



Combining Multi-Sequence and Synthetic Images for Improved Segmentation of Late Gadolinium Enhancement Cardiac MRI

Víctor M. Campello¹(✉), Carlos Martín-Isla¹, Cristian Izquierdo¹, Steffen E. Petersen^{3,4}, Miguel A. González Ballester^{2,5}, and Karim Lekadir¹

¹ Dept. Matemàtiques i Informàtica, Universitat de Barcelona, Barcelona, Spain

victor.campello@ub.edu

² BCN-MedTech, DTIC, Universitat Pompeu Fabra, Barcelona, Spain

³ Barts Heart Centre, Barts Health NHS Trust, London, UK

⁴ William Harvey Research Institute, NIHR Barts Biomedical Research Centre, Queen Mary University of London, London, UK

⁵ ICREA, Barcelona, Spain

Abstract. Accurate segmentation of the cardiac boundaries in late gadolinium enhancement magnetic resonance images (LGE-MRI) is a fundamental step for accurate quantification of scar tissue. However, while there are many solutions for automatic cardiac segmentation of cine images, the presence of scar tissue can make the correct delineation of the myocardium in LGE-MRI challenging even for human experts. As part of the Multi-Sequence Cardiac MR Segmentation Challenge, we propose a solution for LGE-MRI segmentation based on two components. First, a generative adversarial network is trained for the task of modality-to-modality translation between cine and LGE-MRI sequences to obtain extra synthetic images for both modalities. Second, a deep learning model is trained for segmentation with different combinations of original, augmented and synthetic sequences. Our results based on three magnetic resonance sequences (LGE, bSSFP and T2) from 45 different patients show that the multi-sequence model training integrating synthetic images and data augmentation improves in the segmentation over conventional training with real datasets. In conclusion, the accuracy of the segmentation of LGE-MRI images can be improved by using complementary information provided by non-contrast MRI sequences.

Keywords: Multi-sequence cardiac MRI · Late gadolinium enhancement MRI · Image segmentation · Image synthesis · Deep learning

1 Introduction

Late gadolinium enhancement magnetic resonance imaging (LGE-MRI) is widely used to assess presence, location and extent of regional scar or fibrotic tissue in

the myocardium. Whilst LGE-MRI is a well-established technique and key to many cardiovascular magnetic resonance (CMR) examinations there are challenges in quantification and interpretation due to a number of factors. Image analysis depends on image quality which can be affected by suboptimal CMR acquisition. Correct inversion times (TI) need to be identified and then TI require appropriate adjustments to allow good ‘nulling’ of remote, unaffected myocardium. This ensures optimal contrast between scar/fibrosis (bright) and normal, remote myocardium (dark). Timing after contrast administration is important to allow not only sufficient wash-out of contrast agent (gadolinium chelate) from the remote myocardium but also from the blood pool. Images acquired too early will leave the blood pool bright which makes differentiating subendocardial infarct from blood pool challenging.

In the existing literature, two main families of techniques have been proposed to automatically segment LGE-MRI data. The first one segments directly the LGE-MRI images by using different techniques such as graph-cuts [1], atlas-based registration [2], or more recently Convolutional Neural Networks (CNNs) [3]. However, these techniques generally lack robustness due to the limited availability of LGE-MRI datasets for training. As a result, the second family of techniques has considered exploiting other cardiac MRI sequences to provide additional signals for guiding more robustly the segmentation process. For instance, some researchers [4,5] proposed to segment first cine-MRI images and to propagate the obtained contours into the LGE-MRI images through image registration. Similarly but by using additional sequences, the authors in [6] implemented an atlas-based segmentation approach combining information from balanced-Steady State Free Processing (bSSFP), LGE and T2 sequences. However, these techniques are highly dependent on the image registration step, which is challenging due to the inherent differences between the cardiac MRI sequences.

In addition, in order to improve segmentation and increase the model robustness over unseen data, image synthesis has been proposed recently. The most common model combines generative adversarial networks (GANs) with a cycle-consistency constrain for image-to-image translation and two segmentation networks, one for each image domain, trained end-to-end in order to benefit from a combined loss function. This model has been applied for cross-modality segmentation improvement [7,8], domain adaptation across scanners [8] or across modalities [9] and segmentation of an unlabeled target modality using only the source ground truth [10,11]. Alternatively, a GAN can be trained to generate synthetic images from masks according to some conditional value, like the dataset style, as in the case of retinal fundus images for vessel segmentation [12].

