CONVOLUTIONAL SIAMESE NETWORKS FOR MYOCARDIAL LESION CLASSIFICATION IN T1 MAPS

M. Golub[⋆†] *C. Santiago*[⋆†] *C. Baleia*[⋆†] *P. Lopes*[⋆] *A. M. Ferreira*[⋆] *R. G. Nunes*[⋆†] *J. C. Nascimento*[⋆†]

[⋆†] Institute for Systems and Robotics - Lisboa, Instituto Superior Técnico, Lisbon, Portugal
[⋆] Serviço de Cardiologia do Hospital Santa Cruz, Centro Hospitalar de Lisboa Ocidental, Portugal

ABSTRACT

This paper presents a novel framework for myocardial lesion classification in T1 maps. Late gadolinium enhancement (LGE) imaging is the gold standard for detecting myocardial lesions, however, it requires contrast injection, which raises two main challenges: i) increased scan times and patient preparation, and ii) acute side effects in a certain class of patients. It would therefore be desirable to use a less invasive method, such as T1 mapping, however, this modality does not always present noticeably increased T1 values in the lesion. Taking into account the above challenges, we propose a two stage framework: i) the approach is able to learn associations between T1 map and LGE image during a training phase; (ii) in a test phase, only the T1 map is used, using the previously learned associations. The associations are learned following three siamese inference methodologies. Our experimental results testify the usefulness of the proposed approach on classification of the myocardium lesion in T1 maps.

Index Terms— CMR, T1, LGE, Siamese Network

1. INTRODUCTION

Late gadolinium enhancement imaging (LGE) in cardiac magnetic resonance (CMR) relies on the injection of a gadolinium based contrast agent (GBCA) that freely distributes in the extracellular space and presents increased accumulation in regions with scar tissue [1]. Regions with accumulated GBCA have shorter T1 relaxation times and produce hyper-intense signals. LGE imaging has been validated and shown to be sensitive to myocardial lesion detection in various types of cardiovascular diseases [2, 3, 4]. However, the use of this modality is not advised in patients known to have acute side effects to GBCA [5] or those with impaired kidney function [6]. Circumventing the need for LGE would allow the diagnosis and study of cardiomyopathies in a broader group of patients, and also reduce scan time and associated costs in CMR.

Native (no contrast) T1 mapping could provide an adequate candidate for tissue characterization without the need for a contrast agent. This modality measures the spin-lattice relaxation time, a fundamental tissue property sensitive to multiple myocardial abnormalities [7, 8], however, its use in a clinical setting is impeded by a lack of specificity and possible confounding factors [7].

