PROCEEDINGS OF SPIE

SPIDigitalLibrary.org/conference-proceedings-of-spie

Automatic whole-heart segmentation in 4D TAVI treatment planning CT

Bruns, Steffen, Wolterink, Jelmer, van den Boogert, Thomas P., Henriques, José, Baan, Jan, et al.

Steffen Bruns, Jelmer M. Wolterink, Thomas P. W. van den Boogert, José P. Henriques, Jan Baan, R. Nils Planken, Ivana Išgum, "Automatic whole-heart segmentation in 4D TAVI treatment planning CT," Proc. SPIE 11596, Medical Imaging 2021: Image Processing, 115960B (15 February 2021); doi: 10.1117/12.2581020

SPIE.

Event: SPIE Medical Imaging, 2021, Online Only

Automatic Whole-Heart Segmentation in 4D TAVI Treatment Planning CT

Steffen Bruns^{a,b}, Jelmer M. Wolterink^{a,c}, Thomas P.W. van den Boogert^d, José P. Henriques^d, Jan Baan^e, R. Nils Planken^f, and Ivana Išgum^{a,b,f}

^aDepartment of Biomedical Engineering and Physics, Amsterdam UMC, Amsterdam, The Netherlands

^bAmsterdam Cardiovascular Sciences, Amsterdam UMC, Amsterdam, The Netherlands

^cDepartment of Applied Mathematics, Technical Medical Centre, University of Twente, Enschede, The Netherlands

^dHeart Centre, Academic Medical Centre, Amsterdam Cardiovascular Sciences, University of Amsterdam, Amsterdam, The Netherlands

^eDepartment of Clinical and Experimental Cardiology, Amsterdam Cardiovascular Sciences, Amsterdam UMC, Amsterdam, The Netherlands

^fDepartment of Radiology and Nuclear Medicine, Amsterdam UMC, Amsterdam, The Netherlands

ABSTRACT

4D cardiac CT angiography (CCTA) images acquired for transcatheter aortic valve implantation (TAVI) planning provide a wealth of information about the morphology of the heart throughout the cardiac cycle. We propose a deep learning method to automatically segment the cardiac chambers and myocardium in 4D CCTA. We obtain automatic segmentations in 472 patients and use these to automatically identify end-systolic (ES) and end-diastolic (ED) phases, and to determine the left ventricular ejection fraction (LVEF). Our results show that automatic segmentation of cardiac structures through the cardiac cycle is feasible (median Dice similarity coefficient 0.908, median average symmetric surface distance 1.59 mm). Moreover, we demonstrate that these segmentations can be used to accurately identify ES and ED phases (bias [limits of agreement] of 1.81 [-11.0; 14.7]% and -0.02 [-14.1; 14.1]%). Finally, we show that there is correspondence between LVEF values determined from CCTA and echocardiography (-1.71 [-25.0; 21.6]%). Our automatic deep learning approach to segmentation has the potential to routinely extract functional information from 4D CCTA.

Keywords: Whole-heart segmentation, 4D cardiac CT angiography, transcatheter aortic valve implantation, deep learning, left ventricular ejection fraction

1. INTRODUCTION

Transcatheter aortic valve implantation (TAVI) is the preferred treatment for high-risk patients with severe aortic stenosis requiring intervention and recent evidence suggests that it might have potential for lower risk patients as well.¹ Acquisition of retrospectively ECG-triggered 4D cardiac CT angiography (CCTA) is a prerequisite for TAVI treatment planning. Such data consist of 10-20 3D CCTA images acquired at fixed time intervals over the cardiac cycle that allow the necessary measurements for intervention planning: valve prosthesis sizing, quantification of valvular calcification, and determination of the optimal access route for valve implantation.^{2,3}

Furthermore, 4D CCTA provides a unique view of the cardiac anatomy throughout the cardiac cycle. Hence, segmentations of the cardiac chambers and the myocardium in these studies could potentially be used to inspect myocardial motion and deformation, and to quantify cardiac function. Specifically, computing the volume of the left ventricular (LV) cavity in the end-systolic (ES) and end-diastolic (ED) phases allows computing the left

Figure 1. 4D CCTA data for one patient undergoing TAVI treatment planning, with automatic segmentations of the cardiac chambers and left ventricular (LV) myocardium at 15%, 35% (end-systole [ES]), 55%, 75%, and 95% (end-diastole [ED]) of the cardiac cycle. Derived volumes of the LV cavity are used to determine the ES and ED phases, and to calculate the LV ejection fraction, which is 71.8% in this patient.

