Synthetic Balancing of Cardiac MRI Datasets

Carles Garcia-Cabrera[1,2], Eric Arazo Sánchez[3], Enric Moreu[4], Kathleen M. Curran[2], Noel E. O'Connor[1,4], and Kevin McGuinness[1,4]

[1] ML-Labs, Dublin City University (DCU), Dublin, Ireland
[2] University College Dublin (UCD), Dublin, Ireland
[3] Centre for Applied AI (CeADAR), University College Dublin (UCD), Ireland
[4] Insight Centre for Data Analytics, Dublin City University (DCU), Dublin, Ireland
carles.garcia-cabrera@ucd.ie

Abstract. Possessing a balanced dataset to train the segmentation models used in clinical products is hugely important to enhance the performance in any given subset of patients. Due to demographics and uneven distribution of conditions of the subjects, this is not usually the case. In this work, we propose a novel method that synthetically balances a training dataset by applying deformations to an atlas. Once the atlas is processed, we obtain different slice cuts from it and apply style transfer to make it appear as a real short-axis MRI scan. We then add those synthetic scans to our training set for the segmentation network. We found that using synthetic scans to balance the dataset resulted in up to a 0.05 increase in the DICE score.

Keywords: Segmentation · Medical Imaging · Cardiac MRI.

1 Introduction

Cardiac MRI has become an increasingly popular imaging modality in clinical practice due to its ability to provide high-resolution images of the heart. The application of machine learning algorithms to these images has the potential to aid in diagnosis and treatment planning. However, the success of such algorithms depends on the availability of high-quality training data, which can be challenging to obtain due to the limited availability of real-world datasets [1].

Segmenting cardiac MRI images poses a specific challenge related to achieving a balanced performance across various cardiac pathologies and normal cases. Imbalanced datasets can lead to biased machine learning models that perform poorly on underrepresented classes [2]. Various approaches, such as data augmentation and domain adaptation, have been proposed [3] to address this issue.

In this work, we explore the problem of dataset balancing in the context of semantic segmentation for cardiac MRI. We investigated the effectiveness of synthetic dataset balancing in improving the performance of machine learning algorithms on CMR data with pathologies that are unseen to the model. These synthetic datasets are generated by modelling the underlying physics of the imaging process and simulating variations in patient anatomy, physiology, and

imaging parameters. While synthetic datasets have the advantage of being easily scalable and customizable, they often suffer from a lack of diversity and realism.

Our proposed method combines modifications of an existing atlas [4] that represents a healthy patient with image-to-image style transfer [5] to produce synthetic subsets of data that constitute additional images to train segmentation networks.

1.1 Related Work

There are three main components we integrated into this research: (1) heart modelling from 2D slices and cardiac atlases, (2) image style transfer and (3) cardiac MRI segmentation.

Heart Modelling The field of heart modelling from slices has made significant progress recently, driven by statistical models and computational efficiency [6]. Recent works have focused on improving shape modelling by refining boundary extraction [7]. Those steps focus on the misalignment correction of the statistical shape model (SSM) extracted from the contours of the 2D images. From the SSM, a 3D mesh is generated and evaluated to fit plausible heart shapes [8], building the final reconstruction of the heart from the scanned structure.

Cardiac Atlas In the context of heart modelling and statistical parametric mapping (SPM), a method and an atlas were proposed by [4]. The authors provided an analysis of the number of subjects and methods necessary to represent a healthy and anatomically normal heart. This representation can be seen in the end-diastolic frame in Figure 1.

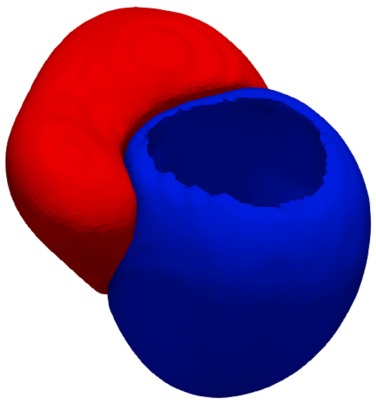


Fig. 1. 3D model of the atlas [4] at ED.

Style transfer Image-to-image translation has attracted a lot of attention during the last 5 years, and in particular, the topic of style transfer has remained an important application within the computer vision community.

