AUTOMATIC SEGMENTATION OF THE ECHOCARDIOGRAPHIC IMAGES USING MACHINE LEARNING

A Thesis

by

TAEOUK KIM

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Chair of Committee,	Iman Borazjani
Committee Members,	Vinayak Krishnamurthy
	Michael Moreno
Head of Department,	Andreas A. Polycarpou

May 2020

Major Subject: Mechanical Engineering

ABSTRACT

Two-dimensional echocardiography (2D echo) is the most widely used cardiac imaging techniques in clinical applications. Boundary delineation of the heart, especially the left-ventricle (LV), is essential to calculate the clinical parameters. Currently, LV segmentation from 2D echo is conducted manually or using semi-automatic techniques. In this study, machine learning techniques were employed. U-net, which is a fully convolutional network, and segAN, which is a generative adversarial network, were trained and evaluated. Training was conducted on the in-house dataset, which consists of 2108 porcine images from 10 different subjects. This dataset was the first dataset, which consists of six standard projections of 2D echo over the entire cardiac cycle. Transfer learning was used for long-axis projections to compensate the limitation of in-house dataset using Cardiac Acquisitions for Multi-structure Ultrasound Segmentation dataset. The models were evaluated on test images by computing metrics such as the dice metric. U-net and segAN models outperformed the level-set method, a traditional segmentation technique. The average dice metric of U-net was 0.903 for LV cavity and 0.787 for LV myocardium. The average dice metric of segAN was 0.912 for LV cavity and 0.801 for LV myocardium. Previous reconstruction algorithm was improved and validated to generate the 3D LV geometry from segmented images. Physiological parameters were calculated from reconstructed geometries with about 15% error which is similar to the previous methods using 2D echo compared to the gold standard MRI. In this study, fully automated algorithm to generate 3D LV geometry from 2D echo images was introduced

by combining machine learning segmentation technique and 3D reconstruction algorithms. This algorithm facilitates the patient-specific LV modeling and simulation without expert's knowledge and effort.

ACKNOWLEDGEMENTS

I would like to thank my committee chair, Dr. Iman Borazjani, and my committee members, Dr. Krishnamurthy and Dr. Moreno, for their guidance and support throughout the course of this research.

Thanks also go to my friends and colleagues and the department faculty and staff for making my time at Texas A&M University a great experience.

Finally, thanks to my family for their encouragement and love.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supervised by a thesis committee consisting of Professor Iman Borazjani, Vinayak Krishnamurthy, and Michael Moreno of the Department of Mechanical Engineering.

The data of Section 2.1 was provided by Professor Marek Belohlavek and Assistant Veronica Vaitkus of the Department of Internal Medicine of Mayo Clinic.

All other work conducted for the thesis was completed by the student independently.

Funding Sources

There are no outside funding contributions to acknowledge related to the research and compilation of this document.

TABLE OF CONTENTS

Page

LIST OF FIGURES

Page

LIST OF TABLES

Page

1. INTRODUCTION

Heart disease are one of the major reasons for the human death in United States of America [1]. Diagnosis of heart disease is essential for fighting this disease, which requires analysis of cardiac morphology and functions from medical images. The function of heart, especially left ventricle (LV), is highly related to its shape, size, and the load exerted on blood during the myocardium shortening [2]. In fact, LV shape and volume changes with time, e.g. ejection fraction (EF) and cardiac volume (CV), are good indicators of some heart diseases like functional mitral regurgitation (MR) [3] and acute myocardial infarction (AMI) [4]. In order to calculate these clinical parameters, the boundary of LV and myocardium need to be identified from low-level images using the techniques such as segmentation and tracking and reconstruct into the 3D geometries.

In this study, images of the LV at six standard projections were segmented using machine learning techniques (U-net and segAN) over one cardiac cycle and reconstructed using 3D reconstruction algorithm. The results of clinical parameters and 3D reconstruction models using automated segmentation versus expert delineation were compared against each other to validate. The rest of the introduction is organized as follows: Section 1.1 describes the advantages of two-dimensional echocardiography. Section 1.2 discusses the previous LV segmentation methods. Section 1.3 describes the application of machine learning techniques on biomedical image segmentation tasks. Section 1.4 emphasizes the dataset on the machine learning techniques and introduces the

previous datasets related to LV segmentation tasks. Finally, Section 1.5 discusses about the 3D reconstruction algorithms using 2D images.

