Automated Torso Contour Extraction from Clinical Cardiac MR Slices for 3D Torso Reconstruction*

Hannah J. Smith, Abhirup Banerjee, Robin P. Choudhury and Vicente Grau

Abstract— **Whilst the electrocardiogram (ECG) is an essential tool for diagnosing cardiac electrical abnormalities, its characteristics are dependent on anatomical variability. Specifically variation in torso geometry affects relative positions of the leads with respect to the heart. We propose a novel pipeline that uses standard cardiac magnetic resonance images to reconstruct the torso and heart, and recreate the ECG considering torso and cardiac anatomy. This requires automated extraction of the torso contours. Our method combines an initial u-net segmenter with a second network that refines contours and removes spurious segments. The networks were evaluated on a cross validation study dataset and an independent test set. The use of two-channel input, including both original image and initial segmentation, in the refinement network significantly improved performance on the independent test set, reducing the Hausdorff distance from 9.1 pixels to 4.3 pixels and increasing Dice coefficient from 0.75 to 0.93.**

Clinical Relevance— **This method can be utilized to allow ECG simulations with personalized torso geometry, which has previously been demonstrated to significantly effect QRS parameters. A clinical tool can be developed using this method that accounts for torso geometry in ECG interpretation.**

I. Introduction

The electrocardiogram (ECG) is a fundamental tool in diagnosing cardiac pathology, stratifying risk, informing treatment plans in emergency settings and for the management of chronic conditions. However, its accuracy is currently fundamentally limited by often overlooked implicit assumptions made in its calculation. The goal of the ECG is to understand the form of electrical impulses in the heart itself, such that arrhythmias, conduction abnormalities and the like can be identified. Measuring these potentials directly via a catheter or surgical intervention is prohibitively invasive, risky, and costly for the vast majority of routine investigations. Thus, leads to measure the electrical potential are placed on the surface of the body, and the potential at the heart is inferred by making assumptions concerning the anatomy of both the heart and torso, and the relative position of the heart within the torso. These factors have been shown to cause significant differences in QRS morphology, duration, and amplitude, both in vivo and in silico [1-5].

In our current research work, we develop a novel solution to this problem by combining non-invasive images showing the surfaces of the heart and torso with a patient's ECG, to understand what a normal reading would look like for the patient's individual anatomy. This could then be developed into a clinical tool to aid in detecting ECG abnormalities in a personalized framework. In our work, Cardiac Magnetic Resonance (CMR) images are segmented to extract the boundaries of both the heart and the torso. A series of two-dimensional (2D) contours is obtained from different CMR views, giving a sparse representation that is used to create a three-dimensional (3D) reconstruction.

For both research and clinical purposes, the extraction of torso contours from CMR images must be automated. This paper focuses on addressing the problem of torso contour extraction, using a machine learning approach. Our method is based on a u-net convolutional neural network [6], following its success in similar biomedical image segmentation tasks [7-8].

Our scans present significant challenges that do not allow us to use standard segmentation methods. The CMR images used are obtained using clinically standard protocols not explicitly designed for torso detection. Therefore, there are anatomical structures and CMR artifacts that would confound our reconstruction, and the region of interest is not always fully within frame. Thus, in this work, we propose a two-step automated framework consisting of a segmentation network followed by a two-channel refinement network, in order to remove these unwanted features before performing the final 3D torso reconstruction.

II. Methods

A. Dataset

The UK Biobank dataset [9] was used as it contains a large collection of high quality CMR images with varying demography. A random sample of 110 healthy subjects was used, including 100 for the cross-validation experiment and 10 for testing on an independent set. Each participant had on

*Research supported by the Wellcome Trust, the British Heart Foundation, and the CompBioMed 2 Centre of Excellence in Computational Biomedicine (European Commission Horizon 2020 research and innovation programme).

H. J. Smith is with the Computational Cardiovascular Science Group, Department of Computer Science, University of Oxford, Oxford OX1 3QD, UK (corresponding author, phone: +44 1865 273838, email: hannah.smith@msdtc.ox.ac.uk).

A. Banerjee is with the Institute of Biomedical Engineering, Department of Engineering Science, University of Oxford, Oxford OX3 7DQ, UK and the Division of Cardiovascular Medicine, Radcliffe Department of Medicine, University of Oxford, Oxford OX3 9DU, UK (e-mail: abhirup.banerjee@cardiov.ox.ac.uk).