In particular, in unpaired image-to-image translation, recent efforts have been made to improve the results and make them realistic [5]. CUT [5] uses a single direction mapping based on patches of images from both domains and benefits from the contrastive representation learning recent advances. In Figure 2 we can see the logic that the patch-wise contrastive loss follows.

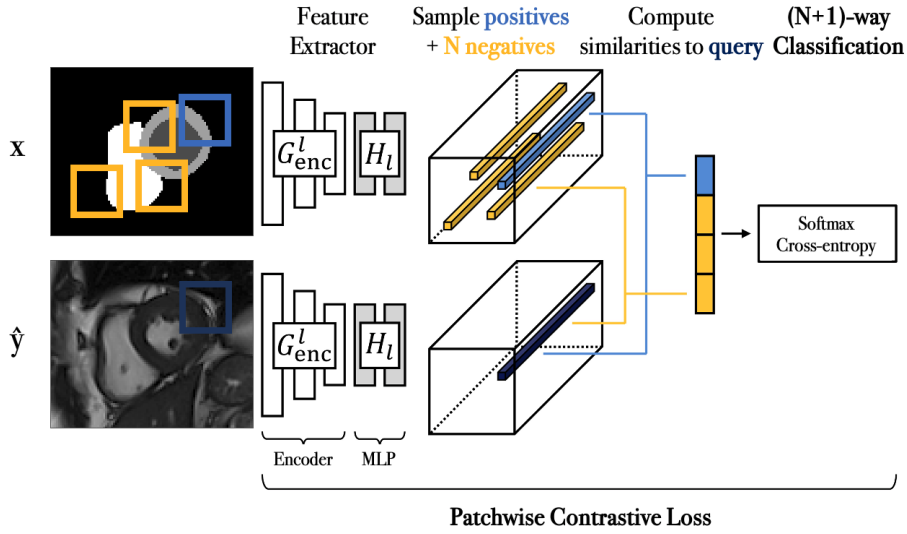


Fig. 2. Patchwise contrastive loss of CUT [5]. Figure adapted from [5].

While style transfer has found applications in regular images, only recently have a few endeavours harnessed these advances to enhance clinical applications, as evidenced by previous research such as [9] and [10]. Notably, while these works have successfully applied this technique, few have specifically explored its potential in the context of cardiac MRI applications or its role in addressing dataset imbalances [11, 12, 13, 14]. Indeed, none have tackled this imbalance without incorporating real data into their pipeline, whereas our approach relies solely on a single atlas.

Cardiac MRI segmentation In the last 5 years, a new paradigm has been established in the topic of cardiac magnetic resonance segmentation, working to extend the success of homogeneous datasets to more heterogeneous and diverse datasets [3, 2]. While the U-shaped architectures have remained a usual and powerful backbone for deep learning-based models [15, 16], data processing has significantly seen a prolific production of new methods [17, 18, 19, 20]. Addi-

tionally, there have been efforts of extending the available data to increment the number of images for training by using image registration [21]. Despite these efforts, there are no works exploring how we can benefit from atlases that represent one type of subject to complement a dataset by applying deformations to its 3D model.

1.2 Contributions

The contributions of this work are threefold: (1) we propose a method to modify a representative cardiac atlas and a sampling technique to obtain new subsets of data, in particular data from synthetic subjects with hypertrophic cardiomyopathy (HCM) and dilated right ventricle (DRV); (2) we built on recent advances in the image-to-image style transfer domain to obtain realistic MRI scans from its labels; (3) we successfully integrated synthetic and real data to address the imbalance in the dataset and trained segmentation networks to improve performance in patients with diseases that were not present in the training set. These three points constitute a novel pipeline that can be implemented in deep learning models to complement available data, similar to available data augmentation techniques.

2 Method

The proposed method is founded on the premise that inaccuracies in model predictions stem from the unique cardiac morphology observed in patients with pathologies. Our hypothesis is that the absence of patients with specific pathologies in the training data can cause the model to overfit to the pathologies that are present. To mitigate this, we introduce scans that exhibit morphology similar to the pathologies not originally represented in the training data. This augmentation increases the dataset's diversity, resulting in enhanced segmentation model performance. The overall structure of the method's pipeline is depicted in Figure 3.