1.1. 2D echocardiography

Among various imaging techniques, two-dimensional echocardiography (2D echo) is one on the most widely used methods for evaluation of heart disease in clinical practice. Following advantages are the reason why 2D echo is the most widely used non-invasive method. First, 2D echo generates high temporal resolution images (50-250 frames per second (fps)) which contain more accurate LV motion information within the heart cycle [5]. Also, 2D echo needs short acquisition time which allows real time analysis [6]. Standardized scanning locations for both long-axis and short-axis which can act as the criteria planes facilitates the clinical analysis from the 2D echo images [7]. Finally, 2D echo is relatively inexpensive and does not need advance preparation for the testing. 2D echo provides a gray scale image with anatomical features, e.g., LV cavity, LV myocardium, and heart valves, are identified to calculate cardiac functions. The first step to quantification of the cardiac functions from the 2D echo images is segmentation of the LV from the gray scale 2D echo images.

1.2. LV segmentation task

At present, manual delineation conducted by experts and semi-automatic segmentation methods are the main techniques which used for LV segmentation from 2D echo images [8]. However, these methods are time-consuming and based on the experts'

subjective judgment, which is prone to intra- and inter-observer variability [9]. To address these problems, automate segmentation methods are necessary to accelerate and generate the consistent result. However, applying automatic segmentation method to 2D echo is harder than other imaging techniques such as magnetic resonance imaging (MRI) and computerized tomography (CT).

The difficulty of automatic segmentation task of 2D echo images stems from inherent properties of 2D echo images [10]. First, 2D echo represents poor contrast between LV myocardium and blood pool inside the LV. For this reason, previous segmentation methods which use contrast difference between LV myocardium and cavity as their criteria do not work well. Second, 2D echo has brightness inhomogeneities in the same structures. Third, there is a variation in the speckle pattern along the cardiac structures due to the different probe orientation from different measuring experts. Also, 2D echo generates similar intensities of myocardium, trabeculae, papillary muscles, and heart valves. Finally, there is variability of shape, intensity, and heart motion across the patients.

To date, several methods have been proposed for the automatic segmentation of a LV. A review of these methods can be found in [11]. They can be generally classified as: pixel classification [12], image-based methods [13], deformable methods [14], active appearance and shape models (AAM/ASM) [15] and atlas models [16]. Pixel classification, image-based and deformable methods require an extensive user interaction and suffer from a low robustness and accuracy [17]. Alternatively, model-based methods such as AAM/ASM and atlas models can overcome the problems with previous methods and

reduce user interaction at the expense of a large training set to build a general model. However, it is very difficult to build a model that is general enough to cover all possible shapes and dynamics of the heart chambers [11]. In this study, state-of-the-art deep learning segmentation method is applied to generate a robust and accurate segmentation from 2D echo images.

1.3. Machine learning techniques on biomedical image segmentation

Recently, medical image analysis has been revolutionized with successful adoption of machine learning techniques [17]. This success has primarily been powered by supervised machine learning methods, which map an input to desired output by training with input-output pair data. Three essential elements of supervised machine learning technique are enough data, computational ability, and appropriate machine learning model. Rapid development of Graphic Processor Unit (GPU) has offered high computation ability for training the machine learning models. Also, lots of large open datasets related to biomedical images, e.g., brain, lung, and heart, have been made public in competitional way [18-20]. Finally, powerful machine learning models have been developed with convolutional neural networks (CNNs). CNNs typically operate on image data and provide one prediction per one image sample, e.g., an image class label or quantitation of disease burden [21]. Recently, most of the medical image analyses using machine learning techniques have employed CNN models [22, 23]. For the biomedical image segmentation tasks, U-net has achieved a high performance [24]. In this study, a new dataset for LV segmentation from 2D echo images was created and two different CNN models, U-net and

segAN [25], were applied on it. Their segmentation performance were compared to each other and previous segmentation methods.

1.4. Dataset for the machine learning LV segmentation technique

As mentioned at the previous section (Section 1.3), a large and proper dataset, which is used to train the CNN model, is one of the most important prerequisites for applying CNNs to medical images. Currently, most of the open LV segmentation datasets for training CNN models are based on the MRI and CT images because they are considered as the standard format [26]. Followings are the open dataset for the LV segmentation tasks. The Left Ventricle Segmentation Challenge (LVSC) dataset, offered by the Medical Image Computing and Computer Assisted Intervention Society (MICCAI), consists of 100 fully delineated MRI images of the short-axis LV [27]. Automated Cardiac Diagnosis Challenge (ACDC) dataset, also organized by the MICCAI, consists of 100 patients' short-axis MRI images [28]. Each patient's images cover the whole LV geometry from the base to the apex with a thickness of 5 mm interval. Besides the challenge datasets, individual researchers have generated their own in-house dataset to train their models with different objects. Bai et al. generated a large dataset, which consists of 4,875 patients' total 93,500 images including both short- and long-axis MRI LV images acquired from the UK Biobank [29]. Majd et al. created in-house dataset with cardiac CT angiography (CCTA) short-axis images from 60 patients and trained their CNN model which combined with voxel classification method [30]. Lieman et al. generated their in-house dataset which consisted of 1,143 short-axis MRI images and applied their ENet model for the

