Multi-Task Deep Convolutional Neural Network for The Segmentation of Type B Aortic Dissection

Jianning Li \(^1\), Long Cao \(^2\), Yangyang Ge\(^2\), Cheng Wang\(^3\), Bowen Meng\(^3\), and Wei Guo\(^*\)\(^2\)

\(^1\)Department of Biomedical Engineering, School of Medicine, Tsinghua University
\(^2\)Department of Vascular and Endovascular Surgery, Chinese PLA General Hospital
\(^3\)Huiying Medical Technology Co., Ltd

Abstract

Type B aortic dissection (TBAD) is a rare but life threatening disease. Segmentation of the entire aorta and true-false lumen is crucial for the planning and follow-up of endovascular repair of TBAD. Manual segmentation in a slice-wise manner is time-consuming and requires expert experience. Current computer-aided methods have several limitations like focusing only on a specific part of the aorta at a time or requiring human interaction. Most importantly, these methods can not segment the entire aorta and detect true-false lumen at the same time. We report in this study a fully automatic approach based on multi-task deep convolutional neural network that segments the entire aorta and true-false lumen from CTA images in a unified framework. For training, we built a database containing 254 CTA images from both pre-operative and post-operative TBAD patients. These images are from multiple manufacturers. Slice-wise manual segmentation of the entire aorta and true-false lumen of each 3-D CTA image is also provided. Our method is evaluated on 7 CTA data (5 preoperative and 2 postoperative) whose ground truth segmentation is provided by experienced vascular surgeons. Results show that our method can segment type B aortic dissection with robustness and accuracy. Furthermore, our method can be easily extended to the segmentation of the entire aorta without dissection.

Keywords: Multi-task Deep Learning, CTA Database, Type B Aortic Dissection, Segmentation

1 Introduction

For patients with TBAD, more and more evidence shows that thoracic endovascular aortic repair (TEVAR) is an optimal treatment option. Treatment planning (e.g., selection of patient-specific stent-graft) as well as follow-up and long-term outcome prediction of TBAD are based on quantitative analysis of the geometrical and biomedical properties of dissection such as maximum transverse diameter of the thoracic aorta, maximum diameter of the true-false lumen and the length of proximal landing zone. But currently, these features are mostly extracted manually or semi-automatically using image post-processing software which is a time-consuming and experience-dependent task. Automatic segmentation of aortic dissection is the prerequisite for developing automatic feature extraction methods and we present here the first step towards this goal. To our best knowledge so far, no methods are reported in literature that can segment the entire aorta and true-false lumen simultaneously in a single model. \([1,2]\) present several different methods to segment the aorta lumen, detect dissection membrane and identify true-false lumen respectively. \([3]\) reports a generative-discriminative model matching method in combination with wavelet analysis for true-false lumen

*Corresponding author: Wei Guo<guowei@301hospital.com.cn>. † authors contribute equally to this work.
segmentation. For the segmentation of the entire aorta, some studies focus only on a specific section of the aorta at a time such as the descending-ascending aorta or the aortic arch and some methods require human initialization. [9] introduces an automatic method for the segmentation of the entire aorta based on Hough transform and 3D level set which is time-consuming. Deep convolutional neural network is another widely used method in medical image segmentation but few published researches report its application in aorta and true-false lumen segmentation. This is partly because of the limited CTA dataset with proper manual segmentation as ground truth for supervised learning. Early neural-network based methods does not generalize well or require multiple network to perform the task. [12] We report in this paper a fully automatic approach based on multi-task deep convolutional neural network that segments the entire aorta and true-false lumen in a unified framework in near real-time. To this end, we first built a database containing 254 CTA images from both preoperative and postoperative TBAD patients. Manual segmentation of the entire aorta and true-false lumen of these data is also provided. Our model originates from 3D U-net where part of the network parameters are shared by three segmentation tasks. We tested our methods on 7 CTA data from 5 preoperative and 2 postoperative patients and dice-coefficient score is used to quantify the model performance. We also evaluate how the number of shared parameters in the proposed multi-task framework influence the model performance.

2 Dataset

2.1 Training Set

We build a dataset containing 254 3D CTA images acquired with anonymity. The data are from 210 preoperative and 44 postoperative TBAD patients and multiple manufactures including GE MEDICAL SYSTEMS, SIEMENS, Philips and Toshiba. These data vary in image quality, noise level, slice thickness and voxel size. As shown in Fig.1(left), CT scans start from neck to pelvis for all patients and stent-graft is visible around true lumen in CTA images from postoperative patients (middle).

