

# Neural Network Based Whole Heart Segmentation from 3D CT Images

Original Scientific Paper

## Marija Habijan

J.J. Strossmayer University of Osijek,  
Faculty of Electrical Engineering, Computer Science  
and Information Technology Osijek  
Kneza Trpimira 2b, Osijek, Croatia  
marija.habijan@ferit.hr

## Hrvoje Leventić

J. J. Strossmayer University of Osijek,  
Faculty of Electrical Engineering, Computer Science  
and Information Technology Osijek  
Kneza Trpimira 2b, Osijek, Croatia  
hrvoje.leventic@ferit.hr

## Irena Galić

J. J. Strossmayer University of Osijek,  
Faculty of Electrical Engineering, Computer Science  
and Information Technology Osijek  
Kneza Trpimira 2b, Osijek, Croatia  
irena.galic@ferit.hr

## Danilo Babin

imec-Ghent University,  
imec-TELIN-IPI, Faculty of Engineering and Architecture  
St. Pietersnieuwstraat 33, Ghent, Belgium  
danilo.babin@ugent.be

**Abstract** – The most recent research is showing the importance and suitability of neural networks for medical image processing tasks. Nonetheless, their efficiency in segmentation tasks is greatly dependent on the amount of available training data. To overcome issues of using small datasets, various data augmentation techniques have been developed. In this paper, an approach for the whole heart segmentation based on the convolutional neural network, specifically on the 3D U-Net architecture, is presented. Also, we propose the incorporation of the principal component analysis as an additional data augmentation technique. The network is trained end-to-end, i.e., no pre-trained network is required. Evaluation of the proposed approach is performed on CT images from MICCAI 2017 Multi-Modality Whole Heart Segmentation Challenge dataset, delivering in a three-fold cross-validation an average dice coefficient overlap of 88.2% for the whole heart, i.e. all heart substructures. Final segmentation results show a high accuracy with the ground truth, indicating that the proposed approach is competitive to the state-of-the-art. Additionally, experiments on the influence of different learning rates are provided as well, showing the optimal learning rate of 0.005 to give the best segmentation results.

**Keywords** – CT, data augmentation, medical image segmentation, neural networks, volumetric segmentation, whole heart segmentation

## 1. INTRODUCTION

One of the biggest causes of health problems are cardiovascular diseases (CVDs). Most recent research provided by European Heart Network presents a high mortality rate of 45% for patients with CVDs, while simultaneously giving a shocking percentage of 11.2 for the population living with disorders in the cardiovascular system [1]. The severity of the problem prompted numerous joint researches between clinical and technical experts. Accelerated detection of CVDs permits physicians to line the proper treatment arrangement, which considerably improves and saves the patients' life. The advancements in commonly used acquisition techniques like computed tomography (CT) and magnetic resonance imaging (MRI) enable quick and less invasive body scans. The raw scans from imaging devices often include plenty of unnecessary information. The segmentation process allows the

examination of particular organs, such as the heart, from previously mentioned scans. This can be done either manually by a clinical expert or by using specialized medical software. Such software includes various image processing algorithms that provide easy three-dimensional visual inspections of the heart and serve as a treatment planning and diagnosis setting tool. The high variability and complexity of anatomical heart structure make the development of such medical software a very challenging task.

The paper has the following structure. State of the art methods and challenges of heart segmentation are the main subjects of Section 2. Our proposed approach is explained in Section 3. Data description and implementation details are provided in Section 4. The segmentation results, as well as analysis of conducted experiments, are presented in Section 5. Finally, the conclusion and future work are given in Section 6.

## 2. RELATED WORK

Manual delineation of the heart is an intensive and time-consuming process. More importantly, it is subjective and prone to variations in observers' expertise and knowledge. To overcome this issue, the various semi-automatic [2] and automatic approaches [3] for cardiac segmentation using different image processing and machine learning techniques are developing rapidly. In the beginnings of cardiac research, the methods based on partial differential equations, where internal or external evolving curves are differentiating heart region from other non-interest parts, are majorly investigated [4]-[8]. However, the formulation of the segmentation process as a process from a previous model to the unseen objects promoted specific registration and atlas-based approaches [9]-[14]. Although they are highly accurate, due to heavy computations on the algorithms of registration, they lack in performance. Commonly known, K-means and Fuzzy clustering [15]-[17] algorithms are showing average accuracy for heart segmentation tasks, as well as classification using random forest [18].

