Communications on Applied Electronics (CAE) - ISSN : 2394 - 4714
Foundation of Computer Science FCS, New York, USA
Volume 7 - No. 29, June 2019 - www.caeaccess.org

Aortic Valve Segmentation using Convolutional Neural Network with Skip Mechanism

H. M. Zabir Haque
University of Saskatchewan
Department of Computer Science
Saskatoon, Saskatchewan, S7N 5C9, Canada

ABSTRACT

Segmentation is a method which can be implemented inside the verge of Artificial Intelligence World. In this approach, each pixel of an image is required to be labeled to yield the final segmented result. In this paper, a novel method has been proposed which is conducted following by Convolutional Neural Network (CNN) with skip mechanisms for Segmentation. In this method, the original 3D medical image captured as a 2D slice to pass through multiple image channels along with Ground Truth in the last channel which work as an input of CNN. This sub-sample, however, gradually generate the segmentation mask for the corresponding input image. The proposed methods were tested to perform segmentation for the CT image of the human organ (Aortic Valve) which show a significant amount of accuracy with very few numbers of dataset. Here, the result has been compared with existing methods. Such a system, hence, will support many experiments to help better understanding of Humankind in the perspective of Artificial Visualization.

Keywords

Segmentation, Convolutional Neural Network, Artificial Intelligence, Deep Learning, Medical Image, Human Organ, Aortic Valve, Image Channels, Ground Truth, Segmentation Mask

1. INTRODUCTION

The importance of Semantic Segmentation is immeasurable in the Artificial Intelligence field. The approach requires the sole labeling of pixels to yield the final segmented result. The 2D slice and Ground-Truth have been considered as input to pass through multiple image channels to the network. This proposed methods tested on the CT image of the human organ: Aortic Valve and a significant amount of accuracy have been observed in the experiment. Hence, The ultimate goal of this work is to extend the platform to maximize the likelihood of the desired description for a complex and diverse scenario of various environments.

1.1 Segmentation of Medical Image

Medical images can be collected implying a variety of imaging technologies such as computed tomography (CT), ultrasound, X-ray and magnetic resonance imaging (MRI). Segmentation is used for the medical image to represent the different anatomical structures all through the human body, including bones, blood vessels, vertebrae, and major organs. In computer vision, intelligent image-based technique enhanced the research through neural network followed by deep learning paradigm. More concretely, the medical field also influenced by deep learning technique to scrutinize images, ensuring precise segmentation. Nonetheless, medical images suffer from noise, artifacts, the impulsiveness of organs shape, size orientation. However, different traditional methods for segmentation of Human Organs do not outperform the accuracy of experts. For example, the challenging part for segment in an axial slice of a cardiac CT volume is such as the separation boundary between the aortic valve (AV) and left ventricle (LV) is indistinguishable as the intensity of neighboring area are quite similar whereas intensity of image is one of the key features for image segmentation.
To accomplish segmentation of medical image the most profound-advanced Deep Learning tool is semantic segmentation using Deep Convolutional Neural Networks. A prerequisite of large amounts of suitable and labeled data is essential; however, to evade the system from over-fitting during the convergence.

1.2 Scope of the Study

Currently, most prominent deep artificial neural networks which alleviate the structure of computer vision is renowned as Convolutional Neural Networks (CNN) [13]. Impact of CNNs is extensively perceptible in various tasks, such as image classification [13, 19, 10, 20] super resolution [9] and semantic segmentation [15]. The glimpse impact of CNN in medical image segmentation and classification is acknowledged on recent publications [2, 17, 16, 1, 8, 6].
Traditional CNN-based approaches for semantic segmentation consider image patches to pass through the CNN as they improve the accuracy. Due to the time consumption issue and the redundancy in the architecture run [5], recently developed state-of-the-art is to utilize fully convolutional approach [15]. From the architecture viewpoint, after the convergence, in this approachs localizing is done by encompassing higher order distinctive feature along with yield results of upsampling layer. Hence, this information helps successive convolution layer to learn to accumulate more accurate output. A huge channel is maintained to restore the vast amount of feature which helps to broadcast the distinguish characteristics throughout the further whole network that is a unique modification on FCN. No fully connected layers are used rather the valid portion of individually participated convolution is used i.e. the full context of the input image is available at any time instead of containing individual pixels information, to this particular architecture. This proposal allows the segmentation approach to work seamlessly even for a randomly chosen bulky file. Pixels which are very near to boundary portion is predicted and by learning from input image the architecture extrapolate the disappeared contexts which are helpful to maintain the higher resolution of the output unless otherwise, the resolution of the large size images may cause a problem within the available allocated size of GPU.