segmentation task [31]. With these large and open datasets, different CNN models have been developed and applied to LV segmentation tasks on MRI and CT images. Patravali et al. showed that 3D U-net [32] outperformed the simple 2D U-net [24] for segmenting the stacked volumetric images based on the results tested on ACDC data. Tan et al. used a CNN regression model to both segment the LV and calculate the radius of LV endocardium and epicardium in short-axis using LVSC data [33].

However, the large dataset for the 2D echo LV segmentation is hard to find. Current, there is only one open dataset, Cardiac Acquisitions for Multi-structure Ultrasound Segmentation (CAMUS) dataset [10]. CAMUS dataset consists of 450 patients' long-axis 2-chamber and 4-chamber projection images. Each projection images contain end-diastolic and end-systolic images. As a result, CAMUS dataset consists of 1800 different images. For this reason, researchers have constructed in-house 2D echo datasets for their specific studies. Veni et al. constructed in-house dataset which consists of 69 images from 4-chamber projection and trained their revised U-net, which combined with the anatomical priors [34]. Zhang et al. generated their in-house dataset using three long-axis, 214 images of 2-chamber, 141 images of 3-chamber, and 182 images of 4-chamber projections, and one short-axis projections, 124 images [35]. They trained their U-net model for segmentation and used the segmented results for auto-diagnosis of the cardiac disease. The limitation of the previous datasets is that they only were concentrated on end-diastolic and end-systolic images and consisted of several projections, e.g., one projection from Veni et al. and four projections from Zhang et al. However, to obtain more precise information about cardiac function, more projections and more time steps' images

are needed. In this study, an in-house dataset, which consists of six standard projections, three long-axis and three short-axis, images over the whole cardiac cycle, was generated for the LV segmentation task. To compensate the small amount of training dataset, transfer learning was conducted on training long-axis projections using CAMUS dataset.

1.5. 3D reconstruction

To analyze the comprehensive dynamic 3D LV geometry and to be a model for the computational fluid dynamic (CFD) simulation, 3D geometry is essential. Currently, MRI, 3D echocardiography (3D echo), and CT methods are used to get 3D LV geometry directly from the medical devices [36, 37]. However, these methods have several weak points. First, they need long acquisition times and high costs. Furthermore, they typically generate dynamic 3D model with 20-30 frames per second (fps) temporal resolution [38]. This is not enough to get accurate quantitative LV motion information with different time resolution in cardiac cycle [5, 39]. In this respect, a 3D reconstruction model based on the 2D echo images can be a good alternative.

Dekker et al. introduced the first 3D reconstruction process with 2D echo images [40]. Since then, several 3D reconstruction approaches using 2D echo images have been proposed. Sawada et al. [41] and Raichlen et al. [42] reconstructed the 3D model based on the multiple short-axis projections which were made at some intervals from apex to base. Also, Ghosh et al. generated LV geometry using rotation of long-axis projections[43]. However, if the number of 2D images was not enough, there was accuracy limitation with using only one kinds of sections, long- or short axis [44]. Gustavsson et al. used three

long-axis and two short-axis projections for the 3D LV reconstruction [45]. However, a major purpose of these previous 3D reconstruction researches was calculating the cardiac volume and visualizing the LV motion. In this study, the in-house 3D reconstruction algorithm, which creates LV models for CFD simulation, was improved and used to generate the 3D LV geometries [38].

2. METHODS

Six standard projections of 2D echo images, three long-axis: 3-chanber, 2-chamber, and 4-chamber projection, and three short-axis: base, mid, and apex projections were segmented using CNN models. 3D LV geometries were reconstructed from the segmented results.

2.1. 2D echo image

2.1.1. Image acquisition

In this study, 2D echo images from porcine hearts were used for the in-house dataset. In cardiovascular research, pig hearts alternate the human hearts due to their similarity in shape, size, and physiological parameters, such as pumping capacity, with human hearts [46]. 2D echo of in situ porcine images were captured over whole heart cycle (R-wave to R-wave) during the open chest pig surgery using a Vivid 7 ultrasound system (GE Vingmed Ultrasound AS, Horton, Norway) and an M4S transducer operating at 1.7/3.4MHz (fundamental/harmonic) frequency by researchers at Mayo clinic. To obtain six different projections' images, the ultrasound transducer was place directly on the proper LV surface location. A transmission gel was added between the transducer and the LV surface to guarantee the acoustic coupling. The frame rate of obtained images was between 36 to 55 fps.