R. P. Choudhury is with the Division of Cardiovascular Medicine, Radcliffe Department of Medicine, University of Oxford, Oxford OX3 9DU, UK (e-mail: robin.choudhury@cardiov.ox.ac.uk).

V. Grau is with the Institute of Biomedical Engineering, Department of Engineering Science, University of Oxford, Oxford OX3 7DQ, UK (e-mail: vicente.grau@eng.ox.ac.uk).

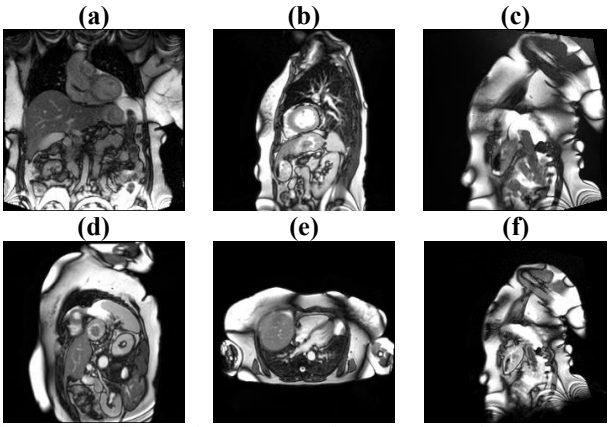

Figure 1. Localizer and scout CMR view subtypes and approximate percentage of the total dataset: (a) coronal localizer (front on, 5% of images), (b) sagittal localizer (side on, 20%), (c) cine images from a sagittal view (20%), (d) heart scout images down the long axis of the heart (30%), (e) long axis definition (5%), (f) short axis definition (20%).

average 60 localizer and scout CMR images, unevenly split between 6 types of views, as shown in Fig. 1.

The scans follow established guidelines for CMR, which does not aim at the reconstruction of the torso. They thus present some shortcomings when used for torso contouring. The views are focused on the heart itself, so there is limited information on the contour of the upper and lower parts of the torso. The torso contour is not always fully in frame as in Fig. 1(a). There are unwanted features that are not relevant for the torso reconstruction such as the head, neck, and arms. MRI artifacts are common – often appearing as a shadow region above the shoulder, as in Fig. 1(c), or an artificial cut off before the edge of the frame, as in the lower section of Fig. 1(f). The latter issue would be particularly problematic in the final 3D reconstruction as it may result in significantly misplaced points in the final torso contour.

B. Segmentation and Contouring

A custom-made MATLAB program was used for the creation of training sets by experts. The training was split into two sections. Firstly, a training set was made by drawing continuous boundaries to segment the image into torso (white) or background (black) as in Fig. 2(b). This strategy was employed rather than initially using a closed contour (as in Fig. 2(c)) as the former approach allows any extra contours to be removed by simply selecting the largest continuous contour. A relatively small training set of 15 patients (900 individual images) was required in this step. However, contours extracted from these segmentations may include undesirable imaging artifacts and sections of the head, neck, and arms. To remove these, we used the previously trained network to create a training set of contours from which we manually removed the undesirable features. In this way we efficiently created a larger training set of 100 subjects with 5993 individual images. A refinement network was trained by supplying the full contour, without features removed, along with the raw image, as a second channel input. This helped to prevent the network from creating any new, undesirable contours (for example in the lungs). As the result is not a

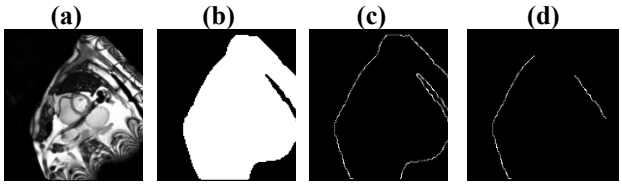

Figure 2. (a) Raw MR image – input for both networks, (b) torso segmentation – output for first network, (c) full contour – extracted from the output of the first network and used as additional channel input for second network, (d) contour segmentation with arm and false lower boundary removed – output for second network.