1.3 Contribution

In this proposed method, like as other Semantic Segmentation, this also requires the individual labeling of pixels to generate the segmented object. Here, 3D medical images have been divided into 2D Slice along with Ground-Truth those are works as the inputs to the network. The network will generate a segmentation mask. These methods had been tested on a CT image for the Aortic Valve. Summary of our contributions have been specified as follows:

(1) This approach introduce and evaluate a relative stable, robust and precise (boundary) image-processing for segmentation.
(2) Spatial detail has been preserve using skip connection to improve the accuracy.
(3) A few numbers of dataset has been used to stabilize the accuracy.

The main deep learning tool CNN which is commonly used for segmentation is discussed along with current state-of-art of Medical Image Segmentation for Aortic Valve, in Section 2. In Section 3, the proposed method is described along with optimization and proposed an architecture which constructed the network. Experimental Results with analysis is depicted in Section 4. The conclusion part is finished with the further idea to implement in the future in Section 5.

2. RELATED WORK

2.1 Current Deep Learning Methods

A method worthy to be mentioned here is Convolutional Neural Networks (CNNs) recently widely used in dealing with visual recognition problems. Traditionally CNNs show remarkable performance in classification tasks on images only to a single label class in the output. Significant performance is observed effectively in learning features in order at multiple scales from the input. In supervised semantic segmentation CNNs recently have shown remarkable performance. Combined Conditional Random Fields (CRFs) and Convolutional Neural Networks (CNNs) approach proposed by Lin et al. [14] perform better in exploring pixel to pixel spatial correlations [14]. Although the limitation of the approach is to be needed to implement a dense CRF to refine the output from the CNN thus increases the computational time. There are few enhanced techniques available in literature [18] where a classifier output is proposed to consider the features from the multiple layers which enable the network to achieve better localization accuracy and simultaneously allow to use information.
U-net [22] is capable of showing some amazing performance to help medical image analysis. However, it has its drawbacks such as being “lazy” to find itself able to handle a problem in low-level layers, the high-level layers will not bother to learn anything. For example, if we train the U-net architecture to perform a simple task like “copying image”, then the loss value will drop to 0 immediately. This happens due to the reason why simply the first layer of the encoder determines that it can transmit all the extracted features to the last layer of the decoding layer directly as it skips connection to minimize the loss value. That is why even the U-net would be trained for many several times, layers of mid-level will not perceive any gradient. Usually, in each decoding layers of U-net, features can be attained from higher layers or skip connected layers. During the training process, in every iterations decoding layers select the output from other layers with nonlinear activation to minimize the loss value. Even if one initializes the U-net with Gaussian random numbers to copy an image, the example discussed earlier, the output of the first layer of an encoder is informative enough to generate the full input map whereas the output of the second to last layers in decoding appears to be very noisy. Hence the “lazy” U-net architecture gives up the relatively noisy features.

2.2 Segmentation Challenge for Medical Images: Aortic Valve

The practical perspective of artery segmentation inspires scholars to contribute to this field of research. There are many publications specific to the topic of aortic segmentation in literature [3, 7, 11]. Among the approaches addressing the aortic root segmentation from CT scans, the approach of Zheng et al. [23] has presented a complete automatic segmentation technique where marginal space learning technique is used for pre- or post-operative planning [23]. 2D watershed-based algorithm is proposed by Lavi et al. [12] to segment medical images like CT. The technique is semi-automatic which is unlike to Zheng’s [23] automatic process. Moreover, it has drawbacks of poor performance in handling low-quality volumes. Another application of segmenting heart chambers and the aortic valve is presented in [21] where model-based segmentation technique is utilized. A similar application of heart valve segmentation is made by Grbic et al. from 4D CT scans. All the approaches are limited in terms of classification which is an essential key measure in practice. An operation of replacing the aortic valve with a prosthetic valve is very crucial as it may cause to block the coronary circulation many times observed during postoperative care. This may occur due to the miscalculation the distance in between coronary ostia and the hinge plane of the aortic valve. Thus measuring the diameter of the aortic valve annulus is very important and critical for patients. Hence, segmentation from the Contrast-enhanced coronary CT angiography (CCTA) is a widely used medical imaging technique for pre-operative processing.

