Proceedings of the Joint Austrian Computer Vision and Robotics Workshop 2020 DOI: 10.3217/978-3-85125-752-6-24

A Centerline-Guided Approach for Aorta and Stent-Graft Segmentation

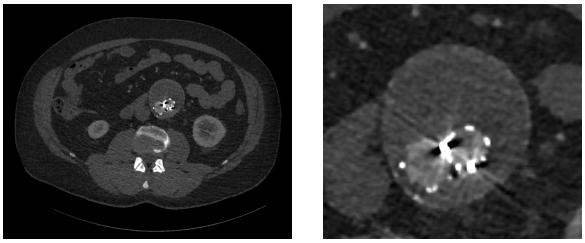
Bertram Sabrowsky-Hirsch, Stefan Thumfart, Richard Hofer, Wolfgang Fenz,
Research Unit Medical Informatics, RISC Software GmbH, Johannes Kepler University, Austria
{bertram.sabrowsky-hirsch,stefan.thumfart,richard.hofer,
wolfgang.fenz}@risc-software.at

Pierre Schmit, Franz Fellner
Central Radiology Institute, Kepler University Hospital, Austria
{pierre.schmit,franz.fellner}@kepleruniklinikum.at

Abstract. *Monitoring of patients after Endovascular aortic repair (EVAR) is a clinical necessity due to the high re-intervention rate associated with the treatment. The risk assessment could be greatly enhanced by the inclusion of metrics based on the aortic blood-flow and stent-graft changes. A preliminary step to this endeavour is, however, the automatic reconstruction of the relevant structures: aortic blood-lumen and the stent-graft wire frame. In this paper we present a centerline-guided approach that leverages knowledge about the target structures through a combination of two 3D U-Nets for efficient automated segmentation of both structures. We evaluate our approach on a real-world clinical dataset yielding Dice similarity coefficients of 0.942 and 0.841 for the blood lumen and stent-graft metal wire, respectively.*

1. Introduction

The abdominal aorta is the largest artery in the human body, with the descending branch supplying the lower body with about 4 liters of blood per minute [2]. Abdominal Aortic Aneurysms (AAAs) are critical as a rupture causes massive blood loss that quickly leads to death at a mortality rate of 85% to 90% [11], with half of the patients succumbing before they reach a hospital [1]. Overall, AAAs account for $175\,000$ deaths per year globally [7]. In contrast to open surgery, endovascular aortic repair (EVAR) poses a minimally invasive alternative that significantly reduces the intraoperative stress on the patients, who in turn experience shorter periods of convalescence. As a result, EVAR is the treatment of choice for 60% of patients [3]. These advantages come, however, at the cost of a high re-intervention rate of 20% [18], necessitating post-operative monitoring of the patients. We seek to aid monitoring by automatically calculating risk factors from blood flow simulations, which require prior segmentation of the target structures. In this paper, we present a novel method for segmenting the aortic blood lumen and the stent graft wire frame from post-operative abdominal CT-A scans.



(a) (b)

Figure 1. An abdominal CT-A scan (a) and a close-up view of imaging artifacts caused by the stent-graft wire frame (b). The wire frame of the Medtronic Endurant stent-graft encompasses the blood lumen in the two iliac bifurcations and is itself surrounded by the thrombosis.

2. Related Work

Blood vessel segmentation is an active field in research [19] and the variety of approaches reflects the diversity of both the targeted anatomical regions and the available imaging modalities. For clinical monitoring of the abdominal aorta after EVAR, CT-A is the modality of choice [22]. However, unique