ventricular ejection fraction (LVEF). The LVEF is a well-known biomarker for cardiac function that is standardly determined using cardiac MRI or echocardiography.⁴

Obtaining manual segmentations of cardiac chambers in 3D CCTA images is extremely tedious and time-consuming, and this challenge is exacerbated twenty-fold in 4D CCTA studies. Therefore, automatic methods have been developed for CCTA segmentation.^{5–9} Thus far, such methods have not been used to automatically segment 4D TAVI treatment planning CCTA. TAVI is planned only for patients with a severe aortic valve stenosis and high co-morbidity and thus, images frequently contain imaging artifacts caused by metal implants, large atherosclerotic calcifications, or other pathology making automatic analysis challenging. To address this, we propose a deep learning method for automatic and robust segmentation of the cardiac chambers and myocardium in 4D TAVI treatment planning CCTA images throughout the cardiac cycle (Fig. 1). We show that this method can be used to accurately segment the cardiac chambers and myocardium and thereby derive volumes of these cardiac structures in different cardiac phases. Moreover, we use segmentations and volumetric measurements to automatically identify the ES and ED phases, and assess LV function by means of the LVEF. Finally, we compare automatically obtained LVEF values to those obtained using echocardiography.

2. MATERIAL AND METHODS

2.1 Data

We retrospectively collected data from patients that underwent TAVI treatment planning at the Amsterdam University Medical Centers, location AMC (Amsterdam, The Netherlands) between 2013 and 2019. We included 472 patients in total (250 men, age 79.2 ± 6.9 [range 36–93]). The data was acquired on two different CT scanners (SOMATOM Force, Siemens Healthcare, Erlangen, Germany; Brilliance 64, Philips Healthcare, Best, Netherlands) with 70–120 kVp and tube currents of 190–2598 mAs. The images were reconstructed with an in-plane spacing ranging from $0.30 \times 0.30 \text{ mm}^2$ to $0.72 \times 0.72 \text{ mm}^2$ and 0.6–0.9 mm slice thickness. The cohort consisted of a selected training set of 12 patients reflecting the diversity present in the full data set in terms of used scanners, attenuation in the cardiac chambers, patient sex, patient weight, and the occurrence of pacemakers, and an evaluation set of 460 consecutive patients for whom CCTA images were available across the cardiac cycle with 5% cardiac cycle intervals. Thus, up to 21 3D CCTA images were available per patient, for a total of 9771 3D CCTA images. For a subset of 123 patients from the evaluation set, the LVEF derived from echocardiography pre-TAVI acquired within six months of CCTA scanning was available.

Further author information: (Send correspondence to S.B.)
S.B.: E-mail: s.bruns@amsterdamumc.nl

Figure 2. 3D fully convolutional network used in this study. The network takes a $128 \times 128 \times 128$ voxel sub-image as input. The image is processed by a convolutional layer and two strided convolutional (downsampling) layers. The resulting feature maps are processed by six residual blocks (ResNetBlocks) and upsampled by two transposed convolutional (upsampling) layers, resulting in a 6-channel $128 \times 128 \times 128$ voxel probability map. BN = batch normalization, ReLU = rectified linear unit.

2.2 Reference segmentations

For training and validation, 24 full 3D reference segmentations of the LV cavity, LV myocardium, right ventricle, left atrium, and right atrium were obtained: one in the ES and one in the ED CCTA image of the 12 patients in the training set. The ES and ED phases were manually selected by a resident (TPWB, 3 years of experience and level II CCTA reader) under supervision of a cardiovascular radiologist (RNP). Initial reference segmentations were obtained using a previously described automatic method,¹⁰ and subsequently manually corrected by TPWB using 3D Slicer (V4.8.1, <http://www.slicer.org>). These 24 3D CCTA images and reference segmentations were used for training and cross-validation of the method.

For independent testing in a larger number of patients and across different cardiac phases, additional reference segmentations were obtained in an independent test set containing 81 consecutive 4D CCTA studies from the evaluation set. Four phases were identified in each study: ES, ED, the mid-diastolic phase between ES and ED, and the mid-systolic phase between ED and ES. In each of the 3D CCTA images corresponding to these phases, the LV cavity, LV myocardium, right ventricle, left atrium, and right atrium were manually segmented in the same centrally located 2D axial image slice for each patient. These reference segmentations were acquired fully manually by a medical student by voxel-wise painting using 3D Slicer and corrected by TPWB if necessary.