3. PROPOSED METHOD

For bio-medical images, there are many constraints to be considered as they are not simple like defining a single class on an image rather it desired several objectives such as localization, segmentation for each pixel to be classed and so on. From such strong motivations to improve the medical image analysis, a network for image segmentation have been proposed.

3.1 Architecture

In this section, first, briefly explain the idea of how these architectures encompass the spirit CNN for segmentation. Here, 2-channels for pre-processing has been used where the original image and ground-truth is assembled through 1^{st} and 2^{nd} channel. The image pixel scale normalized in between (0 to 255).

3.2 Architecture: CNN with Conv-Deconv Mechanism

The down-sampling and up-sampling path have 4 convolutional and deconvolutional blocks, respectively. The filter size of convolutional layers filter size is 5×5 with stride is 1 in bidirectional and rectifier activation. The max-poolings stride 3×3 of down-sampling is used to the every block, so feature maps change from 240×240 to 15×15. On the other direction, deconvolutional blocks have the same filter and stride size of the convolutional layer which convert the segmented image into the original image. However, unlike the U-net Architecture, in this method zero paddings is used to keep the output dimension for all the convolutional layers of both down-sampling and up-sampling path. No fully connected layer is invoked in the network (see Figure 1).

3.3 Optimization

The optimization method has been implemented using Soft Dice Metric for cost function instead of the cross-entropy or quadratic cost function which is quite different from the original Dice Similarity Coefficient (DSC). In this proposed method, Adam optimizer has been implemented (as stochastic gradient-based) to measure the parameters. The Adam optimizer learning rate is fixed on 0.0001 and epochs 350. The Xavier weights have been used for initial weight and biases were initialized as 0. To evaluate the segmentation result using the DSC is calculated. The segmentation results of the automatic method that is:

$$DSC = \frac{2TP}{FP + 2TP + FN}.$$

Communications on Applied Electronics (CAE) - ISSN : 2394 - 4714
Foundation of Computer Science FCS, New York, USA
Volume 7 - No. 29, June 2019 - www.caeaccess.org