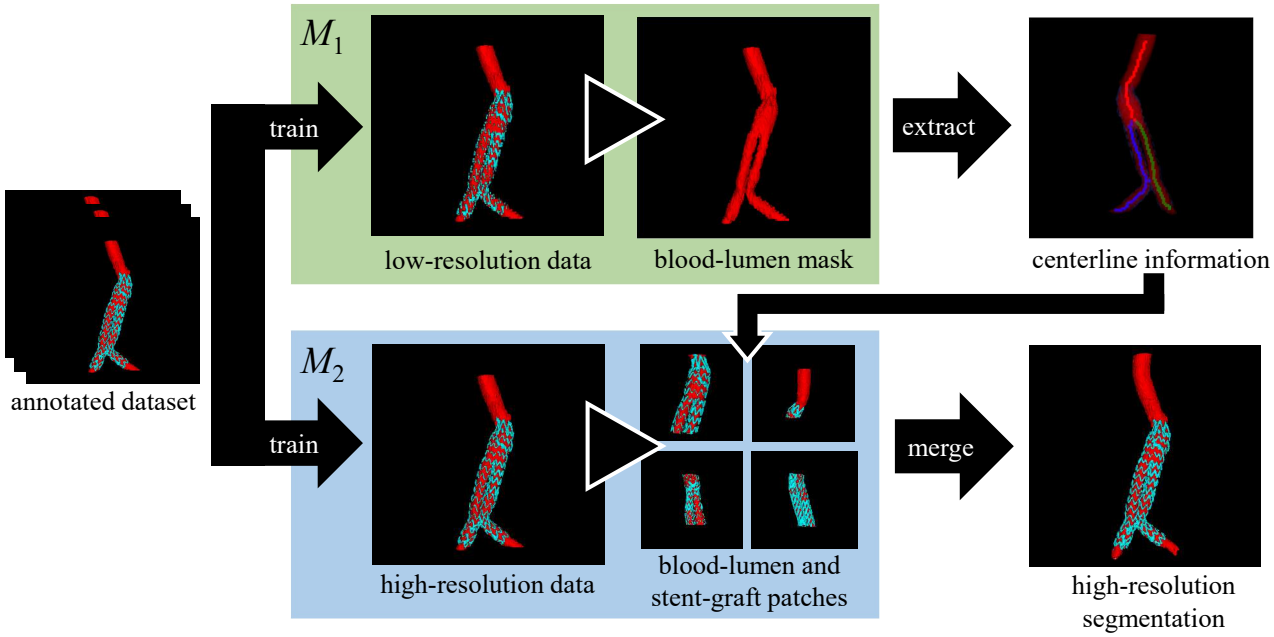
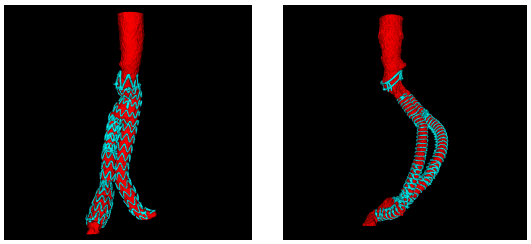


Figure 2. Outline of our method: The top branch shows the centerline extraction step (model M_1) and the bottom branch the patchwise segmentation step (model M_2).

challenges arise due to considerable imaging artifacts caused by the stent-graft wire frame and the distinct boundaries between blood lumen and thrombus. While there are a number of publications on the segmentation of the abdominal aorta, very few have focused on stent segmentation. Klein *et al.* [12] used a graph-based method to create a geometric model of the stent-graft, disregarding the aorta entirely. To the best of our knowledge, there is not a single approach segmenting both structures simultaneously. For the segmentation of the abdominal aorta, traditional approaches include graph-based methods [6, 4, 23] and deformable-models [13, 14] which require user interaction to varying degrees and have predominantly been evaluated on pre-operative scans. A common problem with graph- and deformable-model-based approaches is the introduction of many parameters optimized for the respective dataset, limiting the robustness and applicability of the methods in clinical settings [17]. With the introduction of the convolutional neural network (CNN) the field of medical image analysis changed significantly. Today the U-Net [21] and its 3D equivalent [8] are the most widely models used for medical image segmentation. Both models have been applied to the task of the abdominal aorta segmentation, Zheng *et al.* [26] reporting a Dice similarity coefficient (DSC) of 0.82 for the aneurysm thrombus and Li *et al.* [16] reporting a DSC of 0.92 for the aorta blood lumen. For the segmentation of blood lumen and stent graft wire frame we will therefore likewise rely on the (3D) U-Net architecture. The distinguishing challenge to other segmentation tasks is in our case the fine structure of the stent-graft, with a diameter as small as $0.4\,\mathrm{mm}$ [24], which requires an exceptionally high resolution for accurate reconstruction, pushing the limitations of modern hardware.