2.3 Automatic segmentation method

We developed a multi-class 3D fully convolutional network (FCN) architecture for segmentation of each of the 3D CCTA images acquired as part of a 4D CCTA TAVI planning study. This architecture was based on a previously proposed 2D architecture.¹¹ The FCN consists of an encoding path with two downsampling layers with strided convolutions, six residual ResNetBlocks, and a decoding path with two upsampling layers with transposed convolutions (see Fig. 2). Prior to analysis, 3D CCTA images were resampled to an isotropic resolution of $0.8 \times 0.8 \times 0.8$ mm³. Moreover, all intensities were linearly rescaled between -1024 and 3071 Hounsfield units to a [0, 1] intensity range. FCN training was performed using mini-batches of 8 randomly sampled $128 \times 128 \times 128$ voxel patches for 10000 iterations in total with the Adam optimizer, a base learning rate of 0.001, and a 70% learning rate decay after every 4000 iterations. To optimize the parameters of the FCN, the negative sum of soft Dice similarity coefficients over all classes was minimized as the loss function. During inference, overlapping 3D image patches were processed by the FCN and subsequently combined using a stitching procedure to obtain class probability maps, which were used to obtain the final output segmentation.

Figure 3. Dice similarity coefficients (a, b) and average symmetric surface distances (c, d) for automatic segmentation of the left ventricular (LV) myocardium, LV cavity, right ventricle (RV), left atrium (LA), and right atrium (RA) obtained in the leave-one-patient-out cross-validation with 3D end-systolic (ES) and end-diastolic (ED) CCTA images of 12 patients. Square markers indicate results for an image with lower contrast enhancement than the other images in the training set, shown in Fig. 4b.

3. EXPERIMENTS AND RESULTS

3.1 Automatic segmentation

The training set consisting of 12 ES and 12 ED 3D CCTA images and corresponding full 3D reference segmentations was used in a leave-one-patient-out cross-validation experiment. For each patient, two networks were trained. The first network was trained with only the ES images of all other patients. The second network was trained with only the ED images of all other patients. We investigated the ability of both networks to segment 3D images of the validation patient in the same phase and in the contrasting phase. The Dice similarity coefficient (DSC) and average symmetric surface distance (ASSD) values in Fig. 3 show that FCNs trained on ES images can be used to segment structures in ED images and vice versa. Nevertheless, for all structures, and particularly for the LV myocardium, which undergoes substantial deformation in patients with sufficient LVEF, the segmentation performance is better when a network is trained with images from the same phase. We also segmented each validation image based on an ensemble of the two networks trained on either ES or ED in which the output probabilities were averaged. These ensembles yield improved segmentation performance across almost all structures compared to the networks only trained on images of the corresponding phase. Figure 4a shows an example of a very accurate segmentation of all five cardiac structures in both the ES and the ED phase (mean DSC=0.923 and mean ASSD=0.925 mm over all structures in ES and ED). In one of the 12 patients, however, the contrast enhancement was lower than in the other training images, and the automatic method failed to segment the cardiac structures accurately (Fig. 4b).

To assess segmentation performance in a larger data set and across the cardiac cycle, thus in other phases than only the ES and ED phases, an ensemble of all 24 trained FCNs obtained in the cross-validation experiment

Figure 4. Automatic segmentations of the cardiac structures in end-systolic (ES) and end-diastolic (ED) phases of two different patients in the leave-one-patient-out cross validation. a) Very accurately segmented images, b) images with very low contrast enhancement in comparison with the data in the training set and poor automatic segmentation (indicated with square markers in Fig. 3). All images are displayed with the same window-level setting.

was formed and used to segment the independent test set of 81 patients. Figure 5 shows the resulting DSC and ASSD values obtained across four different cardiac phases.

3.2 Automatic ES and ED selection

Accurate identification of the ES and ED phase in the cardiac cycle is crucial for the computation of the LVEF. We evaluated to what extent our automatic segmentations allow identification of these phases. We automatically segmented all 3D CCTA images in the 460 patients in the evaluation set, and derived the LV volume in each phase. We identified ES as the phase in which the LV volume was minimal and ED as the phase in which the LV volume was maximal. As a reference, TPWB inspected images of all cardiac phases of these patients and manually selected the ES and ED phases. We compared our automatically identified phases with the manually selected phases using Bland-Altman analysis (see Fig. 6). The bias [limits of agreement] were 1.81 [-11.0; 14.7]% for the ES phase and -0.02 [-14.1; 14.1]% for the ED phase, respectively.