To rectify the imbalance in our training data, we introduced what we refer to as "synthetic patients." These synthetic patients are created by applying deformations to an atlas, essentially simulating the heart structures of the two previously mentioned diseases, Hypertrophic Cardiomyopathy (HCM) and Dilated Right Ventricle (DRV). These simulations are achieved by virtually slicing the 3D model of the heart to generate these synthetic patient representations.

The incorporation of these synthetic patients into the dataset involves a crucial style transfer phase. During this step, images that resemble labels or structural representations are transformed into realistic MRI scans. We accomplished this transformation using a technique called "CUT" as described in the reference [5].

Finally, a segmentation network is trained using real and synthetic data.

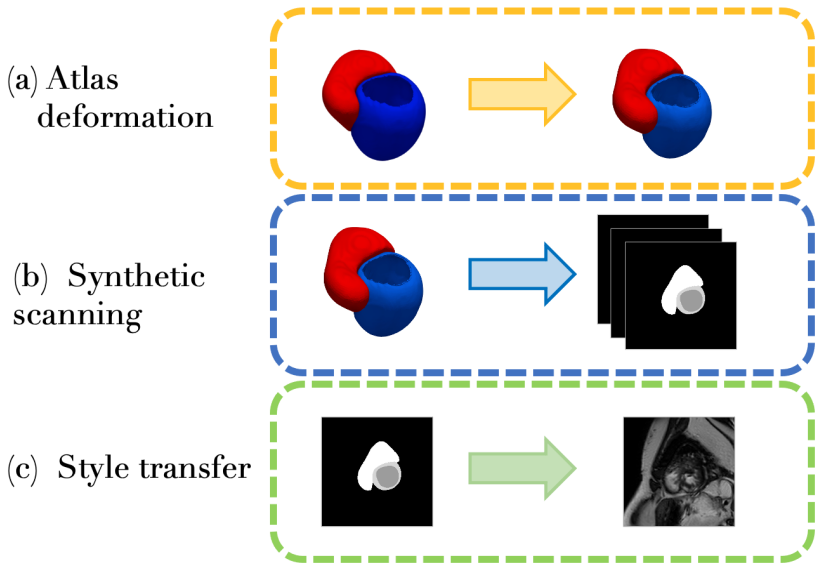


Fig. 3. The pipeline of the proposed method. The three main components correspond to: (a) the atlas deformation part, (b) the synthetic scanning component, and (c) the style transfer part.

2.1 Atlas-based data representation

Rather than creating a model based on the existing data, which would be computationally costly and could introduce errors during the 3D reconstruction steps, we chose to utilize the atlas described in [4]. This atlas represents over 1000 healthy subjects, providing a comprehensive and reliable reference for our purposes.

2.2 3D Deformations and Virtual Scanning Techniques

For this work, we created three different deformations that represent three different types of subjects: (1) a healthy subject, (2) 20 subjects with heart structures similar to patients with a dilated right ventricle, and (3) 20 subjects with heart structures similar to patients with hypertrophic cardiomyopathy. The visual representation of these diseases is presented in Figure 4.

The synthetic patients are obtained by performing a spherical cast operation (see Figure 5) to the cavities of the heart model. The spherical cast consists of a *To Sphere* transformation that will generate different results depending on the number and arrangement of the elements selected. Depending on the number of selected parts in each region (epicardium, endocardium or right ventricle) and the range of the deformation, we obtained a more or less prominent disease in the synthetic subject. For HCM, we enlarged the epicardium and endocardium,