![Figure 1](image1.png)

Figure 1: (left) From neck to pelvis, sagittal and coronal view. (middle) CTA image with and without stent-graft of the same patient in axial view. (right) Inhomogeneous contrast agent distribution in false lumen

Inhomogeneity of contrast agent flow The variation of contrast agent distribution in aorta lumens adds to the difficulty of segmentation. As is illustrated in Fig.1(right), contrast agent distribution in false lumen (brown) varies greatly among slices.

Slice-wise Manual Segmentation For each manual segmentation in training set, we define 4 classes for all the voxels in a CT volume: 0 for background, 1 for aorta wall (including the dissection membrane), 2 for true lumen and 3 for false lumen so that the combination of class 1, 2 and 3 is the entire aorta. For each CTA image, the ground truth mask for the entire aorta (green) and true-false (yellow and brown) lumen is manually prepared using 3D Slicer in a slice-wise manner. As is illustrated in Fig.2(Top), our manual segmentation starts from the ascending aortic artery (left coronary artery origination level) and ends before the iliac bifurcation.

Various dissection morphologies We analyse the total 254 data in training set and group them into 8 classes based on their dissection morphologies and relative position of true-false lumen. Fig.3 illustrates the 8 dissection morphologies in axial view. The morphologies include the true-false lumen in right-left position (1 and 6), posterior-anterior position (2, 4 and 7), false lumen surrounding true lumen (3), multichanneled aortic dissection (8) and dissection with aneurysm (5). Data with true-false
Figure 2: Slice-wise manual segmentation. From left to right. (Top) Axial, sagittal and coronal view of mask for descending-ascending aorta, aortic arch and aorta wall. (Middle) True-false lumen view in 3D. Aorta wall and the dissection membrane in 3D and 2D. (Bottom) Binary label for the entire aorta lumen, true lumen, and false lumen.

Lumen in right-left and posterior-anterior position are the most common in training set while other morphologies only account for a small fraction.

Figure 3: Dissection morphologies in training set.

2.2 Test Set

The 3D CTA images in test set are from 7 patients including 5 who have gone through the endovascular repair. Ground truth segmentation of these images is provided by experienced vascular surgeon manually in a slice-wise manner as in training set.
3 Methodology

The architecture of our proposed multi-task deep convolutional neural network is illustrated in Fig.4. It originates from the 3D U-Net [14] and the network parameters are shared by three tasks: segmentation of the entire aorta, true lumen and false lumen. To adapt the manually segmented label discussed previously to our multi-task framework, we derive three binary labels from the manual segmentation of each input volume as is shown in Fig.2(bottom). By defining the three labels this way, the aorta wall and the dissection membrane can be obtained by simply subtracting the label of true lumen and false lumen from the entire aorta lumen label in contrast to [15] which introduces a semi-automatic method for the detection of aorta dissection wall. Our model takes as input the original 3D CTA image and each task produces a output corresponding to the mask of entire aorta, true lumen and false lumen respectively. In Fig.4(a),the three tasks originates from different deconvolutional layers while in Fig.4(b) they share the same conv-deconv parameters except that each task has their own separate convolutional assembly in their path.

![Figure 4: The proposed multi-task framework for aorta and true-false lumen segmentation](image)

3.1 Tailored Preprocessing Strategy Based on Statistical analysis

Considering the huge variation of the input data, we adopt a tailored preprocessing strategy based on the statistical analysis of the data in training set. The statistics of the voxel size and value in the volume of interest (the entire aorta) is shown in table 1.

<table>
<thead>
<tr>
<th>Statistics</th>
<th>Max</th>
<th>Min</th>
<th>Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pixel Size (mm)</td>
<td>1.265</td>
<td>0.416</td>
<td>0.738</td>
<td>0.124</td>
</tr>
<tr>
<td>Slice Thickness (mm)</td>
<td>3.000</td>
<td>0.500</td>
<td>1.219</td>
<td>0.347</td>
</tr>
<tr>
<td>Voxel Size (mm³)</td>
<td>2.180</td>
<td>0.254</td>
<td>0.909</td>
<td>0.322</td>
</tr>
<tr>
<td>Voxel Value</td>
<td>3071</td>
<td>0</td>
<td>/</td>
<td>/</td>
</tr>
</tbody>
</table>

**Voxel Size Normalization** Based on the statistics in Table 1, we can see that the voxel size variation can be largely attributed to the difference of slice thickness. For each input volume, we resample the voxel size to 1 mm³ by adjusting the pixel size to 0.707 mm and slice thickness to 2 mm using cubic spine interpolation.