In this paper, we propose an approach to circumvent the need for image registration, while addressing the lack of LGE-MRI images for training. Concretely, we implement a CNN-based approach that is capable of learning key properties of the cardiac structures simultaneously from multiple cardiac MRI sequences. Furthermore, image synthesis and data augmentation are used to generate new examples that take into account both the global appearance of LGE-MRI data and the local appearance of scar tissues. With this approach, direct deep learning

based segmentation of LGE-MRI is enabled without the need for inter-sequence image registration and while exploiting the richness of multi-sequence cardiac MRI.

2 Method

2.1 Dataset

Data Description. The LGE-MRI dataset used in this paper was provided as part of the Multi-Sequence Cardiac Magnetic Resonance Segmentation Challenge (MS-CMRSeg). It consists of 45 patients from Shanghai Renji Hospital that were scanned using three MRI sequences: bSSFP, LGE and T2. Ground truth segmentations of the left ventricle (LV), right ventricle (RV) and myocardium (MYO) were provided for some of the cases according to the distribution in Table 1 (second row). Even though all sequences were acquired and selected for the end-diastolic cardiac phase, there were differences in the shape of the cardiac boundaries consistently between the three sequences for the same patient. Moreover, the slices were not aligned between the sequences in the direction of the ventricular axis, which further complicates the application of image registration. Note that all patients in the sample suffer from cardiomyopathies and that every LGE-MRI image presents a scar of variable size within the myocardial wall.

Table 1. MS-CMRSeg sequences details.

	bSSFP	LGE	T2
Number of patients	45	45	45
Segmented patients	35	5	35
Number of slices	8–12	10–18	3–7
Slice thickness (mm)	8–13	5	12–20
TR/TE (ms)	2.7/1.4	3.6/1.8	2000/90
In-plane resolution (mm)	1.25×1.25	0.75×0.75	1.35×1.35

Data Pre-processing. As a first step, intensity bias correction was applied to all sequences to correct for potential artifacts and the intensity histograms of all images were matched to a common one to obtain coherent appearances across images. Furthermore, before the training process, all images were interpolated and cropped so that they had a pixel size of 256×256 and the same resolution. They were also normalised such that the mean intensity and the standard deviation equal 0.5, thus ensuring most of the input values to be positive in between 0 and 1 for convenience in later representation of the images.

2.2 Increasing Training Sample

Before describing the CNN model implemented in this paper for LGE-MRI segmentation, this section presents two methods used to increase the number of training data and obtain higher LGE-MRI variability.

Data Augmentation. By using the provided segmentations, a set of 50 landmarks were evenly placed around the epicardium and endocardium. With these, the myocardium and left ventricle were rotated relative to the rest of the image, as shown in the examples in Fig. 1, in order to obtain an augmented dataset with varying locations of the scar tissues. Since the contour of the epicardium is not perfectly round in general, a Gaussian filter of size 3×3 was applied around the outer boundary to smooth the transition between the rotated and fixed regions, thus preventing image intensity discontinuities. A total of twenty 7.2 degrees rotations were applied. Thus, the LGE-MRI dataset was multiplied by a factor of 20 and the location of the scar in the myocardium ranged between the initial position and 144 degrees clockwise. This augmentation technique increases the variability in the scar locations within the myocardial wall that was otherwise very low due to the small number of patients available for training. Furthermore, further data augmentations were obtained by applying small rotations of the input images up to 15 degrees before training.