3. Dataset

Our dataset consists of 76 abdominal CTA scans of 36 patients treated with EVAR that we received from the Kepler University Hospital Linz. Each scan consists of 155 to 873 axial slices with 512×512 voxels. There are large differences in the resolution with a minimum voxel spacing ranging from $0.404\,\mathrm{mm}$ frontal/sagittal and $0.8\,\mathrm{mm}$ longitudinal



(a) (b)

Figure 3. Examples of two ground truth segmentations: Medtronic Endurant (a) and Anaconda (b). In total the dataset contains 5 different types of stent-grafts.

to $0.977\,\mathrm{mm}$ frontal/sagittal and $3\,\mathrm{mm}$ longitudinal. We used the Active-Contour/Snake-Mode of the Software *ITK-Snap* [25] to semi-automatically create the initial ground truth segmentation of the aortic blood lumen from below the heart to the second iliac bifurcation. The stent-graft segmentation was further added by applying a threshold to a region of interest around the blood lumen. The segmentation was then revised using the Paintbrush Mode of *ITK-Snap*. Figure 3 shows examples of the final ground truth segmentations used for training and validation. The dataset was split into 5-folds using a grouping criterion on the patient number to avoid having multiple scans of the same patients assigned to different folds.

4. Method

We use a two step approach in our segmentation method that is outlined in Figure 2. First we extract the aortic centerlines from a coarse blood-lumen segmentation and subsequently use them to extract high resolution patches along the entire span of the aorta. In the second step we segment the blood lumen and the stent-graft wire frame for each patch and merge the results to a final segmentation. The entire setup is tuned to work with an *NVIDIA GeForce 1080 Ti* (11 GB RAM).

4.1. Centerline Extraction

We use a full-image segmentation model M_1 to create a low resolution segmentation of the aortic blood-lumen. We resample the scans and ground truth to a voxel spacing of $1\,\mathrm{mm}$ frontal/sagittal and $3\,\mathrm{mm}$ longitudinal and crop them to a large region of interest of 192×192 voxels and 128 slices (i.e., a physical extent of $192\,\mathrm{mm}$ frontal/sagittal and $384\,\mathrm{mm}$ longitudinal). The largest connected region of blood lumen voxels in the resulting segmentation is then selected and skeletonized using homotopic thinning [15]. Using the python library *Skan* [20] we extract the centerline graph from the skeletonized images, which is essential for the patchwise segmentation step. Figure 4 outlines the intermediate results of the centerline extraction step and an example patch.

4.2. Patchwise Segmentation

A patchwise segmentation model M_2 is used to segment the aortic blood lumen and the stent-graft wire frame in high resolution patches. We resample the scans and ground truth to a voxel spacing of $0.35\,\mathrm{mm}$ frontal/sagittal and $0.75\,\mathrm{mm}$ longitudinal before extracting patches of size 160×160 voxels and 128 slices. Choosing a high resolution (i.e., small voxel spacing) significantly reduces the amount of distortion introduced by resampling, especially considering the varying voxel spacing in the dataset. However, this results in a rather small physical extent of $56\,\mathrm{mm}$ frontal/sagittal and $96\,\mathrm{mm}$ longitudinal that we seek to use as efficiently as possible by centering the patches at equally distributed locations along the entire centerline graph. In our experiments 100 patches per scan proved more than sufficient to cover the aorta and introduce a significant overlap between the patches. The patches are merged into a final segmentation using a Gaussian-weighted kernel that attenuates voxels at the patch boundaries, where the segmentation results are less reliable.

