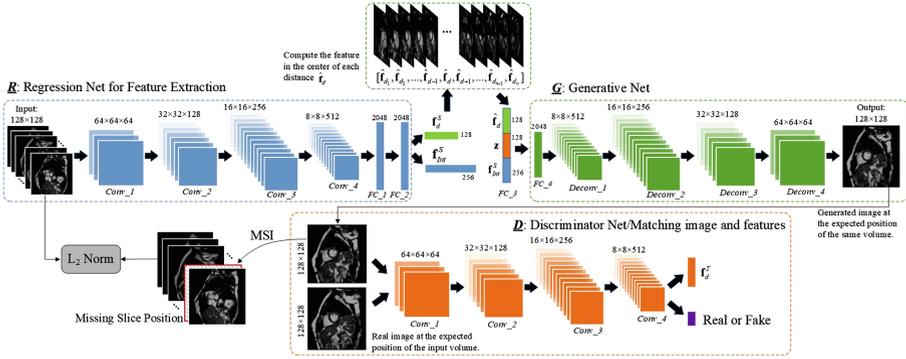


Fig. 1. Structure of the proposed MSIGAN network for cardiac missing slice imputation. The regressor R maps each slice of the input volume to a vector containing intensity and position features. The central point feature of each position cluster over the whole training set can be obtained and used for generator G . Concatenating the intensity feature and random noise to the inferred position cluster centre feature, the new latent vector FC_3 is fed to G . Both R and G are updated based on the L_2 loss between the original and synthetic volumes. The discriminative net D forces the output slice to be realistic and plausible for a given position label.

2.2 MSI Conditional GAN (MSIGAN)

Cardiac Feature Learning and Slice Position Estimation: To generate CMR slices in SAX view, a deep regression network R aims to learn CMR image features including slice position \mathbf{f}_d and intensity \mathbf{f}_{Int} . Formally, \mathbf{X}_S denotes the input cardiac image stack with full ventricular coverage. The trunk architecture of the regression net consists of 4 convolutional layers (kernel size = 5, padding = 2 and stride = 2) and 2 fully-connected layers. The Leaky-ReLU is set after each layer. We then configure two layers for learning the 256-dimensional \mathbf{f}_{Int} with the intrinsic intensity features and 128-dimensional \mathbf{f}_d with the inferred slice position regression separately, since we expect slice position information weakened in \mathbf{f}_{Int} , but strengthened in \mathbf{f}_d . While training the regression net, the loss function can be fast and well converged. Thus, we can easily learn each position’s feature cluster from all the training data by k-means clustering, and compute the feature in the centre of each cluster, \mathbf{f}_{dc} , as a condition to generate slices in the missed position.

Conditional Cardiac GAN: Instead of generating real images using regular GANs, our model aims to transform the features from CMR volumes with full ventricular coverage into the query CMR volumes, which miss slices in certain positions, using a conditional generative model. The conditional generator is defined as $G: \mathbb{R}^F \times \mathbb{R}^Z \times \mathbb{R}^T \rightarrow \mathbb{R}^S$, where F is the dimension of the intrinsic cardiac intensity, Z is for random noise, T is the dimension of the inferred slice position and S is the cardiac slice. Besides, the discriminator is denoted as $D: \mathbb{R}^S \rightarrow \{0, 1\} \times \prod l_i$, where $i = \{1 : \mathbf{f}_{Int}, 2 : \mathbf{f}_{dc}\}$. l_i denotes the range of each

label. The optimization of the G and D can be reformulated as:

$$\mathcal{L}_D = E_{\mathbf{x} \sim p_{data}(\mathbf{x})} [\log D(\mathbf{x})] - \sum_{i=1}^2 \|l_i - D(\mathbf{x})\|_2^2 \quad (2)$$

$$\mathcal{L}_G = E_{A;B;C} [\log(1 - D(G(\mathbf{f}_{dc}^T, \mathbf{z}, \mathbf{f}_{Int}^S)))] \quad (3)$$

where $A \rightarrow \mathbf{f}_{dc}^T \sim p_{data}(\mathbf{f}_{dc}^T)$, $B \rightarrow \mathbf{z} \sim p_{\mathbf{z}}(\mathbf{z})$, and $C \rightarrow \{\mathbf{f}_{Int}^S\} \sim p_{data}(\mathbf{f}_{Int}^S)$. The input of generator G is the concatenation of \mathbf{f}_{Int}^S , \mathbf{f}_{dc}^T and a random noise prior $\mathbf{z} \sim \mathcal{N}(0, 1)$. \mathbf{f}_{Int}^S and can be regarded as intrinsic intensity features from the *original slices*, while \mathbf{f}_{dc}^T is the *target* position feature in the centre of the cluster. A fully-connected layer is set for better fusing the three vectors and then four deconvolutional layers are adopted for generating synthesized slice samples. The architecture of the generative net is set as the reverse of the regression net R .

