2021 43rd Annual International Conference of the
IEEE Engineering in Medicine & Biology Society (EMBC)
Oct 31 - Nov 4, 2021. Virtual Conference

Heart Region Segmentation using Dense VNet from Multimodality Images

Aparna Kanakatte[1], Divya Bhatia[1] and Avik Ghose[2]

Abstract— **Cardiovascular diseases (CVD) have been identified as one of the most common causes of death in the world. Advanced development of imaging techniques is allowing timely detection of CVD and helping physicians in providing correct treatment plans in saving lives. Segmentation and Identification of various substructures of the heart are very important in modeling a digital twin of the patient-specific heart. Manual delineation of various substructures of the heart is tedious and time-consuming. Here we have implemented Dense VNet for detecting substructures of the heart from both CT and MRI multimodality data. Due to the limited availability of data we have implemented an on-the-fly elastic deformation data augmentation technique. The result of the proposed has been shown to outperform other methods reported in the literature on both CT and MRI datasets.**

I. Introduction

According to the world health organization (WHO), cardiovascular death (CVD) is the leading cause of death globally, taking an estimated 17.9 million lives each year. CVD's are a group of disorders of the heart and blood vessels including coronary heart disease, cerebrovascular disease, and rheumatic heart disease. The morphological and pathological information from the three-dimensional (3D) medical imaging modalities like magnetic resonance imaging (MRI), computed tomography (CT) or computed tomography angiography (CTA) enables non-invasive qualitative and quantitative assessment of cardiac anatomical structures and functions providing support for diagnosis disease monitoring, treatment planning, and prognosis.

Whole heart segmentation (WHS) which delineates substructures of the heart is very valuable for modeling and analysis of the anatomy and functions of the heart. Typically anatomical substructure of interest in the heart region includes Left Ventricles (LV), Right Ventricles (RV), Left Atrium (LA), Right Atrium (RA), Aorta, Myocardium (Myo), and Pulmonary Artery (PA). Fig 1 shows the key anatomical regions of the heart [1]. Cardiac image segmentation and identification of various regions of the heart is very important in obtaining quantitative measures like myocardial mass, wall thickness, left and right ventricle volume, and ejection fraction (EF).

Manual delineation of these substructures is labor-intensive, tedious, and time-consuming for a single subject [2]. More importantly, manual delineation is subject to intra and inter-observer variations. Hence, automation of segmentation from multi-modality images is highly sought out but very challenging due to the following complexity [3]. Firstly, the shape of the heart with multiple chambers and great vessels is complex in geometry and varies largely in different subjects or even for the same subject at various cardiac phases, especially with pathological and physiological changes. Secondly, the image quality can be variable like the enhancement patterns of CT images can vary significantly for different scanners or acquisition sessions. Finally, due to the complex motions and blood flow within the heart, the imaged data may contain motion artifacts, intensity inhomogeneity, poor contrast-to-noise ratio, and signal-to-noise ratio which can significantly deteriorate the image quality. In this paper, we attempt to delineate the substructures of the heart from the multi-modality CT and MR 3D images.

[1]Aparna Kanakatte and Divya Bhatia is with TCS Research and Innovation, Bengaluru, India (aparna.kg, bhatia.divya)@tcs.com
[2]Avik Ghose is with TCS Research and Innovation, Kolkata, India avik.ghose@tcs.com

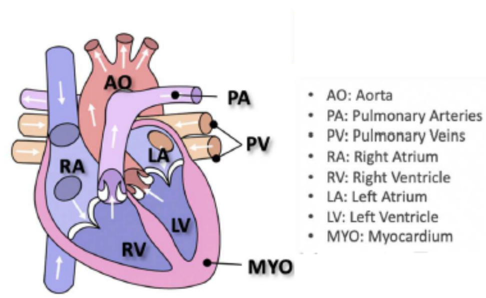


Fig. 1: Anatomy of human heart. Source [1]

II. Related work

Various approaches are proposed in the literature to delineate various substructures of the heart. The detailed review of previously published algorithms for heart segmentation from CT and MRI datasets can be found in [3]. Olivier *et al* [4] and Xuan *et al* [5] have used deformable models to segment specific anatomical structures of the heart. Zhuang *et al* [6] used registration-based propagation framework using atlas for MR images. Xiahai *et al* [7] used multi-scale patch and multi-modality atlases for whole heart segmentation.