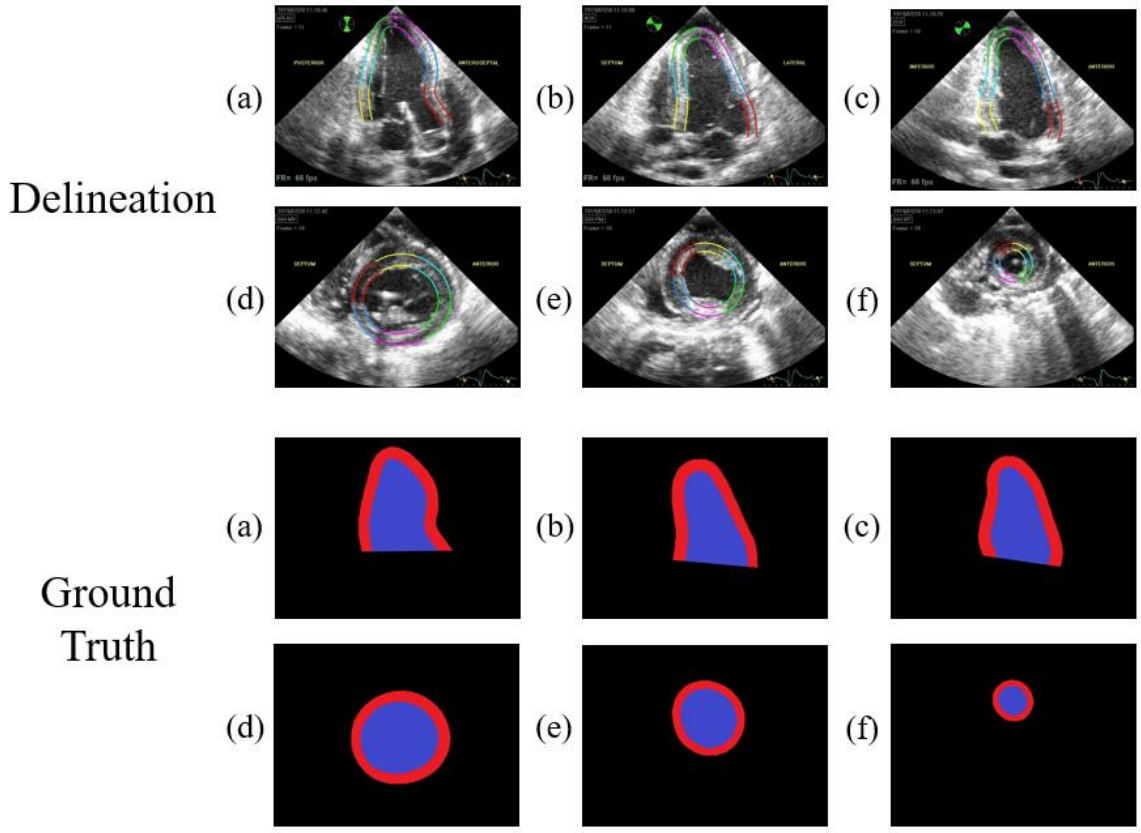

Figure 2.1 The delineated and ground truth 2D echo images for one time instant. LV boundary is delineated by the experts using the speckle tracking software and converted into the CNN proper images. (a), (b), and (c) are long-axis 3-chamber, 4-chamber, and 2-chamber projections. (d), (e), and (f) are short-axis base, mid and apex projections.

2.1.2. Reference delineation and contouring protocol

Supervised machine learning techniques need both input data, for the segmentation task of the original 2D echo images, and output data, for the segmentation task of the delineated results. To obtain a delineated images for the output data, acquired 2D echo images were delineated by experts using an interactive speckle tracking software (EchoPAC, GE Healthcare). Experts manually chose LV inner and outer boundaries at the

first time step images. Then, speckle tracking software offered the images' delineation for the whole cycle. Manual detection of the LV myocardium and cavity over the whole images is error-prone and subjective due to inherent 2D echo image dropouts and noise. In this reason, a semi-automatic speckle tracking band replaced the LV boundary. In cardiovascular research, speckle tracking methods have been used to obtain LV properties [47, 48]. *Figure 2.1 - Delineation* shows the six standard projection images' delineation results at the given time instance. These delineated images were converted into the proper images to be used for the ground truth (reference) output data (*Figure 2.1 – Ground Truth*). Red region represents the LV myocardium and blue region represents the LV cavity.