Nowadays, neural networks and deep learning-based approaches are cutting-edge technologies that replace conventional methods in the field of cardiac segmentation. The multi-planar deep learning [19] has been used to segment pulmonary veins (PV) and left atrium (LA) from MRI. Similarly, a recurrent, fully convolutional neural network (RCNN) is proposed for left ventricle (LV) segmentation in [20]. Furthermore, diverse combinations of encoder-decoder architectures acknowledged and induced the importance of skip connections. The combination of the highly granular feature maps from the decoder with finer granular feature maps from the encoder allows the creation of the segmentation masks while preserving details on noisy background. The U-Net [21] architecture links up-sampled features and adds convolutions between each sampling stage. Skip connections are essential in their ability to restore spatial resolution at the performance of the network. By adding dense and nested skip links, Zhou et al. [22] further improved the segmentation accuracy. This design reduces the discrepancies between maps of encoder features and decoder feature maps, simplifying the problem of optimization.

It is crucial to determine a few main challenges of cardiac image structure and its processing. First, the heart has great geometric complexity as it contains multiple chambers and great vessels where some substructures are overlapping, making it hard to find real boundaries between them. Secondly, its shape significantly varies through different cardiac conditions. In other words, cardiac images of a healthy patient can be very different in shape from pathological cases. Therefore, the differences in shape are representing the major challenge for automating the whole heart segmentation process. To overcome this problem of huge anatomical variability and distinctive boundaries in cardiac images, ap-

proaches incorporating prior knowledge and artificial neural networks are obtaining the highest accuracy scores for segmentation purposes [23].

The main focus of this paper is on the whole heart segmentation using 3D U-net architecture [24]. Our research primarily contributes to the data augmentation process. We add the principal component analysis (PCA) to the initially proposed on-the-fly elastic deformations and smooth, dense field deformation.

## 3. THE PROPOSED METHOD

This section explains the proposed network architecture and the process of data augmentation - mainly the effects of principal component analysis (PCA) on the input and output of the network. The PCA-based data augmentation represents the main scientific contribution of this paper.

### 3.1. NETWORK ARCHITECTURE

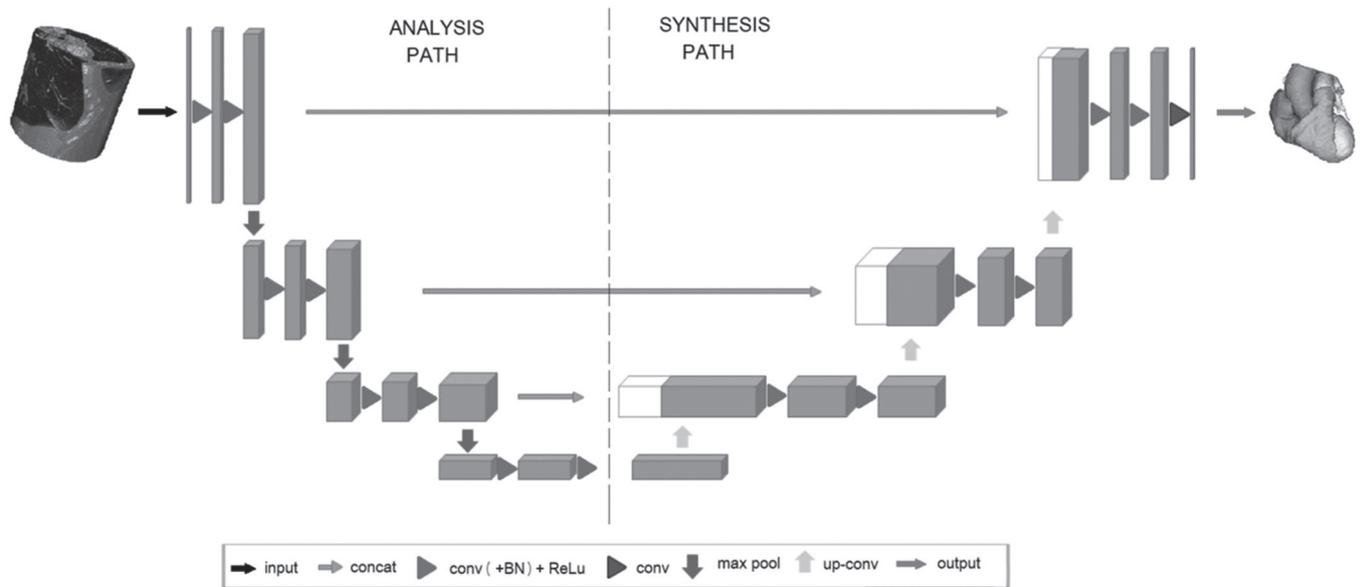
The proposed method uses the 3D U-Net neural network architecture that comprises of two main paths; analysis and synthesis path as shown in Fig. 1. The first path is a contracting encoder with the main task of providing the analysis of the whole image. The second path is the consecutive expanding decoder that produces a full-resolution segmentation.

