Automatic aorta and left ventricle segmentation for TAVI procedure planning using convolutional neural networks

Adriana Złahoda-Huzior[1], Maciej Stanuch[1], Jan Witowski[2] and Dariusz Dudek[3]

***Abstract*— Transcatheter aortic valve implantation (TAVI) is a minimally invasive procedure which is performed on patients with aortic valve defects that are posing a high-risk for conducting a surgical treatment. Preoperative surgical planning and valve sizing play a crucial role in reducing surgery complications and adverse effects such as paravalvular leakage or stroke. Planning process incorporates performing measurements, detecting landmarks and visualizing relevant structures in 3D. To automatize this process, a segmentation is required. Due to the lack of methods enabling parallel aorta and left ventricle segmentation we propose a fully automatic neural network approach based on 2D U-Net architecture. Convolutional neural network architecture was trained on 44 studies (22 raw CTA datasets and 22 elastic deformed scans) and tested on another 18 stacks of data. During every epoch of network learning process cross validation was performed on 8 stacks. As a result, we achieve 0.95 mean Dice coefficient score with standard deviation 0.02 determining high precision of predicted aorta and left ventricle label maps.**

I. INTRODUCTION

Computed tomography angiography imaging (CTA) is commonly used for planning surgical procedures. Thanks to the CT volumetric scans of the inner body organs it is possible for the medical specialists to prepare the procedure specifically to the anatomy of the patient, taking into consideration all of the potential risks and difficulties that may occur during the implantation. An example of a very popular non-invasive surgical procedure that uses the CT scans for the purpose of planning is the transcatheter aortic valve implantation (TAVI). During the TAVI procedure defective human aortic valve is replaced via a catheter system. To perform this action accurate aorta and left ventricle measurements are necessary for the proper landmark detection and the preoperative valve size prediction. Errors in the measurements of only few millimeters may cause surgery complications and adverse effects such as paravalvular leakage, valve embolization or a stroke [1]. That emphasises the significance of proper patient data preparation and visualization. For this reason, the main aim of the work is to propose a fully automatic aorta and left ventricle segmentation method being the crucial part of the automatic image preparation for the TAVI planning. Recently, a lot of different algorithms for fully-automatic heart clustering are developed because of time-consuming manual segmentation process. There are two most popular approaches to face this problem - deterministic and based on the artificial intelligence.

II. STATE-OF-THE-ART

Primarily, automatic aorta segmentation algorithms were developed for the analysis of the aortic aneurysm [2] [3]. It was also used for instance for the purpose of the aortic calcium detection [4].

Along with the increasing popularity of the TAVI procedure there were more and more applications employing the automatic aorta segmentation [5]. The basic approach assumes segmentation based on the thresholding and region growing clustering where the threshold level is adjusted by the user [6]. In [7] the authors proposed watershed segmentation and an adaptive erosion technique for aortic root segmentation. Another approach for aorta and the left ventricle segmentation for TAVI planning assume model-based segmentation. The method finds the position of the heart with an adapted generalized Hough transform and then uses this information to adapt the model to the given data. The algorithm was tested on 20 different CT volumes [8]. Instead of using a pre-prepared model, the pre-computed anatomy label maps (ALM) approach was proposed in [9] with Dice score of 0.933. Twenty scans were used as a base for setting the parameters of the algorithm and it was later tested on the 339 low-dose non-contrast CT images from VIA-ELCAP and LIDC public image databases. A fuzzy classification algorithm used along with thresholding, morphological operations and connected component analysis, proposed in [10] was tested on 20 CTA cases and the obtained Dice coefficient was 0.95.

Lately, deep convolutional neural networks were proposed for the task of object detection. It was used already as a way of fully automatic detection and segmentation of different structures in the CT images. After the procedure of training the neural network no additional tuning of the parameters for a certain CT image is needed. An example of neural network usage for segmentation of abdominal aortic thorumbus in post-operative CTA images was proposed in [11]. What is important, both 3D [12] and 2D neural networks [13] can be used. There is an approach to do segmentation in low-dose, non-contrast-enhanced chest CT scans using convolutional neural network and the mean Dice coefficient for different segments of the aorta ranges from 0.83 to 0.88. This was

*This work was funded by NAWA-PROM, grant no. PPI/PRO/2018/1/00026/U/001.

**All experiments involving human participants were in accordance with the ethical standards of the institutional research committee and with the 1964 Helsinki Declaration and its later amendments.