2.2. Machine learning segmentation

2.2.1. Dataset

In this study, two different datasets were employed to train CNN models. First one is fully annotated CAMUS dataset [10]. The CAMUS dataset consists of 1800 images. Images are obtained from 450 patients' long-axis 2- and 4- chamber projections. Each projection has end-diastolic and end-systolic images. The ground truth image has four labels - 0: background, 1: LV cavity, 2: LV myocardium, and 3: left atrium. Second dataset is in-house dataset, which consists of three long- and three short-axis projection images. The ground truth images (*Figure 2.1 – Ground Truth*) has three labels - 0: background, 1: LV cavity, and 2: LV myocardium. Table 2.1 shows the total number of image and used porcine of each projections. Among total datasets, 80% of the images were assigned to the

Table 2.1 Information of the in-house dataset

Projections	Number of image	Number of porcine
3-chamber	374	10
4-chamber	302	8
2-chamber	306	8
Base	377	10
Mid	375	10
Apex	374	10

training set, which is used to train the CNN models and remaining 20% of images were allotted to the testing set, which were used for validating the trained CNN models. Among 8 to 10 porcine, 2 pigs' data, 77 to 81 images, were assigned to the testing set for each projections. Exact training and testing set information is at APPENDEX A.

2.2.2. Data augmentation

Ronneberger et al. showed that aggressive image augmentation techniques facilitates CNN models, especially U-net, to generate promising segmentation results with little amount of training data [24]. In this study, three data augmentation techniques – rotating, cropping, and blacking out - were employed to supplement the limitation of in-house dataset. The images were rotated with 0.7 probability. They rotated with arbitrary angle between -25 and 25 degree. The background blank part which generates after rotating was filled using a spline interpolation method. Cropping filled out a random size

of left, right, and bottom parts of the images with 0.7 probability. Lastly, black out technique, which was introduced by Zhang et al., was used [35]. They generated circular areas inside the images which have zero intensities. In this study, circular area, which has random amount of diameter and random amount of intensities between the maximum and minimum values, was created inside the images with 0.5 probabilities.

2.2.3. CNN algorithm and training details

The most famous machine learning segmentation algorithm, U-net [24], and one kinds of generative adversarial networks (GAN), segAN [25], CNN models were used for the LV segmentation. Each projections were trained separately for the both models. All final weights of six projections were trained based on the in-house database. However, before training the long-axis projections, transfer learning technique was employed using CAMUS dataset. Transfer learning is the technique that applies the knowledge, which is gained from other related dataset, during training. The effects of transfer learning are the trained model performs better while using less training data and needs less computation time. In this study, first the models were trained on the CAMUS dataset, and then calculated weights were used as the base of the training with in-house dataset.

All images were resized to 128×128 pixels for CNN models. Also, segmented result images had 128×128 pixels. Gray 2D echo images were resized using OpenCV's cubic interpolation function. Ground truth images were resized using OpenCV's nearest interpolation function to maintain the label values, i.e., 0, 1, and 2 values for the in-house dataset.

The Adam optimizer was used to optimize each model's loss with 0.0005 learning rate. Both CNN models were trained for 1000 epochs on the CAMUS dataset and trained for 500 epochs on the in-house dataset. The training was conducted with Python and TensorFlow. The experiments were performed on NVIDIA TESLA K80 GPU and NVIDIA GeForce GTX 1050 GPU. The batch sizes for training were 64 and 16 for the TESLA K80 GPU and GeForce GTX 1050 GPU, respectively.

2.2.4. U-net

U-net deep CNN architecture was proposed by Ronneberger et al. in 2015 [24]. Since the algorithm was developed, it has generated quite successful results in biomedical image segmentation tasks [17]. *Figure 2.2* represents the employed U-net model. U-net consists of two main parts, encoder and decoder. Encoder, which places the left side of the *Figure 2.2*, consists of the Convolution layers and Maxpool layers and performs the contracting images. The features of input images were extracting during the contracting. The encoder has of 4 Maxpool layers and each Maxpool layer downsamples the images with a 2×2 pooling kernel, 2 strides, and same padding method. Similarly, decoder places the right side of the *Figure 2.2* and has the Convolution layers and transposed convolution layers. The images were recovered original sizes during the upsampling. The decoder consists of 4 transposed convolution layers and each layer upsamples with a 4×4 kernel, 2 strides, and same padding method. During the contracting, some information was lost. To recover these fine features, after each transposed convolution layer upsamples the images, same sized images from encoder were concatenated. All