In the analysis path, each layer consists of two  $3 \times 3 \times 3$  convolutions followed by a rectified linear unit (ReLU). Further comes the  $2 \times 2 \times 2$  max-pooling layer with strides of two in each dimension. Similarly, in the synthesis path there are two  $2 \times 2 \times 2$  up-convolutions, followed by two  $3 \times 3 \times 3$  convolutions where ReLU follows each convolution. The last,  $1 \times 1 \times 1$ , convolution layer reduces the output channels to match the number of labels, which is in our case 7. Additionally, avoiding bottlenecks in both paths is enabled using doubled values of the channels before each max-pooling [25].

Furthermore, the network uses batch normalization before every ReLU for faster convergence. The learning process involves the generation of the dense volumetric segmentation while only requiring two-dimensional annotated slices for training. That is possible because of the weighted softmax loss function, which sets weights of the unlabeled voxels to zero, consequently allowing learning from only labeled ones.

### 3.2. DATA AUGMENTATION WITH PCA

Artificial neural networks require vast amounts of annotated data for providing efficient learning. The data augmentation is commonly adopted for improving training performance where there is a limited size of the training data. Essentially, the data augmentation process increases the size of the training dataset through a series of image transformations. The data augmentation process initially proposed by 3D U-Net architecture includes scaling, rotation, and smooth, dense deformation field tech-



**Fig.1.** The 3D U-Net architecture

niques. In this paper, we propose an implementation of 3D U-Net architecture with a modification to the initially proposed data augmentation process by incorporating an additional PCA-based image transformation.

The data augmentation transformation can be simply represented as a sequence of operations on the training dataset. Let  $G$  denotes the set of all training samples, where each data sample consists of two 3D volumes with same dimensions: the input CT volume  $I$  with voxel gray values in range  $[0,4095]$  and the corresponding labeled CT volume  $L$  with voxel gray values in range  $[0,7]$ , i.e. labels. Furthermore, let  $\theta$  be the set of three transformation; scaling, rotation and smooth dense deformation. The transformations are iteratively applied to both an input  $I$  volume and the labeled  $L$  volume as follows:

$$h: X \rightarrow X', h \in \theta, X \in \{I, L\} \quad (1)$$

where transformation parameters for each  $h \in \theta$  are chosen randomly.

Thus, the entire process of data augmentation itself can be described with the following mapping:

$$\theta: G \rightarrow AG \quad (2)$$

Where  $G$  represents original training dataset and  $AG$  represents the augmented, i.e. transformed dataset of  $G$ . The new, i.e. inflated training dataset is thus defined as:

$$G \rightarrow G \cup AG \quad (3)$$

This means that the newly obtained dataset contains both; images from the original training dataset and images obtained after applying all the respective transformations defined with. Therefore, this whole dataset is used for network training. To further enable inflation of

the training dataset, the transformation set is extended with an additional transformation that performs principal component analysis using the singular value decomposition. Let  $\mathbf{v} \in \mathbb{Z}^3$  denote the voxel position in an input image. Let  $P_v$  denote a set of grayscale values representing the vector of principal components after performing the singular value decomposition of an input image  $X$  at position  $\mathbf{v}$ . Our proposed transformation  $h_{pca}$  is defined with the following mapping:

$$h_{pca} = X \rightarrow X', X \in \{I, L\}, \quad (4)$$

where  $X'$  denotes the resulting image after transformation. The proposed transformation modifies the grayscale values of every voxel  $\mathbf{v}$  in the input image in a following manner:

$$X'(\mathbf{v}) = \frac{1}{s_p} P_v [\alpha_1 \lambda_1]^T, \quad (5)$$

where  $\alpha_1$  denotes a random variable drawn from a Gaussian with  $mean = 0$  and  $\sigma = 0.1$ ,  $\lambda_1$  denotes the  $i^{\text{th}}$  eigenvalue corresponding to the eigenvector  $P$  and  $s_p$  denotes the scaling parameter initialized to  $5 \cdot 10^6$ .