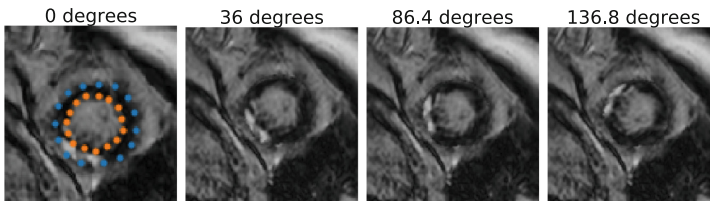


Fig. 1. Example of three rotations of the myocardial wall with respect to the whole image by using the landmarks provided in the leftmost image. This shows the changes in the location of the scar tissues

Image Synthesis. The rationale behind the proposed image synthesis is that there are many more segmented cine-MRI datasets available open-access or in clinical registries for training CNN models. Thus, to increase the number of annotated LGE-MRI cases for training, image synthesis from cine-MRI images sequences is proposed. To achieve this, the CycleGAN method [13] was implemented using the PyTorch library provided at this link¹.

This method translates images from one domain to another without the need for image registration or for the sequences to be from the same patients. It consists of a pair of generators G_{LGE} , G_{bSSFP} and a pair of discriminators D_{LGE} ,

¹ <https://github.com/junyanz/pytorch-CycleGAN-and-pix2pix>.

D_{bSSFP} that have opposed goals. The generator G_{LGE} (G_{bSSFP}) transforms the bSSFP (resp. LGE) sequence into a realistic LGE (bSSFP) image, while the discriminator D_{LGE} (D_{bSSFP}) attempts to distinguish between real and fake LGE (bSSFP) sequences. To achieve a good image translation between the two sequences, the loss function contains two terms: (1) an adversarial loss for each target domain that accounts for the similarity between the generated and real images, and (2) a cycle consistency loss that ensures that the transformed image $G_{LGE}(X)$ ($G_{bSSFP}(Y)$) is transformed back to X (Y) through G_{bSSFP} (G_{LGE}).

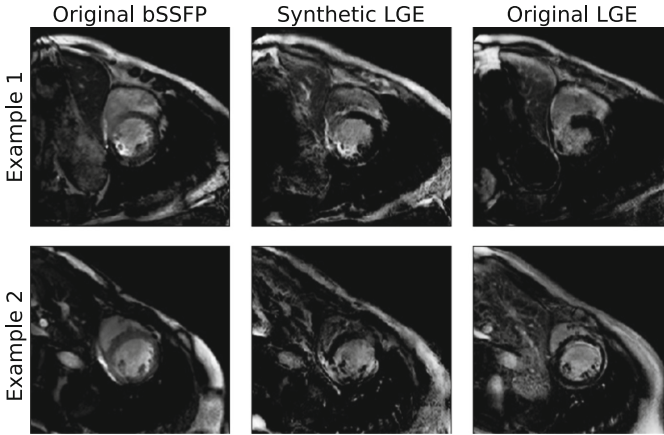


Fig. 2. Examples of synthetic LGE-MRI images. The leftmost column are the original cine images, the central column shows the transformed images to the LGE domain and the rightmost column is the most similar slice from the real LGE sequences, since they were not registered/aligned.

For the training of the CycleGAN model, all slices from the 45 patients for the LGE and bSSFP sequences were used during 200 epochs. The training took 12 h on a NVIDIA 1080 GPU with a batch size of 1. The Adam optimizer was used with learning rate of 2×10^{-4} , with first and second moment decay rates of 0.5 and 0.999, respectively. Some examples for the generated images are shown in Fig. 2.

In order to evaluate the quality of the generated images, two segmentation models (like the one described in the next subsection) were trained using the bSSFP images and the synthetic LGE images separately. The obtained results are presented in Table 2. In particular, the synthetic LGE images, that are anatomically similar to the original bSSFP, provide more information for the task of LGE segmentation.

2.3 CNN-based LGE Segmentation

Once a large set of training sample was obtained from the original, augmented and synthetic images, a modified U-Net architecture [14] was used for the image

Table 2. Average and standard deviation for the Dice score of segmentation results over the five labeled LGE volumes.

	LV		MYO		RV	
	avg.	std.	avg.	std.	avg.	std.
model trained w. bSSFP	0.503	0.406	0.370	0.301	0.515	0.434
model trained w. synthetic LGE	0.809	0.116	0.688	0.145	0.820	0.065

segmentation by integrating two techniques: (1) a deep supervision term in the upsampling path as proposed in [15] that will act as lower-resolution masks that are convolved to condition the final predictions; and (2) a reduction of the number of filters after each upsampling operation to match the number of labels as proposed by [16]. Each image in the dataset was provided as a single channel input, thus forcing the model to differentiate between sequences with a unique set of weights. Additionally, in order to avoid overfitting given the sample size, dropout was used after every max pooling and upsampling operations, except for the high level features in the architecture, as shown in Fig. 3.