The discriminator D takes the generated samples and the real images in the target CMR volume as inputs. The main structure of the Discriminator has a similar structure to the regression net. To match the inferred slices with the same intensity features and correct slice position in the query volumes, we add a fully-connected layer and simultaneously optimize the whole discriminative net by slice position regression. The position label for the synthetic slice is same as the expected position label in the query volume. Batch normalization and ReLU are also adopted for all the layers in the discriminator. Meanwhile, to ensure the output slice shares the intensity of the input image (during training), the input image and output image are expected to be similar as expressed in Eq. (4), where $L(\cdot)$ denotes L_2 norm.

$$\mathcal{L}_{L2N} = L(\mathbf{x}, G(R(\mathbf{x}))) \quad (4)$$

2.3 Optimisation

The training scheme for MSIGAN consists of three steps. In the first step, R , for slice feature learning, is trained using the deep regression net. The computed different slice position features are then obtained. Next, G is fed by the learned real position features from different cardiac volumes, which is fused with the intensity features \mathbf{f}_{Int}^S and the random noise \mathbf{z} . Four deconvolutional layers are adopted for G to generate synthesized slice samples. Both the regressor and the generator are updated based on the L_2 loss between the input and output volumes to ensure they are similar. In the following step, the discriminative net D employs a general fully convolutional network to distinguish the real images from the generated ones. Rather than maximizing the output of the discriminator for generated data, the objective of feature matching [6] is employed to optimize G to match the statistics of features in an intermediate layer of D . The objective

function is defined in the following equation:

$$\begin{aligned} \mathcal{L}_{MSI} = \min_G \max_{D,R} E(\log(1 - D(G(\{\mathbf{f}_{Int}^S, \mathbf{z}, \mathbf{f}_{dc}^T\})))) + \sum_{i=1}^2 \|\mathbf{l}_i - D(\mathbf{x})\|_2^2 \\ + \|E(D_k(\{\mathbf{f}_{Int}^S, \mathbf{f}_d^S\})) - E(D_k(G(\{\mathbf{f}_{Int}^S, \mathbf{z}, \mathbf{f}_{dc}^T\})))\|_2^2 - L(\mathbf{x}, G(R(\mathbf{x}))) \end{aligned} \quad (5)$$

where k means the k^{th} layer in D ($k = 4$ in our setting). Moreover, D is trained with slice position regression to better match generated position features with the identity of the input volume. We apply one more *conv* layer to output the final position features. For all the *conv* layers in G and D . The conditioned G and D nets can be optimized by \mathcal{L}_{MSI} to infer the missing features from query input volumes.

3 Experiments and Analysis

Materials and Position Label Generation. Quality-scored CMR data is available for circa 5,000 volunteers of the UK Biobank imaging (UKBB) resource. Following visual inspection, manual annotation for SAX images was carried out with a simple 3-grade quality score [1]. 4,280 sequences correspond to quality score 1 for both ventricles, which indicates full coverage of the heart from base to apex, and form the source datasets to construct the ground-truth distance label for our experiments. Note that having full coverage should not be confused with having the top/bottom slices corresponding exactly to base/apex [12].

The slice position labels are generated from the distances to apex point and base point. To obtain the apex point, we take last 2 apical 2D manual delineations and fit a spline curve to extrapolate the location of the apex, before measuring the distance to this point for all image slices. To obtain the base point, we use left atrium (LA) manual delineations on the 4CH long-axis image view to define the centre of the mitral valve (MV) as the base, and measure the distance from this point to all image slices. To make each slice label represent the distance to apex and base simultaneously, we normalize the distance from base to apex 1 for all cases (base as 0 and apex as 1), and label the middle slices with values equally increased from 0 to 1. For the slices above the base or below the apex, we also use the equal interval to label them.