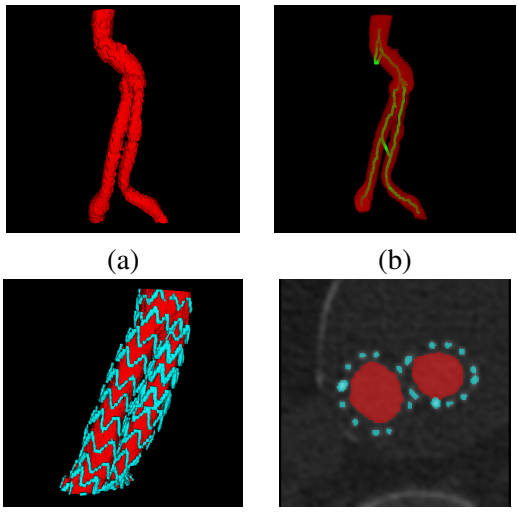


Figure 4. Extracting patches along the segmented aorta lumen: (a) coarse low resolution segmentation of the aortic blood lumen, (b) centerlines approximated via skeletonization of the blood-lumen, (c) ground-truth segmentation for a patch sampled along the centerlines, (d) axial slice of the same patch (scan overlayed with the ground-truth segmentation).

5. Implementation

In this section we discuss the implementation details, i.e., the operations used for preprocessing the dataset, the model architecture and configuration and the training routine.

5.1. Preprocessing

Preprocessing of the dataset consists, in addition to the resampling mentioned in Section 4, of clipping and normalization. As the voxel spacing varies

between the models, the entire preprocessing is done separately for each model. First of all, the dataset is resampled to the respective voxel spacing using a third order B-spline interpolation for the scans and a label-linear interpolation for the ground truth. Next, the intensity values are clipped to the 0.5th and 99.5th percentile over the entire training dataset of the fold. Furthermore, the scans are normalized by subtracting the mean and the standard deviation over the clipped training dataset.

5.2. Architecture

We use the architecture described by Isensee *et al.* [9] and implemented in the Github project 3DUnetCNN [5] as a basis for our experiments. We adjusted the following model parameters: input size, model-depth (number of layers), number of segmentation levels (used for deep supervision) and base-filters (filters in the first convolution kernel). For M_1 (input size of $192 \times 192 \times 128$) we selected a model-depth of 5 with 3 segmentation levels and base-filters set to 8. For M_2 on the other hand (input size of $160 \times 160 \times 128$), we chose an increased model-depth of 6 with 4 segmentation levels and base-filters set to 16. The changes to M_2 were made in order to account for the larger patch size (compared to 128^3 used by Isensee *et al.*) and increase the receptive field of the model. These changes were omitted for M_1, which encompasses a simpler segmentation task, creating only a coarse segmentation of the blood lumen label, while M_2 segments both the blood lumen and the stent-graft wire frame.

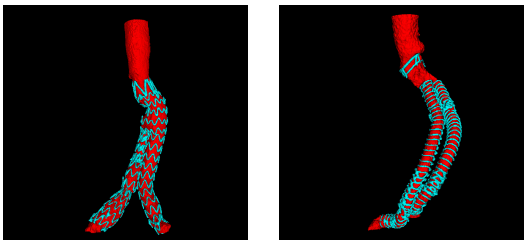
5.3. Training

We trained both models using a weighted multi-class Dice loss [9] in combination with an Adam optimizer. The initial learning rate was set to $\eta_0 = 5 \cdot 10^{-4}$ with a learning rate drop criterion and early stopping after 50 epochs. The training ran for 70 to 120 epochs with 200 training samples per epoch. Due to the 5-fold cross validation used for evaluation, the following statistics are averaged over all folds, where for each fold both models M_1 and M_2 were trained as follows. M_1 was trained first for blood lumen segmentation on the low resolution large regions. The training reached a DSC of 0.978 and 0.898, on average, for the training and validation items, respectively. M_1 was then used to create the blood lumen segmentations for centerline extraction. The resulting centerline graphs were subsequently used during the training of M_2 as the high resolution patches were extracted at random positions along the graph. The average training and validation DSCs for the blood lumen are 0.954 and 0.943, respectively, and 0.843 and 0.841 for the stent-graft.