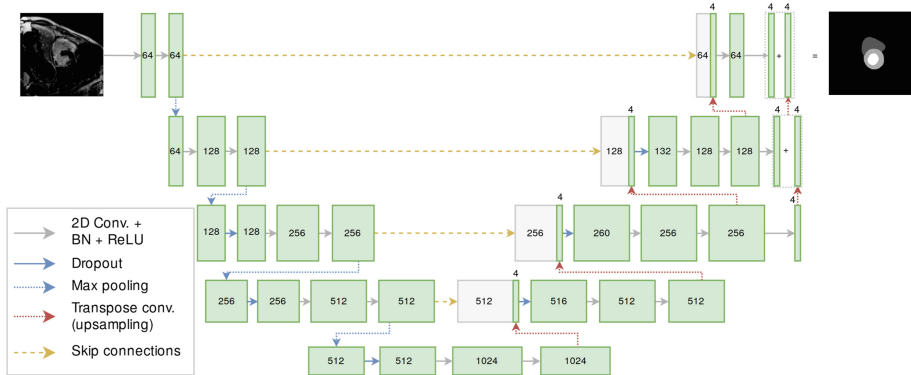


Fig. 3. Detailed architecture of the CNN model used for LGE segmentation. The numbers in the boxes correspond to the number of channels. Convolution operations have a kernel size of 3×3 and stride of 1, while transpose convolutions have a kernel size of 4×4 and stride of 2.

During training, 20% of the patients for each dataset was reserved for validation and early stopping. With a batch size of 8 images, this model took less than 36 h to achieve the best accuracy on the validation set after almost 90 epochs on a NVIDIA TITAN X GPU. The Adam optimizer was used with a learning rate of 10^{-4} , with first and second moment decay rates equal to 0.9 and 0.99, respectively.

3 Results

In order to define the best trained CNN model for LGE-MRI segmentation, various training sets were used by varying the input sequences and combinations of image synthesis and scar augmentation, as follows:

1. LGE sequences only;
2. LGE and bSSFP sequences;
3. All sequences (LGE, bSSFP and T2);
4. All sequences plus MYO and LV rotations in LGE sequences;
5. Number 1 plus synthetic LGE sequences;
6. Number 2 plus synthetic LGE sequences;
7. Number 3 plus synthetic LGE sequences;
8. Number 4 plus synthetic LGE sequences.

When evaluated on the validation set, the training set number 8 resulted in the best segmentations, showing the added value of image synthesis and data augmentation for LGE-MRI segmentation. Thus, we applied the corresponding CNN model to the test dataset composed of 40 LGE-MRI cases. The obtained segmentations were sent to the organizers of MS-CMRSeg Challenge for evaluation. The obtained results are summarized in Table 3, showing average dice scores of 90% (LV), 87% (MYO) and 81% (RV).

Table 3. Average and standard deviation for results over the test set.

	LV		MYO		RV	
	avg.	std.	avg.	std.	avg.	std.
Dice score	0.898	0.045	0.810	0.061	0.866	0.051
Jaccard index	0.817	0.072	0.685	0.084	0.768	0.078
Surface distance (mm)	2.0	0.8	1.8	0.5	2.3	0.9
Hausdorff distance (mm)	11	4	12	4	16	7

Two remarks are important to note regarding the results reported in Table 3: (1) Despite the high variability in the LGE-MRI datasets, especially in the presence, extent and location of the scar tissues, relatively consistent results are obtained with standard deviations for the dice scores around 5%. (2) Despite the availability of only 5 LGE-MRI volumes for training, the proposed approach was able to achieve comparable results to very recent deep learning techniques, which reported dice scores of 0.915 ± 0.052 (LV), 0.812 ± 0.105 (MYO) and 0.882 ± 0.084 (RV) based on 5 times more training cases (25 LGE-MRI images). [3]. This indicates the value of the proposed inter-sequence synthesis and scar augmentation for generating richer training samples.