Experimental Settings. We performed two groups of experiments in this work. In the first experiment, we aim to evaluate the quality of the images generated by MSIGAN. The averaged peak signal-to-noise ratio (PSNR) and structural similarity index (SSIM) are used to measure the image quality of the ground-truth and synthetic MR images. In the second group, we evaluate the imputed cardiac volumes with corresponding ground-truth for the tasks of LV segmentation and cardiac function measurement based on blood volumes. Four parameters are used for performance evaluation, including two commonly used indexes of the cardiac function derived from such volumes viz. stroke volume (SV) and ejection

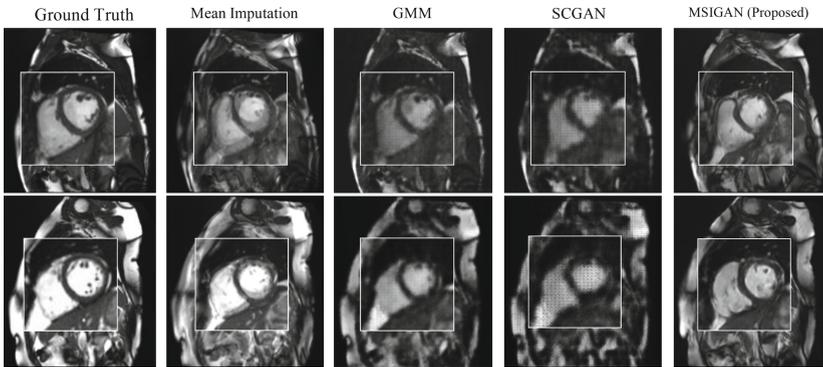


Fig. 2. Example of synthesized images (*left*) generated by Mean, GMM, SCGAN and MSIGAN, compared to the ground truths (*right*) in each pairs.

Table 1. Quantitative results for missing cardiac MRI synthesis based on PSNR and SSIM. Higher values indicate better performance. Values in brackets represent the standard deviation. Highest performing results seen in bold.

	Mean	GMM [5]	SCGAN [11]	MSIGAN (Proposed)
PSNR	20.49 ± 5.21	22.17 ± 3.75	17.49 ± 3.46	24.49 ± 3.69
SSIM	0.547 ± 0.21	0.686 ± 0.24	0.512 ± 1.71	0.703 ± 0.11

fraction (EF), and similarly report the differences between the real and imputed coverages.

Performance of Image Synthesis Model. To evaluate the quality of the images generated by MSIGAN, we first train the MSIGAN model using 3,280 complete subjects from the 4,280 cases with quality score 1 in UKBB, and test our model on the remaining 1,000 subjects. We take the 1,000 testing subjects for which the ground-truth slices are available and randomly remove slices to generate incomplete volumes, before using our MSIGAN to synthesize the missed slices. We also impute the missing slice using mean, mixture of factor analyzers based GMM model [5] and the SCGAN model [11]. Several typical images with real and synthetic slices are shown in Fig. 2. From Fig. 2, we can observe that the synthetic slices produced by MSIGAN show the best image quality against the Mean, GMM and SCGAN methods. The mean and standard deviation of PSNR and SSIM values of synthetic slice are listed in Table 1. MSIGAN significantly outperforms the other three methods based on PSNR and SSIM for missing slice synthesis. These results imply that our trained MSIGAN model is reasonable, and the synthetic cardiac MRI scans have acceptable image quality.

Results of Cardiac Functional Parameters Calculation. To assess the impact of synthetic images in real applications, such as measurement of cardiac function based on blood volumes, we design an experiment where incom-

Table 2. Effect of incomplete cardiac coverage (MBS) on the left ventricle end-diastolic volume (LVEDV), end-systolic volume (LDES), stroke volumes (LVS) and ejection fraction (LVEF). Values are shown as Mean \pm standard deviations.

	Ground truth	Missing Basal Slice (MBS)	Effect (%)	Synthetic image	Effect (%)
LVEDV (ml)	155.8 \pm 35.6	136.1 \pm 33.4	-12.6%	151.7 \pm 33.7	-2.6%
LDES (ml)	66.8 \pm 21.2	53.0 \pm 19.0	-20.7%	61.3 \pm 22.3	-8.2%
LVS (ml)	89.1 \pm 19.8	83.1 \pm 19.7	-6.7%	90.4 \pm 18.7	+1.5%
LVEF (%)	57.1 \pm 0.06	61.0 \pm 0.06	+6.8%	59.6 \pm 0.06	+4.4%