**Voxel Value Normalization** We first calculate the maximum and minimum value of voxels in the volume of interest and set the voxel value out of that range to zero given an input volume which can substantially reduce computation and disturbance from other tissues. We then normalize the voxel value to zero mean and unit variance to deal with value variation. Fig.5 shows the result.

**Data augmentation based on dissection morphologies** In order to balance the number of data with different dissection morphologies, we assign a higher possibility of data augmentation to the data with less common dissection morphologies. Data augmentation is done by randomly rotating the input volume to a random angle each time.
3.2 Multi-task Objective function

Inspired by [16], we adopt a hybrid loss function to guide the back-propagation for each task. From the viewpoint of dissection segmentation, we not only hope to maximize the global similarity between the binary ground truth and the model output but to minimize their distance. Dice coefficient and Hausdorff distance are metrics to indicate such topological relations. However, these metrics are not differentiable from their original definition and cannot be minimized as loss function directly. We adopt the differentiable approximations to the two metrics provided by [16, 17]. In [16], the dice similarity coefficient between two 3-D point clouds $\varsigma$ and $\rho$ is approximated by Eq. 1.

$$
\ell_{\text{dice}} = -\sum_{c \in C} \frac{2}{N} \sum_{i} N \varsigma_{ic} \rho_{ic}^c \left( \sum_{i} N \varsigma_{ic} + \sum_{i} N \rho_{ic}^c \right)
$$

where $\varsigma_{ic}$ and $\rho_{ic}^c$ represents the $i^{th}$ voxel in the $c^{th}$ volume (where $C = 2$ in our case) of the two 3-D shape.

In [17], the Hausdorff similarity measure between two 2-D points sets $\Omega_1$ and $\Omega_2$ is defined as Eq. 2.

$$
\rho([\Omega_1]_d, [\Omega_2]_d) = \|d_{\Omega_1} - d_{\Omega_2}\|_{C(D)} = \sup_{x \in D} |d_{\Omega_1}(x) - d_{\Omega_2}(x)|
$$

where $C(D)$ represents sets of continuous functions on measurable points $D$ and sup is the least upper bound.

By defining the Hausdorff distance in the way as Eq. 2, the differentiable Hausdorff metric between two 3-D points sets $\Gamma_1$ and $\Gamma_2$ can be written as Eq. 3.

$$
\ell_{\text{hausdorff}}(\Gamma_1, \Gamma_2) = \sum_{i} (|d_{\Gamma_1} - d_{\Gamma_2}|)^2_D
$$

$$
(f)_{\varphi}^2 = \varphi^{-1}\left(\frac{1}{m(D)} \int_D \varphi(f(x))dx\right)
$$

where $\Gamma_{i_1}$ is the contour (2-D point set) of the $i^{th}$ slice (S slices in total) in the label volume $\Gamma_1$ and $\Gamma_{i_2}$ is the convex hull of the $i^{th}$ slice in the output volume $\Gamma_2$ by the model.

we form the loss function for each task by blending the two metrics together as in Eq. 5.

$$
\ell = -\ell_{\text{dice}} + \alpha \ell_{\text{hausdorff}}
$$

Finally, our multi-objective loss function can be written as Eq. 6.

$$
\ell_{\text{total}} = \sum \ell_i
$$

where $i = \text{entire, true, false}$. 
3.3 Postprocessing

The output segmentation is further processed by a joint smoothing strategy to remove extrusions, isolated voxels and smooth surface.

4 Experiments And Results

Training Details  Our multi-task 3D CNNs are implemented in Tensorflow (version 1.4.1). We run 4000 training epochs with learning rate set to 0.001. Training stops when the sum of negative dice score and Hausdorff distance reaches minimum. The process takes approximately 25 hours for both the networks described in Fig.4 on a machine with 125G memory, Intel Xeon E5-4600v4 CPU and two NVIDIA GeForce GTX1080 Ti GPUs. The initialization parameters of the network are drawn from normal distribution and batch size is set to 1 considering the limited GPU Memory. Data augmentation by randomly rotating the input to a random angle is also applied.

