Ventricular segmentation and quantitative assessment in cardiac MR using convolutional neural networks

Joshua V. Stough[a], Joseph DiPalma[a], Zilin Ma[a],
Brandon K. Fornwalt[b], and Christopher M. Haggerty[b]

[a]Computer Science, Bucknell University, Lewisburg, PA;
[b]Imaging Science and Innovation, Geisinger Health System, Danville, PA

ABSTRACT

Segmentation of heart substructures in cardiac magnetic resonance (CMR) is an important step in the quantitative assessment of the impact of cardiovascular disease. Manual delineation of these structures, over many patients and multiple time phases, is time consuming and prone to human error and fatigue. In this work we use a deep fully convolutional neural network architecture to automatically segment heart substructures in CMR, achieving state of the art results on a recent benchmark dataset. We further apply our process to a much larger study of CMR subjects, automatically segmenting both left and right ventricular endocardia (LV, RV) with full thirty-phase time resolution, and LV epicardium (Epi) at end-diastole. We validate our automatically obtained results against manual delineations using Dice overlap and Hausdorff distance, as well as Bland-Altman limits of agreement on the derived blood volumes, ejection fraction, and LV mass. We obtain median Dice overlaps of 0.97, 0.94, and 0.97 on the three structures respectively, and further find small biases and narrow limits of agreement between the two assessments (manual, automatic) of volumes and mass. Our results show promise for the fully automated analysis of the CMR data stream in the near future.

Keywords: Neural networks, Cardiac MR, Segmentation

1. INTRODUCTION

Cardiovascular diseases are a major cause of morbidity and mortality globally.[1] Short-axis cardiac magnetic resonance (CMR) imaging, though common in clinical practice, is underutilized for its quantitative power at large scale, such as in population studies in the pursuit of precision medicine. Quantitative analysis of CMR by humans may require time consuming detailed delineation of the relevant heart substructures over time and over multiple slices within the 3D volume (see Fig.1). Exploiting this important data stream thus requires automatic processing, an open problem for many around the world.[2,3]

Neural network learning has recently shown considerable promise on this front. Zhen et al.[4] use unsupervised representation learning and regression forests for the direct estimation of ventricular volume. Avendi et al.[5] combine convolutional neural network learning and a deformable shape model to segment the left ventricle (LV) for volume and ejection fraction estimation. Emad et al.[6] segment the LV using a multi-scale convolutional neural network. Recently, Romageura et al.[7] proposed a deep fully convolutional neural network (CNN) for LV segmentation based on the work of Tran.[8]

Here we extend the work of Romageura et al., using the TensorFlow[9] deep learning library to implement the neural network architecture, and automatically segment the LV endocardium. We achieve comparable state of the art results on the Sunnybrook CMR dataset.[10] We then apply the same basic neural network architecture to a much larger study of 51 healthy volunteers recruited from Geisinger Health System. In addition to LV endocardium, we segment the LV *epicardium* (Epi) and right ventricle (RV) endocardium. We validate against manual delineations on Dice overlap and Hausdorff distance, as well as derived blood volumes, ejection fraction, and LV mass.

Corresponding author: joshua.stough@bucknell.edu

This work is supported through the Bucknell Geisinger Research Initiative (BGRI) and the Bucknell Program for Undergraduate Research (PUR).

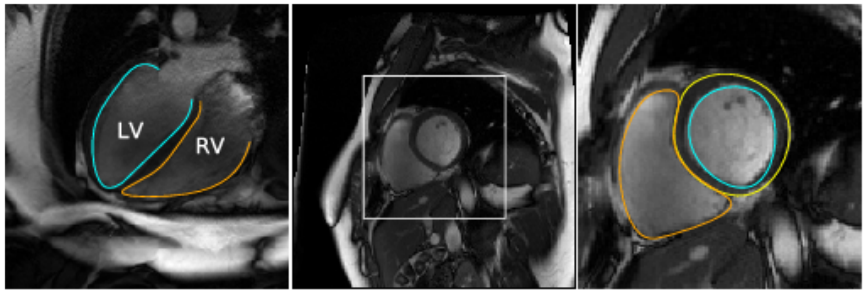

Figure 1. Left: Four Chamber (off axis) view with left and right ventricles (LV, RV) highlighted. Middle and right: example short axis CMR, and zoomed area showing the left ventricular endocardium (LV, cyan) and epicardium (Epi, yellow), and right ventricular endocardium (RV, orange) boundary contours.