The recent success of deep learning has influenced several researchers. There are several works where authors have addressed the whole heart segmentation problem by using deep network-based approaches. A detailed review paper on cardiac image segmentation using deep learning has been provided by Chen *et al* [1]. Wang *et al* [8] used deep learning and shape context to segment the heart regions. Habijan *et al* [17] proposed a convolutional neural network approach for the whole heart segmentation based upon the 3D U-Net architecture incorporating principle component analysis as an additional data augmentation technique. Payer *et al* [10] proposed a pipeline of two fully convolutional networks for automatic multi-label whole heart segmentation from CT and MRI volumes. Mortazi *et al* [22] used a multi-planar deep segmentation network for cardiac substructures from MRI and CT. A combination of faster R-CNN and U-Net (CFUN) was proposed by Xu *et al* for efficient whole heart segmentation [18]. In this paper, we use Dense VNet to segment various substructures of the heart. As the amount of data is limited we use the elastic deformation data augmentation technique on-the-go to reduce storage requirements. As mentioned earlier the heart undergoes deformation during various cardiac phases and having elastic deformations will help in better data generation for training the Dense network.

III. Proposed Method

For performing semantic segmentation and identification of various substructures of the heart we use Dense VNet, a fully connected convolutional neural network that has demonstrated success in establishing voxel-voxel connections between input and output images. Dense VNet was proposed by Gibson [12]. It consists of three layers of dense feature stacks [11] whose outputs are concatenated after a single convolution in the skip connection and bilinear upsampling. The Dense VNet architecture is shown in fig 2. A 72^3 feature maps are computed using a strided convolution. Then a cascade of dense feature stacks and strided convolutions generate activation maps at three resolutions. Then a convolution unit is applied at each resolution to reduce the number of features. Next a bilinear upsampling is applied to get back to 72^3, the maps are concatenated and a final convolution generates the likelihood logits. Finally, these are added to the upsampled spatial prior to generate the segmentation logit [12].

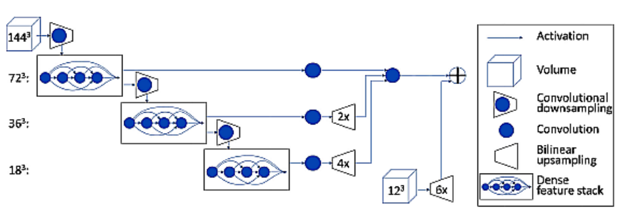


Fig. 2: Architecture of Dense VNet

It has to be noted that annotated medical volumes are not easily available as one or more experts are required to manually trace a reliable ground-truth annotation which is highly time-consuming and costly. In this work, due to less training dataset availability, we have augmented the original training dataset to increase precision and robustness on the test dataset. Instead of using the regular linear augmentation techniques like translating, rotating, flipping, stretching, etc, we have used the elastic deformation technique. It is very similar to stretching but with increased freedom as elastic deformation allows to change the image almost like kneading a stress ball [15]. However, care should be taken such that overdoing this can result in almost unrecognizable training images. We have chosen this elastic deformation considering the cardiac motion life cycle that includes continuous blood-pumping action of the heart which results in movement of different regions of the heart. These movements can be mimicked by elastic deformation to some extent by generating additional training inputs with deformations which will assist in increasing the efficiency of the trained model.

During each training iteration, the deformed versions of the training images are randomly fed as input to the network using a dense deformation field obtained through a $2 \times 2 \times 2$ grid of control-points and B-spline interpolation [16]. To reduce the excessive storage requirements, the augmentation has been performed "on-the-go", before each optimization iteration.

IV. Results and Discussion

The proposed method with elastic data augmentation is evaluated on the MM-WHS challenge dataset in conjunction with MICCAI 2017. This dataset consists of 20 CT and 20 MRI volumes with corresponding manual segmentations of seven whole heart substructures LV, RV, LA, RA, myocardium of the left ventricle, pulmonary artery and the ascending aorta. The volumes were acquired in clinical settings with different scanners, with varying image quality, resolution and voxel spacing. The maximum physical size of the input volumes for CT is 300 x 300 x 188 mm^3 while for MRI it is 400 x 360 x 400 mm^3.

A. Implementation details

Ubuntu workstation with Intel Core i7, 32GB RAM, and NVIDIA GeForce 1080i GPU card is used for training and testing. The proposed approach is implemented using the tensorflow [13] based deep learning platform called NiftyNet [14] version 0.6.0. NiftyNet is a modularly structured deep learning platform specifically built for medical image analysis. It is capable of handling various medical imaging formats like Nifti (nii), mha, dicom, etc. This is an end-to-end platform that includes various modules for pre-processing, data augmentation, training, evaluation, and inference. This platform also includes prebuilt algorithms and models which can be modified to suit the requirements. The necessary pre-processing of dataset normalization for both CT and MR data has been applied before feeding into the algorithm for training.