Evaluation on Test Set  We evaluate the two models on the test set stated in Section 2. For each test volume, dice coefficient score for the entire aorta, true lumen and false lumen is calculated. We also conduct experiments to quantitatively assess the impact of the location of task branch (Model-1 and Model-2) on model performance. Results are shown in Table 2 and 3.

<table>
<thead>
<tr>
<th>Table 2: Dice Coefficient by Model-1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dice Coefficient</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>-----------------------------</td>
</tr>
<tr>
<td>Entire Aorta</td>
</tr>
<tr>
<td>True Lumen</td>
</tr>
<tr>
<td>False Lumen</td>
</tr>
<tr>
<td>Mean</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 3: Dice Coefficient by Model-2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dice Coefficient</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>-----------------------------</td>
</tr>
<tr>
<td>Entire Aorta</td>
</tr>
<tr>
<td>True Lumen</td>
</tr>
<tr>
<td>False Lumen</td>
</tr>
<tr>
<td>Mean</td>
</tr>
</tbody>
</table>

From the results above we can see that the dice coefficient score for true lumen and false lumen is significantly improved by moving the corresponding task branches backwards (Model-2). A sensible explanation of this phenomenon may be that there are more shared parameters in model-2 than in model-1 for the three segmentation tasks so that the intrinsic relations among these tasks can be fully learned. We also visualize our automatic segmentation result in Fig.6 and Fig.7 which shows that our model is robust against the inhomogeneous distribution of contrast agent in false lumen and can learn the shape variance in difference sections of the aorta lumens. The prediction time for each input volume is less than 0.7 seconds on our machine excluding preprocessing and postprocessing. More segmentation results can be found in appendix.

True-false lumen segmentation and recognition  Based on the segmentation results by our methods, true lumen and false lumen in each slice can be recognized at the same time. Fig.7 illustrates the recognition of true lumen (yellow) and false lumen (brown) in axial view.

Applying to the segmentation of aorta without dissection  While the three tasks in our multi-task framework are closely linked to each other in training, they work independently when applied to test data. Therefore, when a 3-D CTA image without dissection is given, the model can still output the
Figure 6: Example of automatic segmentation result of the entire aorta in 2D (a) and aorta and true-false lumen in 3D (b).

Figure 7: True-false lumen segmentation and recognition

5 Conclusion

We report a fully automatic method based on multi-task deep neural network for the segmentation of type B aortic dissection. The 3D CTA image database we built for training is the first of its kind in China. We adopt a hybrid loss function strategy inspired by previous work to blend differentiable dice similarity coefficient and hausdorff distance together in the loss function of each task. The methods produce satisfactory results on test set regardless of the strong inhomogeneity of false lumen. This work is the first step towards automatic measurement of clinically relevant parameters of aortic dissection and we will be focusing on developing automatic methods for the extraction of these parameters in our following works and explore its potential clinical applications.
Acknowledgements

The TBAD CTA database is built through the joint effort by Huiying Medical Technology Inc(Beijing) and Chinese PLA General Hospital(Beijing).

References


## Appendix

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Entire Aorta</td>
<td>0.934</td>
<td>0.869</td>
<td>0.904</td>
<td>0.890</td>
<td>0.902</td>
<td>0.912</td>
<td>0.914</td>
<td>0.893</td>
<td>0.897</td>
<td>0.902</td>
</tr>
<tr>
<td>True Lumen</td>
<td>0.818</td>
<td>0.868</td>
<td>0.864</td>
<td>0.835</td>
<td>0.837</td>
<td>0.817</td>
<td>0.878</td>
<td>0.875</td>
<td>0.816</td>
<td>0.845</td>
</tr>
<tr>
<td>False Lumen</td>
<td>0.878</td>
<td>0.821</td>
<td>0.780</td>
<td>0.740</td>
<td>0.832</td>
<td>0.821</td>
<td>0.788</td>
<td>0.781</td>
<td>0.754</td>
<td>0.799</td>
</tr>
<tr>
<td>Mean</td>
<td>0.877</td>
<td>0.862</td>
<td>0.849</td>
<td>0.822</td>
<td>0.857</td>
<td>0.850</td>
<td>0.860</td>
<td>0.850</td>
<td>0.822</td>
<td>0.849</td>
</tr>
</tbody>
</table>

Figure 8: More testing segmentation results