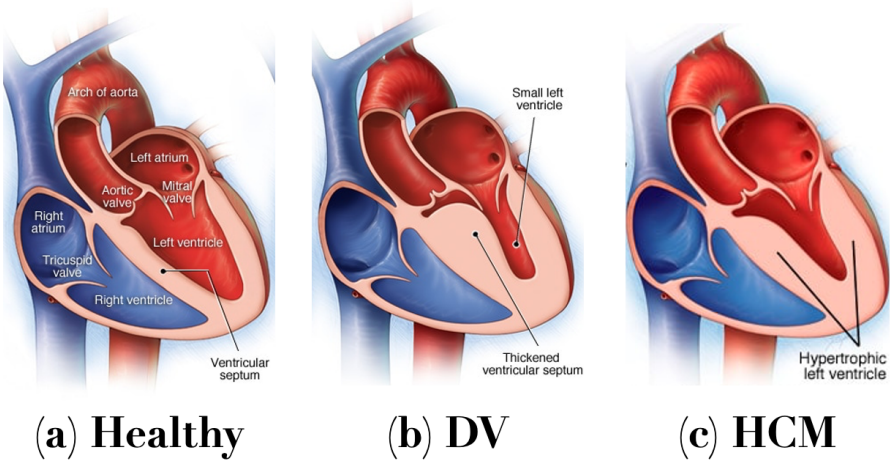


Fig. 4. Anatomical structure of the heart with: (a) healthy heart, (b) dilated ventricle, and (c) ventricular hypertrophy. Figure adapted from [22].

while for the DRV, we enlarged the RV. The factor of the spherical cast ranged between 0 and 0.5.

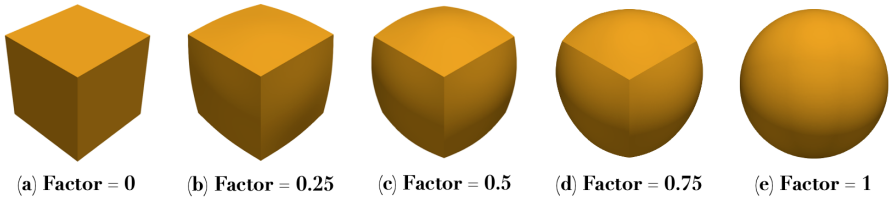


Fig. 5. Example of spherical cast applied to a cube with different factors.

The resulting 3D models are then processed as if they were scanned by obtaining slices throughout the longitudinal axis, to sample the equivalent to short axis labels. An example for each subset of data is presented in Figure 6.

2.3 Style transfer

We trained our own CUT [5] model with the available data from M&Ms 2 [2], where one domain is the MRI scan and the other is the label from the MRI scan.

CUT was our preferred style transfer model because, besides being based on CycleGAN [23], it has a contrastive term that encourages spatial consistency in

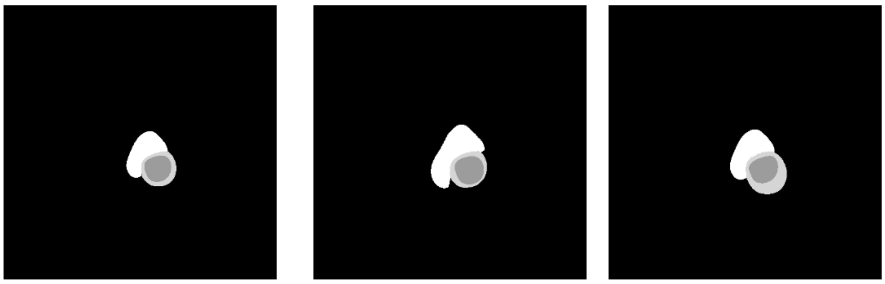


Fig. 6. Representations of the sampled slices from the 3D heart model: (a) no deformation, (b) deformations on the right ventricle representing a dilated right ventricle, and (c) deformations on the left ventricle representing hypertrophic cardiomyopathy in the left ventricle.

the generated image. The CUT loss function is presented by:

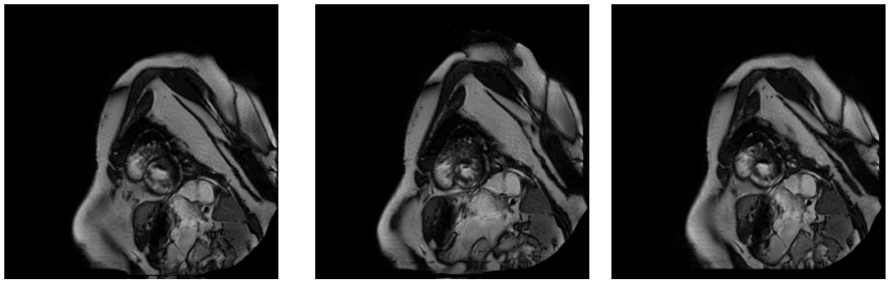
$$\mathcal{L}_{\text{GAN}}(G, D, X, Y) + \lambda_X \mathcal{L}_{\text{PatchNCE}}(G, H, X) + \lambda_Y \mathcal{L}_{\text{PatchNCE}}(G, H, Y) \quad (1)$$