## 4. IMPLEMENTATION DETAILS

This section includes all the necessary information about the implementation of our work. The dataset description used in our research is presented as well as the framework of our proposed method that includes input pre-processing and convolutional neural networks for localization and segmentation.

### 4.1. DATASET DESCRIPTION

The medical images contain private patient information, and they are strictly confidential properties. In the

last few years, different algorithm development challenges were conducted that are using advanced data anonymization methods, which allowed the public sharing of medical datasets.

We used a dataset provided by the MICCAI 2017 Multi-Modality Whole Heart Segmentation Challenge. It comprises 120 multi-modality CT and MRI volumetric cardiac images, acquired from the real clinical environment. The dataset is divided into 60 CT, and 60 MRI images, where for each modality, 20 volumes have their corresponding ground truth (labels), and 40 for testing purposes within the challenge. We used only 20 CT volumes since the testing phase is closed at the moment of writing this paper. The provided ground truths contain all heart substructures from the upper abdomen to the aortic arch including the left ventricle (LV), the myocardium of the left ventricle (Myo), the right ventricle (RV), the right atrium (RA), the left atrium (LA), the ascending aorta (AA) and the pulmonary artery (PA).

Furthermore, the images were collected with different scanner types, using cardiac CT angiography, and were stored as NIfTI file format. The slices were acquired in the axial view with the pixel resolution of  $512 \times 512$ . The average slice thickness is 1.60 mm and the average in-plane resolution is about  $0.78 \times 0.7$ , leading to volumetric data consisting of 350 to 500 two-dimensional slices.

#### 4.2. DATA PREPROCESSING

To overcome differences in resolutions and voxel sizes all volumes used for training were pre-processed. We use the intensity normalization by applying a following linear transformation:

$$out_p = (in_p - in_{min}) \cdot \frac{out_{max} - out_{min}}{in_{max} - in_{min}} + out_{min} \quad (6)$$

where  $in_p$  represents all input pixel values,  $out_p$  denotes output pixel values, while  $out_{min}$  and  $out_{max}$  are user defined parameters having values 0 and 20000, respectively. The intensity normalization is applied on each CT volume, individually. Hence, the input to network are resampled three-dimensional volumes of the voxel size  $144 \times 144 \times 144$ , concurrently returning the same size voxels at the final layer in  $x$ ,  $y$  and  $z$  directions. This simple pre-processing step significantly reduces the computational time as well as simplifies the network training.

#### 4.3. LOCALIZATION AND SEGMENTATION NETWORKS

The developed framework, presented in Fig. 2., comprises of two 3D U-Net architectures. The first one localizes the whole heart from the raw CT data, giving the finer region of interest and input for the second network. The second network is used for whole heart segmentation, and it provides only the necessary in-

formation for the spatial regions of the interest from the whole volume. Furthermore, it is essential to mention that used ground-truth bounding boxes represent the whole heart rather than its specific substructures such as the left atrium or the pulmonary artery that are strongly connected.

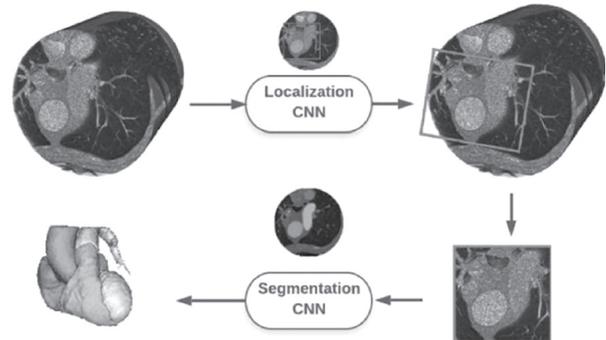


Fig. 2. The flowchart of the proposed method

### 5. EXPERIMENTS AND RESULTS

The approach explained in previous sections is developed using Python, Keras and Tensorflow. The training of huge 3D networks is computationally expensive, so we also include cuDNN to increase memory efficiency. Data augmentation is done on-the-fly, which results in as many different artificial training images as the training iterations. We ran in total 120000 training iterations, simultaneously on two NVidia Geforce Titan V GPUs, which took approximately 25 hours. Furthermore, the segmentation of the new dataset took on average only 7.4 seconds on the same hardware.