Finally, for visual illustration, Fig. 4 shows three segmentation examples as obtained in this study. Model number 3 (second column) introduces errors that

are corrected when adding synthetic images (model number 7 in the third column). The last column shows that the segmentation further improves when integrating the scar tissue augmentation as proposed in this paper (model 8).

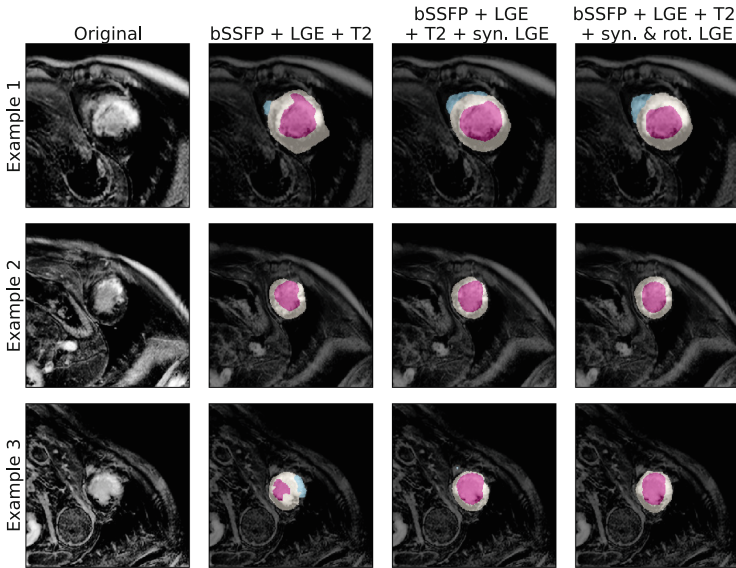


Fig. 4. Three segmentation examples as obtained by using different training combinations, showing the improvement achieved by integrating inter-sequence image synthesis (column 3) and scar tissue augmentation (column 4) during training.

4 Conclusions

This paper proposed to address the limited availability of training samples for LGE-MRI segmentation by enriching the CNN models using two complimentary methods. Firstly, since samples of annotated cine-MRI sequences are more commonly available, image synthesis of LGE-MRI images was implemented using a CycleGAN approach, thus obtaining a larger number of LGE-MRI cases during training. Secondly, we performed LGE-specific data augmentation through shape-guided rotations of the myocardium, which increases the variability related to the location of the scar tissues in the myocardium. The validation shows consistent results across the datasets, indicating the potential of this approach for enhancing the richness and generalization of LGE-specific CNNs.

Future work include the extension of the image synthesis to take into account local cardiac motion abnormality for synthesizing scar tissue, as well as the use of elastic deformations of the myocardium and scar to augment non-rigidly the LGE-MRI examples. Furthermore, extensive validation will be performed

to assess in detail the relative importance of the different steps and sequences (bSSFP, T2) in enriching the CNN models for LGE segmentation.

Acknowledgements. This work was partly funded by the European Union’s Horizon 2020 research and innovation programme under grant agreement No 825903 (euCan-SHare project). SEP acts as a paid consultant to Circle Cardiovascular Imaging Inc., Calgary, Canada and Servier. SEP acknowledges support from the National Institute for Health Research (NIHR) Cardiovascular Biomedical Research Centre at Barts, from the SmartHeart EPSRC programme grant (EP/P001009/1) and the London Medical Imaging and AI Centre for Value-Based Healthcare. SEP and KL acknowledge support from the CAP-AI programme, London’s first AI enabling programme focused on stimulating growth in the capital’s AI Sector.