6. Evaluation

Having trained two models M_1 and M_2 for each fold, we use our method to create high resolution segmentations. Just like during training, M_1 is used to segment the blood lumen used for centerline extraction. The resulting centerline graph is again used to place patches at, however, not randomly but rather at equally distributed positions along the entire span of the graph, as described in Section 4.2. In a post-processing step, the largest connected region of non-background voxels was selected. To compare the results to the ground truth, the segmentations where furthermore resampled to their original voxel-spacing. The last step may be skipped when using the results for further processing rather than evaluation (e.g., mesh generation for blood-flow simulations). Using our method, the cross validation yields an average DSC of 0.961 for the blood lumen and 0.841 for the stent-graft label. Two examples are shown in Figure 5.

To evaluate the effectiveness of our patch extraction method, we further conducted an experiment using only M_2, which was trained using a traditional patch extraction method (see Isensee et al. [10]). Rather than placing the patches along the aorta centerlines, they where placed in a sliding-window fashion, where the patches are aligned in a regular grid of overlapping tiles. The overlap was set to 32 voxels in each dimension (corresponding to $11.2\,\mathrm{mm}$ frontal/sagittal and $24\,\mathrm{mm}$ longitudinal). While this technique was used both during training and inference, the remaining setup (including pre- and post-processing) was left unchanged. We evaluated



(a) (b)

Figure 5. Evaluation results for the two scans shown in Figure 3.

the experiment on the first fold and compared the results to those of our method for the same fold: Although the blood-lumen segmentation is with a DSC of 0.963 slightly better than our method, yielding 0.951 for the same fold, the more complex stent-graft segmentation does not compare well, with a DSC of 0.785 versus our method's score of 0.852.

7. Discussion

The strength of our method is the centerline-guided segmentation method using the aortic centerlines to optimize the patch locations during training and inference. While our method yields better results than a comparable model using a traditional sampling setup, it also reduces the computational cost significantly. For traditional patching using a grid of overlapping tiles (when not allowing the patches to contain regions outside the image) the number of patches calculates as follows:

$$n_{patches} = \prod_{d=1}^{n_d} \lceil \frac{|I|_d - \theta_d}{|P|_d - \theta_d} \rceil \tag{1}$$

where n_d is the number of dimensions, $|I|_d$ the size of an image, $|P|_d$ the size of a patch and θ_d the overlap in dimension d. For our setup and a chosen overlap of 32 voxels this results in 343 patches on average per scan ($|I| = (990, 990, 678)$, $|P| = (160, 160, 128)$, $\theta_d = (32, 32, 32)$). Increasing this overlap to improve the model's performance quickly raises this number, e.g., an overlap of half the patch size (as used by Isensee *et al.* [10]) would result in 1440 patches on average per scan ($\theta_d = (80, 80, 64)$). The majority of the patches are irrelevant for the result, as they do not intersect with the target structure. By using the aorta centerline information, our method is able to greatly reduce the number of patches, while also optimizing their content for training and inference. As a result, we can target a smaller voxel spacing (which effectively reduces the physical extent of the patches) without the disadvantages of excessive computational costs and poor model performance.