3.3 Functional analysis

For the assessment of the LV function, we derived the LVEF from the automatic segmentations: We computed the LV volume in the ES and ED images from the automatic segmentations and calculated $LVEF = \frac{V_{ED} - V_{ES}}{V_{ED}}$. CT-derived LVEF values were compared to echocardiography LVEF values in 123 patients. Figure 7 shows a Bland-Altman analysis of the LVEF comparison. The bias [limits of agreement] were -1.71 [-25.0; 21.6]%. Furthermore, it shows ES and ED images with very accurate automatic segmentations for the two patients with the highest differences between the LVEF derived from echocardiography and the LVEF automatically derived from CCTA.

Figure 5. Dice similarity coefficients and average symmetric surface distances for automatic segmentation of the left ventricular (LV) myocardium, LV cavity, right ventricle (RV), left atrium (LA), and right atrium (RA) obtained in axial slices in 81 patients across four different cardiac phases: end-systolic (ES), mid-diastolic, end-diastolic (ED), and mid-systolic.

Figure 6. Bland-Altman plots of the manually vs. automatically selected end-systolic (ES) and end-diastolic (ED) phases. Darker colors correspond to more overlapping data points.

4. DISCUSSION

We have presented an automatic deep learning method for whole-heart segmentation in 4D CCTA for TAVI treatment planning. We have shown that 3D FCNs trained on the ES and ED CCTA images can be used to accurately segment 3D CCTA images across the whole cardiac cycle. These segmentations can be used to automatically identify the ES and ED phase, to derive cardiac chamber volumes over time, and to assess cardiac function by means of the LVEF.

Our cross-validation experiments show that in the majority of images all structures were segmented accurately with DSC above 0.85 and ASSD below 2 mm. Images in which the segmentation performance was worse were often affected by lower contrast enhancement (Fig. 4b) or artifacts caused by metal implants which were not represented in the training set. To further increase segmentation performance, additional images with such characteristics could be added to the training data and additionally, augmentation strategies could be applied. In independent test images of 81 patients, the method showed comparable segmentation performance for all cardiac structures across different cardiac phases.

Figure 7. Bland-Altman plot of the left ventricular ejection fraction (LVEF) manually derived from echocardiography vs. the LVEF automatically derived from 4D cardiac CT angiography (CCTA). Axial slices with automatic segmentations of the end-systolic (ES) and end-diastolic (ED) phases for the two patients with the highest differences between LVEF from echocardiography and CCTA shown on the right.

Automatic segmentation enabled robust automatic identification of the ES and ED phases with respect to the manual identification. Large deviations between the manual and the automatic selection were mainly found in patients with very low LVEF. In these patients, volume differences between phases are small and small errors in the automatic segmentation can lead to an incorrect selection of the ES and ED phases. However, these deviations might cause volume errors within an acceptable range. Consequently, the LVEF could be automatically derived from the CCTA images. The limits of agreement between the LVEF derived from echocardiography and CCTA were relatively wide with $[-25.0; 21.6]\%$. Note, however, that echocardiography is not considered the reference standard for extraction of the LVEF as inadequate acoustic windows can cause erroneous LVEF computations.⁴ This might be the reason why large differences between the LVEF from echocardiography and the LVEF automatically derived from CCTA occurred, even in patients in which the automatic segmentations were very accurate. Two such examples are shown in Fig. 7, where the automatic segmentations in CCTA were very accurate, yet differences with LVEF values determined in echocardiography were large. Previous studies have shown that there are no significant differences between LVEF determined in CT and the gold standard cardiac MRI¹² and further validation of our method against this gold standard is required.

The proposed method enables an accurate analysis of the morphology and function of the heart over the cardiac cycle. In future work, the volumes of the cardiac structures could be used to estimate the progression of damage induced by the aortic stenosis.¹³ Moreover, the motion and deformation of the chambers and the LV myocardium throughout the cardiac cycle could be studied to assess cardiac function beyond the LVEF.

5. CONCLUSION

We have presented an automatic deep learning method trained on end-systolic and end-diastolic images for accurate and robust whole-heart segmentation in 4D cardiac CT angiography for transcatheter aortic valve implantation planning.

ACKNOWLEDGMENTS

This study was funded by the Dutch Technology Foundation (STW, perspectief, P15-26) with participation of Philips Healthcare, Haifa, Israel.

