



Published in final edited form as:

Comput Biol Med. 2020 October ; 125: 104019. doi:10.1016/j.compbimed.2020.104019.

Improving Multi-Atlas Cardiac Structure Segmentation of Computed Tomography Angiography: A Performance Evaluation Based on a Heterogeneous Dataset

Vy Bui^{1,2}, Li-Yueh Hsu^{1,3}, Sujata M. Shanbhag¹, Loc Tran², W. Patricia Bandettini¹, Lin-Ching Chang², Marcus Y. Chen¹

¹National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, MD 20892 USA

²Department of Electrical Engineering and Computer Science, Catholic University of America, Washington, DC 20064 USA

³Department of Radiology and Imaging Sciences, National Institutes of Health, Bethesda, MD 20892 USA

Abstract

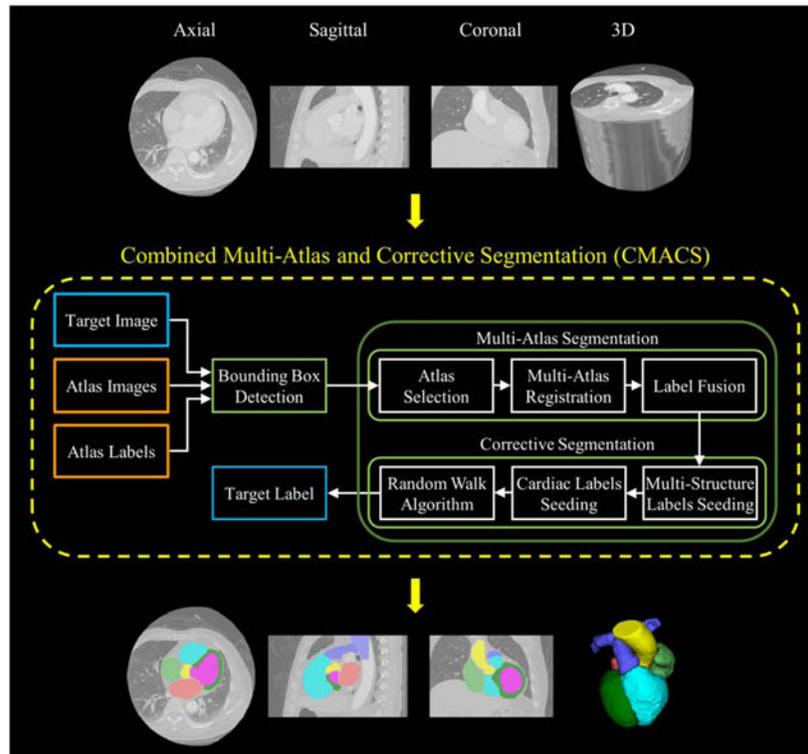
Multi-atlas based segmentation is an effective technique that transforms a representative set of atlas images and labels into a target image for structural segmentation. However, a significant limitation of this approach relates to the fact that the atlas and the target images need to be similar in volume orientation, coverage, or acquisition protocols in order to prevent image misregistration and avoid segmentation fault. In this study, we aim to evaluate the impact of using a heterogeneous Computed Tomography Angiography (CTA) dataset on the performance of a multi-atlas cardiac structure segmentation framework. We propose a generalized technique based upon using the Simple Linear Iterative Clustering (SLIC) supervoxel method to detect a bounding box region enclosing the heart before subsequent cardiac structure segmentation. This technique facilitates our framework to process CTA datasets acquired from distinct imaging protocols and to improve its segmentation accuracy and speed. In a four-way cross comparison based on 60 CTA studies from our institution and 60 CTA datasets from the Multi-Modality Whole Heart Segmentation MICCAI challenge, we show that the proposed framework performs well in segmenting seven different cardiac structures based upon interchangeable atlas and target datasets acquired from different imaging settings. For the overall results, our automated segmentation framework attains a median Dice, mean distance, and Hausdorff distance of 0.88, 1.5 mm, and 9.69 mm over the entire datasets. The average processing time was 1.55 minutes for both datasets. Furthermore, this study shows that it is feasible to exploit heterogeneous datasets from different imaging protocols and institutions for accurate multi-atlas cardiac structure segmentation.

Corresponding author: Li-Yueh Hsu. D.Sc., Building 10, Room 1C370, 10 Center Drive, Bethesda, Maryland, 20892, USA. lyhsu@nih.gov, Tel: (301) 496-7133.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

The authors declare that there is no conflict of interest regarding the publication of this paper.

Graphical Abstract



Keywords

Cardiac Computed Tomography; Heart Segmentation; Multi-Atlas Segmentation

1. Introduction

Cardiac computed tomography angiography (CTA) is an important imaging modality for assessing the morphology of the heart and coronary arteries for diagnosing cardiovascular disease. Accurate heart segmentation in cardiac CTA images is an important task for quantitative assessment of different cardiovascular structures. The quantification usually involves volumetric measurement of the ventricles, atria, and left ventricular myocardium, as well as the great vessels from the CTA images. Conventionally, these 3D volume sizes are measured by manual delineation on 2D images, which is tedious and subject to user variability, and the process requires domain expertise to perform such a task. Automatic segmentation is desirable but challenging due to large shape variations of the cardiac anatomy among different subjects, changes of the heart shape as it contracts and relaxes, and indistinct boundaries between different cardiac substructures (e.g. right ventricle and right atrium) or between the heart and surrounding tissues (e.g. liver, ribs, sternum). In addition, the image quality and tissue appearance may be influenced by imaging protocol variations such as different contrast enhancement timing and concentration, radiation exposure dose, acquisition field-of-view, matrix size, and slice thickness. Different scanner detector arrays and vendors may further affect the robustness of automated image segmentation tasks.

Several automatic methods have been developed for cardiac CTA image segmentation. We refer the reader to a comprehensive review of existing works presented by Bui et al. and Zhuang et al. [1,2] which includes several deep learning-based methods [3–7] as well as multi-atlas segmentation methods [8–16] that are widely used for multi-structure cardiac image segmentation. In the multi-atlas based approach, a set of atlases, called an atlas library, are created which include pre-segmented labels of various cardiac structures from the images of different subjects. These atlases are used as templates to segment a target image by registering each image in the atlas library with the target. The corresponding labels in the atlas library are then transformed after the image registration, and then combined to obtain the final segmentation of the target image.

It is known that the segmentation accuracy of multi-atlas based approaches relies heavily on the registration algorithm and the label fusion strategy of multiple atlases to obtain an optimal segmentation result. In addition, the performance of multi-atlas based segmentation can be affected by the similarity of the images in the atlas library versus the target image. Different acquisition protocols that possess large discrepancies in image quality, volume coverage and orientation, or tissue contrast appearance between the atlas and target image dataset may potentially lead to image misregistration and segmentation errors.

In this paper, we investigate the impact of using a heterogeneous atlas versus a target image dataset in multi-atlas cardiac structure CTA image segmentation. Our work is the first to systemically study the interplay between atlas and target image datasets acquired from distinct imaging protocols. In a four-way cross comparison using different atlas and target datasets collected from two independent institutions, we evaluate the robustness of using a generalized multi-atlas segmentation framework that is based on different image and label configurations.

This current work contributes to and extends our previous CTA segmentation framework [1] by introducing a novel bounding box detection technique to localize the whole heart before performing multi-atlas based segmentation. This technique addresses the often encountered differences in image field-of-view and volume coverage among different clinical sites and improves the overall segmentation accuracy. Our technique is based upon a supervoxel algorithm [18] combined with a generalized rule-based analysis to detect a bounding box around the whole heart region. Furthermore, we also augment the label fusion strategy in our previous work [1] to select optimal atlas labels based on a structural similarity index which leads to an improvement of overall segmentation accuracy.

2. Material and methods

We have previously developed a fully automatic method based upon a multi-atlas approach to segment various cardiac structures from CTA images [1,17]. The performance of the method was evaluated on datasets acquired under the same imaging protocol from our institution. In this current study, we extend our method and evaluate its performance on heterogeneous datasets acquired from two independent institutions with different scanner vendors and imaging settings.

2.1. Clinical dataset

The first dataset consisted of 60 clinical CTA scans of patients with suspected coronary artery disease referred to the National Heart, Lung, and Blood Institute (NHLBI). All CTA exams were performed under procedures and protocols approved by the Institutional Review Board of the National Institutes of Health. Written informed consent was obtained from all subjects prior to participating in the study. All CTA studies were performed on a 320-detector row scanner (Aquilion One Genesis, Canon Medical Systems, Japan) with 0.5 mm detector collimation, 275 msec gantry rotation time, 100–120 kVp tube voltage, 200–850 mA tube current adjusted according to the patient's attenuation profile determined by the scout image. Contrast material dose was 50–70 mL administered at a flow rate of 5.0–5.5 mL/sec and adjusted for body habitus. Prospective ECG-triggered image acquisition was initiated by a target threshold of 350–400 HU in the descending aorta. For each dataset, images were reconstructed at a 75% phase window around diastole in the cardiac cycle with a matrix size of 512×512 and an average pixel size 0.36×0.36 (from 0.26×0.26 to 0.43×0.43) mm². Each study contained 240 to 520 images with an average slice thickness of 0.33 (from 0.25 to 0.5) mm.