plete coverage is simulated and volume differences between ground-truth, synthetic volumes and incomplete volumes are measured. The experimental results achieved by four cardiac parameters using the LV segmentation method in [8] are reported in Table 2. For this experiment, we compute blood pool volumes at the ED and ES phases, and from these, we obtain $SV = EDV - ESV$ and $EF = SV/EDV$. Then, the average volumes and indexes are computed across the sample, comparing the ground-truth, synthetic volumes and incomplete volumes. Table 2 shows that MBS reduces ED and ES volumes by an average of 12% and 20%, respectively. The synthetic values are much closer than the GT values, with 2.6% and 8.2% reduction in volumes at ED and ES phases. These results demonstrate synthetic images generated by MSIGAN are useful in population imaging applications.

4 Conclusion

In this paper, we proposed a novel deep MSIGAN to implement missing slice generation in cardiac cine MRI and contribute to the missing data imputation problem neglected by the medical imaging community. The MSI adopts a slice position regression model and an adversarial training architecture to impute missing slices based on their corresponding distances to the base and apex, considering the relationship between neighbouring slices scanned for the same subject. Extensive experimental results showed that our model could achieve satisfactory performance on missing slice generation compared to some baseline methods. Our method is also reasonable for practical applications. Only the complete images are used for learning the segmentation models. Using these synthetic slice data could further augment the training samples for improvement of the segmentation models, which will be our future work.

References

1. Carapella, V., et al.: Towards the semantic enrichment of free-text annotation of image quality assessment for UK biobank cardiac cine MRI scans. In: Carneiro, G., et al. (eds.) LABELS/DLMIA -2016. LNCS, vol. 10008, pp. 238-248. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-46976-8_25

2. Dong, Y., Peng, C.Y.J.: Principled missing data methods for researchers. Springer-Plus **2**(1), 222 (2013)
3. García-Laencina, P.J., Sancho-Gómez, J.L., Figueiras-Vidal, A.R.: Pattern classification with missing data: a review. Neural Comput. Appl. **19**(2), 263–282 (2010)
4. Myrtveit, I., Stensrud, E., Olsson, U.H.: Analyzing data sets with missing data: an empirical evaluation of imputation methods and likelihood-based methods. IEEE Trans. Softw. Eng. **27**(11), 999–1013 (2001)
5. Richardson, E., Weiss, Y.: On GANs and GMMs. In: Advances in Neural Information Processing Systems, pp. 5852–5863 (2018)
6. Salimans, T., Goodfellow, I., Zaremba, W., Cheung, V., Radford, A., Chen, X.: Improved techniques for training GANs. In: Advances in Neural Information Processing Systems, pp. 2234–2242 (2016)
7. Schlomer, G.L., Bauman, S., Card, N.A.: Best practices for missing data management in counseling psychology. J. Couns. Psychol. **57**(1), 1 (2010)
8. Tran, P.V.: A fully convolutional neural network for cardiac segmentation in short-axis MRI. arXiv preprint [arXiv:1604.00494](https://arxiv.org/abs/1604.00494) (2016)
9. Williams, D., Liao, X., Xue, Y., Carin, L., Krishnapuram, B.: On classification with incomplete data. IEEE Trans. Pattern Anal. Mach. Intell. **29**(3), 427–436 (2007)
10. Zhang, L., et al.: Automatic assessment of full left ventricular coverage in cardiac cine magnetic resonance imaging with fisher-discriminative 3-D CNN. IEEE Trans. Biomed. Eng. **66**(7), 1975–1986 (2019)
11. Zhang, L., Gooya, A., Frangi, A.F.: Semi-supervised assessment of incomplete LV Coverage in cardiac MRI using generative adversarial nets. In: Tsafaris, S.A., Gooya, A., Frangi, A.F., Prince, J.L. (eds.) SASHIMI 2017. LNCS, vol. 10557, pp. 61–68. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-68127-6_7
12. Zhang, L., Pereañez, M., Piechnik, S.K., Neubauer, S., Petersen, S.E., Frangi, A.F.: Multi-input and dataset-invariant adversarial learning (MDAL) for left and right-ventricular coverage estimation in cardiac MRI. In: Frangi, A.F., Schnabel, J.A., Davatzikos, C., Alberola-López, C., Fichtinger, G. (eds.) MICCAI 2018. LNCS, vol. 11071, pp. 481–489. Springer, Cham (2018). https://doi.org/10.1007/978-3-030-00934-2_54