#### 5.1. RESULTS

To evaluate our approach, we perform a three-fold cross validation on the training images of the MICCAI 2017 Multi-Modality Whole Heart Segmentation challenge for CT imaging modality, such that each image is tested exactly once.

The quantitatively model performance during training is measured with the dice similarity coefficient. The similarity of the predicted segmentation and the ground-truth label is described with the following formula:

$$dice(Y, Z) = \frac{2 \times |Y \cap Z|}{|Y| + |Z|} \quad (7)$$

where  $Y$  is the segmentation result and  $Z$  is the corresponding ground truth label, while  $| \cdot |$  is utilized to indicate the number of foreground voxels in the ground truth and segmentation images. In this manner, we obtained an overall average score for the whole heart segmentation of 88.2%. The table-like representation of all obtained results for the original 3D U-Net architecture is shown in Table 1. In contrast, the results of our proposed approach with improvements in data augmentation technique is presented in Table 2.

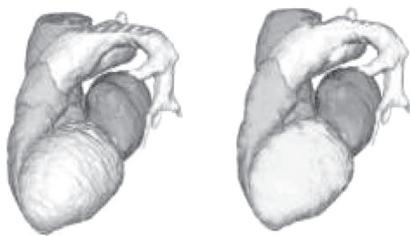
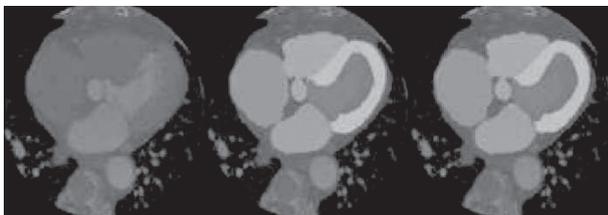
**Table 1.** An obtained Dice Coefficient of the original 3D U-Net architecture

Cardiac structure	Obtained DSC [%]
LV	89.8
RV	86.4
LA	88.2
RA	85.5
PA	83.2
Myo	82.4
Aorta	89.2
WH	87.3

**Table 2.** An obtained dice coefficient of the 3D U-Net architecture with an improved data augmentation

Cardiac structure	Obtained DSC [%]
LV	91.2
RV	87.2
LA	86.5
RA	82.6
PA	83.7
Myo	83.9
Aorta	90.2
WH	88.20

An example of resulting best-segmented dataset, is shown in Fig. 3.



**Fig. 3.** An example of successful segmentation. From the upper left to the right: original 2D slice; slice of the ground truth; 2D slice of the segmentation result; 3D visualization of the ground truth; 3D visualization of the predicted heart

For additional evaluation of overall accuracy, we compared obtained results with similar whole heart segmentation approaches that are based on convolutional neural-networks architectures. The method proposed by Payer et al. [26] is exploring the idea of

individual landmark localization with regressing heat-maps and introducing a novel SpatialConfiguration-Net architecture that is combining local appearance responses with spatial landmark configurations for modelling anatomical variability. Furthermore, Xu, Wu, and Feng in their research [27] show that Faster R-CNN, in combination with the U-Net network, has a significant influence on speed increase. A little different approach is conducted by Wang et al. [28]. They developed a framework that includes the 2.5D segmentation with orthogonal 2D U-nets, shape context estimation, and refining segmentation with U-net and shape context.

The mentioned three approaches, as well as ours, use the same validation dataset and evaluation methods. Therefore, a comparison of the obtained results is highly reliable. The results presented in Table 3. are indicating the competitiveness of our approach to the state-of-the-art.

**Table 3.** The comparison of obtained DSC with the state-of-the-art methods.

Method	Authors	Used approach	DSC [%]
1	Payer et al. [26]	SpatialConfiguration-Net architecture	88.9
2	Xu, Wu and Feng [27]	Orthogonal 2D U-Net	85.9
3	Wang et al. [28]	Faster R-CNN + U-Net	86.6
4	Our method	Localization and segmentation 3D U-Net	88.20

Furthermore, for experimental purposes, we performed training with different initial learning rates and found the optimal learning rate of 0.005 to give best segmentation results as shown in Table 4.

**Table 4.** The different initial learning rates ( $\mu$ ) and obtained dice coefficient for the heart segmentation results.

