Introducing Anatomical Knowledge to A Deep Learning Approach for Segmentation of Cardiac Magnetic Resonance Images

Jinming Duan[1,2], Jo Schlemper[1], Wenjia Bai[1], Timothy J W Dawes[2], Ghalib Bello[2], Carlo Biffi[1,2], Antonio De Marvao[2], Declan O'Regan[2], and Daniel Rueckert[1]

[1]Biomedical Image Analysis Group, Imperial College London, London, UK
[2]MRC London Institute of Medical Sciences, Imperial College London, London, UK
[1]`j.duan@imperial.ac.uk`

Abstract

Deep learning approaches such as convolutional neural networks (CNN) have achieved state-of-the-art performance in cardiac magnetic resonance (CMR) image segmentation. However, it is non-trivial to introduce shape prior knowledge to CNN-based approaches. In this paper, we employ a CNN-based method combined with image registration to develop and evaluate a shape-based bi-ventricular segmentation tool from short-axis CMR images. The method first employs a fully convolutional network (FCN) to learn the segmentation task from manually labelled ground truth CMR images generated in a low long-axis resolution. The FCN segmentation results are then used to perform a non-rigid registration using multiple high-resolution atlases, allowing the shape constraint to be inferred. This approach produces accurate, high-resolution and automatically smooth segmentation results from input images with low long-axis resolution, thus retaining clinically important global anatomical features.

1 Introduction

Cardiac magnetic resonance (CMR) imaging is a gold standard method for assessing cardiac anatomy and function. CMR imaging techniques, together with semi-automated or automated CMR segmentation algorithms, have shown a great impact on studying, understanding and diagnosing cardiovascular diseases. However, due to limitations of imaging modality, CMR images contain several artefacts, including intensity inhomogeneity, respiratory motion, large slice thickness, etc. Building an accurate, motion-free, automatically meaningful bi-ventricular segmentation model therefore remains an open problem. Recently, neural network-based methods have achieved state-of-the-art segmentation performance in the CMR domain. On the other hand, incorporating anatomical shape prior knowledge into image segmentation algorithms has proven very useful to obtain highly accurate and plausible results, which facilitates further statistical shape analysis. As such, in this paper we propose a novel deep learning-based method combined with the anatomical bi-ventricular shape constraint for CMR image segmentation. The method has two steps. First, a FCN model is trained to effectively learn useful segmentation features from manually labelled 2D CMR images which have a low resolution in the long-axis and potential motion artefact. Second, multiple 3D high-resolution CMR atlases are propagated onto the 2D CMR segmentation to form a high-resolution segmentation. This step effectively induces a hard anatomical shape constraint. As such, the proposed method can produce accurate, high-resolution and automatically smooth segmentation results from input images with low long-axis resolution, thus retaining clinically important global anatomical features.

1st Conference on Medical Imaging with Deep Learning (MIDL 2018), Amsterdam, The Netherlands.

2 Methodology

Convolutional neural network: Let us formulate the deep learning problem as follows: we denote the input training data set by $S = \{(U_p, R_p), p = 1, ..., N\}$, where sample $U_p = \{u_j^p, j = 1, ..., |U_p|\}$ is the raw CMR input image, $R_p = \{r_j^p, j = 1, ..., |R_p|\}$, $r_j^p \in \{1, ..., n\}$ represents the ground truth labels ($n = 5$ regions) for image U_p. We denote all network layer parameters as $\mathbf{W}$. In a supervised setting, we propose to minimise the following objective function via standard (back-propagation) stochastic gradient descent (SGD): $\mathbf{W}^* = \mathrm{argmin}(L_S(\mathbf{W}) + \alpha L_D(\mathbf{W}))$, where $L_S(\mathbf{W})$ is the cross-entropy loss associated with softmax probability, while $L_D(\mathbf{W})$ is the Dice index loss. The weight α balances the two losses. By minimising the objective function, the network is able to obtain predictions for the 5 classes simultaneously. In our image-to-image training, the loss functions are computed over all pixels in all training images U and label maps R.