For each of these 60 cases, 12 cardiovascular structures were manually labeled by two trained observers using a custom developed interactive image analysis software and reviewed by experienced cardiologists. These labels included the whole heart (WH) region and seven cardiac structures: left ventricular cavity (LV), left atrial cavity (LA), left ventricular myocardium (LVM), left atrial appendage (LAA), right ventricular cavity (RV), right atrial cavity (RA), ascending aorta (AA); as well as four associated vascular structures including the superior vena cava (SVC), inferior vena cava (IVC), pulmonary artery (PA), and pulmonary vein (PV), all of which carry blood to or from the heart.

The second dataset was provided by the Multi-Modality Whole Heart Segmentation (MMWHS) challenge [2] which included 60 cardiac CTA datasets acquired using two 64-slice CT scanners (Philips Medical Systems, Netherlands) utilizing a standard coronary CT angiography protocol at two sites in Shanghai, China. All of the data covered the whole heart from the upper abdomen to the aortic arch. The slices were acquired in the axial view with a matrix size of 512×512 and an average pixel size of 0.43×0.43 (from 0.28×0.28 to 0.59×0.59) mm². Each study contained 177 to 363 images with an average slice thickness of 0.60 (from 0.45 to 0.63) mm.

For the MMWHS dataset, CTA studies were split into 20 training and 40 test datasets for the performance evaluation. For the training datasets, image data and manual segmentation of seven whole heart substructures LV, RV, LA, RA, LVM, AA, and PA were made available online [19]. For the test datasets, only the image data were available, but the segmentation results could be self-evaluated by an executable tool and encrypted label data that were available at the MMWHS data website [19].

2.2. Multi-atlas segmentation

Fig. 1. shows the flow diagram of our automated processing pipeline for multi-structure CTA image segmentation. This framework adds an automated bounding box detection

scheme to our previous combined multi-atlas and corrective segmentation method [1] to localize a volume of interest encompassing the whole heart. This bounding box detection feature, which is based on the supervoxel technique, is an important addition to our previous method to construct a standardized and confined processing volume from the CTA images before applying the subsequent multi-structure heart segmentation. The following sections will describe each processing step in more details.

2.2.1. Bounding box detection—Fig. 2. shows a series of image processing steps to identify the six faces of a 3D bounding box containing the whole heart in a CTA image volume. This bounding box detection feature improves the registration accuracy for the subsequent multi-atlas segmentation. It also reduces the processing data size and thus increases the multi-atlas segmentation speed.

As a preprocessing step, the CTA image volume is first enhanced by mapping pixels that have greater than -30 Hounsfield Unit (HU) intensity values into an 8-bit dynamic range. This threshold is chosen based on tissue intensity statistics of the whole heart found in our previous study [1]. For the first step (SLIC Supervoxel) in Figure 2, the SLIC algorithm [18], which adapts a k-means clustering approach, is applied to create a supervoxel image from the 3D grayscale image volume. Under this supervoxel image representation, each supervoxel assembles a group of voxels that share similar intensity and are within a close spatial proximity. In our implementation, the compactness parameter is set to 20 to increase the weighting toward the spatial proximity [20].

The second step (Supervoxel Segmentation) is to detect extremum intensity structures from the grayscale images which include low intensity structures such as lung and air regions, as well as high intensity structures such as left-sided cardiac chambers, aorta, and spine and rib bone regions within the image volume. A sliding thin-slab minimum intensity projection of the grayscale image is created to remove small bright vessel regions in the lung, and then a HU threshold window between -400 and -90 is used to generate a low intensity structure mask. Similarly, a sliding thin-slab maximum intensity projection image is created, and then a HU threshold value of 150 is used to generate a high intensity structure mask. Within the high intensity structure mask, the most inferior point of the left ventricular and atrial chambers blood pool region, which is the closest position to the chest wall, is located, and any supervoxel residing above this point is removed. All supervoxels located within the low intensity structure mask are also removed. These procedures effectively separate the remaining supervoxels in the liver region from the lung, chest wall, spine, aorta, and heart regions. The bottom face of the bounding box is then computed from the averaged z-coordinate values of the liver region, or is defaulted to the bottom slice of the image volume if no liver region detected.

The third step (Region Segmentation) is to detect the top face of the bounding box. The eigenvalues of the inertia tensor are calculated from each high contrast structure region including the left ventricular and atrial chambers and ascending aorta. This process is performed on a slice-by-slice basis from the superior to the inferior direction. Since the appearance of the ascending aorta in each axial image slice has a circular disc-like shape and a diameter normally less than 30 mm, such a circular region can be detected by measuring

its major and minor axes length derived from the inertia tensor. As an inertia tensor measures covariance of a region along its axes, the circularity of the region can be readily quantified by the relative magnitude of the tensor's eigenvalues. Here we selected the first slice that has a circular region with a major axis length less than 30 mm, and a minor axis length less than 5 mm difference with the major axis length, as the top face of the bounding box. After allocating the top and bottom faces, the front face of the bounding box is computed from the anterior high intensity chest wall structure region from its average x-coordinate values. Similarly, the back face of the bounding box is computed from the posterior high intensity spine and descending aorta structures by their x-coordinate average.

Finally, the two lateral faces of the bounding box are detected by thresholding the grayscale image with a -30 HU intensity value to create a binary mask within the top, bottom, front, and back boundaries. The largest connected region in this binary mask will encapsulate the whole heart and can then subsequently be used to define the left and right faces of the bounding box.

2.2.2. Combined multi-atlas and corrective segmentation—After the bounding box detection, the confined region that contains the whole heart is processed by an automatic combined multi-atlas and corrective segmentation (CMACS) pipeline that we introduced previously [1]. In summary, this automated framework consists of two core processing blocks, 1) multi-atlas segmentation and 2) corrective segmentation, to perform simultaneous multi-structure heart and peripheral tissue segmentation. Our methods are implemented in a multi-threading architecture to increase computational efficiency of all processing blocks. We have evaluated the CMACS framework on a homogeneous clinical dataset with the same acquisition settings and have shown that the results matched closely with manual reference segmentation [1].

For a brief description of the methods, the multi-atlas segmentation in the first block begins with a fast strategy to select an optimal set of atlases from an atlas library. The structural similarity index (SSI) performs a structural similarities assessment [21], comparing the given target image to all images in the atlas library in order to select an optimal set of seven atlas images. A pairwise atlas-to-target deformable image registration is then performed on each selected atlas to obtain a non-linear transformation to warp the associated atlas label into the target image space [22]. After the multi-atlas registration, an improved label fusion scheme is implemented by comparing the structural similarities between all warped images and the target image. Those warped labels with structural similarity values above 0.3 are preserved and merged into a target label for the next processing block.

The corrective segmentation block is designed to refine the cardiovascular labels obtained from the previous block and to separate the intrathoracic tissue structures surrounding the heart. It begins with automated image processing steps to extract representative seed voxels from non-cardiac structures that include lung, chest wall, liver, spine, and descending aorta. Together with the previously obtained cardiac structures, a random walk algorithm [23] is then performed on each seed region in a multiple-pass fashion to improve the segmentation result for each structure. The final segmentation is obtained by additional post-processing steps for further refinement.

2.3. Evaluation methods

To evaluate the interplay of different configurations of atlas versus target datasets based on our proposed framework, the two previously mentioned datasets, namely NHLBI and MMWHS, were used. Each dataset consisted of 60 pairs of image studies and reference labels obtained by manual segmentation, in which 20 cases were used for training and 40 cases were preserved for testing. In our multi-atlas image segmentation framework, the 20 training cases were retained as the atlas library and the 40 testing cases were used as the target images for independent assessment. We performed a four-way cross comparison by using the atlases from each dataset to segment the target images on both datasets. Additionally, we compared our results with 10 independent methods as evaluated in the MMWHS challenge study [2].

All algorithms were implemented in Python (www.python.org) and Interactive Data Language (Harris Geospatial Solutions). The registration method was developed in C++ by Heinrich et al. [22] and compiled to a dynamic link library under Microsoft Visual Studio in our framework. All studies were processed with the same parameter settings on a computer with an Intel Core i9-7980XE 2.6GHz CPU and 128GB RAM.

For the performance evaluation of the automated segmentation, three quantitative metrics that measure the differences between the automatic results versus the corresponding manual labels were computed which include the Dice coefficient (Dice), mean surface distance (MSD), and Hausdorff distance (HD). Summary statistics of the results were expressed as the median and 95% confidence interval for non-normally distributed data and as the mean and standard deviation for normally distributed data. The processing image volume size (in Liter) and quantitative metrics from different dataset configurations were compared using a nonparametric Mann-Whitney rank test. The SSI values of selected atlases with and without the bounding box detection were compared using a paired Student's T-test. A p-value >0.05 indicated a statistically nonsignificant (NS) difference.

3. Results

For the 60 NHLBI CTA studies, the 3D imaging volume for the entire dataset comprised an average in-plane field-of-view of $181 \times 181 \text{ mm}^2$ (from 131×131 to 221×221) and an average through-plane depth of 120 mm (from 100 to 140) for an average volumetric coverage of 4.13 L (from 2.22 to 6.33). For the 60 MMWHS dataset, the 3D imaging volume comprised an average in-plane field-of-view of $222 \times 222 \text{ mm}^2$ (from 144×144 to 300×300) and an average through-plane depth of 153 mm (from 110 to 188) for an average volumetric coverage of 8.07 L (from 2.76 to 15.92). On average, the 3D volume size of MMWHS dataset was 95% larger than the NHLBI size ($p < 0.01$).