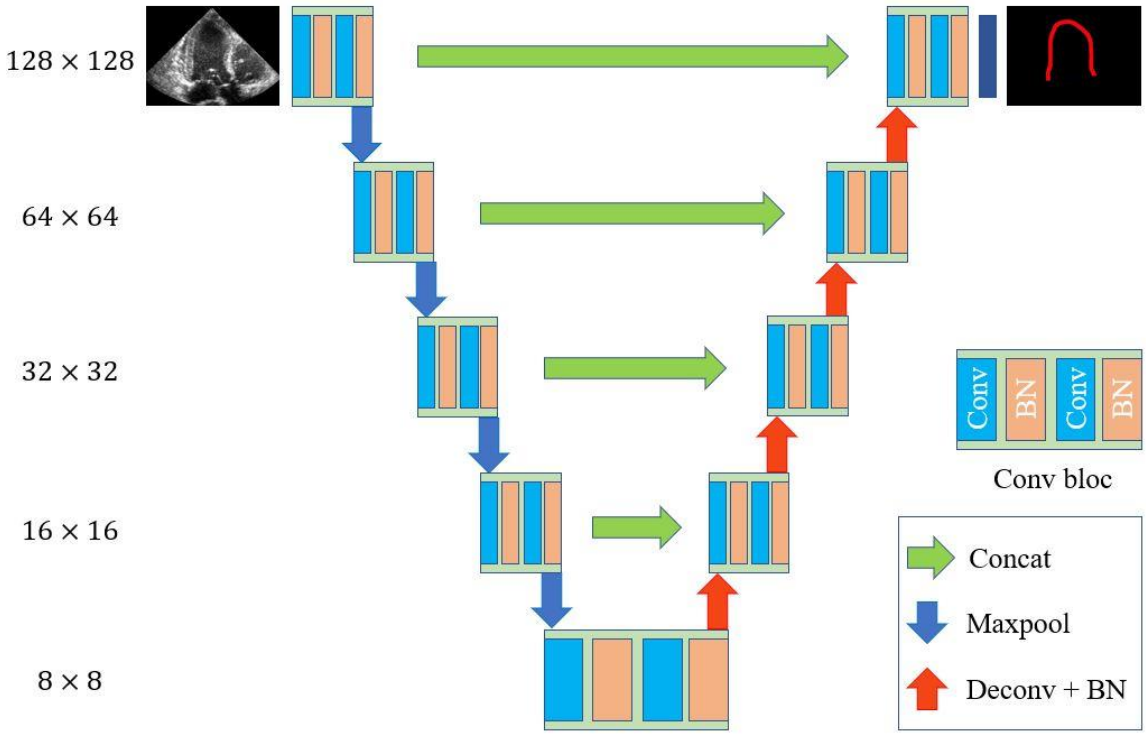

Figure 2.2 U-net architecture. Each layer consists of two convolution layers followed by batch normalization layers. Blue and red arrows indicate max pooling and deconvolution layer followed by batch normalization. Green arrows are concatenation operations.

convolution layers used 3×3 kernel, rectified linear unit (ReLU) activation function, and L2 regularization. Batch normalization layers were used for normalization scheme. After decoder recovered the images, final convolution layers, which used 1×1 kernel, generated feature maps to match segmentation labels (4 for the CAMUS dataset and 3 for the in-house dataset). Multi-dimensional dice loss was used to calculate the loss during the training.