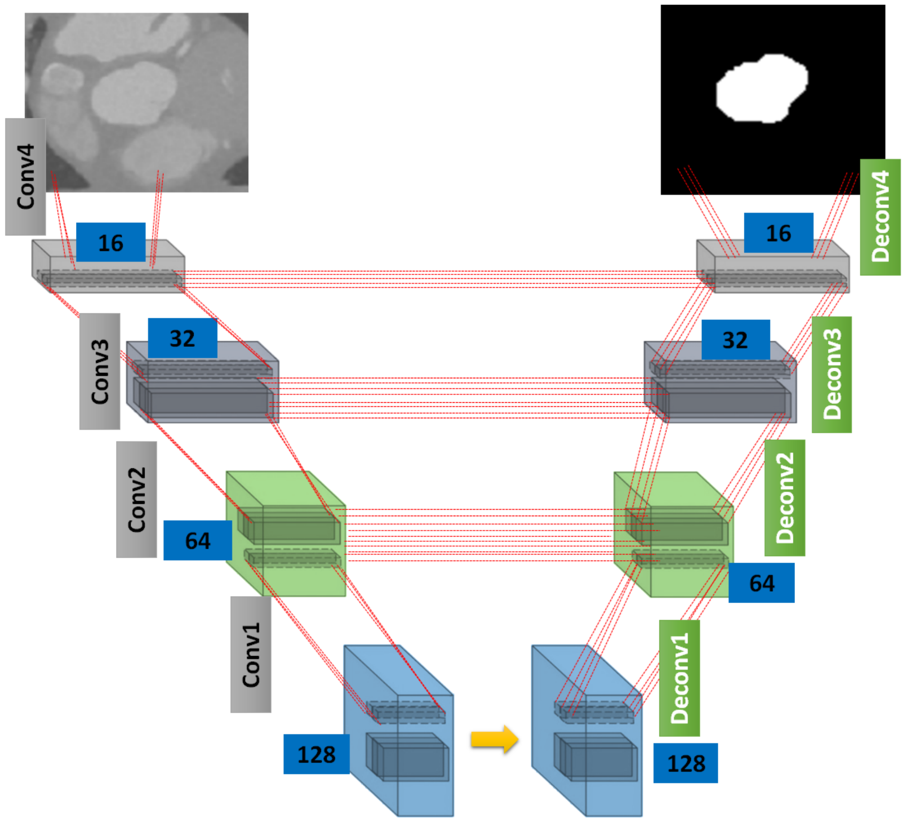

Fig. 1. Architecture: CNN with Conv-Deconv Mechanism.

Table 1. Detail of Aortic Valve Data-set

	Vol	Type	Height	Width	Depth	Availability
Aortic Valve	50	CT	512	512	200 $\sim$220	Private

Here, TP, FP, FN are denoted as true positive, false positive and false negative, respectively. In addition, Sensitivity is used to evaluate the number of and that is:

$$Sensitivity = \frac{TP}{TP + FN}.$$

The mean DSC results of the four-fold cross-validation are reported here.

3.4 Dataset

The detail information of dataset has been depicted in Table 1. It can be observed that the Aortic Valve collected from a private data source and experts mark the ground-truth.
The height and width of the Aortic Valve are fixed to 512, and depth is $200 - 220$. However, this dataset is not available in any public database. For this experiment 50 Volume has been used for both training and testing (training: 40 volume and testing: 10 volume).

4. EXPERIMENTAL EVALUATION

Segmentation is perform in 3D Medical datasets such as Aortic Valve in Computed Tomography (CT) image. Here, 2-channels are used for real image and Ground-truth. For this kind of deep learning based experiment required GPU support. GPU: "GeForce GTX 780" is used to accomplished this experiment.
In this experiment, one 3D volume (Height: 512, Width: 512, No. of Slice varies: $200 \sim 220$) is segmented in less than 4 seconds. To segment Aortic Valve, CT volume is used. A sample segmented output of different position of Aortic Valve has been shown in Figure 2.
The Aortic Valve's clearly visible in the beginning Column and gradually move to the end of the volume, successfully. Most of the traditional methods suffer from segmenting the end portion of the volume. However, as the Aortic Valve have very low resolution, similar 3 prior intensity and the invisible boundary between Aortic Valve and Left-Atrium, so these datasets require more depth channel to extract features. The previous paper on Aortic Valve Segmentation [4] acquired 0.95 of DSC. However, after closely observing their result, it is easily noticeable that they did not go-through for sharp edges during marking the ground-truth. However, in this experiment, very shape edge is also considered on ground-truth. Even, we do not have any reliable source to compare with others due to lack of exact same datasets, but our Aortic Valve dataset also marked by some expert. Due to the lack of exact same datasets, and any reliable sources to compare with others, the Aortic Valve dataset is then marked by some expert. Also some time this result outperforms the ground-truth for some cases which is shown in the Figure 3. Even though the limitation to compare proposed result with the previous paper on Aortic Valve Segmentation [4] cause of different dataset, but as the proposed method consider the sharp edge compare to them, so the accuracy of proposed method's DSC value along with previous paper's accuracy result Table 2 is attached here. This particular experiment observed the four-folder orientation of 10 random volume as a test set.

5. CONCLUSION

Semantic Segmentation is the most prominent tool to accomplish the true virtue of Artificial Intelligence in terms of Human Perspective. Here, a novel segmentation method based on CNN with

Communications on Applied Electronics (CAE) - ISSN : 2394 - 4714
Foundation of Computer Science FCS, New York, USA
Volume 7 - No. 29, June 2019 - www.caeaccess.org