[1]AGH University of Science and Technology, Department of Measurement and Electronics, Cracow, Poland zlahodahuzior@agh.edu.pl, stanuch@agh.edu.pl

[2]Jagiellonian University Medical College, 2nd Department of General Surgery, Cracow, Poland jwitos@gmail.com

[3]Jagiellonian University Medical College, Institute of Cardiology, Cracow, Poland mcdudek@cyfronet.pl

978-1-5386-1311-5/19/$31.00 ©2019 IEEE

tested on 20 CT scans, 10 for the training and 10 as a test group [14].

A progress has been made to utilize not only CT, but also intraoperative C-arm imaging in TAVI planning, and the proposed algorithm was Marginal Space Learning for 3D Object Segmentation [15].

III. METHODS

A. Data augmentation

One of the most severe problems faced while working with neural networks is the need for huge datasets on which the parameters of every layer may be configured in such a way that they generalize the problem. The medical data are often hard to acquire due to the numbers of regulations, privacy concerns and often because of the limited number of the cases which are taken in the account. Annotating the data, including ground truth segmentation or gathering data from patients history is an expensive and time-consuming process. There are many known ways of augmenting the datasets for the purpose of neural network learning. A single computed tomography volume can be cut into slices not only by the axial plane but also sagittal and coronal one. Only by preparing the data in this way can we get 3 times the original slice count. Furthermore, it is possible to add rotation and translation to the data. This operation can enlarge the dataset by another multiplier which we can define by the number of different rotations and translations applied. For the neural networks most of these changes which can be defined as affine transformations are perceived as different data from which they can learn to generalize the given task.

There are also more sophisticated ways of expanding the datasets like elastic distortions which no longer belong to the affine transformations and are no longer easily reversible [16]. This approach synthesizes the new information based on the original data but it adds invariance to it. This deformation is achieved by computing a new target location for every pixel with the respect to the original location. The variability is achieved by generating two images of the same size as the original slice and populating it with pseudo random numbers acquired with Mersenne Twister random number generator. The outcome is then smoothed with a gaussian filter of size $4 * \sigma + 1$ where $\sigma = imageWidth * 0.06$ in order to establish the size of areas that should be displaced in the same manner. Later on, these images are multiplied by $\alpha = imageWidth * 2$ which is a scaling factor that enlarges the vectors generated by the previous operations.

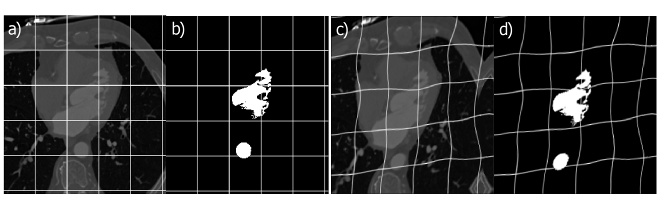

Fig. 1. Example of CT data augmentation: CT axial image (a,c) with matching binary label mask (b,d) before (left) and after (right) applying elastic deformation.

The parameters mentioned above were suited to this solution on the basis of the quality analysis of the outcoming images. The obtained vectors are often not discrete that is why an interpolation is often required. In this way we can generate images that are stimulating the neural networks to generalize the given task on the larger dataset with the variability that is higher than the one acquired using the affine transformations.

B. Convolutional Neural Network Architecture

We propose a 2D U-Net based [13] convolutional neural network architecture illustrated in Figure 2. The network structure is prepared for fully automatic aorta and left ventricle area joint segmentation assuming landmark detection for TAVI as a next crucial procedure planning step. The network predicts labeled binary mask for every image in single dataset as an output. It has been proved that there is no need to use a big amount of data for U-Net based image processing.

For CT scans we use the bilinear interpolation to downsample the input images to the size of 256 x 256 pixels to overcome memory limitations. The network architecture consists of 31 layers: 19 convolutional, 4 convolutional with layer transposition, 4 max pooling, and 4 concatenate. All convolutional layers except the last one have the same kernel size 3 x 3, rectified linear unit (ReLU) activation function and the same padding with different filter sizes. Similarly convolutional layers with transposition and max pooling layers have the same kernel size of 2 x 2. Additionally, convolutional layers with transposition have strides parameter set to 2 x 2. The network output - convolutional layer with kernel size 1 x 1 and sigmoidal activation function generate the predicted label masks. Final network model consists of 485 673 parameters. We compile it using *Adam* optimizer with a learning rate of 0.0001 and *binary crossentropy* loss function.