References

1. Alba, X., Figueras i Ventura, R.M., Lekadir, K., Tobon-Gomez, C., Hoogendoorn, C., Frangi, A.F.: Automatic cardiac LV segmentation in MRI using modified graph cuts with smoothness and interslice constraints. *Magn. Reson. Med.* **72**(6), 1775–1784 (2014)
2. Kurzendorfer, T., Forman, C., Schmidt, M., Tillmanns, C., Maier, A., Brost, A.: Fully automatic segmentation of left ventricular anatomy in 3-D LGE-MRI. *Comput. Med. Imaging Graph.* **59**, 13–27 (2017)
3. Yue, Q., Luo, X., Ye, Q., Xu, L., Zhuang, X.: Cardiac segmentation from LGE MRI using deep neural network incorporating shape and spatial priors. arXiv preprint [arXiv:1906.07347](https://arxiv.org/abs/1906.07347) (2019)
4. Wei, D., Sun, Y., Chai, P., Low, A., Ong, S.H.: Myocardial segmentation of late gadolinium enhanced MR images by propagation of contours from cine MR images. In: Fichtinger, G., Martel, A., Peters, T. (eds.) MICCAI 2011. LNCS, vol. 6893, pp. 428–435. Springer, Heidelberg (2011). https://doi.org/10.1007/978-3-642-23626-6_53
5. Tao, Q., Piers, S.R., Lamb, H.J., van der Geest, R.J.: Automated left ventricle segmentation in late gadolinium-enhanced MRI for objective myocardial scar assessment. *J. Magn. Reson. Imaging* **42**(2), 390–399 (2015)
6. Zhuang, X.: Multivariate mixture model for myocardial segmentation combining multi-source images. *IEEE Trans. Pattern Anal. Mach. Intell.* **41**(12), 2933–2946 (2018)
7. Zhang, Z., Yang, L., Zheng, Y.: Translating and segmenting multimodal medical volumes with cycle-and shape-consistency generative adversarial network. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 9242–9251 (2018)
8. Cai, J., Zhang, Z., Cui, L., Zheng, Y., Yang, L.: Towards cross-modal organ translation and segmentation: a cycle-and shape-consistent generative adversarial network. *Medical Image Anal.* **52**, 174–184 (2019)
9. Chen, C., Dou, Q., Chen, H., Qin, J., Heng, P.A.: Synergistic image and feature adaptation: towards cross-modality domain adaptation for medical image segmentation. arXiv preprint [arXiv:1901.08211](https://arxiv.org/abs/1901.08211) (2019)
10. Huo, Y., et al.: Synseg-net: synthetic segmentation without target modality ground truth. *IEEE Trans. Med. Imaging* **38**(4), 1016–1025 (2018)

11. Zhang, Y., Miao, S., Mansi, T., Liao, R.: Task driven generative modeling for unsupervised domain adaptation: application to X-ray image segmentation. In: Frangi, A.F., Schnabel, J.A., Davatzikos, C., Alberola-López, C., Fichtinger, G. (eds.) MICCAI 2018. LNCS, vol. 11071, pp. 599–607. Springer, Cham (2018). https://doi.org/10.1007/978-3-030-00934-2_67
12. Zhao, H., Li, H., Maurer-Stroh, S., Guo, Y., Deng, Q., Cheng, L.: Supervised segmentation of un-annotated retinal fundus images by synthesis. *IEEE Trans. Med. Imaging* **38**(1), 46–56 (2018)
13. Zhu, J.Y., Park, T., Isola, P., Efros, A.A.: Unpaired image-to-image translation using cycle-consistent adversarial networks. In: Proceedings of the IEEE international conference on computer vision, pp. 2223–2232 (2017)
14. Ronneberger, O., Fischer, P., Brox, T.: U-Net: convolutional networks for biomedical image segmentation. In: Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F. (eds.) MICCAI 2015. LNCS, vol. 9351, pp. 234–241. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-24574-4_28
15. Isensee, F., Jaeger, P.F., Full, P.M., Wolf, I., Engelhardt, S., Maier-Hein, K.H.: Automatic cardiac disease assessment on cine-MRI via time-series segmentation and domain specific features. In: Pop, M., et al. (eds.) STACOM 2017. LNCS, vol. 10663, pp. 120–129. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-75541-0_13
16. Baumgartner, C.F., Koch, L.M., Pollefeys, M., Konukoglu, E.: An exploration of 2D and 3D deep learning techniques for cardiac MR image segmentation. In: Pop, M., et al. (eds.) STACOM 2017. LNCS, vol. 10663, pp. 111–119. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-75541-0_12