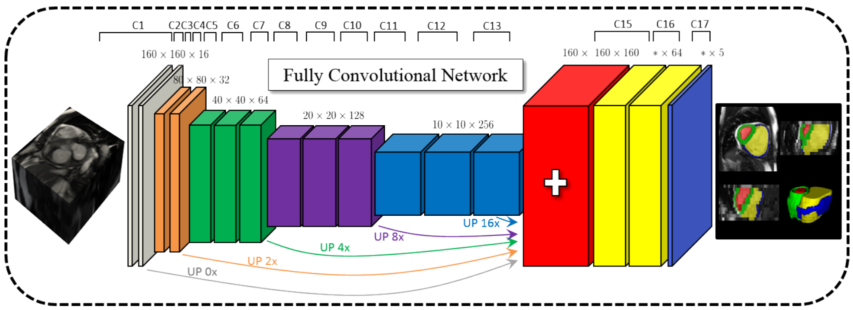

Figure 1: The architecture of a fully convolutional network with 17 convolutional layers. The network takes the CMR image as input, applies a branch of convolutions, learns image features from fine to coarse levels, concatenates multi-scale features and finally segments the image into 5 disjoint regions.

In Fig 1, we show the proposed network architecture for automatic CMR segmentation, which is a fully convolutional network (FCN). It is adapted from [1] and similar to the U-net architecture [2]. In contrast to U-net, our network predicts feature maps in the original resolution. Batch-normalisation (BN) is used after each convolutional layer, and before a rectified linear unit (ReLU) activation. The last layer is followed by the channel-wise softmax function. In the FCN, input images have pixel dimensions of 160×160. Every layer whose label is prefixed with 'C' performs the operation: convolution $\rightarrow$ BN $\rightarrow$ ReLU, except C17. The (filter size/stride) is (3×3/1) for layers from C1 to C16, excluding layers C3, C5, C8 and C11 which are (3×3/2). The arrows represent (3×3/1) convolutional layers (C14a−e) followed by a transpose convolutional (up) layer with a factor necessary to achieve feature map volumes with size $160 \times 160 \times 32$, all of which are concatenated into the red feature map volume. Finally, C17 applies a (1×1/1) convolution with a softmax activation, producing the blue feature map volume with a depth 5, corresponding to 5 segmented regions of an input image.

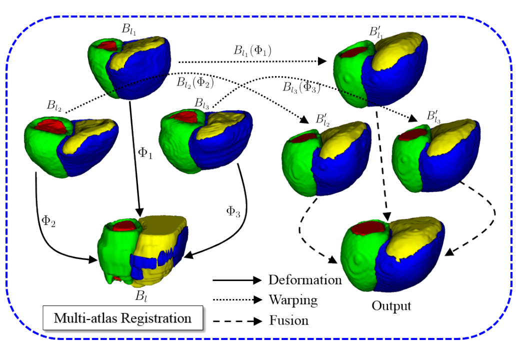

Figure 2: Incorporation of shape constraints by multi-atlas registration. B_{l_1}, B_{l_2} and B_{l_3} are high-resolution bi-ventricular atlas shape models. They are warped to B'_{l_1}, B'_{l_2} and B'_{l_3} using the transform between the FCN result B_l and each atlas. The warped results are finally fused together to generate a smooth high-resolution output.

Multi-atlas registration: As Fig 1 shows, the segmentation produced by FCN is influenced by respiratory motion artefact. Moreover, as the CMR images are low resolution in the long axis, the 3D segmentation model is not smooth. By incorporating shape constraints, these drawbacks can be overcome. Specifically, we perform multi-atlas image registration to induce such shape constraints. The objective function for registration is defined as: $\Phi_n^* = \arg\min S\left(B_l, B_{l_n}(\Phi_n)\right)$ where Φ_n is the transformation between the FCN result B_l and the nth atlas B_{l_n}. It is modelled by a free-form deformation based on B-splines [3]. S denotes the label consistency between two binary labels. Note that the atlases we used here are 3D bi-ventricular binary labels derived from motion-free CMR images with a high long-axis resolution. Gradient descent is used to minimise the objective function. After the optimal Φ^* is found, we transform the atlases to the FCN segmentation space and perform label fusion in that space to generate an accurate, motion-free and smooth bi-ventricular model. See Fig 2 for the illustration of the process.