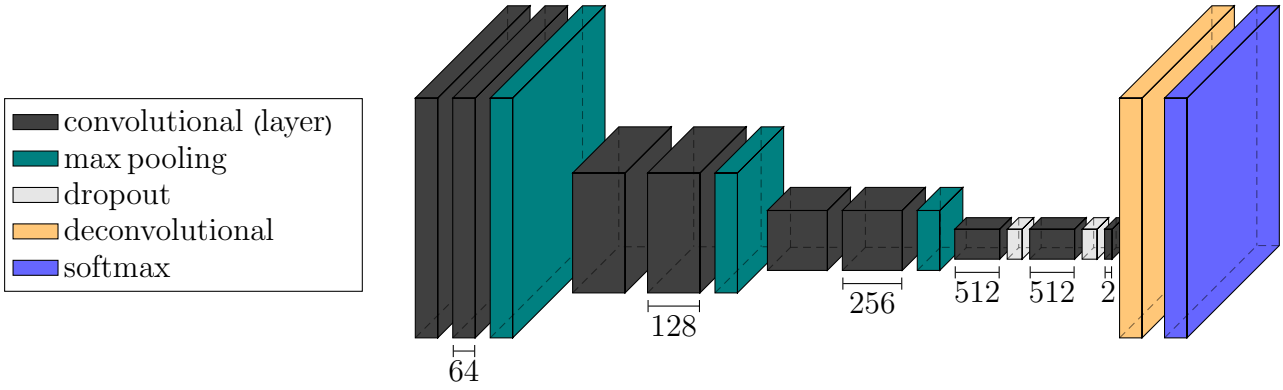

Figure 2. Deep fully convolutional neural network architecture.[7] Inference takes place left-to-right, as learned image filtering operations provide larger-scale semantic information about the input image. Learning takes place right-to-left through backpropagation, where the learned parameters are modified according to gradient descent in order to minimize the inferred error in the training cases.[11]

2. MATERIALS AND METHODS

Using the TensorFlow[9] deep learning library, we implemented the CNN architecture recently proposed by Romageura et al.,[7] which learns to segment 2D short-axis slices of the typical CMR study (see Fig. 1). As with other recent CNN work,[5,8] this network consists of successive stages each of which incorporates multiple convolutional layers for feature learning followed by a max pooling layer for translational invariance and to reduce network dimensionality (Fig. 2). At the end of three of these stages, our network outputs a two-channel tensor representing foreground and background measures at one-eighth of the original image resolution. This tensor is then upsampled through a learned deconvolutional layer to the original image resolution, finally followed by a softmax layer for differentiability and numerical considerations in the backpropagation learning process.[11]

We validated the established network architecture on the Sunnybrook CMR dataset.[10] Given a five-fold cross validation experimental setup, we obtained an average Dice overlap of 0.920 ($\pm$ 0.122) on LV, in line with previous state of the art results on this dataset.[5–7,12]

3. EXPERIMENTAL RESULTS

Our data consist of 51 healthy subjects recruited from the Geisinger Health System following an Institutional Review Board approved protocol. All subjects provided informed consent. CMR evaluation was performed on a 3 Tesla Siemens scanner, and included 8–10 short-axis slices spanning the ventricles with 1.17mm isotropic in-plane resolution and 8mm slice thickness. LV and RV endocardial contours were manually delineated on each slice. Additionally, contour propagation through time (30 phases per slice) was performed using an optical flow algorithm, with manual corrections made as needed. Finally, the LV epicardium (Epi) was manually delineated