The loss function used is the weighted sum of L2 regularization loss with label-smoother probabilistic dice scores for each organ. To mitigate the extreme class imbalance (organ occupancy in an image), dice score hinge loss penalizing dice scores thresholds were introduced after warm up periods of 25 and 100 iterations respectively. The network is trained separately for CT and MR data using Adam optimizer with a learning rate of 0.001 for 10000 iterations.

B. Performance comparison

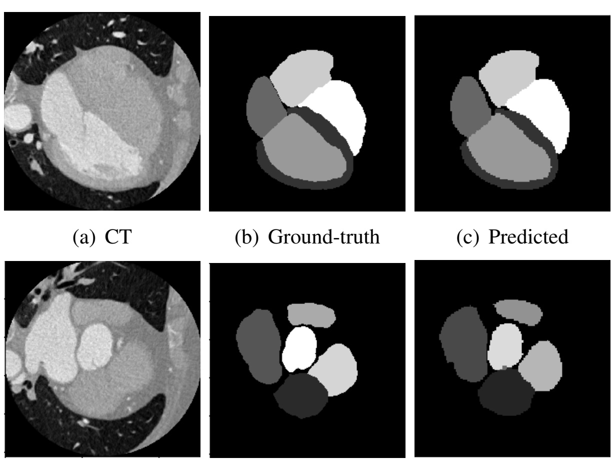
The performance comparison of the algorithm in segmenting various heart substructures is evaluated quantitatively by using Dice scores. The Dice similarity coefficient of the predicted segmentation and the ground-truth (GT) label is calculated by

$$Dice(X,Y) = \frac{2\,|X \cap Y|}{|X| + |Y|} \quad (1)$$

where X is the segmentation result and Y is the corresponding ground-truth label, while $||$ is to indicate the number of foreground voxels in ground-truth and segmentation images.

1) Model evaluation with and without elastic data augmentation: The performance of the model with and without elastic data augmentation is provided in the Table I. It can be seen that performing elastic data augmentation has increased the overall segmentation accuracy for both CT and MRI datasets. Due to the pumping nature of the heart, the various chambers of the heart get deformed and as a result, we have used elastic deformation to mimic the deformations of organs in images. As the data augmentation is performed on-the-fly, this reduces the storage requirement.

2) Model evaluation on CT data: Fig 3 shows the output of the proposed model with the original image and GT. The different shades of gray in segmented result suggests different substructures of heart. The quantitative evaluation and comparison of the proposed Dense VNet with other methods reported in the literature on CT data is provided in Table II. It can be seen that the proposed method outperforms other reported methods. Our method has a better dice score for LV, RV, and RA segmentation. The whole heart segmentation (WHS) average including background is better for our method with 91.0 compared to the previous best of 90.8 reported by Payer *et al* [10]. The 3D model of the predicted and GT is shown in fig 4. Different colors in the 3D model indicate segmented substructures.

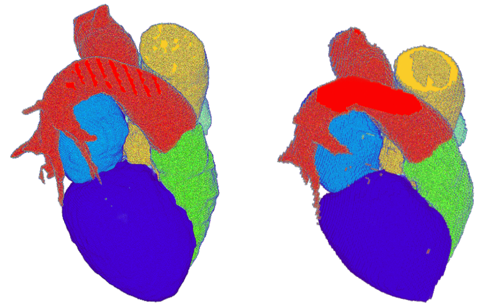


(a) CT (b) Ground-truth (c) Predicted

(d) CT (e) Ground-truth (f) Predicted

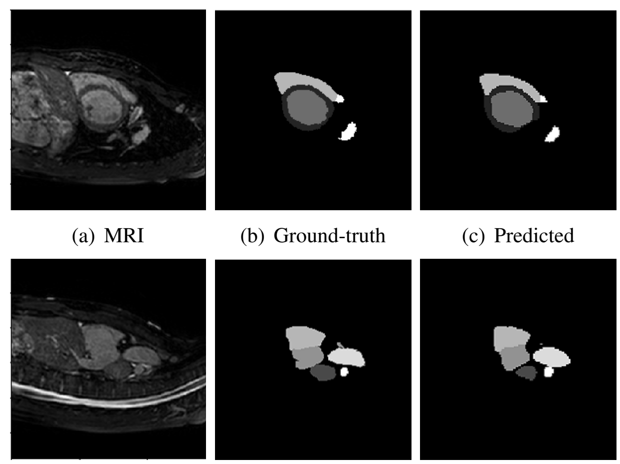
Fig. 3: Results of the proposed method on CT data

3) Model evaluation on MRI data: Fig 5 shows the output of the proposed model with the original image and GT. The quantitative evaluation and comparison of the proposed with other methods reported in the literature on MRI data is provided in Table III. It can be seen that the proposed method outperforms other reported methods. Our method has a better dice score for LV, Myo, RV, PA segmentation and on par with Payer *et al* [10] for LA. The WHS average including background is better for our method with 88.1 compared to the previous best of 86.3. The 3D model of the predicted and GT is shown in fig 6. It has to be noted that different colors in the 3D model indicate segmented substructures.



