Semi-supervised Learning of Cardiac MRI using Image Registration

Carles Garcia-Cabrera[1], Kathleen M. Curran[2], Noel E. O'Connor[1], and Kevin McGuinness[1]

[1]*Dublin City University*

[2]*University College Dublin*

Abstract

In this work, we propose a method to aid the 2-D segmentation of short-axis cardiac MRI. In particular, the deformation fields obtained during the registration are used to propagate the labels to all time frames, resulting in a weakly supervised segmentation approach that benefits from the features in unlabelled volumes along with the annotated data. Experimental results over the M&Ms datasets show that the addition of the synthetically obtained labels to the original dataset yields promising results in the performance and improves the capability of the network to generalise to scanners from different vendors.

Keywords: Cardiac MRI, Image Segmentation, Semi-Supervised Learning, Image Registration, Medical Imaging.

1 Introduction

Cardiac image segmentation is an important first step for many approaches to quantitative analysis for cardiac diagnostic assessment. This process requires partitioning the image into a number of clinically meaningful regions such as left ventricle (LV), right ventricle (RV), or myocardium (MYO). Acquiring this information allows clinicians understand important features such as the ejection fraction and the volume that the heart is managing at different times. Those features are later used to determine if there is any possible pathology and how bad it is [Kawel-Boehm et al., 2015].

Data is a key challenge when trying to use off-the-shelf algorithms in this area, specifically the limited amount of annotated data available, but also its quality. Many researchers report struggling to achieve improved results with existing annotated data, especially when working with open datasets [Chen et al., 2020]. Difficulties range from low availability of data, to domain shift using data from a certain scanner vendors, to images from patients with rare conditions.

There is growing interest in the community in understanding how to transfer models that work well for specific scanners to unseen ones. In this work we address these challenges using the M&Ms challenge dataset, which we believe to be the representative dataset for this issue [Campello et al., 2020].

Convolutional neural networks (CNN) are the most common type of deep neural networks for image analysis and have advanced the state-of-the-art in many object segmentation tasks, including in the medical imaging domain. In particular, U-Net [Ronneberger et al., 2015] is the architecture with the best results over most of the challenges that the cardiac MRI currently faces [Chen et al., 2020]. Hence, we have selected it to run our experiments.

Recently, [Zhang et al., 2021] proposed a method where labels can be propagated using image registration in an unsupervised manner and those labels are used to enhance the process where a 3D U-Net learns features. Even though results ranked second in the M&Ms challenge [Campello et al., 2020], the study did not test those propositions in 2D networks, a key novelty of our work, nor explored different tools for the registration part.

Contributions: this work proposes a method to enhance the segmentation of short-axis cardiac MRI by synthetically labelling volumes without annotations. The proposed system works by registering the labelled volumes of the end-systolic (ES) and the end-diastolic (ED) time-frames and using the warping field over their ground truth. This helps the network to learn robust features from those volumes, which are important towards increasing the performance over different scanners. Our study extends recent works where the propagation of labels is used in a 3D U-Net and using a different registration tool. In contrast to that work, we investigate whether this technique is valid for 2D U-Net applications that can be used in less computationally powerful machines and expect to have better performances in anisotropic datasets.

2 Related Work

The clinical interest in ventricle segmentation has pushed the community towards improving the performance for this task. With the advent of deep learning, a succession of different approaches were investigated. One of the first, was the usage of a fully convolutional network [Tran, 2016]. From there, many works have improved the networks, increasing the learning capacity for segmentation [Khened et al., 2019]. In that regard, the low through-plane resolution and the motion artifacts limited the applicability of 3D networks [Baumgartner et al., 2017].

Annotated data is the foundation of fully supervised approaches but at the same time it is scarce and costly to obtain. Moreover, annotating medical images requires significant expertise and manual effort and even then can lead to noisy labels and/or biases. Additionally, there is an imbalance between the amount of available data regarding different source scanner, target pathology, or scanning parameters. All the above-mentioned challenges point strongly towards the need to design methods where the accuracy of the networks is maintained compared to fully supervised approach, even when there is significantly less data available.