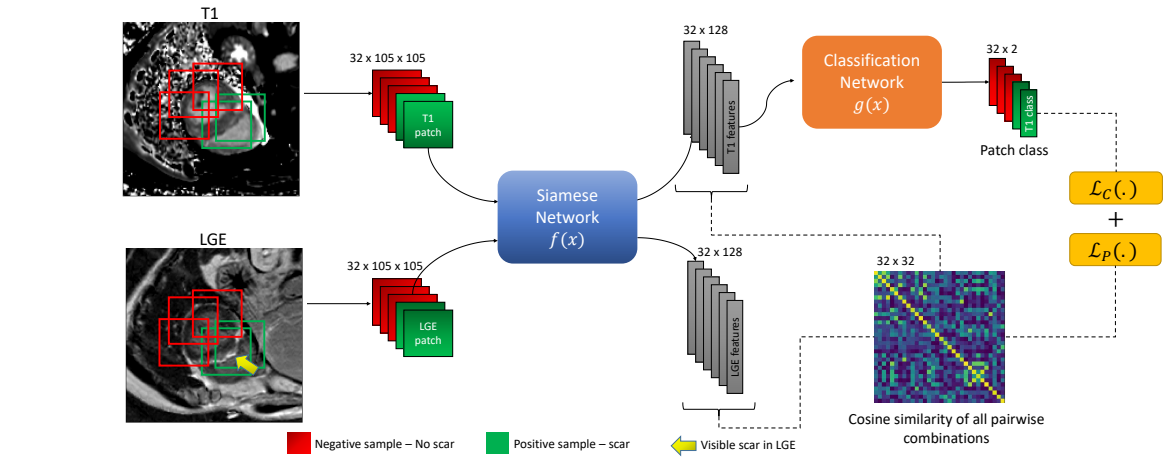
The main goal of this study is to assess if there is any intrinsic information in native T1 maps that correlates to LGE images, and if it is possible to use that information to classify myocardial lesions in the absence of LGE. For this purpose, a convolutional siamese network was used to learn from both modalities. Since these types of networks have been successfully used to learn similarity between two inputs in problems such as facial recognition [9] and fake signature detection [10], we hypothesized that it would be an adequate approach to extract a representation that contains information common to T1 maps and LGE images, so that it could be used in a classifier network to detect scar tissue. The designed network and two proposed variations were trained on pairs of T1 maps and corresponding LGE images of healthy individuals and patients diagnosed with hypertrophic cardiomyopathy (HCM). HCM is a common disorder characterized by hypertrophy and fibrosis in the left ventricle (LV), known to present focal hyper-intensities in LGE [11, 12]. Changes in T1 maps are more subtle but can be valuable to detect diffuse changes and are becoming increasingly more used in the clinical setting [7, 13]. In a subsequent validation phase, the trained network was used to detect scar tissue using only the T1 map.

2. METHODS

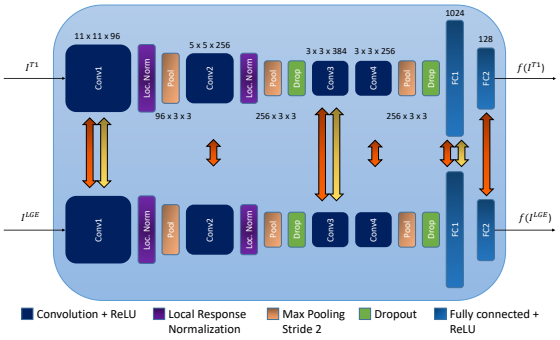
A description of used data and its pre-processing is provided in section 2.1 and details on model architecture in section 2.2.

2.1. Data and Pre-processing

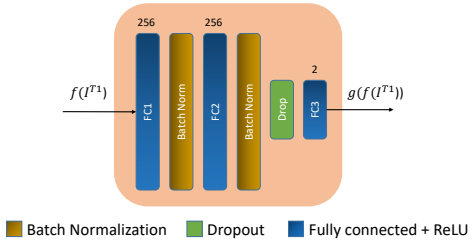
T1 maps and LGE image pairs were acquired from 15 subjects, two short-axis slices of the LV were selected from each (medial and basal), resulting in 30 image pairs, out of which 17 presented hyper-intensities in LGE. Since LGE images are typically acquired 10 minutes after the T1 maps, to reduce the impact of eventual subject movement, an intra-subject image registration was performed. Directly registering T1 maps to LGE space did not result in a good alignment. Each T1 map was hence transformed to LGE space, by first aligning binary myocardium masks, which were manually segmented. A lesion mask was also obtained from the LGE image, identifying hyper-intense pixels compared to a healthy reference region. Pixels with intensities higher than six-standard devi-



(a) Model overview



(b) Siamese network architecture



(c) Classification network architecture

Fig. 1. Proposed methodology for lesion classification from T1 map patches, by taking advantage of LGE information.

ations above the mean signal intensity of the healthy region were labeled as lesion [14].

The data was divided into train and test sets (four image pairs). For data augmentation, a sliding window was used to obtain pairs of square patches centered on pixels inside the myocardium mask (see Fig. 1 for an illustration). During this process, some patch pairs were also randomly flipped to obtain new samples. Each patch pair was labeled as positive/negative if it contained at least 25% of the total lesion present in the entire image plane from where the patch was taken. A total of 920 training samples (371 positive, 549 negative) and 222 test samples (117 positive, 105 negative) were obtained.