After applying the proposed bounding box detection, the cropped NHLBI 3D volumes comprised an average in-plane field-of-view of $144 \times 158 \text{ mm}^2$ (from 114×123 to 169×193) and an average through-plane depth of 108 mm (from 90 to 126) for an average volumetric coverage of 2.46 L (from 1.66 to 3.32). Likewise, the cropped MMWHS 3D volumes had an average in-plane field-of-view of $152 \times 154 \text{ mm}^2$ (from 115×120 to 186×233) and an average through-plane depth of 130 mm (from 95 to 168) for an average

volumetric coverage of 3.08 L (from 1.76 to 6.25). The average 3D volume size of MMWHS dataset was 25% larger than the NHLBI size ($p<0.01$).

Qualitative examples of the 3D image volume before and after the bounding box detection are shown in Fig. 3a. and Fig. 3b. for NHLBI and MMWHS datasets respectively. For the group data overall, the bounding box detection significantly reduced the volume size of the NHLBI dataset by 39% ($p<0.01$) and the MMWHS dataset by 56% ($p<0.01$). There was a more significant size reduction in MMWHS volumes than the NHLBI dataset ($p<0.01$).

In our CMACS framework for automated CTA image segmentation, an important step to reduce the computation time and improve the segmentation results is the proposed optimal atlas selection strategy based upon a structural similarity index (SSI) score. Fig. 4 shows the effectiveness of this SSI score to measure the similarity between target and atlas images. Pixelwise measurements of SSI maps are computed from a target and two different atlas images and displayed in a calibrated color scale, where red corresponds to higher SSI values and blue indicates lower SSI values. The atlas image in Fig. 4a clearly shows more similar anatomical structures to the target image than the one in Fig. 4b. The corresponding SSI map in Fig. 4c also displays more red and yellow pixels, while the one in Fig. 4d shows more blue and cyan pixels.

Another important feature in our automated segmentation framework is the inclusion of the bounding box detection step as described in the Methods section. Restricting the processing image volume size not only reduces the overall segmentation time but also improves the performance of both optimal atlas selection and non-rigid image registration steps in the framework as demonstrated in Fig. 5. and Fig. 6. For the group data overall, the average SSI value calculated among the selected optimal atlases was increased significantly from 0.13 ± 0.03 to 0.16 ± 0.04 after the bounding box detection ($p<0.01$). Fig. 7. shows a qualitative comparison of the transformed atlas labels with and without using the bounding box detection. The transformed atlas label in Fig. 7b. based on the bounding box show a better match to the target image cardiac structure than the one in Fig. 7c. without using it.

A comparison of the runtime performance between our original CMACS method [1] versus the new framework with bounding box detection is shown in Table I. There was a significant increase of the computational speed by 24% for the NHLBI dataset ($p<0.01$) and 26% for the MMWHS dataset ($p<0.01$) based upon the new framework. A longer computational time in segmenting the MMWHS dataset was due to its larger volume size compared to the NHLBI dataset. Overall, the computational time averaged 1.55 ± 0.32 minutes for the new CMACS framework to segment seven cardiac structures from a CTA study.

For a qualitative evaluation of the proposed automated segmentation versus the manual reference labels, Fig. 8. displays three cases with the maximum, median, and minimum averaged Dice indices obtained from the 40 NHLBI test dataset. Overall, our automatic segmentation shows well-maintained results on all cases, even on the one with the lowest Dice index.

Table II and III summarize the segmentation performance using the Dice index with and without using the bounding box detection (BBD) method. For a four-way cross comparison

of interchanging the atlas and the target images between NHLBI and MMWHS datasets, Table IV and V summarize the quantitative results of our automated framework against the manual references in segmenting seven cardiac structures. These comparisons are based upon using an atlas library consisting of either 20 NHLBI or 20 MMWHS datasets to segment the independent test datasets of 40 NHLBI (Table IV) and 40 MMWHS (Table V) CTA studies. Finally, for a comparison of our results with other published methods in the MMWHS challenge [2], Table VI summarizes the overall quantitative scores of segmenting seven heart substructures from the 40 MMWHS test dataset.

4. Discussion

To our best knowledge, this is the first study to investigate the interplay between different atlas and target image datasets acquired from distinct imaging protocols in multi-atlas based cardiac CTA image segmentation. We present a generalized automated segmentation framework that can process intermixed datasets with distinct image volume coverage. In a four-way cross comparison based on our institutional NHLBI and publicly available MMWHS datasets, the proposed CMACS framework successfully processed the entire group of CTA studies without exclusion. Our quantitative evaluation of the proposed method shows excellent results among different target and atlas data combinations.

In routine contrast enhanced cardiac CTA examinations, the image field-of-view and volume coverage sizes may vary significantly among different sites depending on the image acquisition protocols and scanner parameter settings. For example, the MMWHS image shown in Fig. 3a. has a volume size of 15.8 L which is almost 3 times bigger than the NHLBI image volume of 5.3 L. This discrepancy between volumes is due to a radiation reduction protocol that was used in our institution to minimize the body x-ray exposure during the CTA scan. The volume coverage in our dataset generally starts from the pulmonary trunk to a small extent beyond the apex of the heart. In contrast, the volume coverage in MMWHS datasets often starts from the aortic arch above the pulmonary trunk into a large portion of the liver well-below the heart. Furthermore, our institutional dataset is reconstructed with a restricted field-of-view to avoid inclusion of the lung, whereas the MMWHS dataset covers the whole transverse chest including most of the lung.

These differences in image field-of-view and volume coverage introduce large inconsistencies in captured anatomical structures between the two datasets and highlight the importance of incorporating a volume of interest, such as a bounding box, detection to improve anatomical similarity for both atlas selection and image registration in multi-atlas based segmentation methods. Instead of warping a large volume to a small volume, or vice versa, with inconsistent anatomical structures, the bounding box detection localizes a similar whole heart region among different datasets before applying the non-rigid image registration, thus improving its accuracy. As shown in Fig. 5 to Fig. 7, the bounding box detection improves the segmentation labels as both the optimal atlas selection and the images' registration procedures do not have to consider most of the non-cardiovascular structures outside the whole heart. This is also evident from SSI values improved by incorporating the bounding box detection step for optimal atlas matching.

In the image volume size comparison, we showed that the bounding box detection effectively reduced the NHLBI volume by 39% and the MMWHS volume by 56%. However, we also observed that the cropped MMWHS volumes were still on average 20~25% larger than the cropped NHLBI volumes. This difference was due to a more disparate body volume coverage in some of the MMWHS cases that included a large air space above the chest or included a large portion of the liver. This disparity leads to a relatively small whole heart region compared to the entire 3D image and results in a larger cropped volume due to incomplete separation of the surrounding non-cardiac structures after the bounding box detection.

In the previously published work for cardiac CTA segmentation, Payer *et al.* [4] used a U-net framework and Yang *et al.* [14] used a registration approach to locate the heart region before performing structural segmentation. Our proposed bounding box detection technique performs in an unsupervised fashion that requires no training set [4] nor pre-registration [14] and only takes an average of 10~15 seconds to run.

In the survey of other medical image processing applications, automated bounding box or region of interest (ROI) extraction techniques have also been widely used to improve the segmentation or registration accuracy, and to reduce the amount of data to be processed. For example, Foruzan *et al.* used maximum intensity projection (MIP) images for pre-alignment and liver ROI extraction before applying CT and MR image registration [24]. Liu *et al.* proposed an ROI extraction method based on anatomical knowledge to segment the sternum from low-dose chest CT images [25]. Wu *et al.* also used MIP images and thresholding methods to extract a liver ROI, and then used the SLIC supervoxel method to partition the ROI before applying a graph cuts algorithm to segment the liver from abdominal CT images [26]. Commandeur *et al.* extracted image features such as intensities, gradient vectors, and Haar-like features, and then applied a random forest classifier to estimate a prostate ROI from CT images [27].

Quantitatively comparing our segmentation results with and without using the bounding box detection, Table II and III show a higher Dice index was obtained in most of the structures when segmenting cross-institutional datasets with the bounding box detection. For the segmentation of NHLBI target images based upon the MMWHS atlas, Table II shows an average of 2.9% improvement in the Dice index by applying the bounding box. Similarly, an average of 5.2% improvement in the Dice index was observed in Table III for segmenting the MMWHS target images based upon the NHLBI atlas with the bounding box detection. Furthermore, there was an average of 2.8% Dice improvement in LV, LA, LVM, and RA regions when segmenting the intra-institutional dataset of MMWHS target images based upon the MMWHS atlas with the bounding box detection (see Table III). This result may be explained by the fact that the registration can perform better with a more precise field-of-view containing the whole heart as it does not need to account for optimizing the deformation field outside of the heart. In contrast, segmenting the intra-institutional dataset of NHLBI target images based upon the NHLBI atlas did not show much improvement (see Table II) as the NHLBI dataset was already acquired in a more focused field-of-view around the whole heart; thus the additional bounding box detection did not further improve the segmentation results.