C. Metrics

The proposed method was evaluated using three types of metrics: Dice coefficient, Jaccard index and Hausdorff distance [17]. All of the chosen methods were computed in 3D for a whole CT data stack. Afterwards, a mean, median and a standard deviation were counted from the metrics results for every stack in the test dataset. For the Hausdorff distance not only the maximum distance among a stack was calculated but also the maximum value of the mean values from a single stack was used.

IV. EXPERIMENTS AND RESULTS

A. Dataset

The proposed method was evaluated through challenging CT data segmentation task. We established two major areas which should be segmented before TAVI procedure to allow visualization and simulations: aorta (aortic root and partially ascending aorta) and left ventricle. A dataset of 31 CTA volumes were used in experiment (fig. 3a). Volumes were registered with various CT scanners to avoid overfitting the network to one specific data source and to a specific, institute-limited population. Each dataset contained 235 - 640

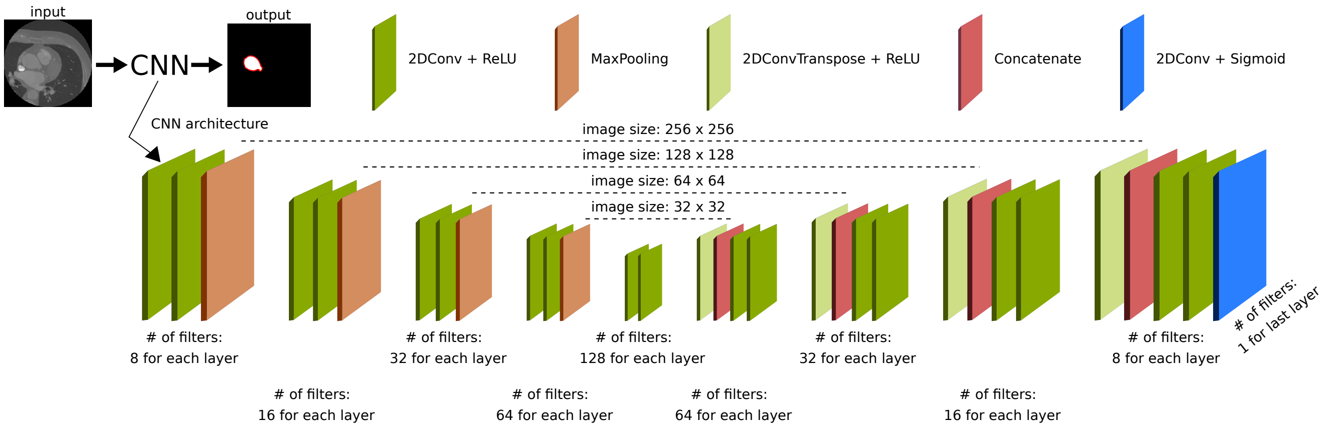

Fig. 2. Proposed architecture of Convolutional Neural Network (CNN). Model consists of 30 hidden layers composed of Convolutional, MaxPooling and Concatenate layers. Each part of the architecture is described with different filter size. The network predicts labeled mask of aorta and left ventricle area for every CT image of size 256 x 256 used as the network input.

slices of size 512 x 512 pixels (pixel spacing between 0.323 - 0.617 [mm]).

B. Image Preprocessing

Binary masks of aorta and left ventricle area (fig. 3b) were manually segmented using Slicer3D software [18] by two experienced experts - radiologist and radiology technician. In preprocessing step all images were normalized and set to respective value in range between 0 and 1. To reduce network overfitting we augmented the database using elastic deformation with parameters mentioned above which provides virtually new unique images without the loss of medical information for neural network training. As a result, we collected 62 volumetric stacks which were divided into train and test sets in a 70 to 30 ratio (44 stacks in train set and 18 in test set). Every image was resized using bilinear interpolation method to size 256 x 256 for computational complexity reduction. During experiments we chose network parameters setting batch size as 1 and validation split to 0.15 which means that 7 random stacks were chosen during each learning process iteration to perform cross validation. Proposed model was trained on 100 epochs.