A number of works in the literature proposed unsupervised or semi-supervised techniques to overcome the scarcity of labels. One approach was to use a scribble annotation, that consists of a set of quickly drawn labels, and recursively re-training the network using the output segmentation, including a conditional random field and an uncertainty estimator [Can et al., 2018]. The advances in the calculation of optical flow inspired the usage of these algorithms for motion estimation and this has been useful in multi-task approaches where a Siamese network performs the motion estimation and the segmentation simultaneously, exploiting the information contained in unlabelled data [Qin et al., 2018].

Voxelmorph [Balakrishnan et al., 2018] is an image registration tool for alignment and registration that can also model deformations. It has been used for atlas based registration [Dalca et al., 2019] and for probabilistic diffeomorphic registration [Dalca et al., 2018]. Given its proven reliability, we use it for the registration part of our study.

3 Method and Experiment

This section describes the proposed method, detailing the different steps and implementations of all the tools involved in the process. Distinguishing two first steps corresponding to the registration and label propagation, and two last parts corresponding to training and testing the two different models i.e. the original dataset and the one with the addition of the synthetic labels.

3.1 Data

For our experiments we used the M&Ms dataset, released during the Multi-Centre, Multi-Vendor and Multi-Disease Cardiac Image Segmentation Challenge. This dataset consists of 345 patients with hypertrophic and dilated cardiomyopathies and healthy subjects and it also permits the opportunity to measure performance over different scanners, and to evaluate the resilience of our approach towards different data sources. In particular the training set includes 150 cases (75 from vendor A and 75 from vendor B). The test set includes cases from two more vendors (C and D).

The CMR images have been segmented by experienced clinicians from the respective host institutions, including contours for the LV and RV blood pools, as well as for the left ventricular MYO.

3.2 Label Propagation

Our selected image registration technique was Voxelmorph [Balakrishnan et al., 2019], and given that the pre-trained models were trained with images from other tissues, we decided to train our own with our available data. We also included the ground truth segmentation labels in the training process to produce a model that later we used for label propagation. For this part of the process we used the standard configuration of Voxelmorph (Tensorflow version)[1].

After training the model we computed all the warping fields for the time-frame between ED and ES, always in an intra-subject way. With the warping fields we modified the ground truth for the ED time-frame, obtaining the synthetic labels for out target time-frames.

Lastly, we join in a single dataset all the volumes with annotations, including the the original and the synthetic.

3.3 Segmentation

In our approach, we used a 2D U-Net to segment the end-diastolic and end-systolic volumes in the target dataset. In particular, our U-Net model used 32 feature channels with kernel size equal to three in the first level of the convolutional layers, where batch normalization was also applied and the activation function was ReLU, followed by a 2×2 max-pooling layer. The final 1×1 convolution is set with four channels that match the four target regions: background, right ventricle, left ventricle, and myocardium.

To train the model, we used an Adam optimizer (learning rate = 10^{-3}) with a plateau learning rate scheduler, and as a loss function we calculated the cross-entropy loss. In addition to this, we computed the Dice score for all the regions in each epoch for the validation set.

4 Results

Table 1 shows the results obtained for both datasets, including regions and vendors (M&Ms). The proposed method performs 2.6% better than the baseline, where the region that benefits the most from this change is the corresponding to the myocardium. When we compare the difference in the performance of both models in the different scanners, we see how the performance of unseen scanners C and D improves in the model resulting from the proposed method while it sees its performance almost untouched for the scanners present in the training set (A and B).

	Regions			**Vendors**				**Total**
	LV	**Myo**	**RV**	**A**	**B**	**C**	**D**	**Dice**
Baseline (M&Ms)	0.475	0.386	0.394	0.553	0.583	0.419	0.277	0.418
Proposed (M&Ms)	0.504	0.430	0.398	0.548	0.573	0.443	0.338	0.444

Table 1: Results (dice scores) on the M&Ms datasets. Higher is better.

5 Conclusions

Experimental results using synthetic labels generated from VoxelMorph showed an improvement over the baseline, demonstrating the potential for using diffeomorphic image registration as a label propagation technique. Future work will investigate using this technique with a stronger baseline. In particular, we plan to adopt nnUNet [Isensee et al., 2021], an automatic segmentation framework for medical images based on the U-net architectures, which has been demonstrated to give state-of-the-art results. Finally, we believe the results can be improved by leveraging research on unreliable or noisy labels (e.g. [Arazo et al., 2019]), weighting them differently to reflect the level of trust in the annotations, or post-processing the synthetic annotations. Future work will investigate this.