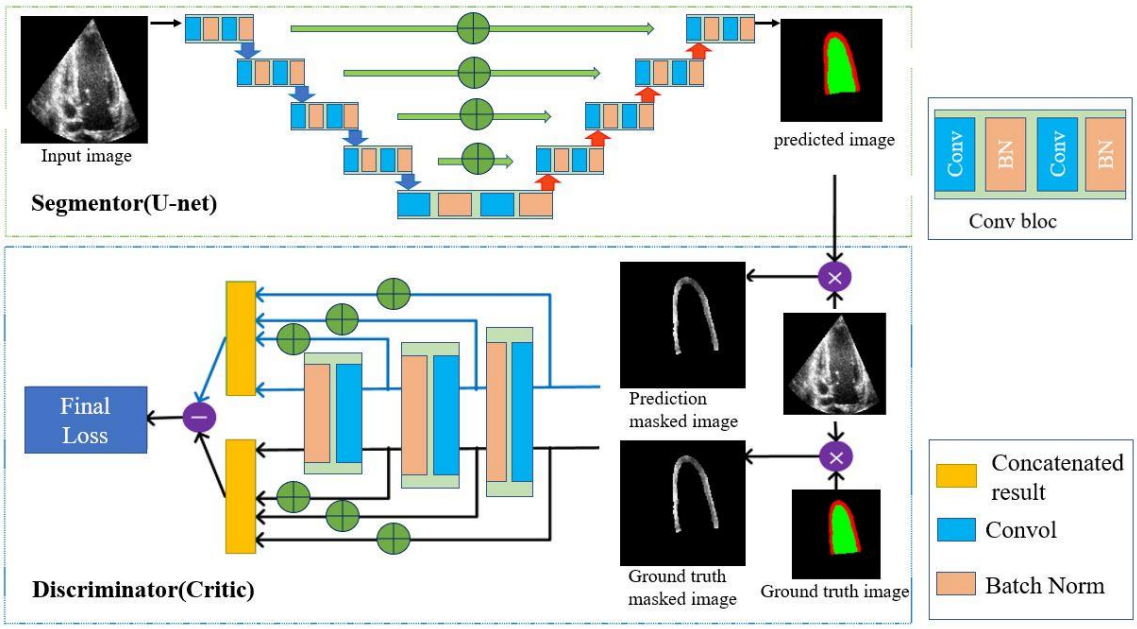

Figure 2.3 The segAN architecture. In segmentor, U-net architecture is used for segmentation. In critic, two masked images are used to calculate the multi-scale loss. Masked image shows the segmented part of the input image (here, LV myocardium).

2.2.5. SegAN

SegAN was used as second algorithm for LV segmentation. Xue et al. proposed the medical image segmentation generative adversarial network (GAN) model, segAN, in 2018 [25]. They showed that segAN model outperformed the previous segmentation models, including U-net, on MRI brain tumor segmentation task using BRATS dataset. *Figure 2.3* shows the segAN structure. The common GAN consist of two neural networks. One neural network, called a generator, creates new data instances, and the other, discriminator, evaluates the authenticity of the generative data. The entire GAN model's performance is improved by the min-max game between generator and discriminator.

Similarly, in the segAN, the segmentor (which performs as a generator, *Figure2.3* upper box) generated the segmented images from the original images with encoder-decoder algorithm. In this study, same U-net structure which described above, was employed for the segmentor architecture. The critic (which performs as a discriminator, *Figure2.3* lower box) compared the two inputs, segmented results from the segmentor and ground truth images, and calculated the loss between them. The final loss was calculated by the multi-scale feature loss function based on the mean absolute error (MAE) and dice loss. During the training, the segmentor tried to minimize the final loss, which means created segmented results as similar as ground truth images, whereas the critic aimed to maximize the final loss. After training, segmentor was used to LV segmentation task.

2.2.6. Post-process

Some segmented images from both CNN models contained independent island-like structures at the background and an incomplete LV myocardium structure. To reduce these problems, post-process was conducted on the segmented images using Morphological Transformation technique, which is based on the erosion and dilation function [49]. Based on the vector subtraction and addition, erosion and dilation combine two sections. To remove the island-like structures, erosion followed by dilation was used. Inversely, to complete the LV boundary line, dilation followed by dilation was used. The diameters of erosion and dilation were manually added during the post-process.

2.2.7. Evaluation of segmentation results

To compare the segmentation performance of CNN models and previous non-machine learning methods, level-set, which is widely used semi-automatic segmentation algorithm [50], was tested on same testing set. The Creaseg, open software of level-set method, was employed to segment the LV [51]. For the testing set which consists of 2 pigs, 5 images were selected with same time interval per one pig. Total 10 images were tested with level-set method for each projections.

To assess and compare the segmentation results quantitatively, four widely used metrics in the segmentation researches were employed: the dice metric, precision, sensitivity, and Hausdorff distance. Let P and G represent the predictive segmentation result and ground truth image, respectively. The dice metric [52] calculates the overlap between P and G. The dice metric is defined as:

$$Dice = \frac{2|P \cap G|}{|P| + |G|} \tag{2.1}$$

The dice value varies from 0, which means no overlap, to 1, which means perfect overlap. The good segmentation results have high dice metric. The precision and sensitivity (also known as recall) are other metrics used to evaluate the relevance between P and G [53]. Precision and sensitivity are defined as:

$$Precision = \frac{TP}{TP + FP} \tag{2.2}$$

$$Sensitivity = \frac{TP}{TP + FN} \tag{2.3}$$

Where TP, FP, and FN are the true positives, false positives, and false negatives. In image segmentation task, TP, FP, and FN are the number of pixels, which are correctly classified as denoted labels, incorrectly classified as denoted labels, and incorrectly classified as not labels, respectively. Also, precision and sensitivity vary from 0 to 1. Good segmentation results represent high values from both. The Hausdorff distance, which is widely used to compare two different geometries, represents the maximum distance between the contours of P(δP) and G(δG) [54]. Hausdorff distance is defined as:

$$Hausd.dist. = max\left(\underset{i\in\delta P}{max}\, d(i,\delta G), \underset{i\in\delta G}{max}\, d(i,\delta P)\right) \tag{2.4}$$

Where d(i,δ) is the shortest distance from specific point i to contour δ. Good segmentation result has low Hausdorff distance value, which means the difference between predict segmentation and ground truth images is small. The Hausdorff distance is calculated on the [mm] dimension.