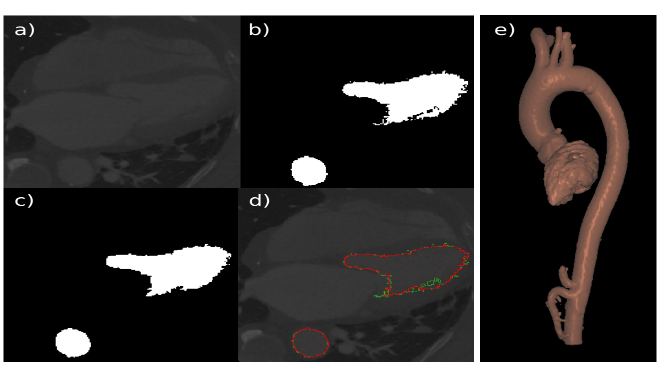

Fig. 3. Example of qualitative segmentation evaluation: (a) Input CT image; (b) ground truth label mask of aorta and left ventricle; (c) output binary mask; (d) Segmentation contours - green color indicates the ground truth segmentation, red color - the result of automatic segmentation; (e) 3D vizualization of segmentation results.

C. Evaluation

Quantitative results evaluation was performed using three metrics: Dice coefficient, Jaccard index and Hausdorff distance [17]. Final summary was presented in table I. Hausdorff distance was calculated to be at the level of 7 mm. Such a high value of this metric occured most likely due to the fact that there are observable differences between ground truth and automatic segmentation as it comes to the lower part of the abdominal aorta where many little blood vessels were often omitted by the neural network but this part of the scan is not crucial as it comes to the TAVI procedure.

TABLE I

AUTOMATIC SEGMENTATION RESULTS

Applied metric	**Statistical parameters**		
	Mean	***STD***	***Median***
Dice coefficient	0.95	0.02	0.96
Jaccard index	0.91	0.03	0.93
Hausdorff distance [mm]	7.04	1.42	7.83

Comparing obtained results with actual state-of-the-art solutions (table II) proposed method is characterized by one of the highest Dice coefficient scores. What is important, the standard deviation we obtained is relatively low which indicates the high algorithm robustness.

TABLE II

COMPARISION OF METHODS

Method	**Dice coefficient**	**Modality**
Isgum et al. [19]	0.87 +/- 0.03	non-contrast CT
Xie et al. [9]	0.93 +/- 0.01	non-contrast CT
Kurugol et al. [20]	0.93 +/- 0.01	non-contrast CT
Noothout et al. [14]	0.91 +/- 0.04	non-contrast CT
Elattar et al. [10]	0.95 +/- 0.04	CT angiography
Proposed method	0.95 +/- 0.02	CT angiography

V. CONCLUSIONS

In this article, we proposed a method for fully automatic, simultaneous aorta and left ventricle segmentation. We designed a convolutional neural network architecture based on 2D U-Net approach which was trained on 44 and tested on 18 datasets. All scans were normalized and resized to 256 x 256 pixels arrays. To enlarge the database we used elastic deformation method which allowed to expand 31 CTA scans to the number of 62. We performed both quantitative and qualitative evaluation, achieving mean Dice Similarity score at the level of 0.95 with standard deviation 0.02.

Compared to clinically used software, our model does not need any manual input or additional preprocessing. On the other hand, output of the neural network should be evaluated each time by the radiologist - the algorithm still needs some improvements to present fully credible results.

Further works might focus on enlarging the database used for the training and evaluation of the implemented neural network. It is known that CNNs benefit from larger training datasets allowing them to better generalize knowleadge. Especially, due to the fact that it has reached the results that are comparable or even slightly better as the state-of-the-art on a relatively small dataset. What is more, it would be valuable to experiment with different architectures, e.g. 2.5D, 3D, dense or probabilistic U-Nets which can potentially lead to better structure spatial recognition. Future evaluation should also take into consideration surface distance metrics, giving more complete understanding of network performance.

The results are promising in terms of using this solution in clinical setting, and can potentially reduce time and effort needed for the planning of the TAVI procedure. Nevertheless, further evaluation is essential as it comes to employing this method into everyday use of a clinician.

ACKNOWLEDGMENT

This work was funded by NAWA-PROM, grant no. PPI/PRO/2018/1/00026/U/001. We would like to thank MedApp S.A. for sharing CTA datasets. Special thanks should also be given to Andrzej Skalski from AGH University of Science and Technology, for the support and technical guidance during algorithm tests.