Conclusions

We presented a novel centerline-guided method for fully automated segmentation of the aortic blood-lumen and the stent graft wire frame in abdominal CT-A scans. Using our method, both training and inference can be conducted more efficiently. The evaluated DSC of 0.961 for the blood lumen and 0.841 for the stent graft wire frame suggest results that are suitable for medical analysis. In the future, we plan to use the results of our method for the analysis of risk factors for post-EVAR patients. Futhermore, we plan to extend the use of our method to other medical segmentation tasks.

Acknowledgements

This work was funded by the FFG (Austrian Research Promotion Agency) under the grants 851461 (EndoPredictor), 872604 (MEDUSA) and 867536 (vizARd). This project was supported by the strategic economic and research programme "Innovatives OÖ 2020" of the province of Upper Austria. RISC Software GmbH is a Member of UAR (Upper Austrian Research) Innovation Network.

References

[1] S. Aggarwal, A. Qamar, V. Sharma, and A. Sharma. Abdominal aortic aneurysm: A comprehensive review. *Experimental & Clinical Cardiology*, 16(1):11–15, 2011.

[2] M. Amanuma, R. H. Mohiaddin, M. Hasegawa, A. Heshiki, and D. B. Longmore. Abdominal aorta: characterisation of blood flow and measurement of its regional distribution by cine magnetic resonance phase-shift velocity mapping. *European Radiology*, 2(6):559–564, Dec. 1992.

[3] A. W. Beck, A. Sedrakyan, J. Mao, M. Venermo, R. Faizer, S. Debus, C.-A. Behrendt, S. Scali, M. Altreuther, M. Schermerhorn, B. Beiles, Z. Szeberin, N. Eldrup, G. Danielsson, I. Thomson, P. Wigger, M. Björck, J. L. Cronenwett, and K. Mani. Variations in abdominal aortic aneurysm care: A report from the international consortium of vascular registries. *Circulation*, 134(24):1948–1958, 2016.

[4] J. Egger, B. Freisleben, R. Setser, R. Renapuraar, C. Biermann, and T. O'Donnell. Aorta segmentation for stent simulation. *Computing Research Repository - CORR*, 2011.

[5] D. G. Ellis. 3DUnetCNN. `github.com/ellisdg/3DUnetCNN`, 2018. Accessed: 2020-02-11.

[6] M. Freiman, S. J. Esses, L. Joskowicz, and J. Sosna. An iterative model-constrained graph-cut algorithm for abdominal aortic aneurysm thrombus segmentation. In *2010 IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, pages 672–675, 2010.

[7] D. P. J. Howard, A. Banerjee, J. F. Fairhead, A. Handa, L. E. Silver, and P. M. Rothwell. Age-specific incidence, risk factors and outcome of acute

abdominal aortic aneurysms in a defined population. *The British Journal of Surgery*, 102(8):907–915, July 2015.

[8] z. Çiçek, A. Abdulkadir, S. S. Lienkamp, T. Brox, and O. Ronneberger. 3d U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation. In *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2016*, Lecture Notes in Computer Science, pages 424–432. Springer, Cham, 2016.

[9] F. Isensee, P. Kickingereder, W. Wick, M. Bendszus, and K. H. Maier-Hein. Brain tumor segmentation and radiomics survival prediction: Contribution to the brats 2017 challenge. In A. Crimi, S. Bakas, H. Kuijf, B. Menze, and M. Reyes, editors, *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, pages 287–297, Cham, 2018. Springer International Publishing.

[10] F. Isensee, J. Petersen, S. A. A. Kohl, P. F. Jäger, and K. H. Maier-Hein. nnU-Net: Breaking the Spell on Successful Medical Image Segmentation. *arXiv:1904.08128 [cs]*, 2019. arXiv: 1904.08128.

[11] K. C. Kent. Abdominal Aortic Aneurysms. *New England Journal of Medicine*, 371(22):2101–2108, Nov. 2014.

[12] A. Klein, J. A. van der Vliet, L. J. Oostveen, Y. Hoogeveen, L. J. S. Kool, W. K. J. Renema, and C. H. Slump. Automatic segmentation of the wire frame of stent grafts from ct data. *Medical image analysis*, 16 1:127–39, 2012.