2.2. Network design

The network architecture used in this study consists of two sub-networks: i) a siamese convolutional network inspired by [15] and ii) a fully connected classifier network. The siamese network has two identical branches, each branch takes one type of input, T1 or LGE patches, and consists mainly of four convolutional layers and two linear fully connected layers. The branches of the siamese network have shared weights that are updated simultaneously. In practice, this means that the two inputs are forwarded consecutively through a single network before back propagation. Each branch produces a 128 size feature encoding and every pairwise combination of feature vectors in a mini-batch are compared with a normalized-temperature cross entropy loss function. Concretely, for a given pair of T1 and LGE image patches (I_i^{T1}, I_j^{LGE}) in a mini-batch of size N, the pairing loss is given by

$$\mathcal{L}_P(I_i^{T1}, I_j^{LGE}) = -\log \frac{\exp(\text{sim}(f(I_i^{T1}), f(I_j^{LGE}))/\tau)}{\sum_{k=1}^{N} \exp(\text{sim}(f(I_i^{T1}), f(I_k^{LGE}))/\tau)} y_i \quad (1)$$

where $f(I_i^{T1})$, $f(I_j^{LGE})$ are the feature vectors produced by each siamese branch and $\text{sim}(x, y)$ is the cosine similarity between them, $\tau = 0.07$ is the temperature hyper-parameter and y_i is the class for a given I_i^{T1}, that indicates which is the corresponding I_j^{LGE} (i.e. extracted from the same anatomical region) in that batch. This loss function enables the network to learn to discriminate matched pairs of T1 and LGE patches, *i.e.*, pairs of patches extracted from the same anatomical region of the same subject, from mismatched pairs.

The vector encoding produced by the T1 branch of the siamese network is then passed to a classifier network, which

	Classification Accuracy			Classification Loss	Pairing Loss
	balanced	positive	negative		
pure classifier	50	100	0	5.11	-
siamese	69.51	41.88	97.14	1.31	1.63
partial	78.96	59.83	98.1	0.91	1.82
decoupled	**81.61**	71.79	91.43	0.51	1.78

Table 1. Test set accuracies and losses

consists of three fully connected layers. The final output of this subnet is passed to the binary cross entropy loss function

$$\mathcal{L}_C(I_i^{T1}) = -\log \frac{\exp(g(I_i^{T1}))}{\sum_k^K \exp(g(I_{i,k}^{T_1}))} y_i \quad (2)$$

where $g(I_i^{T_1})$ is the classification of the ith T1- image patch, $K = 2$, and y_i is the ground-truth label that indicates if a given patch contains lesion. Note that this subnet does not have a twin branch for LGE encoding and no weight sharing.

The total network (siamese plus classifier) aims to minimize the total loss function

$$\mathcal{L}_T(I_i^{T1}, I_j^{LGE}) = \mathcal{L}_P(I_i^{T1}, I_j^{LGE}) + \mathcal{L}_C(I_i^{T1}) \quad (3)$$

With this approach, the network will produce a vector from T1, that encodes features common to LGE. Thus, containing enough information for scar tissue classification.

3. EXPERIMENTAL SETUP

In our experimental setup, we introduced two additional variations to the architecture described in Sec. 2.2, the three models are as follows: (i) *full weight sharing*, (ii) *partial weight sharing*, and (iii) *decoupled* architectures.

In (i), all layers with learnable parameters in each branch of the siamese subnet share their weights, as illustrated in orange double arrows in Fig. 1 (a), referred to as *siamese* network. In (ii), only half of the learnable layers have weight sharing, shown in yellow arrows in Fig. 1 (a), and referred to as *partial* network. In (iii) the network has no shared weights at all, the *decoupled* network. These three networks were compared against a naive approach: *pure classifier*, implemented by taking the *siamese* network and dropping the LGE branch, and thus, dropping the $\mathcal{L}_P$ term from Eq. 3.