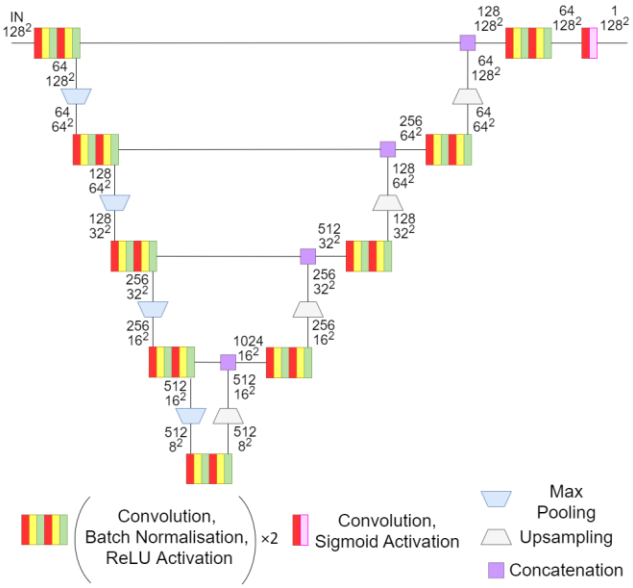

Figure 3. Implementation of u-net. Upper numbers represent the number of channels, where IN is the number of input channels (1 for the segmentation network, 2 for refinement), lower represent the size of the images in pixels.

closed contour, these extra contours could not easily be post-processed out by deleting all but the largest continuous contour. An independent test set of 10 subjects was used to evaluate the final refinement network's performance.

C. Model Architecture

An implementation of the u-net [6], as shown in Fig. 3, was chosen for both the segmentation and refinement networks. A sigmoid activation layer was added due to the binary nature of the desired output, along with bilinear upsampling. The optimizer used was Adam [10] and loss function was Dice loss. The segmentation network had the MR image as its single input, whilst for the refinement network, both the MR image and the full contours extracted from the first network's output were included as input channels.

D. Run Parameters and Testing

The train:validate ratio was maintained at 80:20, during our experiment over 100 subjects. A 5-fold cross validation was

used during both segmentation and refinement to select the final network for testing, considering Hausdorff distance, Dice coefficient, and qualitative results. The training was concluded at 100 epochs for the first network and 50 for the second. The Dice coefficient and Hausdorff distance were computed for each network. A comparison was performed between using one and two input channels for the refinement network. The six subsets of image views shown in Fig. 1 were compared, to identify whether any were not worth including in the final reconstruction.

E. 3D Torso Reconstruction

Over the automatically extracted sparse torso contours in 3D space, a statistical shape model (SSM), generated from 4308 scans of human body shapes [11], was applied to generate the initial 3D torso meshes. After optimally fitting the SSM by varying its first 40 principal components along with a rigid transformation, a deformation using approximate thin plate splines was applied to produce the final 3D torso meshes minimising the distance to the extracted contours [12].

III. RESULTS

A. Segmentation

Typical results of the segmentation network are shown in Fig. 4(a) and 4(b). The median performance over the validation subjects for each of the cross-validation networks was a Hausdorff distance of 3.2 pixels and Dice coefficient of 0.99, with the breakdown over each view shown in Fig. 5(a). One of the most common problems was segmenting particularly lighter areas of the torso with sharp gradients to darker areas such as the lungs, as background, as shown in Fig. 4(c). This error was fixed by automated postprocessing. Another typical source of error was in the armpit region where there is some subjectivity as to where exactly this boundary should be placed. This discrepancy became effectively irrelevant when the arms were removed by the refinement network.

B. Refinement

The median performance of a subject's contouring over the cross-validation sets was a Hausdorff distance of 5.1 pixels and Dice coefficient 0.93, and over the independent test set a Hausdorff distance of 4.3 pixels and a Dice coefficient of 0.93, with the breakdown over each view shown in Fig. 5(b) and 6 respectively. The addition of the second channel significantly improved performance, increasing the median test set Dice coefficient from 0.75 to 0.93 and decreasing the Hausdorff distance from 9.1 to 4.3 pixels. An example of the errors which it best corrected is shown in Fig. 7. The small additional contours on Fig. 7(b) that the single input network placed on areas of the lungs with particularly sharp boundaries to the rest of the torso are not present in the results from the network with two input channels, shown in Fig. 7(c).

C. Reconstruction

From the automatically extracted contours of 10 subjects in the independent test set, Subject 1 - male, 54, body mass index (BMI) 22 and Subject 2 - female, 58, BMI 36, were chosen for the 3D torso reconstruction. The generated torso meshes attained the average contour-to-surface distance of 2.02 ± 2.47 mm and 2.58 ± 3.09 mm (median ± inter-quartile range), as shown in Fig. 8.