[13] T. Kovács, P. Cattin, H. Alkadhi, S. Wildermuth, and G. Székely. Automatic Segmentation of the Vessel Lumen from 3d CTA Images of Aortic Dissection. In H. Handels, J. Ehrhardt, A. Horsch, H.-P. Meinzer, and T. Tolxdorff, editors, *Bildverarbeitung für die Medizin 2006*, Informatik aktuell, pages 161–165, Berlin, Heidelberg, 2006. Springer.

[14] F. Lalys, V. Yan, A. Kaladji, A. Lucas, and S. Esneault. Generic thrombus segmentation from pre- and post-operative cta. *International Journal of Computer Assisted Radiology and Surgery*, 12(9):1501–1510, 2017.

[15] T. C. Lee, R. L. Kashyap, and C. N. Chu. Building Skeleton Models via 3-D Medial Surface Axis Thinning Algorithms. *CVGIP: Graphical Models and Image Processing*, 56(6):462–478, 1994.

[16] J. Li, L. Cao, Y. Ge, C. Wang, B. Meng, and W. Guo. Multi-task deep convolutional neural network for the segmentation of type B aortic dissection. *CoRR*, abs/1806.09860, 2018.

[17] K. López-Linares, I. García, A. García-Familiar, I. Macía, and M. Á. G. Ballester. 3d convolutional neural network for abdominal aortic aneurysm segmentation. *CoRR*, abs/1903.00879, 2019.

[18] M. Malina. Reinterventions after open and endovascular AAA repair. *The Journal of Cardiovascular Surgery*, 56(2):257–268, 2015.

[19] S. Moccia, E. De Momi, S. El Hadji, and L. S. Mattos. Blood vessel segmentation algorithms — Review of methods, datasets and evaluation metrics. *Computer Methods and Programs in Biomedicine*, 158:71–91, 2018.

[20] J. Nunez-Iglesias, A. J. Blanch, O. Looker, M. W. Dixon, and L. Tilley. A new Python library to analyse skeleton images confirms malaria parasite remodelling of the red blood cell membrane skeleton. *PeerJ*, 6:e4312, Feb. 2018.

[21] O. Ronneberger, P. Fischer, and T. Brox. *U-Net: Convolutional Networks for Biomedical Image Segmentation*. Springer International Publishing, Cham, 2015.

[22] A. M. Rozenblit, M. Patlas, A. T. Rosenbaum, T. Okhi, F. J. Veith, M. P. Laks, and Z. J. Ricci. Detection of Endoleaks after Endovascular Repair of Abdominal Aortic Aneurysm: Value of Unenhanced and Delayed Helical CT Acquisitions. *Radiology*, 227(2):426–433, May 2003.

[23] T. Siriapisith, W. Kusakunniran, and P. Haddawy. Outer wall segmentation of abdominal aortic aneurysm by variable neighborhood search through intensity and gradient spaces. *Journal of Digital Imaging*, 31, 2018.

[24] R. Winder, Z. Sun, B. Kelly, P. Ellis, and D. Hirst. Abdominal aortic aneurysm and stent graft phantom manufactured by medical rapid prototyping. *Journal of medical engineering & technology*, 26(2):75–78, 2002.

[25] P. A. Yushkevich, Yang Gao, and G. Gerig. ITK-SNAP: An interactive tool for semi-automatic segmentation of multi-modality biomedical images. *Conference proceedings : ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual Conference*, 2016:3342–3345, 2016.

[26] J.-Q. Zheng, X.-Y. Zhou, Q.-B. Li, C. Riga, and G.-Z. Yang. Abdominal Aortic Aneurysm Segmentation with a Small Number of Training Subjects. *arXiv:1804.02943 [cs]*, 2018. arXiv: 1804.02943.