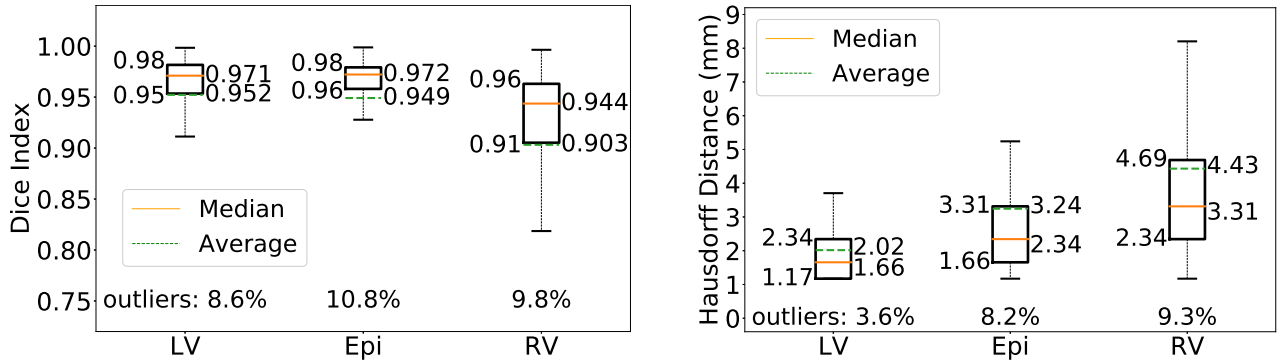

Figure 3. Box-Whisker plots showing the distribution of Dice overlap (left) and Hausdorff distance (right) for each of the three heart substructures segmented. Numbers to the left of a box are (top to bottom) the 75^{th} and 25^{th} percentiles.

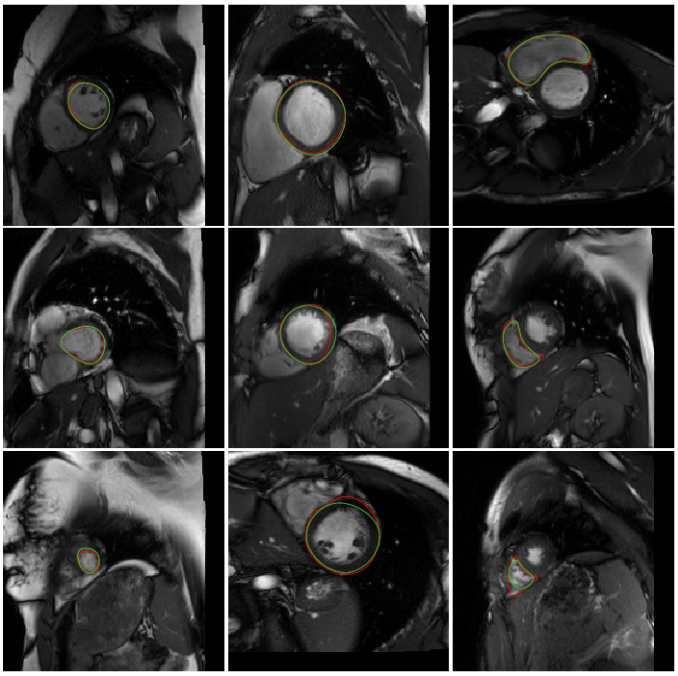

Figure 4. Good, median, and bad example segmentations showing manual (red) and automatically obtained (green) contours. Left to right: LV, Epi, RV. Top to bottom: rows are sampled from the 90^{th}, 50^{th}, and 10^{th} percentiles in Dice overlap, approximately 0.98 (0.97 RV), 0.97 (0.94) and 0.91 (0.81) respectively.

at end-diastole. In all, our data thus include approximately 15000 segmented short-axis images for LV, 14000 for RV, and 528 for Epi.

We perform a five-fold cross validation study, training on $4/5^{\text{th}}$ of subjects and validating on the left out $1/5^{\text{th}}$. The neural network architecture (see Sec. 2) is trained separately for each region of interest. We validate our automatically obtained segmentations against manual delineations using Dice overlap and Hausdorff distance, summarized over all validation sets in Figure 3. Our average Dice overlaps of 0.95 (LV, Epi) and 0.90 (RV) and limited outliers compare favorably against recent work.[12–14] Figure 4 shows example segmentations over the range of Dice scores for each region of interest.