In our four-way cross comparison to evaluate the segmentation performance among different configurations of atlas and target image datasets, both Table IV and Table V showed a better segmentation accuracy by using comparable atlas and target images from the same dataset. For the NHLBI test dataset, Table IV showed a slightly higher overall Dice index of 0.91 by using the comparable NHLBI atlas versus 0.88 from the MMWHS atlases ($p < 0.01$). Five of the seven cardiac structures also showed slightly better Dice or distance scores based upon the NHLBI atlas ($p < 0.01$). Similarly, for the MMWHS CTA test dataset, Table V showed a slightly higher overall Dice index of 0.86 by using the comparable MMWHS atlases versus 0.85 from the NHLBI atlases ($p = \text{NS}$). Four of the seven cardiac structures also showed slightly better Dice or distance scores based upon the MMWHS atlases ($p < 0.01$).

These results demonstrate that the consistency of atlas versus target datasets can affect the segmentation accuracy in our multi-atlas based approach. As the image quality, tissue HU intensity, in-plane and through-plane volume coverages may be influenced by differing clinical imaging protocols, including changes in variables such as contrast enhancement timing and concentration, radiation exposure dose, acquisition field-of-view, volume coverage, slice thickness, and matrix size settings, etc., a change of these imaging parameters may inevitably alter the intrinsic signal intensity across different tissues as well as the image volume coverage, and subsequently affect the robustness of an automated image segmentation method. The proposed bounding box detection addresses the field of view and volume coverage differences between multi-institutional datasets. Another factor that may also impact the segmentation performance is the quality of manual annotation which may be subjective and dependent upon the experience and attention of human observers.

Furthermore, the label generation protocol for manual segmentation is another factor that could influence the segmentation accuracy. In the NHLBI dataset, the manual segmentation of LV, LA, RV, RA, and LVM regions was performed in a similar way as the MMWHS dataset [2] except for the AA and PA structures. Due to a more restricted field of view in our acquisition protocol used to minimize patient radiation exposure, the volume coverage in the NHLBI dataset generally starts from the pulmonary trunk, rather than from the aortic arch above the pulmonary trunk as seen in the MMWHS dataset. Therefore, the AA and PA labels in the NHLBI atlas only cover these two regions up to the pulmonary trunk and below the aortic arch, as opposed to the more extensive coverage seen in the MMWHS atlas. Moreover, our NHLBI dataset also includes an additional four vascular structures, the LAA, SVC, IVC, and PV, which are not labeled in the MMWHS dataset. These additional labels may help to improve the localization of other adjacent structures in a multi-atlas based segmentation.

In comparing our CMACS framework with other published methods based upon the 40 MMWHS CTA test dataset, Table VI shows our results are within the range of other methods. There is a slightly lower Dice index (0.86 vs. 0.87) but better MSD (1.56 vs. 2.12 mm) and HD (9.46 mm vs. 37.68 mm) scores comparing ours versus the average results. Nevertheless, using 40 NHLBI test datasets, our framework obtained a higher segmentation accuracy, with an overall Dice index of 0.91, MSD of 1.25 mm and HD 7.81 mm.

For the computational speed comparison in Table I, our improved CMACS framework with the bounding box detection averages about 1.55 minutes for segmenting seven cardiac structures from a CTA study which is about 25% faster than our original framework [1]. This processing speed is comparable to deep learning-based methods as presented in the MMWHS Challenge [2] and is significantly faster than other multi-atlas based methods which have been reported to take around 21 minutes [2].

5. Conclusion

In summary, we evaluated the impact of using a heterogeneous atlas versus a target image dataset in multi-atlas cardiac structure CTA image segmentation. We present an improved version of our combined multi-atlas and corrective segmentation framework [1] for segmenting different cardiac structures from contrast enhanced CTA studies. This new framework expands our previous method by including a bounding box detection step to improve the speed and robustness for segmenting multi-institutional datasets. We showed that the proposed framework can reliably segment two institutions' differently-acquired datasets with interchangeable atlas libraries. Our quantitative comparisons also showed strong agreements between automated segmentation and manual reference standards for all cardiac structures assessed. These results demonstrate that the proposed CMACS framework may be readily generalized for processing large-scale datasets from other institutions, and for a more comprehensive clinical evaluation.

Acknowledgments

This work was supported by the intramural research program of the National Heart, Lung and Blood Institute (ZIA HL006220).

Biographies



VY BUI received the B.S., M.S., and Ph.D. degrees in electrical engineering from The Catholic University of America, Washington DC, USA, in 2012, 2014, and 2020. She is working on medical image computing research within the Cardiovascular CT laboratory at the National Heart, Lung, and Blood Institute since 2016. Her research interests include image processing, computer vision and deep learning for medical and biomedical image analysis.



LI-YUEH HSU received his M.Sc. and D.Sc. degrees in electrical and biomedical engineering from the George Washington University, Washington DC, USA in 1993 and 1999. He has been a Staff Scientist at the National Institutes of Health in Bethesda, Maryland since 2002. His main research interests include multi-dimensional medical imaging, image computing, machine learning, computer vision, pattern recognition, computer-aided detection and diagnosis, computational modeling in biology and physiology.



SUJATA SHANBHAG received her M.D. degree from St. Louis University School of Medicine in 2000 and completed a M.P.H degree at the Harvard T.H. Chan School of Public Health in 2013. She is a Staff Cardiologist at the National Heart, Lung, and Blood Institute of the National Institutes of Health, Bethesda, Maryland. In her clinical practice, she balances her research interests in translating newer cardiovascular MR and CT imaging methods into patient studies and outcomes research. She has a specific interest in the diagnosis, management of non-ischemic cardiomyopathies, myocardial ischemia and infarction, and evaluating rare disease phenotypes using combined advanced imaging techniques.



LOC TRAN received the B.S. degree in Computer Science from The Catholic University of America, Washington DC, USA in 2020. He has been working as a volunteer at the National Heart, Lung, and Blood Institute since 2020. His research interests include machine learning, computer vision, medical image processing.



W. PATRICIA BANDETTINI received her B.A. and M.D. degrees in a combined program at the University of Missouri-Kansas City in 1995. She then completed an Internal Medicine residency and General Cardiology fellowship at Lutheran General Hospital in Park Ridge, Illinois, and an advanced cardiovascular imaging fellowship at the National Institutes of Health in Bethesda, Maryland. Currently, Dr. Bandettini is a Senior Research Physician within the Cardiology Branch of the National Heart, Lung, and Blood Institute at NIH. She

has clinical and research interests in teaching and promoting cardiovascular MR and CT imaging for various clinical applications.



LIN-CHING CHANG received her M.Sc. and D.Sc. degrees in computer science from the George Washington University, Washington DC, USA in 1993 and 1998. She is an Associate Professor in the Department of Electrical Engineering and Computer Science and the Director of Data Analytics Program at the Catholic University of America. Prior to that, she was a postdoctoral fellow within the Intramural Research and Training Award program at the National Institutes of Health working on computational neuroscience projects with a focus on quantitative image analysis of diffusion tensor MRI. Her main research interests include machine learning, data analytics, computer vision and medical informatics.



MARCUS CHEN is the Director of Cardiovascular CT at the National Heart, Lung, and Blood Institute of the National Institutes of Health in Bethesda, MD, USA. He received his B.S. degree in biology from the California Institute of Technology in 1994 and M.D. degree from the University of Wisconsin in Madison in 1998. As a multi-modality imager in computed tomography, magnetic resonance imaging and nuclear techniques, he has clinical and research interests in technical development of new imaging methods and applications.

References

- [1]. Bui V, Shanbhag SM, Levine O, Jacobs M., Bandettini WP, Chang LC, Chen MY, and Hsu LY. “Simultaneous Multi-Structure Segmentation of the Heart and Peripheral Tissues in Contrast Enhanced Cardiac Computed Tomography Angiography.” *IEEE Access*, vol.8, pp. 16187–16202, 1 2020.
- [2]. Zhuang X, Li L, Payer C, Stern D, Urschler M, Heinrich MP et al., “Evaluation of Algorithms for Multi-Modality Whole Heart Segmentation: An Open-Access Grand Challenge,” *Med. Image Anal*, vol.58, 101537, 8 2019. [PubMed: 31446280]
- [3]. Yang X, Bian C, Yu L, Ni D and Heng PA, “Hybrid Loss Guided Convolutional Networks for Whole Heart Parsing,” in: *Int. Workshop on Stat. Atlases Comput. Models Heart*, 2017, vol.10663, pp.215–223.
- [4]. Payer C, Štern D, Bischof H and Urschler M, “Multi-label whole heart segmentation using CNNs and anatomical label configurations,” in: *Int. Workshop on Stat. Atlases Comput. Models Heart*, 2017, vol.10663, pp.190–198.
- [5]. Ye C, Wang W, Zhang S, and Wang K, “Multi-depth fusion network for whole-heart CT image segmentation,” *IEEE Access*, vol.7, pp.23421–23429, 2 2019.