(a) Ground-truth (b) Predicted Dice = 0.94

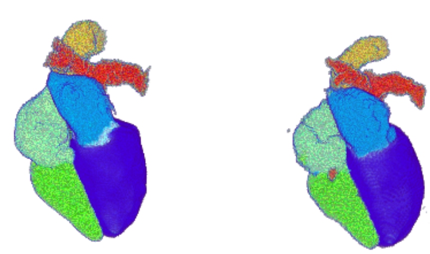
Fig. 4: 3D segmented results on CT data



(a) MRI (b) Ground-truth (c) Predicted

(d) MRI (e) Ground-truth (f) Predicted

Fig. 5: Results of the proposed method on MRI data



(a) Ground-truth (b) Predicted Dice = 0.88

Fig. 6: 3D segmented results on MRI data

TABLE I: Comparison of the Dice score of the proposed with and without elastic data augmentation on CT/MRI data

	LV	Myo	RV	LA	RA	Aorta	PA	WHS Average
Without Data Augmentation	91.4/92.7	83.9/83.3	89.4/90.4	91.3/83.0	89.4/85.5	88.7/79.1	80.9/77.1	89.3/86.2
With Elastic Data Augmentation	93.1/93.8	87.4/83.9	91.0/91.4	92.5/85.5	90.7/87.7	89.5/82.5	83.4/79.7	91.0 /88.1

TABLE II: Comparison of the Dice score of the proposed with other methods reported on CT data

	Method	LV	Myo	RV	LA	RA	Aorta	PA	WHS Average
Payer *et al* [10]	Seg-CNN	91.8	**88.1**	90.9	92.9	88.8	**93.3**	84.0	90.8
Yang *et al* [19]	3D CNN	92.3	85.6	85.7	**93.0**	87.1	89.4	83.5	89.0
Marija *et al* [17]	3D-U-Net	92.8	85.8	88.7	90.5	85.1	92.7	**85.1**	89.0
Mortazi *et al* [22]	Multiplanar deep network	90.4	85.1	88.3	91.6	83.6	90.7	78.4	87.9
Xu *et al* [18]	CFUN	87.9	82.2	90.2	83.2	84.4	91.3	82.1	NA
Proposed	Dense VNet	**93.1**	87.4	**91.0**	92.5	**90.7**	89.5	83.4	**91.0**

TABLE III: Comparison of the Dice score of the proposed with other methods reported on MRI data

	Method	LV	Myo	RV	LA	RA	Aorta	PA	WHS Average
Payer *et al* [10]	Seg-CNN	91.6	77.8	86.8	**85.5**	**88.1**	**88.8**	73.1	86.3
Yang *et al* [19]	3D CNN	75.0	65.8	75.0	82.6	85.9	80.9	72.6	78.3
Wang *et al* [20]	3D U-Net	86.3	74.4	84.9	85.2	84.0	82.4	78.8	83.2
Tong *et al* [21]	3D supervised U-Net	70.2	62.3	68.0	67.6	65.4	59.9	47.0	67.4
Mortazi *et al* [22]	Multiplanar deep network	87.1	74.7	83.0	81.1	75.9	83.9	71.5	81.8
Proposed	Dense VNet	**93.8**	**83.9**	**91.4**	**85.5**	87.7	82.4	**79.7**	**88.1**

V. Conclusion and Future Work

In this paper, we have shown a better heart substructure segmentation network on a multi-modality dataset by using Dense VNet with elastic deformation data augmentation technique. The motivation for using elastic deformation is the deformation of various substructures during the normal cardiac action of the heart. We have shown that the proposed data augmentation technique has increased the accuracy of segmentation. The proposed method helps in generating a patient-specific heart model with a higher WHS dice score accuracy of 91.0 and 88.1 in segmenting various substructures of the heart for both CT and MRI datasets respectively. Our future work includes improving the accuracy and identifying various structural anomalies from the segmented 3D data.