We further used our automatic segmentations to derive clinically relevant quantitative measures for each patient, including the LV and RV blood volumes at end-diastolic and end-systolic phases (EDV, ESV, respectively), LV ejection fraction (LVEF), and LV mass. We compared these quantitative measures against those derived from the manual delineations using the Bland-Altman method, which plots the discrepancy of the mea-

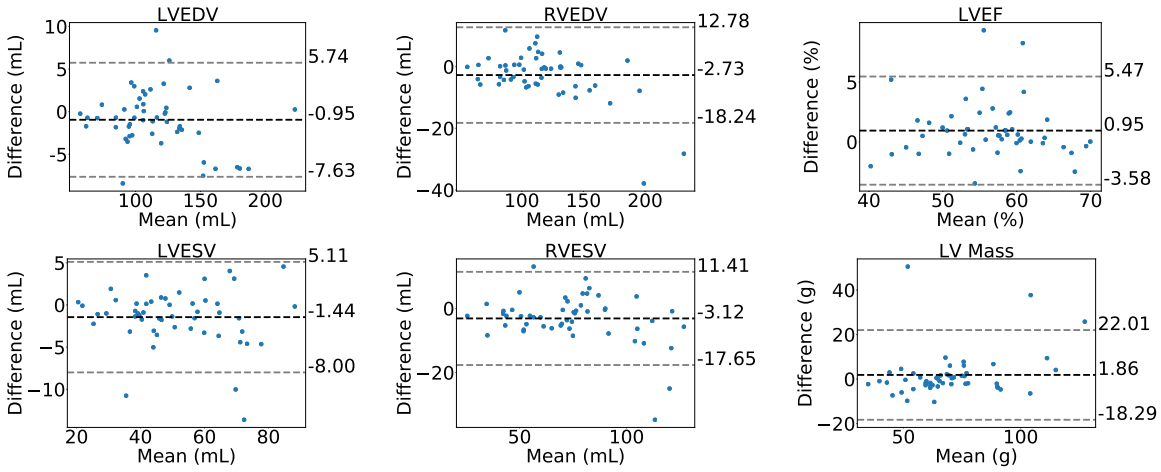

Figure 5. Bland-Altman plots comparing the automatically-obtained heart structure to the manual annotations. Left column: left ventricle end-diastole (top) and end-systole (bottom) volume. Middle column: right ventricle corresponding volumes. Right column: Ejection fraction (top) and left ventricle mass (bottom). Results indicate that the presented fully automatic scheme obtains results that are interchangeable with those of the manual rater.[16,17]

sures against the average to quantify the mean bias and 95% limits of agreement (Fig. 5)[15] Given prior studies of inter-rater variability in CMR LV segmentation,[16] our automatically obtained segmentations are comparable with the manual in derived LVEF and LV[EDV,ESV].

4. DISCUSSION

Our work is the first to use convolutional neural nets to segment both LV and RV heart substructures over all time phases of short-axis CMR. With the additional epicardium substructure, our state of the art results indicate the automated analysis of CMR in the near future. However, though simple quantifiers such as ejection fraction and volumes may be tractable to our processing now, more detailed quantitative assessments will require addressing the outlying failure cases of our automatic method. These cases are principally slices in the apical region and in images with artifacts. Shape priors[5] and the re-incorporation of spatial and temporal coherence should help alleviate these issues.

In the short term we will add pathology cases to our initial study. Longer term, we will consider our automated analyses in conjunction with other electronic health record data and genetic data for large-scale population studies such as in Bai et al.,[18] though using the less than ideal clinically-acquired historical CMR data.

The full length submission of this work will include more general prior work in Sec. 1, mathematical background in Sec. 2, computational aspects of validation in Sec. 3, and unabbreviated references.