3 Experimental Results

Experiments were performed using short-axis CMR images from 430 patients with pulmonary hypertension. For each patient 10 to 16 short-axis slices were acquired roughly covering the whole heart. Each short-axis image has resolution of $1.5 \times 1.5 \times 8.0 \text{ mm}^3$. Due to the large slice thickness of the short-axis slices and the inter-slice shift caused by respiratory motion, we train the FCN in a 2D fashion and apply the trained FCN to segment each slice separately. The segmentation is then passed to the multi-atlas non-rigid registration to form a high-resolution smooth shape model. The high-resolution atlases are derived from healthy subjects, scanned at Hammersmith Hospital, Imperial College London. The image resolution acquired is improved to $1.25 \times 1.25 \times 2.0 \text{ mm}^3$. The reduction of the slice thickness enables to model the cardiac shape in greater detail. In Fig 3, we visually compare the proposed method with the vanilla deep learning method without shape prior knowledge. As evident, the proposed method gives better 3D phenotype results free of respiratory motion artefact. This is due to the application of shape constraints. Our method thus outperforms the vanilla FCN in this regard.

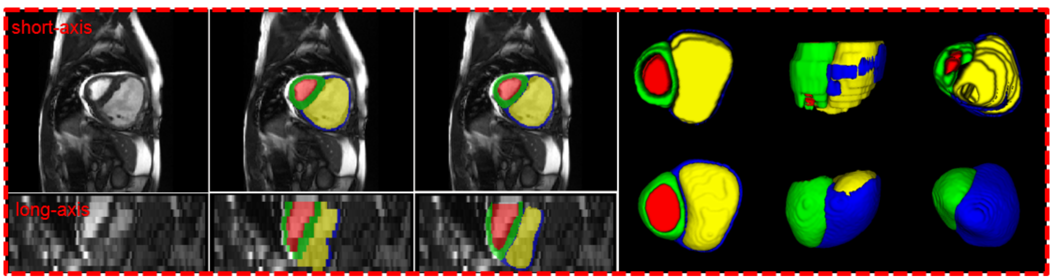

Figure 3: Visual comparison of segmentation results from the vanilla FCN and the proposed method. 1st column: original short- and long-axis CMR images. 2nd column: vanilla FCN results. 3rd column: results by the proposed method. Last column: FCN results (top) and our results (bottom). The 5 segmented regions are respectively right ventricular (RV) blood pool (yellow), left ventricular (LV) blood pool (red), RV wall (blue), LV wall (green) and background.

4 Conclusion

In this paper, we developed a CNN-based method combined with image registration for shape-based bi-ventricular segmentation of short-axis CMR images. The method first employs a novel fully convolutional network (FCN) to segment the images at a low resolution level. Based on the FCN results, the method then performs the non-rigid registration by using multiple high-resolution atlases, thereby imposing shape constraints. Experiments show that the method has capability of producing smooth high-resolution segmentation results that follow the global anatomical properties of the underlying anatomy, even though the input images are low resolution in the long-axis. Future work will focus on the development of a method to quantify the segmentation results produced by our shape-based method and using the results for statistical shape analysis for clinical use.

References

[1] Wenjia Bai, Matthew Sinclair, Giacomo Tarroni, Ozan Oktay, Martin Rajchl, Ghislain Vaillant, Aaron M Lee, Nay Aung, Elena Lukaschuk, Mihir M Sanghvi, et al. Human-level cmr image analysis with deep fully convolutional networks. *arXiv preprint arXiv:1710.09289*, 2017.

[2] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, pages 234–241. Springer, 2015.

[3] Daniel Rueckert, Luke I Sonoda, Carmel Hayes, Derek LG Hill, Martin O Leach, and David J Hawkes. Nonrigid registration using free-form deformations: application to breast mr images. *IEEE transactions on medical imaging*, 18(8):712–721, 1999.