where $\lambda_Y \mathcal{L}_{\text{PatchNCE}}(G, H, Y)$ is the contrastive term that enhances spatial consistency with the source image in X (synthetic image in our case). This term promotes the proximity of input-output patches from a specific location in an image within the feature space while ensuring that such patches are distant from other patches in the image. The discriminator part of the GAN is represented by D and the generator part as a G. The weights of a two-layer perceptron that projects the patches to the feature space are denoted as H, and the hyperparameters λ_X and λ_Y regulate the influence of the respective contrastive terms.

The results of applying a CUT trained on the available data to the synthetic data obtained from the previous steps are presented in Figure 7. The two domains are: (1) the MRI scans, and (2) the labels of the MRI scans. In other words, our CUT model learns to transfer the style of an MRI scan to a label, resulting in a realistic MRI scan with the structure displayed in a label. We used the resulting data as annotated data for the segmentation models.

2.4 Segmentation Network

To test the performance of the models trained with the different datasets we used the same architecture and the same preprocessing and normalization techniques. In particular, we implemented a standard U-Net [15] with 32 filters in the first convolutional layer. Data normalisation was performed with z normalisation and histogram standardisation. The preprocessing consisted in centre-cropping with padding.



(a) Healthy (b) DRV (c) HCM

Fig. 7. Slices from Figure 6 processed with our trained CUT [5]: (a) no deformation, (b) deformations on the right ventricle representing a dilated right ventricle, and (c) deformations on the left ventricle representing hypertrophic cardiomyopathy in the left ventricle.

Short axis Cardiac MRI Data We used the M&Ms2 dataset [2], released as part of the Multi-Center, Multi-View & Multi-Disease Right Ventricular Segmentation in Cardiac MRI Challenge. We evaluated the performance in all chambers for the short-axis view. In particular, we compared the performance of the different pathologies that the subjects of the dataset have. The pathologies are dilated left ventricle (LV), hypertrophic cardiomyopathy (HCM), congenital arrhythmogenesis (ARR), tetralogy of fallot (FALL), interatrial communication (CIA), dilated right ventricle (DRV), and tricuspidal regurgitation (TRI).

3 Experimental Methodology and Results

Our experiment consisted of evaluating the performance of four models trained with three different sets of data: (1) normal (healthy) cases from M&Ms2 plus the atlas sampled without deformation, (2) normal (healthy) cases from M&Ms2 plus 20 dilated right ventricle deformations applied to the atlas, (3) normal (healthy) cases from M&Ms2 plus 20 hypertrophic cardiomyopathy deformations applied to the atlas, and (4) normal (healthy) cases from M&Ms2 and classic data augmentation (affine and elastic transformation). The first model serves as a control model to compare the improvements when adding the synthetic data, and the last one is a comparison to classic data augmentation. The augmentations were applied with an overall probability of 0.5, and then each augmentation had equal probabilities of being applied.

The evaluation of the models was performed over the test split of the M&Ms2. In particular, we present the results for normal (healthy), hypertrophic cardiomyopathy, and dilated right ventricle; a part of overall performance. DICE scores are provided for each region. The results are presented in Table 2 and in Figure 8.

	Normal	HCM	DRV	Augs
Real Patients	40	40	40	40
Synthetic Patients Healthy	1	1	1	1
Synthetic Patients HCM	0	20	0	0
Synthetic Patients DRV	0	0	20	0

Table 1. Number of real and synthetic patients used to train the segmentation model. Synthetic refers to the atlas or its deformations with the style-transfer performed.