REFERENCES

1. Vernikouskaya, I. *et al.* Image-guidance for transcatheter aortic valve implantation (TAVI) and cerebral embolic protection. *International journal of cardiology* **249,** 90–95 (2017).
2. Loncaric, S., Subasic, M. & Sorantin, E. *3-D deformable model for aortic aneurysm segmentation from CT images* in *Engineering in Medicine and Biology Society, 2000. Proceedings of the 22nd Annual International Conference of the IEEE* **1** (2000), 398–401.
3. Olabarriaga, S. D., Rouet, J.-M., Fradkin, M., Breeuwer, M. & Niessen, W. J. Segmentation of thrombus in abdominal aortic aneurysms from CTA with nonparametric statistical grey level appearance modeling. *IEEE transactions on medical imaging* **24,** 477–485 (2005).
4. Kurkure, U., Avila-Montes, O. C. & Kakadiaris, I. A. *Automated segmentation of thoracic aorta in non-contrast CT images* in *Biomedical Imaging: From Nano to Macro, 2008. ISBI 2008. 5th IEEE International Symposium on* (2008), 29–32.
5. Karar, M., Gessat, M., Walther, T., Falk, V. & Burgert, O. *Towards a new image guidance system for assisting transapical minimally invasive aortic valve implantation* in *Engineering in Medicine and Biology Society, 2009. EMBC 2009. Annual International Conference of the IEEE* (2009), 3645–3648.
6. Gessat, M. *et al. A planning system for transapical aortic valve implantation* in *Medical Imaging 2009: Visualization, Image-Guided Procedures, and Modeling* **7261** (2009), 72611E.
7. Lavi, G., Lessick, J., Johnson, P. C. & Khullar, D. *Single-seeded coronary artery tracking in CT angiography* in *Nuclear Science Symposium Conference Record, 2004 IEEE* **5** (2004), 3308–3311.
8. Wächter, I. *et al. Patient specific models for planning and guidance of minimally invasive aortic valve implantation* in *International Conference on Medical Image Computing and Computer-Assisted Intervention* (2010), 526–533.
9. Xie, Y., Padgett, J., Biancardi, A. M. & Reeves, A. P. Automated aorta segmentation in low-dose chest CT images. *International journal of computer assisted radiology and surgery* **9,** 211–219 (2014).
10. Elattar, M. A. *et al.* Automatic segmentation of the aortic root in CT angiography of candidate patients for transcatheter aortic valve implantation. *Medical & biological engineering & computing* **52,** 611–618 (2014).
11. López-Linares, K. *et al.* Fully automatic detection and segmentation of abdominal aortic thrombus in post-operative CTA images using Deep Convolutional Neural Networks. *Medical image analysis* **46,** 202–214 (2018).
12. Dou, Q. *et al. 3D deeply supervised network for automatic liver segmentation from CT volumes* in *International Conference on Medical Image Computing and Computer-Assisted Intervention* (2016), 149–157.
13. Ronneberger, O., Fischer, P. & Brox, T. *U-net: Convolutional networks for biomedical image segmentation* in *International Conference on Medical image computing and computer-assisted intervention* (2015), 234–241.
14. Noothout, J. M., de Vos, B. D., Wolterink, J. M. & Išgum, I. *Automatic segmentation of thoracic aorta segments in low-dose chest CT* in *Medical Imaging 2018: Image Processing* **10574** (2018), 105741S.
15. Zheng, Y. *et al.* Automatic aorta segmentation and valve landmark detection in C-arm CT for transcatheter aortic valve implantation. *IEEE transactions on medical imaging* **31,** 2307–2321 (2012).
16. Simard, P. Y., Steinkraus, D. & Platt, J. C. *Best practices for convolutional neural networks applied to visual document analysis* in *null* (2003), 958.
17. Cárdenes, R., de Luis-García, R. & Bach-Cuadra, M. A multidimensional segmentation evaluation for medical image data. *Computer methods and programs in biomedicine* **96,** 108–124 (2009).
18. Fedorov, A. *et al.* 3D Slicer as an image computing platform for the Quantitative Imaging Network. *Magnetic resonance imaging* **30,** 1323–1341 (2012).
19. Isgum, I. *et al.* Multi-atlas-based segmentation with local decision fusion application to cardiac and aortic segmentation in CT scans. *IEEE transactions on medical imaging* **28,** 1000–1010 (2009).
20. Kurugol, S., Estepar, R. S. J., Ross, J. & Washko, G. R. *Aorta segmentation with a 3D level set approach and quantification of aortic calcifications in non-contrast chest CT* in *Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE* (2012), 2343–2346.