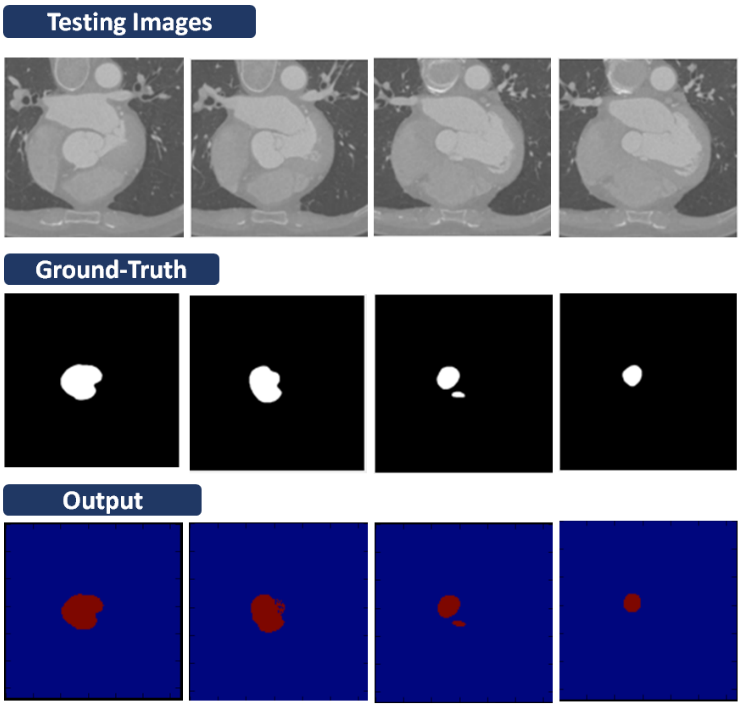

Fig. 2. Aortic Valve Segmentation. (a) Upper Row : Input Images of various positioned Aortic Valve (b) Middle Row : Ground-Truth and (c) Lower Row : Segmented Output.

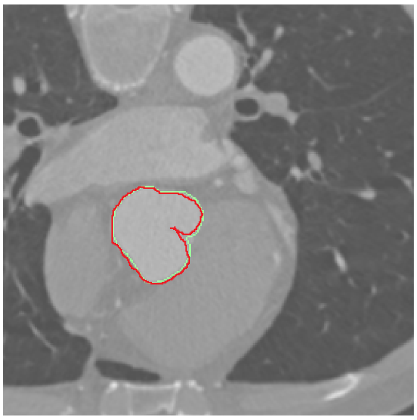

Fig. 3. Example of outperforming the ground-truth by finding the sharp edge. Red is marked for the segmented result where green indicates the ground-truth.

Table 2. Aortic Valve DSC Value Comparison.

DSC Table		
Paper Title	Data Set	DSC (%)
Automatic Segmentation of the Aortic Root in CT Angiography of Candidate Patients for Transcatheter Aortic Valve Implantation	CTA	0.95 ± 0.03
This Experiment	CT	0.95 ± 0.03

skip method is proposed based on core phenomena of Semantic segmentation. It is noticeable that unlike other experiments this method go-through for a very sharp edge and able to stabilize the accuracy comparatively using a very small amount of datasets.

6. REFERENCES

[1] Özgün Çiçek, Ahmed Abdulkadir, Soeren S. Lienkamp, Thomas Brox, and Olaf Ronneberger. 3d u-net: Learning dense volumetric segmentation from sparse annotation. In Sebastien Ourselin, Leo Joskowicz, Mert R. Sabuncu, Gozde Unal, and William Wells, editors, *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2016*, pages 424–432, Cham, 2016. Springer International Publishing.

[2] Dan Ciresan, Alessandro Giusti, Luca M. Gambardella, and Jürgen Schmidhuber. Deep neural networks segment neuronal membranes in electron microscopy images. In F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger, editors, *Advances in Neural Information Processing Systems 25*, pages 2843–2851. Curran Associates, Inc., 2012.

[3] Anthony Adam Duquette, Pierre-Marc Jodoin, Olivier Bouchot, and Alain Lalande. 3d segmentation of abdominal aorta from ct-scan and mr images. *Computerized Medical Imaging and Graphics*, 36(4):294–303, 2012.

[4] M. A. Elattar, E. M. Wiegerinck, R. N. Planken, E. vanbavel, H. C. van Assen, J. Baan, and H. A. Marquering. Automatic segmentation of the aortic root in ct angiography of candidate patients for transcatheter aortic valve implantation. *Medical & Biological Engineering & Computing*, 52(7):611–618, Jul 2014.

[5] C. Farabet, C. Couprie, L. Najman, and Y. LeCun. Learning hierarchical features for scene labeling. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(8):1915–1929, Aug 2013.