Cardiac structure	Obtained DSC [%] for $\mu=0.0001$	Obtained DSC [%] for $\mu=0.001$	Obtained DSC [%] for $\mu=0.005$
LV	87.2	90.4	91.2
RV	84.1	86.6	87.2
LA	85.2	87.8	86.5
RA	82.3	84.6	82.6
PA	81.5	85.7	83.7
Myo	80.3	85.9	83.9
Aorta	88.7	89.8	90.2
WH	85.7	86.9	88.20

## 6. CONCLUSION

In this paper, an approach for automatic whole heart localization and segmentation from CT images is presented. The developed framework combines two 3D U-Net neural network architectures, first used for localization of the bounding box around the heart and second used for the segmentation. The results were evaluated on the five CT volumes from the MICCAI 2017 Multi-Modality Whole Heart Segmentation challenge. With additional improvements to the data augmentation process, we achieve an average dice score for the whole heart segmentation of 88.2%.

The results of our proposed approach show that CNN neural network architectures are a suitable approach for cardiac segmentation tasks. The increase of training data would certainly ensure better results, as well as more sophisticated pre-processing and localization methods. For future work, we will investigate techniques that will allow uniform multi-modality pre-processing, which will enable more accessible inputs for network training.

## 7. ACKNOWLEDGEMENTS

The current archival periodical article is based on the conference presentation [29].

This work has been supported in part by Croatian Science Foundation under the project UIP-2017-05-4968.

## 8. REFERENCES:

- [1] The European Heart Network, HeartNet Connection, European Cardiovascular Disease Statistics 2017, <http://www.ehnheart.org/cvd-statistics.html> (accessed: 2019)
- [2] H. Leventić et al., "Left atrial appendage segmentation from 3D CCTA images for occlude placement procedure", *Computers in Biology and Medicine*, Vol. 104, 2018, pp. 163-174.
- [3] M. Habijan, H. Leventić, I. Galić, D. Babin, "Estimation of the Left Ventricle Volume Using Semantic Segmentation", *Proceedings of the 61<sup>st</sup> International Symposium ELMAR, Zadar, Croatia, 23-25 September 2019*, pp. 39-44.
- [4] M. Kass, A. Witkin, D. Terzopoulos, "Snakes: active contour models", *International Journal of Computer Vision*, Vol. 1, No. 4, 1988, pp. 321-331.
- [5] M. Kass, A. Witkin, D. Terzopoulos, "Active contour models", *International Journal of Computer Vision*, Vol. 1, No. 4, 1987, pp. 321-331.
- [6] V. Caselles, R. Kimmel, G. Sapiro. "Geodesic active contours", *International Journal of Computer Vision*, Vol. 22, No. 1, 1997, pp. 61-79.
- [7] S. Kichenassamy A. Kumar, P. Olver, A. Tannenbaum, A. Yezzi, "Gradient flows and geometric active contour models", *Proceedings of the IEEE International Conference on Computer Vision, Cambridge, MA, USA, 20-23 June 1995*, pp. 810-815.
- [8] A. Yezzi S. Kichenassamy, A. Kumar, P. Olver, A. Tannenbaum, "A geometric snake model for segmentation of medical imagery", *IEEE Transactions on Medical Imaging*, Vol. 16, No. 2, 1997, pp. 199-209.
- [9] X. Zhuang and J. Shen, "Multi-scale patch and multi-modality atlases for whole heart segmentation of MRI", *Medical Image Analysis*, Vol. 31, 2016, pp. 77-87.
- [10] X. Zhuang K. S. Rhode, R. S. Razavi, D. J. Hawkes, S. Ourselin, "A Registration-Based Propagation Framework for Automatic Whole Heart Segmentation of Cardiac MRI", *IEEE Transactions on Medical Imaging*, Vol. 29, No. 9, 2010, pp. 1612-1625.
- [11] X. Zhuang et al., "An Atlas-Based Segmentation Propagation Framework Using Locally Affine Registration – Application to Automatic Whole Heart Segmentation", *Proceedings of the International Conference Medical Image Computing and Computer-Assisted Intervention, New York, NY, USA, 6-10 September 2008*, pp. 425-433.
- [12] X. Zhuang, S. Ourselin, R. Razavi, D. L. G. Hill, D. J. Hawkes, "Automatic Whole Heart Segmentation Based on Atlas Propagation with A Priori Anatomical Information", *Proceedings of the Medical Image Understanding and Analysis Conference, Dundee, Scotland, UK, 2008*, pp. 29-33
- [13] X. Zhuang et al., "Multiatlas whole heart segmentation of CT data using conditional entropy for atlas ranking and selection", *Medical Physics*, Vol. 42, 2015, pp. 3822-3833.
- [14] E. Rikxoort et al., "Adaptive Local Multi-Atlas Segmentation: Application to the Heart and the Caudate Nucleus", *Medical image analysis*, Vol. 14, No. 1, 2009, pp. 39-49.
- [15] Y. Chen, X. Xiao-ming, C. Ken, O. Shan-xin, "CT Image Segmentation based on Clustering and