- [6]. Liu T, Tian Y, Zhao S, Huang X, and Wang Q, “Automatic Whole Heart Segmentation Using a Two-Stage U-Net Framework and an Adaptive Threshold Window,” *IEEE Access*, vol.7, pp.83628–83636, 6 2019.
- [7]. Baskaran L, Maliakal G, Al’Aref SJ, Singh G, Xu Z et al., “Identification and Quantification of Cardiovascular Structures From CCTA: An End-to-End, Rapid, Pixel-Wise, Deep-Learning Method,” *JACC Cardiovasc. Imag.*, vol.13, pp. 1163–1171, 5 2020.
- [8]. van Rikxoort EM, Isgum I, Arzhaeva Y, Staring M, Klein S, Viergever M, et al., “Adaptive local multi-atlas segmentation: Application to the heart and the caudate nucleus,” *Med. Image Anal.*, vol.14, no.1, pp.39–49, 2 2010. [PubMed: 19897403]
- [9]. Kiri li HA, Schaap M, Klein S, Papadopoulou S, Bonardi M, Chen C, et al., “Evaluation of a multi-atlas-based method for segmentation of CCTA data: a large-scale, multicenter, and multivendor study,” *Med. Phys.*, vol.37, no.12, pp. 6279–6291, 12 2010. [PubMed: 21302784]
- [10]. Zuluaga M, Cardoso M, Modat M and Ourselin S, “Multi-atlas propagation whole heart segmentation from MRI and CTA using a local normalised correlation coefficient criterion,” in: *Int. Conf. on Funct. Imaging Model Heart*, vol.7945, pp.174–181, 6 2013.
- [11]. Yang G, Chen Y, Tang L, Shu H & Toumoulin C, “Automatic left ventricle segmentation based on multi-atlas registration in 4D CT images,” in: *IEEE Int. Sym. on Biomed. Img.*, pp.413–416, 4 2014.
- [12]. Zhuang X, Bai W, Song J, Zhan S, Qian X, Shi W, et al., “Multiatlas whole heart segmentation of CT data using conditional entropy for atlas ranking and selection,” *Med. Phys.*, vol.42, no.7, pp.3822–3833, 7 2015. [PubMed: 26133584]
- [13]. Zhuang X, Shen J, “Multi-scale patch and multi-modality atlases for whole heart segmentation of MRI,” *Med. Image Anal.*, vol.31, pp.77–87, 7 2016. [PubMed: 26999615]
- [14]. Yang G, Sun C, Chen Y, Tang L, Shu H and Dillenseger J, “Automatic whole heart segmentation in CT images based on multi-atlas image registration,” in: *Int. Workshop on Stat. Atlases Comput. Models Heart*, vol.10663, pp.250–257, 9 2017.
- [15]. Katouzian A et al., “Hashing-Based Atlas Ranking and Selection for Multiple-Atlas Segmentation,” in: *Int. Conf. on Med. Image Comput. Comput. Assist. Interv.*, vol.11073, pp.543–551, 9 2018.
- [16]. Wang H et al., “Fast Anatomy Segmentation by Combining Coarse Scale Multi-Atlas Label Fusion with Fine Scale Corrective Learning,” *Comput. Med. Imaging Graph.*, vol.68, pp.16–24, 9 2018. [PubMed: 29870822]
- [17]. Bui V, Hsu LY, Chang LC, and Chen MY, “An automatic random walk-based method for 3D segmentation of the heart in cardiac computed tomography images,” in: *IEEE Int. Sym. on Biomed. Img.*, pp.1352–1355, 4 2018.
- [18]. Radhakrishna A, Shaji A, Smith K, Lucchi A, Fua P, and Süsstrunk S. “SLIC superpixels compared to state-of-the-art superpixel methods.” *IEEE transactions on pattern analysis and machine intelligence*, vol. 34, no. 11, pp. 2274–2282, 5 2012. [PubMed: 22641706]
- [19]. “MM-WHS: Multi-Modality Whole Heart Segmentation”. <http://www.sdspeople.fudan.edu.cn/zhuangxiahai/0/mmwhs/>. Accessed 12 Dec. 2019.
- [20]. Van der Walt S, Schönberger J, Nunez-Iglesias J, Boulogne F, Warner J, Yager N, et al., “Scikit-image: Image processing in Python,” *PeerJ*, vol.2, pp. e453, 6 2014. [PubMed: 25024921]
- [21]. Wang Z, Bovik AC, Sheikh H and Simoncelli E, “Image quality assessment: from error visibility to structural similarity,” *IEEE Trans. on Med. Imaging*, vol.13, no.4, pp.600–612, 4 2004.
- [22]. Heinrich M, Jenkinson M, Brady M and Schnabel J, “MRF-based deformable registration and ventilation estimation of lung CT,” *IEEE Trans. on Med. Imaging*, vol.32, no.7, pp.1239–1248, 2 2013.
- [23]. Grady L, “Random walks for image segmentation,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol.28, no.11, pp.1768–1783, 11 2006. [PubMed: 17063682]
- [24]. Foruzan AH, and Motlagh HR, “Multimodality liver registration of Open-MR and CT scans,” *International journal of computer assisted radiology and surgery*, vol.10, no.8, pp.1253–1267, 8 2015. [PubMed: 25556525]

- [25]. Liu S, Xie Y, and Reeves AP, "Segmentation of the sternum from low-dose chest CT images," in: *Medical Imaging: Computer-Aided Diagnosis*, International Society for Optics and Photonics, vol. 9414, pp. 941403, 3 2015.
- [26]. Wu W, Zhou Z., Wu S and Zhang Y, "Automatic liver segmentation on volumetric CT images using supervoxel-based graph cuts," *Computational and Mathematical Methods in Medicine*, vol. 2016, pp. 9093721, 4 2016. [PubMed: 27127536]
- [27]. Commandeur F, Antoine S, Romain M, Mohamed N, Juan DOA, Yan R, Pascal H, Renaud C, and Oscar A. "MRI to CT prostate registration for improved targeting in cancer external beam radiotherapy," *IEEE Journal of Biomedical and Health Informatics*, vol. 21, no. 4, pp.1015–1026, 6 2016. [PubMed: 27333613]

- Evaluate the impact of using a heterogeneous dataset on the performance of multi-atlas segmentation of cardiac CTA images.
- Develop a fully automatic framework to segment CTA datasets acquired from distinct imaging protocols.
- Introduce a generalized technique to detect a bounding box enclosing the heart to improve segmentation speed and accuracy.

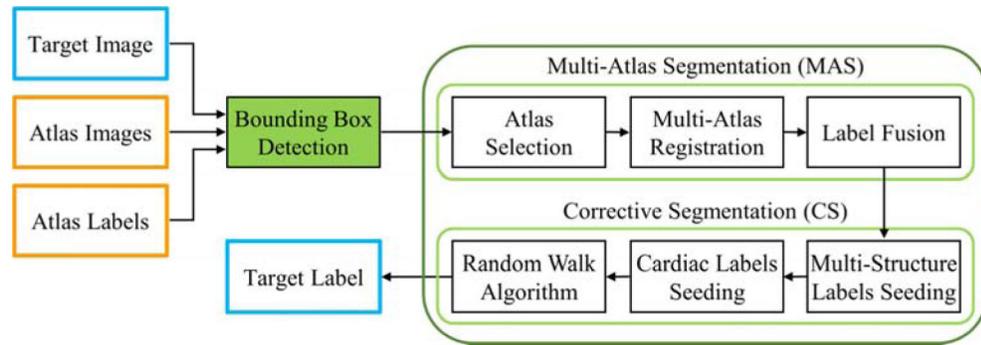


FIGURE 1.

Flow diagram of the proposed new bounding box detection function added to a combined multi-atlas and corrective segmentation (CMACS) framework for fully automatic multi-structure cardiac CTA image segmentation. For a more detailed description of the CMACS framework, we refer the readers to our previous work [1].

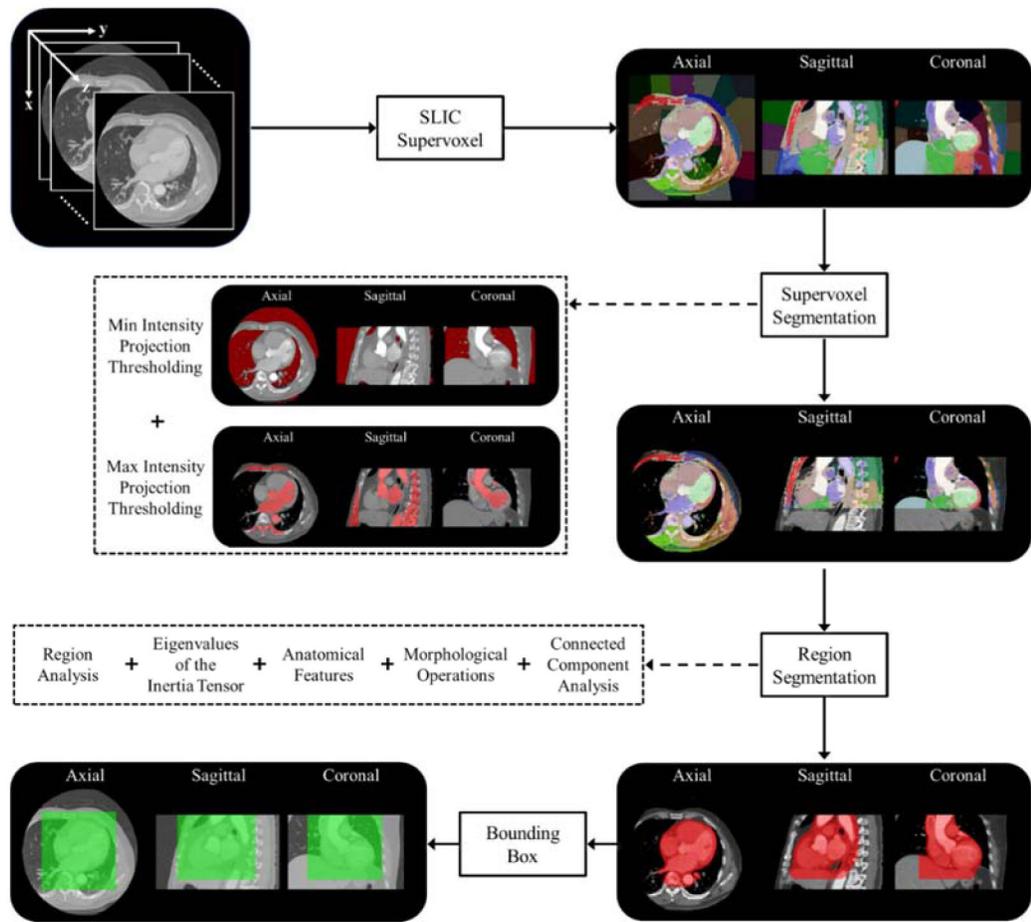


FIGURE 2. Flow chart of the proposed bounding box detection algorithm based on Simple Linear Iterative Clustering (SLIC) supervoxel whole heart segmentation.