[6] Mohammad Havaei, Axel Davy, David Warde-Farley, Antoine Biard, Aaron Courville, Yoshua Bengio, Chris Pal, Pierre-Marc Jodoin, and Hugo Larochelle. Brain tumor segmentation with deep neural networks. *Medical Image Analysis*, 35:18–31, 2017.

[7] Ivana Isgum, Marius Staring, Annemarieke Rutten, Mathias Prokop, Max A Viergever, and Bram Van Ginneken. Multi-atlas-based segmentation with local decision fusion-application to cardiac and aortic segmentation in ct scans. *IEEE transactions on medical imaging*, 28(7):1000–1010, 2009.

[8] Konstantinos Kamnitsas, Christian Ledig, Virginia FJ Newcombe, Joanna P Simpson, Andrew D Kane, David K Menon, Daniel Rueckert, and Ben Glocker. Efficient multi-scale 3d cnn with fully connected crf for accurate brain lesion segmentation. *Medical Image Analysis*, 36:61–78, 2017.

[9] Jiwon Kim, Jung Kwon Lee, and Kyoung Mu Lee. Accurate image super-resolution using very deep convolutional networks. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2016.

[10] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. In F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger, editors, *Advances in Neural Information Processing Systems 25*, pages 1097–1105. Curran Associates, Inc., 2012.

[11] Uday Kurkure, Olga C Avila-Montes, and Ioannis A Kakadiaris. Automated segmentation of thoracic aorta in non-contrast ct images. In *2008 5th IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, pages 29–32. IEEE, 2008.

[12] G. Lavi, J. Lessick, P. C. Johnson, and D. Khullar. Single-seeded coronary artery tracking in ct angiography. In *IEEE Symposium Conference Record Nuclear Science 2004.*, volume 5, pages 3308–3311 Vol. 5, Oct 2004.

[13] Yann LeCun, Léon Bottou, Yoshua Bengio, Patrick Haffner, et al. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324, 1998.

[14] Guosheng Lin, Chunhua Shen, Anton van den Hengel, and Ian Reid. Efficient piecewise training of deep structured models for semantic segmentation. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2016.

[15] Jonathan Long, Evan Shelhamer, and Trevor Darrell. Fully convolutional networks for semantic segmentation. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2015.

[16] F. Milletari, N. Navab, and S. Ahmadi. V-net: Fully convolutional neural networks for volumetric medical image segmentation. In *2016 Fourth International Conference on 3D Vision (3DV)*, pages 565–571, Oct 2016.

[17] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In Nassir Navab, Joachim Hornegger, William M. Wells, and Alejandro F. Frangi, editors, *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, pages 234–241, Cham, 2015. Springer International Publishing.

[18] Mojtaba Seyedhosseini, Mehdi Sajjadi, and Tolga Tasdizen. Image segmentation with cascaded hierarchical models and logistic disjunctive normal networks. In *The IEEE International Conference on Computer Vision (ICCV)*, December 2013.

[19] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv Preprint arXiv:1409.1556*, 2014.

[20] Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, and Andrew Rabinovich. Going deeper with convolutions. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2015.

[21] I. Waechter, R. Kneser, G. Korosoglou, J. Peters, N. H. Bakker, R. v. d. Boomen, and J. Weese. Patient specific models for planning and guidance of minimally invasive aortic valve implantation. In Tianzi Jiang, Nassir Navab, Josien P. W. Pluim, and Max A. Viergever, editors, *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2010*, pages 526–533, Berlin, Heidelberg, 2010. Springer Berlin Heidelberg.

[22] L. Zhang, Y. Ji, X. Lin, and C. Liu. Style transfer for anime sketches with enhanced residual u-net and auxiliary classifier gan. In *2017 4th IAPR Asian Conference on Pattern Recognition (ACPR)*, pages 506–511, Nov 2017.

[23] Y. Zheng, M. John, R. Liao, A. Nottling, J. Boese, J. Kempfert, T. Walther, G. Brockmann, and D. Comaniciu. Automatic aorta segmentation and valve landmark detection in c-arm ct for transcatheter aortic valve implantation. *IEEE Transactions on Medical Imaging*, 31(12):2307–2321, Dec 2012.