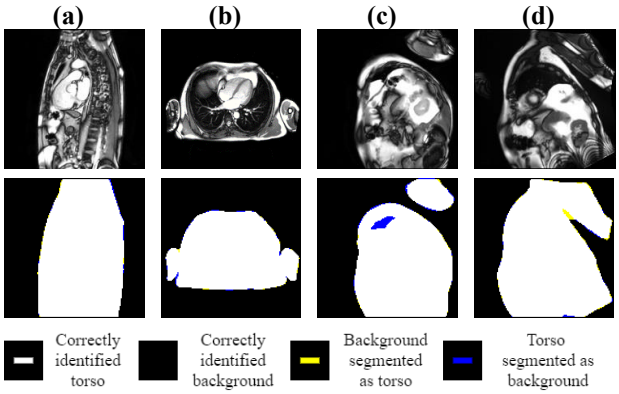

Figure 4. Final segmentation results, upper: raw input images, lower: results, (a) and (b) are typical results, showing a very high degree of accuracy whereas (c) and (d) represent common problems in the remaining minority of cases.

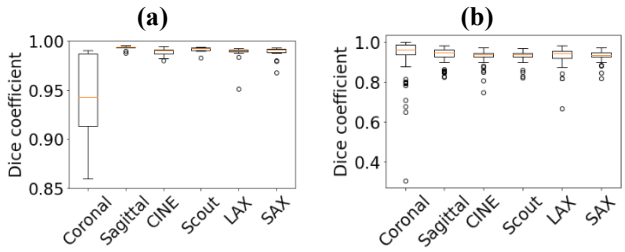

Figure 5. Box plot of mean Dice coefficient of an individual subject's images for each image view evaluated over the validation subjects for each of the cross-validation networks. Left: segmentation network, right: refinement network.

IV. DISCUSSION

The proposed refinement network performed accurately, with the majority of its errors only on the ends of contours where sections such as the neck and arms had been removed, as the exact position that these features start is subjective in manual annotations. The addition of the second input channel significantly improved both performance metrics. Since the first part of the proposed approach aimed on the identification of largest connected torso region, it enabled the refinement network to focus only on the removal of unnecessary segments from the torso boundary. The errors are larger for the refinement network as expected, as it is performing a more challenging task of identifying unnecessary torso segments to be removed, rather than the edge detection or the torso segmentation as the first network.

The coronal view typically had a larger number of outliers in the Dice score, mainly because subjects where most of the contour was off-frame were more likely to be badly contoured. This view was retained for use in the reconstruction however, as when the shoulders were in frame,

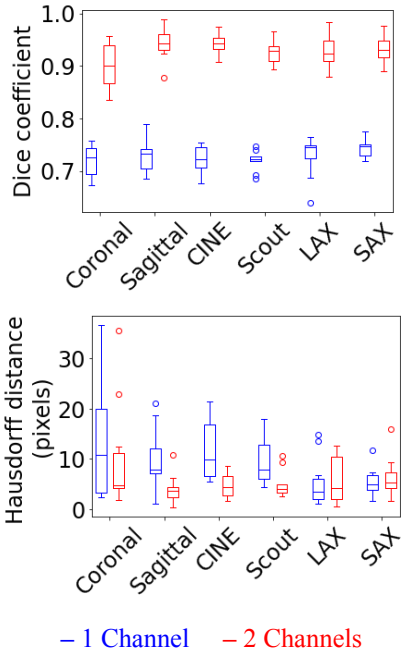

Figure 6. Comparison between using one channel input (raw CMR image only) and two channels (raw image and full contour extracted from segmentation). The additional channel significantly improved Dice coefficient for all views (above), and improved Hausdorff distance for 4 views (below).

these images gave information on the translational position of the body that was not visible in other views.

Hausdorff distance, a measure of the distance between two point clouds, was used to evaluate the network because it is unaffected by the class imbalance between contour and background. However, it tends to over-penalize certain views due to the subjectivity of contour placement. This is particularly relevant for coronal localizer images as in Fig. 1(a), where a subjective judgment has to be made as to whether the lines at the top of the shoulders are sufficiently well defined to be included.

V. Conclusion

The potential applications for this torso reconstruction pipeline are numerous and promising. As it is entirely automated, large datasets could be processed quickly to give thousands of ECG simulations with lead placement specific to each subject. These simulations could be analysed such that the impact of changes in heart and torso geometry on the ECG can be quantified. This represents a step forward in reducing patient specific variation of a normal ECG and facilitates future research into automated ECG interpretation. This opens the possibility of creating a clinical tool that corrects for the effects of heart and torso geometry, leading to improved clinical outcomes from more accurate identification of electrical abnormalities.