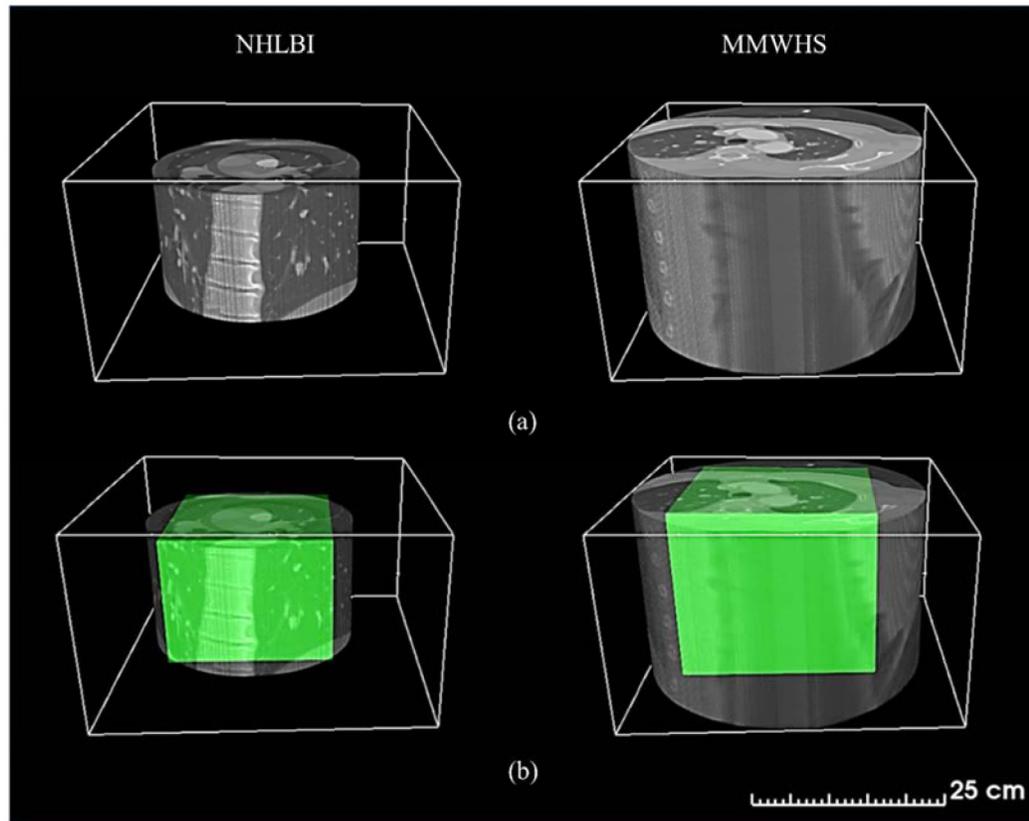


FIGURE 3. Comparison of 3D image volume size between NHLBI and MMWHS datasets. The original volume is shown in row (a). The volume after the bounding box detection is shown in row (b) in green.

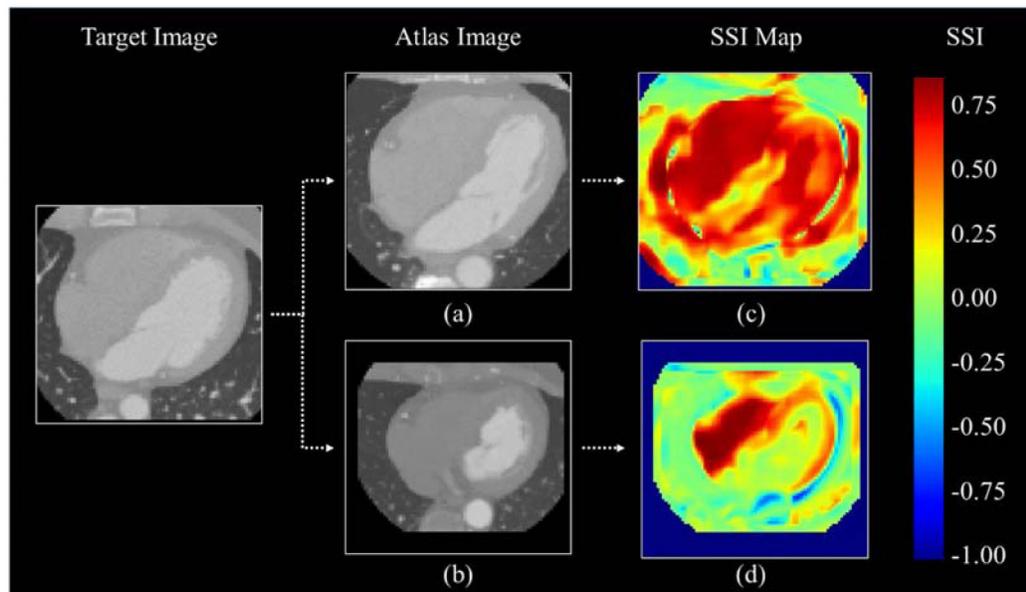


FIGURE 4.

Comparison of structural similarity index (SSI) measurement between a target and two different atlas images. The atlas image in (a) shows more similar anatomical structures with the target image than the atlas image in (b). The corresponding SSI map in (c) also shows higher pixelwise SSI values than the SSI map in (d) with more mixed red and yellow colors.

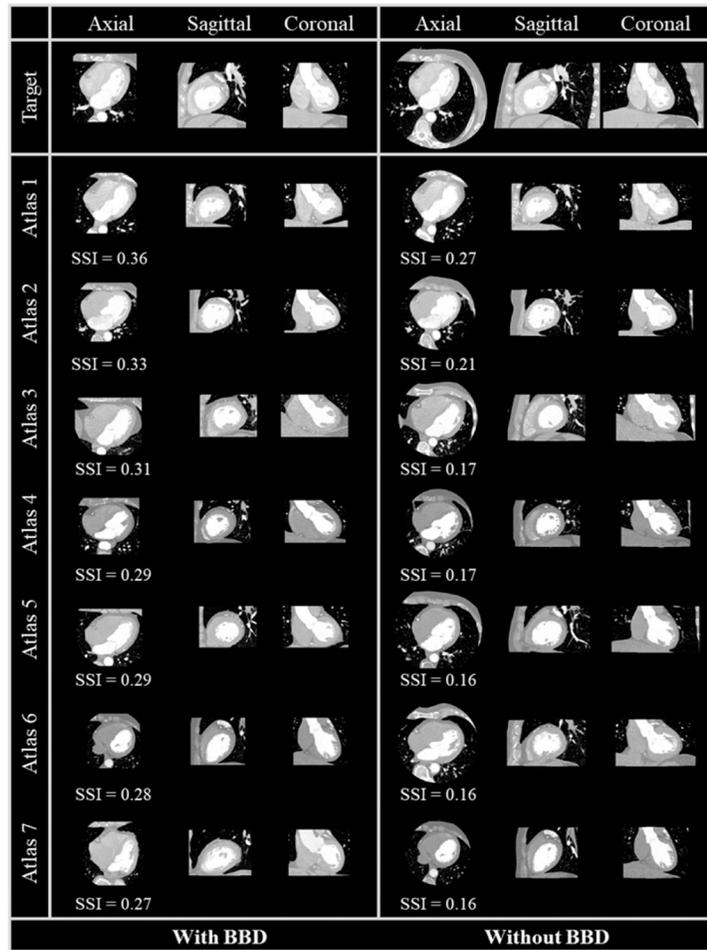


FIGURE 5.

Comparison of optimal atlas selection with and without using the bounding box detection (BBD). The selected seven atlas images based on the BBD (left panel) show a better visual match and increased structural similarity index (SSI) score with the target images (top row) than the atlases without using the BBD (right panel).