- Graph-Cuts”, *Procedia Engineering*, Vol. 15, 2011, pp. 5179-5184.
- [16] H. W. Jo, H. Y. Lee, “Automatic left ventricle segmentation with K-means clustering and graph searching on cardiac magnetic resonance image”, *International Workshop on Image Analysis for Multimedia Interactive Services*, 2011
- [17] L. Wang, Y. Ma, K. Zhan, Y. Ma, “Automatic Left Ventricle Segmentation in Cardiac MRI via Level Set and Fuzzy C-Means”, *Proceedings of the 2<sup>nd</sup> International Conference on Recent Advances in Engineering & Computational Sciences*, Chandigarh, India, 21-22 December 2015, pp. 1-6.
- [18] J. Margeta et al., “Decision Forests for Segmentation of the Left Atrium from 3D MRI”, *Proceedings of the International Workshop on Statistical Atlases and Computational Models of the Heart*, Nagoya, Japan, 26<sup>th</sup> September 2013, pp. 49-56.
- [19] A. Mortazi, R. Karim, K. Rhode, J. Burt, U Bagci, “CardiacNET: Segmentation of Left Atrium and Proximal Pulmonary Veins from MRI Using Multi-View CNN”, *Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention*, Quebec City, QC, Canada, 10-14 September 2017, pp. 377-385.
- [20] R. P. K. Poudel, P. Lamata, G. Montana, “Recurrent fully convolutional neural net-works for multi-slice MRI cardiac segmentation”, *arXiv*, No. 1608.03974, 2016
- [21] O. Ronneberger, P. Fischer, and T. Brox, “U-Net: Convolutional Networks for Biomedical Image Segmentation”, *Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention*, Munich, Germany, 5-9 October 2015, pp. 234-241.
- [22] Z. Zhou, Md M. R. Siddiquee, N. Tajbakhsh, J. Liang, “UNet++: A nested U-Net architecture for medical image segmentation”, *Proceedings of the International Workshop on Deep Learning in Medical Image Analysis*, Granada, Spain, 20<sup>th</sup> September 2018, pp. 3-11.
- [23] O. Oktay et al., “Anatomically constrained Neural Networks (ACNNs): Application to cardiac image enhancement and segmentation”, *IEEE Transactions on Medical Imaging*, Vol. 37, No. 2, 2018, pp. 384-395.
- [24] Ö. Çiçek, A. Abdulkadir, S. S. Lienkamp, T. Brox, O. Ronneberger, “3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation”, *Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention*, Athens, Greece, 17-21 October 2016, pp. 424-432.
- [25] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, Z. Wojna, “Rethinking the inception architecture for computer vision”, *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition*, Las Vegas, NV, USA, 27-30 June 2016, pp. 2818-2826.
- [26] C. Payer, D. Stem, H. Bischof, M. Urschler, “Multi-label whole heart segmentation using CNNs and anatomical label configurations”, *Proceedings of the International Workshop on Statistical Atlases and Computational Models of the Heart*, Quebec City, QC, Canada, 10-14 September 2017, pp. 190-198.
- [27] Z. Xu, Z. Wu, J. Feng, “CFUN: Combining Faster R-CNN and U-net Network for Efficient Whole Heart Segmentation”, *arXiv*, No. 1812.04914, 2018.
- [28] C. Wang, O. Smedby, “Automatic whole heart segmentation using deep learning and shape context”, *Proceedings of the International Workshop on Statistical Atlases and Computational Models of the Heart*, Quebec City, QC, Canada, 10-14 September 2017, pp. 242-249.
- [29] M. Habijan, H. Leventic, I. Galic, D. Babin, “Whole Heart Segmentation from CT images Using 3D U-Net architecture”, *Proceedings of the 2019 International Conference on Systems, Signals and Image Processing*, Osijek, Croatia, 5-7 June 2019, pp. 121-126.