Data subset	Model	DICE				Hausdorff (mm)
		LV	MYO	RV	ALL	All
Normal (Healthy)	Healthy	0.9417	0.8252	0.8542	0.8737	6.311
	HCM	0.9400	0.8397	0.8273	0.8690	**6.2916**
	DRV	0.9333	0.8380	0.8053	0.8589	6.4883
	Augs	**0.9477**	**0.8416**	**0.8790**	**0.8894**	6.4223
HCM	Healthy	**0.9382**	**0.8583**	0.8284	**0.8750**	**6.5544**
	HCM	0.9195	0.8574	**0.8317**	0.8700	6.6080
	DRV	0.9148	0.8450	0.8239	0.8612	6.7190
	Augs	0.9089	0.8198	0.8111	0.8467	6.635
DRV	Healthy	0.7808	0.6036	**0.7742**	0.7195	7.3460
	HCM	0.7625	0.6146	0.7662	0.7149	7.1411
	DRV	**0.8049**	**0.6499**	0.7723	**0.7424**	6.9020
	Augs	0.7916	0.6243	0.7142	0.7101	**6.5030**
All	Healthy	0.8577	0.7149	0.7720	0.7815	6.8173
	HCM	0.8313	0.7137	**0.7856**	0.7769	6.770
	DRV	**0.8672**	**0.7388**	0.7729	**0.7930**	6.9027
	Augs	0.8581	0.7228	0.7667	0.7825	**6.5474**

Table 2. Segmentation performances for each subset of data. Each subset presents the DICE score (higher is better) and Hausdorff distance (lower is better) for each of the four models and each of the regions (Left Ventricle, Myocardium, and Right Ventricle). Models refer to the model trained with healthy patients of the original dataset [2] plus the atlas or its synthetic deformed aliases. The best results are in bold.

The results show how the addition of synthetic data had a positive impact on the performance, with a strong emphasis on the dilated right ventricle deformations. Overall performance increased the DICE score in that model by 0.01, and by 0.05 in the myocardium for the DRV subset of data, where the mean performance was more than 0.02 better than in the healthy subset.

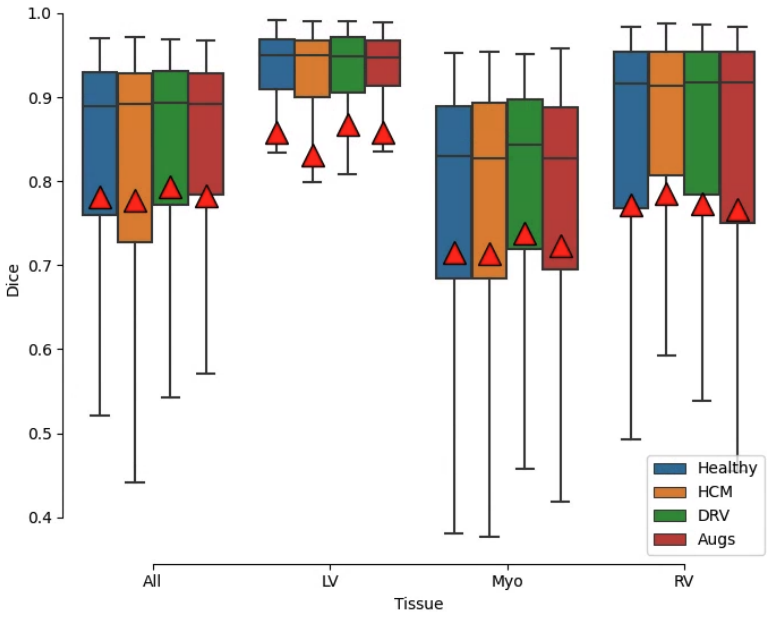


Fig. 8. Segmentation performances for each subset of data. Each subset presents the DICE score (higher is better) for each of the four models (Healthy, HCM, DRV, and Augs) and each of the tissues (Left Ventricle, Myocardium, Right Ventricle and all regions). Models refer to the model trained with healthy patients of the original dataset plus the atlas or its synthetic deformed aliases. The mean is plotted with a red mark.

4 Conclusions

In our study, we successfully trained a style transfer model to effectively generate synthetic cardiac MRI images that are sampled from a deformed atlas. Those synthetic images were successfully used to balance the training dataset. The results show how the addition of this data meant an increase in the DICE score of up to 0.05 in some regions within the target pathologies of the data. Nonetheless, the method was significantly less successful when simulating hypertrophic cardiomyopathies than it was when synthesising dilated right ventricle patients.