REFERENCES

[1] Mendis, S. et al., [*Global Status Report on Noncommunicable Diseases 2014*], Geneva: World Health Organization (2014). who.int.

[2] Petitjean, C. and Dacher, J., "A review of segmentation methods in short axis cardiac mr images," *Medical Image Analysis* **15**(2), 169–184 (2011). `https://www.ncbi.nlm.nih.gov/pubmed/21216179`.

[3] Peng, P. et al., "A review of heart chamber segmentation for structural and functional analysis using cardiac magnetic resonance imaging," *Magn Reson Mater Phy* **29**(2), 155–95 (2016). `https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4830888/`.

[4] Zhen, X. et al., "Multi-scale deep networks and regression forests for direct bi-ventricular volume estimation," *Medical Image Analysis* **30**, 120–129 (2016). `http://dx.doi.org/10.1016/j.media.2015.07.003`.

[5] Avendi, M. et al., "A combined deep-learning and deformable-model approach to fully automatic segmentation of the left ventricle in cardiac mri," *Medical Image Analysis* **30**, 108 – 119 (2016). https://doi.org/10.1016/j.media.2016.01.005.

[6] Emad, O. et al., "Automatic localization of the left ventricle in cardiac mri images using deep learning," in [*Engineering in Medicine and Biology Society (EMBC)*], (2015). https://doi.org/10.1109/EMBC.2015.7318454.

[7] Romaguera, L. V. et al., "Left ventricle segmentation in cardiac mri images using fully convolutional neural networks," in [*Proc. SPIE-Medical Imaging*], (2017). http://dx.doi.org/10.1117/12.2253901.

[8] Tran, P. V., "A fully convolutional neural network for cardiac segmentation in short-axis mri," tech. rep., Cornell University (2017). https://arxiv.org/abs/1604.00494.

[9] Abadi, M. et al., "TensorFlow: Large-scale machine learning on heterogeneous systems," (2015). http://tensorflow.org/.

[10] Radau, P. et al., "Evaluation framework for algorithms segmenting short axis cardiac mri.," *The MIDAS Journal - Cardiac MR Left Ventricle Segmentation Challenge* (07 2009). http://hdl.handle.net/10380/3070.

[11] Goodfellow, I. et al., [*Deep Learning*], MIT Press (2016). http://www.deeplearningbook.org.

[12] Hu, H. et al., "Hybrid segmentation of left ventricle in cardiac mri using gaussian-mixture model and region restricted dynamic programming," *Magnetic Resonance Imaging* **31**(4), 575 – 584 (2013). https://doi.org/10.1016/j.mri.2012.10.004.

[13] Lu, Y.-L. et al., "Automatic functional analysis of left ventricle in cardiac cine mri," *Quantitative Imaging in Medicine and Surgery* **3**(4) (2013). http://qims.amegroups.com/article/view/2552.

[14] Suinesiaputra, A. et al., "A collaborative resource to build consensus for automated left ventricular segmentation of cardiac MR images," *Medical Image Analysis* **18**(1), 50–62 (2014). https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3840080/.

[15] Bland, J. M. and Altman, D. G., "Comparing methods of measurement: why plotting difference against standard method is misleading," *The lancet* **346**(8982), 1085–1087 (1995). http://dx.doi.org/10.1016/j.ijnurstu.2009.10.001.

[16] Jing, L. et al., "Cardiac remodeling and dysfunction in childhood obesity: a cardiovascular magnetic resonance study," *Journal of Cardiovascular Magnetic Resonance* **18**(1) (2016). https://dx.doi.org/10.1186%2Fs12968-016-0247-0.

[17] Giavarina, D., "Understanding bland altman analysis," *Biochemia Medica* **25**(2), 141–151 (2015). https://dx.doi.org/10.11613%2FBM.2015.015.

[18] Bai, W. et al., "A bi-ventricular cardiac atlas built from 1000+ high resolution mr images of healthy subjects and an analysis of shape and motion," *Medical Image Analysis* **26**, 133–145 (2015). https://doi.org/10.1016/j.media.2015.08.009.