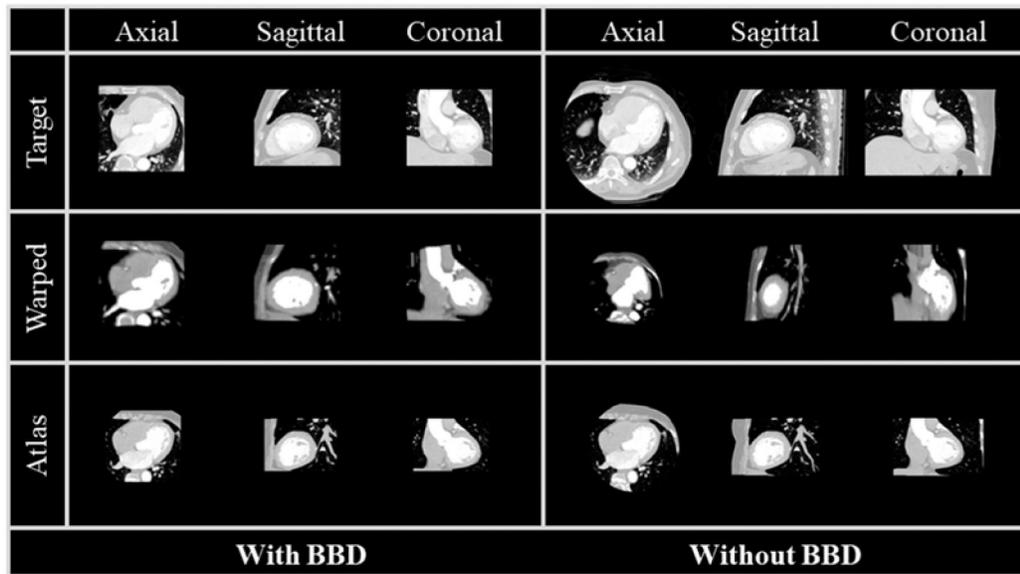


FIGURE 6.

Comparison of atlas-to-target image registration with and with using the bounding box detection (BBD). The warped atlas image based on the BBD (left bottom) show a better visual match to the target images (top row) than the atlas image that does not use the BBD (right bottom).

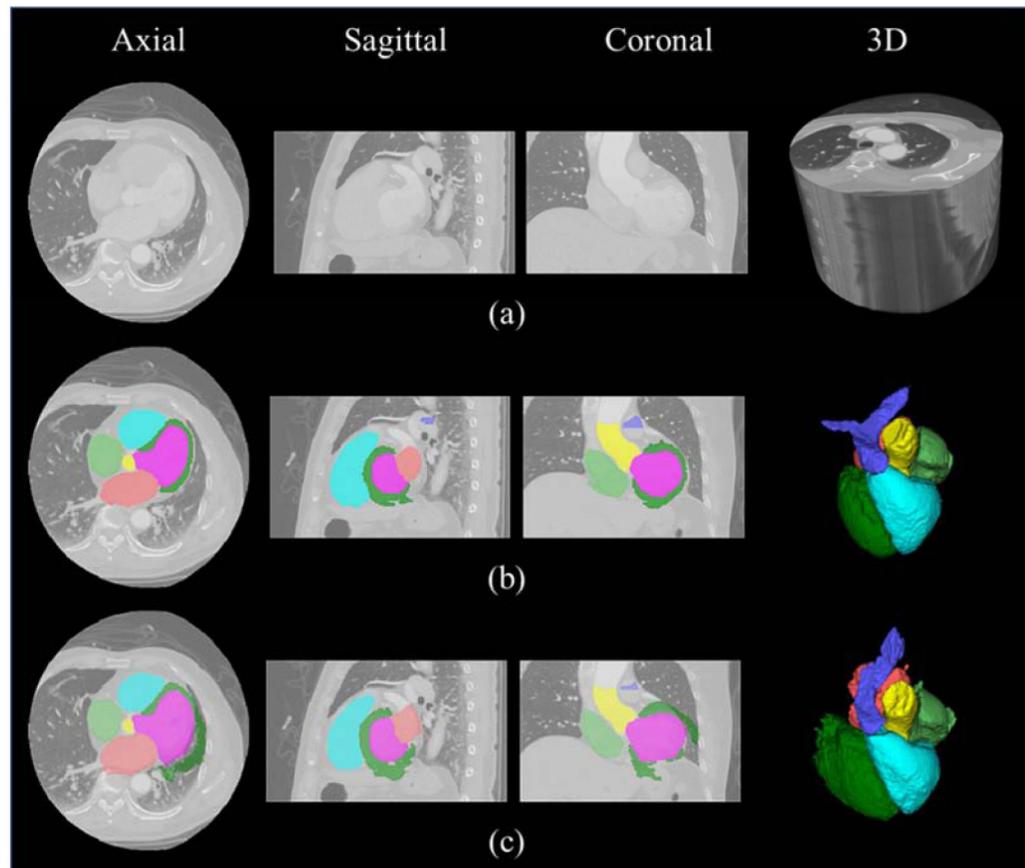


FIGURE 7.

Comparison of the transformed atlas labels (before corrective segmentation) with and without using the bounding box detection. Three orthogonal and 3D rendering views of a target image volume are shown in (a). The transformed atlas labels in (b) based upon the bounding box show a better match to the target image cardiac structures compared to the labels in (c) created without using the bounding box. This improved match is due to improved multi-atlas selection and atlas-to-target image registration that are based upon a more restricted 3D region after the bounding box detection (see Fig. 5 and Fig. 6).

Transformed labels of different cardiac structures are presented in different colors — LV: left ventricular cavity, LA: left atrial cavity, LVM: left ventricular myocardium, RV: right ventricular cavity, RA: right atrial cavity, AA: ascending aorta, PA: pulmonary artery.

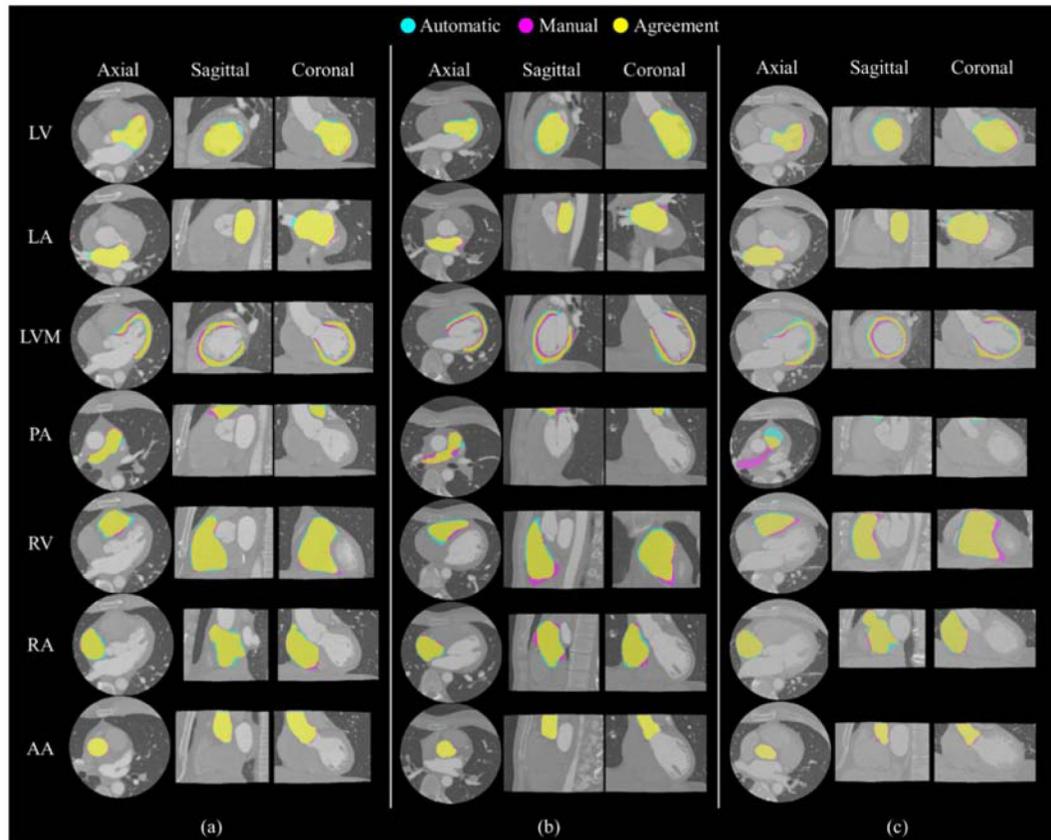


FIGURE 8.

Comparison of automatic vs. manual segmentation results in three orthogonal and 3D views. The cases with (a) maximum, (b) median, and (c) minimum Dice scores from the NHLBI test dataset are shown. Cyan represents automatic segmentation, Magenta represents manual segmentation, and Yellow indicates the agreement areas.

Table I.

Runtime comparison of the automated CMACS framework with and without the bounding box detection (BBD). Results are expressed as mean and standard deviation of segmenting 40 NHLBI and 40 MMWHS CTA datasets.

(minutes)	CMACS (with BBD)	CMACS (without BBD)	p-value
NHLBI	1.43±0.21	1.87±0.31	p<0.001
MMWHS	1.68±0.36	2.28±0.78	p<0.001

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table II.

Performance evaluation of the CMACS framework on the NHLBI CTA dataset using two different atlas libraries. The quantitative Dice index is calculated between the automatic and manual segmentation using two different atlas libraries. Results are expressed as the median and 95% confidence interval.