The incorporation of synthetic data in the DRV model led to improvements of up to 0.01 in the overall dataset and up to 0.03 on the DRV subset compared to classical augmentations. This demonstrates a distinct advantage over conventional techniques.

Future endeavours will delve into incorporating additional deformations to synthesize a more extensive range of diseases, coupled with enhancements in the style-transfer component of the pipeline. Moreover, extending the pipeline's design to generate synthetic patients with pathologies identified through functional analysis would constitute a significant advancement in the field.

References

[1] Chen Chen et al. *Deep learning for cardiac image segmentation: A review.* 2019. arXiv: 1911.03723 [eess.IV].

[2] Víctor M. Campello et al. "Multi-Centre, Multi-Vendor and Multi-Disease Cardiac Segmentation: The M&Ms Challenge". In: *IEEE Transactions on Medical Imaging* 40.12 (2021), pp. 3543–3554. DOI: 10.1109/TMI.2021.3090082.

[3] Víctor M. Campello et al. "Multi-Centre, Multi-Vendor and Multi-Disease Cardiac Segmentation: The M&Ms Challenge". In: *IEEE Transactions on Medical Imaging* 40.12 (2021), pp. 3543–3554. DOI: 10.1109/TMI.2021.3090082.

[4] Wenjia Bai et al. "A bi-ventricular cardiac atlas built from 1000+ high resolution MR images of healthy subjects and an analysis of shape and motion". In: *Medical Image Analysis* 26.1 (2015), pp. 133–145. ISSN: 1361-8415. DOI: https://doi.org/10.1016/j.media.2015.08.009. URL: https://www.sciencedirect.com/science/article/pii/S1361841515001346.

[5] Taesung Park, Alexei A. Efros, Richard Zhang, and Jun-Yan Zhu. "Contrastive Learning for Unpaired Image-to-Image Translation". In: *European Conference on Computer Vision*. 2020.

[6] Abhirup Banerjee et al. "A completely automated pipeline for 3D reconstruction of human heart from 2D cine magnetic resonance slices". en. In: *Philos. Trans. A Math. Phys. Eng. Sci.* 379.2212 (Dec. 2021), p. 20200257.

[7] Krithika Iyer et al. "Statistical Shape Modeling of Biventricular Anatomy with Shared Boundaries". In: *Statistical Atlases and Computational Models of the Heart. Regular and CMRxMotion Challenge Papers*. Ed. by Oscar Camara et al. Cham: Springer Nature Switzerland, 2022, pp. 302–316. ISBN: 978-3-031-23443-9.

[8] Benjamin Villard, Vicente Grau, and Ernesto Zacur. "Surface Mesh Reconstruction from Cardiac MRI Contours". In: *Journal of Imaging* 4.1 (2018). ISSN: 2313-433X. DOI: 10.3390/jimaging4010016. URL: https://www.mdpi.com/2313-433X/4/1/16.

[9] Enric Moreu, Eric Arazo, Kevin McGuinness, and Noel E. O'Connor. "Joint one-sided synthetic unpaired image translation and segmentation for colorectal cancer prevention". In: *Expert Systems* n/a.n/a (2022), e13137. DOI: https://doi.org/10.1111/exsy.13137. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/exsy.13137. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/exsy.13137.

[10] Mengting Liu et al. "Style Transfer Using Generative Adversarial Networks for Multi-site MRI Harmonization". In: *Medical Image Computing and Computer Assisted Intervention – MICCAI 2021*. Ed. by Marleen de Bruijne et al. Cham: Springer International Publishing, 2021, pp. 313–322. ISBN: 978-3-030-87199-4.

[11] Sina Amirrajab et al. "Label-informed cardiac magnetic resonance image synthesis through conditional generative adversarial networks". eng. In: *Computerized Medical Imaging and Graphics: The Official Journal of the*

Computerized Medical Imaging Society 101 (Oct. 2022), p. 102123. ISSN: 1879-0771. DOI: 10.1016/j.compmedimag.2022.102123.