Acknowledgments

This work was funded by Science Foundation Ireland through the SFI Centre for Research Training in Machine Learning (18/CRT/6183).

[1] VoxelMorph `https://github.com/voxelmorph/voxelmorph`.

References

[Arazo et al., 2019] Arazo, E., Ortego, D., Albert, P., O'Connor, N., and McGuinness, K. (2019). Unsupervised label noise modeling and loss correction. In *International Conference on Machine Learning*, pages 312–321. PMLR.

[Balakrishnan et al., 2018] Balakrishnan, G., Zhao, A., Sabuncu, M. R., Dalca, A. V., and Guttag, J. (2018). An unsupervised learning model for deformable medical image registration. *Computer Vision and Pattern Recognition (CVPR)*.

[Balakrishnan et al., 2019] Balakrishnan, G., Zhao, A., Sabuncu, M. R., Guttag, J., and Dalca, A. V. (2019). Voxelmorph: A learning framework for deformable medical image registration. *IEEE Transactions on Medical Imaging*, 38(8):1788–1800.

[Baumgartner et al., 2017] Baumgartner, C. F., Koch, L. M., Pollefeys, M., and Konukoglu, E. (2017). An exploration of 2d and 3d deep learning techniques for cardiac mr image segmentation.

[Campello et al., 2020] Campello, V. M., Palomares, J. F. R., Guala, A., Marakas, M., Friedrich, M., and Lekadir, K. (2020). Multi-Centre, Multi-Vendor & Multi-Disease Cardiac Image Segmentation Challenge.

[Can et al., 2018] Can, Y. B., Chaitanya, K., Mustafa, B., Koch, L. M., Konukoglu, E., and Baumgartner, C. F. (2018). Learning to segment medical images with scribble-supervision alone.

[Chen et al., 2020] Chen, C., Qin, C., Qiu, H., Tarroni, G., Duan, J., Bai, W., and Rueckert, D. (2020). Deep learning for cardiac image segmentation: A review. *Frontiers in Cardiovascular Medicine*, 7.

[Dalca et al., 2018] Dalca, A. V., Balakrishnan, G., Guttag, J., and Sabuncu, M. R. (2018). Unsupervised learning for fast probabilistic diffeomorphic registration. *Lecture Notes in Computer Science*, page 729–738.

[Dalca et al., 2019] Dalca, A. V., Rakic, M., Guttag, J., and Sabuncu, M. R. (2019). Learning conditional deformable templates with convolutional networks.

[Isensee et al., 2021] Isensee, F., Jaeger, P. F., Kohl, S. A. A., Petersen, J., and Maier-Hein, K. H. (2021). nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. *Nature Methods*, 18(2):203–211.

[Kawel-Boehm et al., 2015] Kawel-Boehm, N., Maceira, A., Valsangiacomo-Buechel, E. R., Vogel-Claussen, J., Turkbey, E. B., Williams, R., Plein, S., Tee, M., Eng, J., and Bluemke, D. A. (2015). Normal values for cardiovascular magnetic resonance in adults and children. *Journal of Cardiovascular Magnetic Resonance*, 17(1):29.

[Khened et al., 2019] Khened, M., Kollerathu, V. A., and Krishnamurthi, G. (2019). Fully convolutional multi-scale residual densenets for cardiac segmentation and automated cardiac diagnosis using ensemble of classifiers. *Medical Image Analysis*, 51:21 – 45.

[Qin et al., 2018] Qin, C., Bai, W., Schlemper, J., Petersen, S. E., Piechnik, S. K., Neubauer, S., and Rueckert, D. (2018). Joint learning of motion estimation and segmentation for cardiac mr image sequences.

[Ronneberger et al., 2015] Ronneberger, O., Fischer, P., and Brox, T. (2015). U-Net: Convolutional networks for biomedical image segmentation.

[Tran, 2016] Tran, P. V. (2016). A fully convolutional neural network for cardiac segmentation in short-axis MRI.

[Zhang et al., 2021] Zhang, Y., Yang, J., Hou, F., Liu, Y., Wang, Y., Tian, J., Zhong, C., Zhang, Y., and He, Z. (2021). Semi-supervised cardiac image segmentation via label propagation and style transfer. In *Statistical Atlases and Computational Models of the Heart. M&Ms and EMIDEC Challenges*, pages 219–227. Springer.