Atlas	BBD	LV	LA	LVM	PA	RA	RV	AA	All
MMWHS	With BBD	0.88 [0.80, 0.93]	0.85 [0.74, 0.91]	0.71 [†] [0.54, 0.81]	0.66 [0.38, 0.80]	0.84 [0.75, 0.88]	0.85 [0.79, 0.90]	0.86 [0.74, 0.91]	0.83 [0.56, 0.90]
	Without BBD	0.87 [0.75, 0.93]	0.85 [0.72, 0.89]	0.66 [0.46, 0.77]	0.61 [0.43, 0.79]	0.83 [0.66, 0.88]	0.85 [0.76, 0.89]	0.84 [0.64, 0.91]	0.81 [0.49, 0.90]
NHLBI	With BBD	0.89 [0.84, 0.93]	0.91 [0.86, 0.94]	0.80 [0.70, 0.85]	0.60 [0.38, 0.77]	0.89 [0.83, 0.91]	0.90 [0.85, 0.93]	0.89 [0.84, 0.93]	0.89 [0.56, 0.94]
	Without BBD	0.93 [0.90, 0.95]	0.91 [0.88, 0.94]	0.80 [0.67, 0.85]	0.63 [0.46, 0.74]	0.88 [0.83, 0.92]	0.89 [0.83, 0.93]	0.89 [0.83, 0.92]	0.88 [0.57, 0.94]

[†] indicates significantly better performance in a higher Dice score ($p < 0.05$).

Table III.

Performance evaluation of the CMACS framework on the MMWHS CTA dataset using two different atlas libraries. The quantitative Dice index is calculated between the automatic and manual segmentation using two atlas libraries. Results are expressed as the median and 95% confidence interval.

Atlas	BBD	LV	LA	LVM	PA	RA	RV	AA	All
MMWHS	With BBD	0.87 [0.74, 0.91]	0.90 [0.77, 0.94]	0.80 [0.56, 0.89]	0.76 [0.54, 0.91]	0.84 [0.61, 0.89]	0.84 [0.69, 0.90]	0.89 [0.78, 0.92]	0.84 [0.60, 0.92]
	Without BBD	0.85 [0.72, 0.91]	0.89 [0.78, 0.93]	0.75 [0.59, 0.88]	0.81 [0.52, 0.90]	0.83 [0.65, 0.89]	0.86 [0.72, 0.91]	0.90 [0.83, 0.92]	0.84 [0.64, 0.92]
NHLBI	With BBD	0.86 [0.61, 0.94]	0.86 [†] [0.68, 0.92]	0.70 [†] [0.48, 0.82]	0.58 [0.29, 0.77]	0.82 [†] [0.62, 0.88]	0.82 [0.62, 0.89]	0.86 [0.75, 0.90]	0.80 [†] [0.50, 0.90]
	Without BBD	0.81 [0.41, 0.93]	0.81 [0.58, 0.90]	0.64 [0.30, 0.82]	0.54 [0.36, 0.69]	0.79 [0.61, 0.87]	0.80 [0.59, 0.88]	0.85 [0.77, 0.91]	0.76 [0.43, 0.88]

[†] indicates significantly better performance in a higher Dice ($p < 0.05$).

Table IV.

Performance evaluation of the CMACS framework on the NHLBI CTA dataset using two different atlas libraries. Quantitative Dice index, 3D Hausdorff distance (HD) and mean surface distance (MSD) are calculated between the automatic and manual segmentation using two different atlas libraries. Results are expressed as the median and 95% confidence interval.

n=40	Atlas	LV	LA	LVM	PA	RA	RV	AA	All
Dice	NHLBI	0.92 [0.91, 0.94]	0.95 [†] [0.91, 0.97]	0.83 [†] [0.77, 0.87]	0.74 [0.53, 0.85]	0.91 [†] [0.85, 0.93]	0.89 [†] [0.85, 0.93]	0.96 [†] [0.90, 0.97]	0.91 [†] [0.66, 0.97]
	MMWHS	0.92 [0.88, 0.94]	0.92 [0.80, 0.95]	0.79 [0.71, 0.85]	0.78 [0.48, 0.89]	0.88 [0.79, 0.92]	0.88 [0.83, 0.91]	0.95 [0.88, 0.97]	0.88 [0.65, 0.96]
MSD (mm)	NHLBI	1.33 [†] [1.04, 1.68]	0.74 [†] [0.50, 1.12]	1.15 [†] [0.97, 1.60]	1.73 [1.09, 7.68]	1.33 [†] [1.05, 2.15]	1.69 [†] [1.35, 2.49]	0.51 [†] [0.33, 0.93]	1.25 [†] [0.46, 2.49]
	MMWHS	1.47 [1.10, 2.19]	1.17 [0.77, 2.39]	1.35 [1.01, 2.09]	1.91 [1.09, 13.6]	1.93 [1.33, 3.08]	1.98 [1.56, 2.66]	0.60 [0.38, 1.28]	1.54 [0.54, 3.29]
HD (mm)	NHLBI	5.74 [†] [4.47, 8.82]	6.20 [†] [3.74, 16.8]	9.19 [5.38, 17.6]	17.1 [8.46, 75.3]	8.69 [†] [4.87, 18.6]	10.3 [7.0, 18.5]	3.46 [†] [2.24, 6.18]	7.81 [3.00, 23.1]
	MMWHS	7.84 [5.0, 13.3]	9.67 [6.52, 17.1]	9.90 [5.37, 14.8]	18.9 [8.32, 77.4]	12.6 [9.16, 19.3]	11.3 [7.78, 17.2]	4.24 [2.80, 7.68]	10.1 [3.61, 25.8]

[†] indicates significantly better performance in a higher Dice, a lower HD, or a lower MSD ($p < 0.05$).

Table V.

Performance evaluation of the CMACS framework on the MMWHS CTA dataset using two different atlas libraries. Quantitative Dice index, 3D Hausdorff distance (HD) and mean surface distance (MSD) are calculated between the automatic and manual segmentation using two atlas libraries. Results are expressed as the median and 95% confidence interval.

n=40	Atlas	LV	LA	LVM	PA	RA	RV	AA	All
Dice	NHLBI	0.89 [0.78, 0.94]	0.91 [0.81, 0.95]	0.80 [0.68, 0.86]	0.73 [0.50, 0.89]	0.85 [0.66, 0.90]	0.84 [0.61, 0.89]	0.93 [0.86, 0.96]	0.85 [0.62, 0.95]
	MMWHS	0.88 [0.77, 0.93]	0.92 [0.84, 0.95]	0.83 [0.72, 0.88]	0.80 [0.53, 0.90]	0.84 [0.60, 0.91]	0.86 [†] [0.73, 0.91]	0.93 [0.84, 0.96]	0.86 [0.64, 0.94]
MSD (mm)	NHLBI	1.46 [†] [1.01, 2.72]	1.28 [0.80, 2.63]	1.51 [1.01, 3.10]	1.67 [0.86, 4.24]	2.21 [1.51, 3.84]	2.31 [1.64, 3.92]	0.71 [0.38, 1.38]	1.65 [0.62, 3.59]
	MMWHS	1.66 [1.16, 2.98]	1.15 [0.83, 2.28]	1.17 [†] [0.73, 2.10]	1.74 [1.01, 3.47]	2.13 [1.41, 4.44]	1.68 [†] [1.25, 3.39]	0.78 [0.43, 1.54]	1.56 [0.65, 3.36]
HD (mm)	NHLBI	8.88 [5.79, 12.4]	11.4 [6.93, 16.8]	11.4 [7.71, 17.5]	14.0 [5.79, 25.7]	13.33 [8.04, 25.9]	12.0 [7.13, 27.8]	4.52 [†] [2.80, 8.55]	11.4 [4.0, 25.6]
	MMWHS	8.43 [6.38, 14.0]	8.91 [†] [5.78, 17.1]	8.92 [5.07, 19.9]	7.71 [†] [4.10, 14.0]	11.9 [6.93, 23.6]	11.0 [6.93, 25.7]	6.67 [4.10, 11.5]	9.46 [†] [5.00, 22.7]

[†] indicates significantly better performance in a higher Dice, a lower HD, or a lower MSD ($p < 0.05$).

Table VI.

Performance comparison of the proposed CMACS framework and other published methods based on the MMWHS CTA test dataset. Quantitative Dice coefficient, 3D Hausdorff distance (HD) and mean surface distance (MSD) measurements of 10 automated CTA segmentation methods reported in the MMWHS challenge [2] versus our results. The 10 methods in MMWHS challenges are: GUT, KTH, CUHK1, CUHK2, UCF, SEU, SIAT, UT, UB1, and UOE.

Methods	Dice	MD	HD
GUT	0.91 ± 0.09	1.12 ± 0.25	25.2 ± 10.8
KTH	0.89 ± 0.03	1.39 ± 0.52	31.2 ± 13.2
CUHK1	0.89 ± 0.05	1.43 ± 0.59	29.0 ± 15.8
CUHK2	0.89 ± 0.05	1.68 ± 0.59	42.0 ± 16.3
UCF	0.88 ± 0.08	1.54 ± 1.01	28.5 ± 11.4
SEU	0.88 ± 0.02	1.71 ± 0.40	34.1 ± 12.5
SIAT	0.85 ± 0.06	1.93 ± 0.92	44.9 ± 16.1
UT	0.84 ± 0.15	4.81 ± 13.6	34.6 ± 12.4
UB1	0.89 ± 0.03	1.44 ± 0.30	55.4 ± 10.9
UOE	0.81 ± 0.16	4.20 ± 7.78	51.9 ± 17.5
CMACS ¹	0.86 ± 0.04	1.66 ± 0.88	10.2 ± 4.86
CMACS ²	0.84 ± 0.04	1.82 ± 1.00	11.4 ± 5.36

¹CMACS: our results based on the MMWHS atlas.

²CMACS: our results based on the NHLBI atlas. All results are expressed as mean \pm standard deviation.