[12] T. Karras, S. Laine, and T. Aila. "A Style-Based Generator Architecture for Generative Adversarial Networks". In: *IEEE Transactions on Pattern Analysis & Machine Intelligence* 43.12 (Dec. 2021), pp. 4217–4228. ISSN: 1939-3539. DOI: 10.1109/TPAMI.2020.2970919.

[13] Youssef Skandarani, Nathan Painchaud, Pierre-Marc Jodoin, and Alain Lalande. *On the effectiveness of GAN generated cardiac MRIs for segmentation*. 2020. arXiv: 2005.09026 [eess.IV].

[14] Fanwei Kong and Shawn C. Shadden. "A Generalizable Deep-Learning Approach for Cardiac Magnetic Resonance Image Segmentation Using Image Augmentation and Attention U-Net". en. In: *Statistical Atlases and Computational Models of the Heart. M&Ms and EMIDEC Challenges*. Ed. by Esther Puyol Anton et al. Lecture Notes in Computer Science. Cham: Springer International Publishing, 2021, pp. 287–296. ISBN: 978-3-030-68107-4. DOI: 10.1007/978-3-030-68107-4_29.

[15] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. *U-Net: Convolutional Networks for Biomedical Image Segmentation*. 2015. arXiv: 1505.04597 [cs.CV].

[16] Fabian Isensee, Paul F Jaeger, Simon A A Kohl, Jens Petersen, and Klaus H Maier-Hein. "nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation". In: *Nature Methods* 18.2 (2021), pp. 203–211. ISSN: 1548-7105. DOI: 10.1038/s41592-020-01008-z. URL: https://doi.org/10.1038/s41592-020-01008-z.

[17] Carles Garcia-Cabrera, Eric Arazo, Kathleen M. Curran, Noel E. O'Connor, and Kevin McGuinness. *Cardiac Segmentation using Transfer Learning under Respiratory Motion Artifacts*. 2022. DOI: 10.48550/ARXIV.2209.09714. URL: https://arxiv.org/abs/2209.09714.

[18] Fanwei Kong and Shawn C. Shadden. "A Generalizable Deep-Learning Approach for Cardiac Magnetic Resonance Image Segmentation Using Image Augmentation and Attention U-Net". In: *Statistical Atlases and Computational Models of the Heart. M&Ms and EMIDEC Challenges*. Ed. by Esther Puyol Anton et al. Cham: Springer International Publishing, 2021, pp. 287–296. ISBN: 978-3-030-68107-4.

[19] Xiaoqiong Huang et al. "Style-Invariant Cardiac Image Segmentation with Test-Time Augmentation". In: *Statistical Atlases and Computational Models of the Heart. M&Ms and EMIDEC Challenges*. Ed. by Esther Puyol Anton et al. Cham: Springer International Publishing, 2021, pp. 305–315. ISBN: 978-3-030-68107-4.

[20] Jun Ma. "Histogram Matching Augmentation for Domain Adaptation with Application to Multi-centre, Multi-vendor and Multi-disease Cardiac Image Segmentation". In: *Statistical Atlases and Computational Models of the Heart. M&Ms and EMIDEC Challenges*. Ed. by Esther Puyol Anton et al. Cham: Springer International Publishing, 2021, pp. 177–186. ISBN: 978-3-030-68107-4.

[21] Carles Garcia-Cabrera, Kathleen M. Curran, Noel E. O'Connor, and Kevin McGuinness. "Semi-supervised learning of cardiac MRI using image registration". In: *Irish Pattern Recognition & Classification Society Conference Proceedings* (Sept. 2021). URL: https://doras.dcu.ie/26161/.

[22] Mayo Clinic. *Diseases and conditions: Cardiomyopathies*. https://www.mayoclinic.org/diseases-conditions/chronic-fatigue-syndrome/symptoms-causes/syc-20360490. Jan. 2023.

[23] Jun-Yan Zhu, Taesung Park, Phillip Isola, and Alexei A Efros. "Unpaired image-to-image translation using cycle-consistent adversarial networks". In: *Proceedings of the IEEE international conference on computer vision*. 2017, pp. 2223–2232.