REFERENCES

- [1] Mack, M. J., Leon, M. B., Thourani, V. H., Makkar, R., Kodali, S. K., Russo, M., Kapadia, S. R., Malaisrie, S. C., Cohen, D. J., Pibarot, P., et al., “Transcatheter aortic-valve replacement with a balloon-expandable valve in low-risk patients,” *N Engl J Med* **380**(18), 1695–1705 (2019).
- [2] Achenbach, S., Delgado, V., Hausleiter, J., Schoenhagen, P., Min, J. K., and Leipsic, J. A., “SCCT expert consensus document on computed tomography imaging before transcatheter aortic valve implantation (TAVI)/transcatheter aortic valve replacement (TAVR),” *J Cardiovasc Comput Tomogr* **6**(6), 366–380 (2012).
- [3] van den Boogert, T., Vendrik, J., Claessen, B., Baan, J., Beijk, M., Limpens, J., Boekholdt, S., Hoek, R., Planken, R., and Henriques, J., “CTCA for detection of significant coronary artery disease in routine TAVI work-up,” *Neth Heart J* **26**(12), 591–599 (2018).
- [4] Asferg, C., Usinger, L., Kristensen, T. S., and Abdulla, J., “Accuracy of multi-slice computed tomography for measurement of left ventricular ejection fraction compared with cardiac magnetic resonance imaging and two-dimensional transthoracic echocardiography: a systematic review and meta-analysis,” *Eur J Radiol* **81**(5), e757–e762 (2012).
- [5] Zheng, Y., Barbu, A., Georgescu, B., Scheuering, M., and Comaniciu, D., “Four-chamber heart modeling and automatic segmentation for 3-D cardiac CT volumes using marginal space learning and steerable features,” *IEEE Trans Med Imaging* **27**(11), 1668–1681 (2008).
- [6] Ecabert, O., Peters, J., Schramm, H., Lorenz, C., von Berg, J., Walker, M. J., Vembar, M., Olszewski, M. E., Subramanian, K., Lavi, G., et al., “Automatic model-based segmentation of the heart in CT images,” *IEEE Trans Med Imaging* **27**(9), 1189–1201 (2008).
- [7] Zhuang, X., Bai, W., Song, J., Zhan, S., Qian, X., Shi, W., Lian, Y., and Rueckert, D., “Multiatlas whole heart segmentation of CT data using conditional entropy for atlas ranking and selection,” *Med Phys* **42**(7), 3822–3833 (2015).
- [8] Zhuang, X., Li, L., Payer, C., Štern, D., Urschler, M., Heinrich, M. P., Oster, J., Wang, C., Smedby, Ö., Bian, C., et al., “Evaluation of algorithms for multi-modality whole heart segmentation: An open-access grand challenge,” *Med Image Anal* **58**, 101537 (2019).
- [9] Myronenko, A., Yang, D., Buch, V., Xu, D., Ihsani, A., Doyle, S., Michalski, M., Tenenholtz, N., and Roth, H., “4D CNN for semantic segmentation of cardiac volumetric sequences,” in *[Stat Atlases Comput Models Heart]*, 72–80, Springer (2019).
- [10] Bruns, S., Wolterink, J. M., van Hamersvelt, R. W., Zreik, M., Leiner, T., and Išgum, I., “Improving myocardium segmentation in cardiac CT angiography using spectral information,” in *[Medical Imaging 2019: Image Processing]*, **10949**, 109492M, International Society for Optics and Photonics (2019).
- [11] Johnson, J., Alahi, A., and Fei-Fei, L., “Perceptual losses for real-time style transfer and super-resolution,” in *[Comput Vis ECCV]*, 694–711, Springer (2016).
- [12] Pickett, C. A., Cheezum, M. K., Kassop, D., Villines, T. C., and Hulten, E. A., “Accuracy of cardiac CT, radionuclide and invasive ventriculography, two- and three-dimensional echocardiography, and SPECT for left and right ventricular ejection fraction compared with cardiac MRI: a meta-analysis,” *Eur Heart J Cardiovasc Imaging* **16**(8), 848–852 (2015).
- [13] Génèreux, P., Pibarot, P., Redfors, B., Mack, M. J., Makkar, R. R., Jaber, W. A., Svensson, L. G., Kapadia, S., Tuzcu, E. M., Thourani, V. H., et al., “Staging classification of aortic stenosis based on the extent of cardiac damage,” *Eur Heart J* **38**(45), 3351–3358 (2017).