References

[1] Chen. C, Chen. Q, Huaqi. Q, Giacoma. T, Jinming. D, Wenjia. B and Daniel. R: Deep learning for cardiac image segmentation: A Review, Frontiers in Cardiovascular medicine, vol 7, article 25, 2020

[2] Zhuang. X and Shen. J: Multi-scale patch and multi-modality atlases for whole heart segmentation of MRI, Medical image analysis, vol 31, pp 77-87, 2016

[3] Zhuang. X: Challenges and methodologies of fully automatic whole heart segmentation: A review, Journal of Healthcare Engineering, vol 4 (3), 371-407

[4] Olivier Ecabert et al: Automatic Model-Based Segmentation of the Heart in CT Images, IEEE Trans on Med Img, vol.27, pp. 1189–1201

[5] Xuan. Z, Yao. W and Gabor. J: Robust shape-constrained active contour for whole heart segmentation in 3-D CT images for radiotherapy planning, ICIP, pp 1-5, doi: 10.1109/ICIP.2014.7024999

[6] Zhuang. X, Rhode K, Razavi R, Hawkes D.J and Ourselin S.: A Registration-Based Propagation Framework for Automatic Whole Heart Segmentation of Cardiac MRI. IEEE Transactions on Medical Imaging, 29 (9): 1612-1625, 2010

[7] Xiahai. Z and Juan. S: Multi- scale patch and multi-modality atlases for whole heart segmentation of MRI, Medical Image Analysis, vol.31, pp.77 - 87, 2016

[8] Wang. C and Smedby. O: Automatic Whole Heart Segmentation Using Deep Learning and Shape Context.,2018, 10.1007/978-3-319-75541-0 26.

[9] Habijan. M, Leventić. H, Galić. I and Babin. D: Whole Heart Segmentation from CT images Using 3D U-Net architecture, IWSSIP, 2019, pp 121-126

[10] Payer. C, Stern. D, Bischof. H and Urschler. M: Multi-label whole heart segmentation using CNNs and anatomical label configurations, International Workshop on Statistical Atlases and Computational Models of the Heart, Springer. pp190-198.

[11] Huang. G, Liu. Z, Van Der. M. L, Weinberger. K Q: Densely connected convolutional networks, CVPR, pp 4700-4708, 2017

[12] Eli. G, Francesco. G, Yipeng. H, Ester. B, Steve. B, Kurinchi. G, Brian R. D, Stephen P. P, Mathew J. C, Dean C. B: Automatic Multi-Organ Segmentation on Abdominal CT With Dense V-Networks. IEEE Trans. Medical Imaging 37(8): 1822-1834 (2018)

[13] Abadi. M, Agarwal. A, Barham. P, Brevdo. E, Chen Z, et. al:, "Tensor-Flow: Large-scale machine learning on heterogeneous systems", 2015

[14] Gibson. E, Li. W, Sudre. C, et al. NiftyNet: a deeplearning platform for medical imaging, Comput Methods Programs Biomed. 2018;158:113-122

[15] Castro. E, Cardoso. J.S and Pereira. J.C: Elastic deformations for data augmentation in breast cancer mass detection, BHI, pp 230 - 234

[16] Milletari. F, Navab. N and Ahmadi. S: V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation, Fourth International Conference on 3D Vision (3DV), 2016, pp. 565-571

[17] Habijan. M, Leventić. H, Galić. I and Babin. D, Whole Heart Segmentation from CT images Using 3D U-Net architecture, IWSSIP, 2019, pp. 121-126

[18] Xu. Z,Wu. Z, Feng. J: CFUN: combining faster R-CNN and U-net network for efficient whole heart segmentation. arxiv (2018), http://arxiv.org/abs/1812.04914

[19] Yang X, Bian C, Yu L, Ni D and Heng P.A: 3D convolutional networks for fully automatic fine-grained whole heart partition, International Workshop on Statistical Atlases and Computational Models of the Heart, Springer. pp181-189.

[20] Wang C, MacGillivray T, Macnaught G, Yang G and Newby D: A two-stage 3D Unet framework for multi-class segmentation on full resolution image. http://arxiv.org/abs/1804.04341

[21] Tong Q, Ning M, Si W, Liao X, Qin J. 3D deeply-supervised U-net based whole heart segmentation. International Workshop on Statistical Atlases and Computational Models of the Heart, Springer. pp 224 - 232

[22] Mortazi. A, Burt. J and Bagci. U. Multi-planar deep segmentation networks for cardiac substructures from MRI and CT. International Workshop on Statistical Atlases and Computational Models of the Heart, Springer. pp 199 - 206