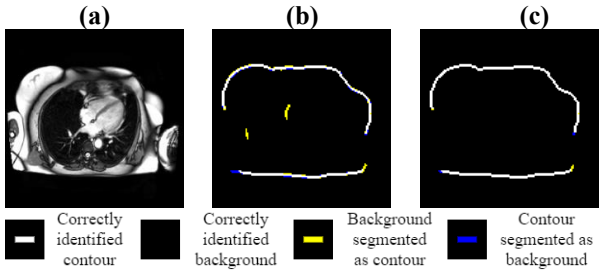

Figure 7. Final results: (a) input CMR image, (b) resulting contour for 1 input channel, showing extra contour on a sharp gradient caused by the lung in the center of torso and minor errors along the length of the contour, (c) resulting contour for 2 input channels, showing only minor misclassification at the sides where arms have been removed.

Figure 8. Reconstructions of two subjects' torsos, with positions of ECG leads represented by circles. Left: a male subject with low BMI, right: a female subject with high BMI. White circles are the positions of ECG leads.

VI. Compliance with Ethical Standards

This is a numerical simulation study for which no ethical approval was required.

Acknowledgment

This research has been conducted using the UK Biobank Resource under Application Number '40161'. The authors express no conflict of interest. The work of A. Banerjee was supported by the British Heart Foundation (BHF) Project under Grant HSR01230. The work of V. Grau was supported by the CompBioMed 2 Centre of Excellence in Computational Biomedicine (European Commission Horizon 2020 research and innovation programme, grant agreement No. 823712).

References

[1] R. Hoekema, G. J. H. Uijen, L. van Erning, and A. van Oosterom, "Interindividual variability of multilead electrocardiographic recordings: influence of heart position," *J. Electrocardiol.*, vol. 32, no. 2, pp. 137–148, Apr. 1999.

[2] R. Hoekema, G. J. H. Uijen, and A. van Oosterom, "Geometrical aspects of the interindividual variability of multilead ECG recordings," *IEEE Trans. Biomed. Eng.*, vol. 48, no. 5, pp. 551–559, May 2001.

[3] A. van Oosterom, R. Hoekema, and G. J. H. Uijen, "Geometrical factors affecting the interindividual variability of the ECG and the VCG," *J. Electrocardiol,* vol. 33, no. Suppl. 1, pp. 219–227, 2000.

[4] A. D. Corlan, R. S. Macleod, and L. De Ambroggi, "The effect of intrathoracic heart position on electrocardiogram autocorrelation maps," *J. Electrocardiol.,* vol. 38, no. 2, pp. 87–94, Apr. 2005.

[5] A. Mincholé, E. Zacur, R. Ariga, V. Grau, and B. Rodriguez, "MRI-based computational torso/biventricular multiscale models to investigate the impact of anatomical variability on the ECG QRS complex," *Front. Physiol.*, vol. 10, pp. 1103, Aug. 2019.

[6] O. Ronneberger, P. Fischer, and T. Brox, "U-net: convolutional networks for biomedical image segmentation," in *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2015*, Munich, Germany, pp. 234-241.

[7] S. Li, Y. Chen, S. Yang, and W. Luo, "Cascade dense-unet for prostate segmentation in MR images," in *Intelligent Computing Theories and Application*, Nanchang, China, 2019, pp. 481-490.

[8] T. Wang *et al.*, "ICA-UNet: ICA inspired statistical UNet for real-Time 3D cardiac cine MRI segmentation," in *Medical Image Computing and Computer Assisted Intervention - MICCAI 2020*, Lima, Peru, pp. 614-622.

[9] Z. Raisi-Estabragh, N. C. Harvey, S. Neubauer, and S. E. Petersen, "Cardiovascular magnetic resonance imaging in the UK Biobank: a major international health research resource," *Eur. Heart J. – Cardiovasc. Imag.,* vol. 22, no. 3, pp. 251–258, Feb. 2021.

[10] D. P. Kingma and J. Ba, "Adam: a method for stochastic optimization," presented at 3rd Int. Conf. for Learning Representations, San Diego, CA, USA, May 7-9, 2015.

[11] L. Pishchulin, S. Wuhrer, T. Helten, C. Theobalt, and B. Schiele, "Building statistical shape spaces for 3D human modelling," *Pattern Recogn.*, vol. 67, pp. 276–286, July 2017.

[12] E. Zacur *et al.*, "MRI-Based Heart and Torso Personalization for Computer Modeling and Simulation of Cardiac Electrophysiology," in *Imaging for Patient-Customized Simulations and Systems for Point-of-Care Ultrasound – BIVPCS 2017,* Québec City, Canada, pp. 61-70.