All networks were trained on mini-batches of size 32, for 300 epochs and learning rate $lr = 1e^{-4}$.

4. RESULTS AND DISCUSSION

The classification accuracy is presented in Tab. 1. The naive approach had zero predictive power and classified all samples as positive due to overfitting. This increases our confidence that the LGE is adding valuable information to the training process. We initially hypothesized that the siamese network with full weight sharing would be the best approach to extract a feature vector for lesion classification. However, we experimentally observed that performance increased with less weight sharing and that the fully decoupled network achieved the best balanced accuracy. This finding is aligned to [16], where it is argued that the decoupled network provides the best results when the inputs come from different sensors, while introducing weight sharing benefits inputs from the same modality. While the T1 map and LGE are technically acquired with the same sensor, the MRI scanner, they are still separate modalities that present very different contrasts, and the siamese network used in our work was originally intended to pair grayscale images of signatures that have mostly very similar contrasts. This could explain why decoupling improved performance. This means that, with weight sharing, the network is constrained to learn features that are common to both T1 and LGE. On the other hand, having a network without weight sharing allows to capture distinct features that are specific to each of the modalities T1 and LGE.

5. CONCLUSION AND FUTURE WORK

We presented a siamese network framework to classify lesion in T1 map patches, by taking advantage of the information provided by LGE during training. Our findings showed that the siamese approach outperforms the naive classifier and, by experimenting with variations of the architecture, we found that decoupling the weights improved performance, possibly due to the dual modality nature of the problem, which benefits from fully independent data streams. While the data sample used in this study could be considered small, the achieved classification accuracies motivate further work in this direction, for this type of problem. Future works should collect a larger set of data and expand the proposed methodology to integrate lesion segmentation besides the classification. This will requires larger training dataset, that can be tackled reducing the patch size so as to decrease the lesion location uncertainty.

Compliance with ethical standards: This research study was conducted retrospectively using human subject data.

Acknowledgments: This work was supported by the Portuguese Science Foundation through grants PTDC/EMD-EMD/29686/2017 and UIDB/50009/2020, and Programa Lisboa 2020 reference LISBOA-01-0145-FEDER-029686.

6. REFERENCES

[1] Raymond J Kim, Edwin Wu, Allen Rafael, Enn-Ling Chen, Michele A Parker, Orlando Simonetti, Francis J

Klocke, Robert O Bonow, and Robert M Judd, "The use of contrast-enhanced magnetic resonance imaging to identify reversible myocardial dysfunction," *New England Journal of Medicine*, vol. 343, no. 20, pp. 1445–1453, 2000.

[2] Hiroshi Satoh, Makoto Sano, Kenichiro Suwa, Takeji Saitoh, Mamoru Nobuhara, Masao Saotome, Tsuyoshi Urushida, Hideki Katoh, and Hideharu Hayashi, "Distribution of late gadolinium enhancement in various types of cardiomyopathies: Significance in differential diagnosis, clinical features and prognosis," *World journal of cardiology*, vol. 6, no. 7, pp. 585, 2014.

[3] Heiko Mahrholdt, Anja Wagner, Robert M Judd, Udo Sechtem, and Raymond J Kim, "Delayed enhancement cardiovascular magnetic resonance assessment of non-ischaemic cardiomyopathies," *European heart journal*, vol. 26, no. 15, pp. 1461–1474, 2005.

[4] Ippei Shimizu, Nobuo Iguchi, Hiroyuki Watanabe, Jun Umemura, Tetsuya Tobaru, Ryuta Asano, Kazuhiko Misu, Masatoshi Nagayama, Masaru Aikawa, Nobusada Funabashi, et al., "Delayed enhancement cardiovascular magnetic resonance as a novel technique to predict cardiac events in dilated cardiomyopathy patients," *International journal of cardiology*, vol. 142, no. 3, pp. 224–229, 2010.

[5] Ashkan Heshmatzadeh Behzadi, Yize Zhao, Zerwa Farooq, and Martin R Prince, "Immediate allergic reactions to gadolinium-based contrast agents: a systematic review and meta-analysis," *Radiology*, vol. 286, no. 2, pp. 471–482, 2018.

[6] Mark A Perazella and Roger A Rodby, "Gadolinium use in patients with kidney disease: a cause for concern," in *Seminars in dialysis*. Wiley Online Library, 2007, vol. 20, pp. 179–185.

[7] Dina Radenkovic, Sebastian Weingärtner, Lewis Ricketts, James C Moon, and Gabriella Captur, "T1 mapping in cardiac mri," *Heart Failure Reviews*, vol. 22, no. 4, pp. 415–430, 2017.

[8] Valentina O Puntmann, Elif Peker, Y Chandrashekhar, and Eike Nagel, "T1 mapping in characterizing myocardial disease: a comprehensive review," *Circulation research*, vol. 119, no. 2, pp. 277–299, 2016.

[9] Sumit Chopra, Raia Hadsell, and Yann LeCun, "Learning a similarity metric discriminatively, with application to face verification," in *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*. IEEE, 2005, vol. 1, pp. 539–546.

[10] Jane Bromley, Isabelle Guyon, Yann LeCun, Eduard Säckinger, and Roopak Shah, "Signature verification using a" siamese" time delay neural network," *Advances in neural information processing systems*, vol. 6, 1993.

[11] Ronen Rubinshtein, James F Glockner, Steve R Ommen, Philip A Araoz, Michael J Ackerman, Paul Sorajja, J Martijn Bos, A Jamil Tajik, Uma S Valeti, Rick A Nishimura, et al., "Characteristics and clinical significance of late gadolinium enhancement by contrast-enhanced magnetic resonance imaging in patients with hypertrophic cardiomyopathy," *Circulation: Heart Failure*, vol. 3, no. 1, pp. 51–58, 2010.

[12] John J Green, Jeffery S Berger, Christopher M Kramer, and Michael Salerno, "Prognostic value of late gadolinium enhancement in clinical outcomes for hypertrophic cardiomyopathy," *JACC: Cardiovascular Imaging*, vol. 5, no. 4, pp. 370–377, 2012.

[13] Jing Xu, Baiyan Zhuang, Arlene Sirajuddin, Shuang Li, Jinghan Huang, Gang Yin, Lei Song, Yong Jiang, Shihua Zhao, and Minjie Lu, "Mri t1 mapping in hypertrophic cardiomyopathy: evaluation in patients without late gadolinium enhancement and hemodynamic obstruction," *Radiology*, vol. 294, no. 2, pp. 275, 2020.

[14] Mateusz Spiewak, Lukasz A Malek, Jolanta Misko, Lidia Chojnowska, Barbara Milosz, Mariusz Klopotowski, Joanna Petryka, Maciej Dabrowski, Cezary Kepka, and Witold Ruzyllo, "Comparison of different quantification methods of late gadolinium enhancement in patients with hypertrophic cardiomyopathy," *European journal of radiology*, vol. 74, no. 3, pp. e149–e153, 2010.

[15] Sounak Dey, Anjan Dutta, J Ignacio Toledo, Suman K Ghosh, Josep Lladós, and Umapada Pal, "Signet: Convolutional siamese network for writer independent offline signature verification," *arXiv preprint arXiv:1707.02131*, 2017.

[16] Lloyd H Hughes, Michael Schmitt, Lichao Mou, Yuanyuan Wang, and Xiao Xiang Zhu, "Identifying corresponding patches in sar and optical images with a pseudo-siamese cnn," *IEEE Geoscience and Remote Sensing Letters*, vol. 15, no. 5, pp. 784–788, 2018.