2.3. 3D reconstruction

2.3.1. Input data

Six 2D echo standard projections' one cardiac cycle videos are the input data of the 3D reconstruction algorithm. To get these input video, 30 to 40 images over one cardiac cycle are concatenated into one video. The segmentation images from CNN

models have fixed 128×128 pixels size. Therefore, the segmented images were resized to their original size, e.g., 648×480 pixels size, for recovering the original LV height to width ratio and pixel to mm ratio.

2.3.2. Previous 3D reconstruction algorithm

3D reconstruction algorithm, which was developed by our group [38], uses six standard projections, i.e., 3-, 2-, and 4-chamber and apex, mid, and base projections. This algorithm consists of seven steps. First, endocardial boundary detection part delineates LV boundaries by the intensity differences between RGB channels. Second, data smoothing part smooths the coarse LV boundaries from previous step with B-spline curve fitting and variable span weighted moving average methods. Third, temporal interpolation section matches six different frame rates using a cubic spline with natural boundary conditions. Fourth, sectional scaling and orientation part arranges the six projections using nominal positions after matching all projections' scaling. Also, positions and angles of long- and short-axis projections are optimized by minimizing the error between difference sections. Fifth, spatial interpolation section interpolates the LV surface points based on frame of six projections. Sixth, temporal smoothing conducts Fourier curve fitting to smoothen the volume flux with considering the periodic nature of the dependent variable. Finally, mesh generation part converts generated surface points into cartesian co-ordinates and triangulated to generated mech. The final reconstructed LV geometries are stored in a standard visualization tool kit (VTK) file format.

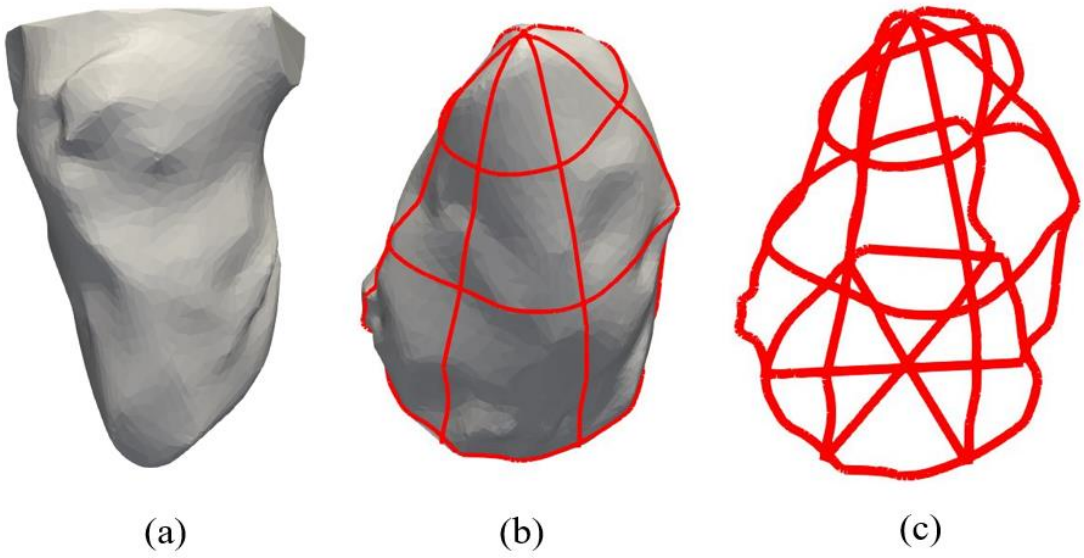

Figure 2.4 Validation reference. (a) represents the LV geometry obtained from the MRI images. (b) shows the extracting process. (c) represents the six extracted projection frame.

2.3.3. Validation process

In this study, two validation processes were conducted for the 3D reconstruction algorithm. First, the reconstruction ability of algorithm was validated against a known geometry from MRI. Second, the number of projections, which were used for the 3D reconstruction, was verified.

3D reconstruction algorithm's capability was validated using reference LV geometry, which was obtained from the MRI images. MRI images of whole LV consists of lots of short-axis projections from apex to base. *Figure 2.4 (a)* shows the 3D geometry from MRI which is considered as the reference geometry. Six projections were extracted from the reference geometry as shown in *Figure 2.4 (b)* and *(c)*. Extracted projections

Table 2.2 Projection combinations. Ch denotes the chamber in the long-axis.

Combinations	Long-axis projections	Short-axis projections
Reference	6: 0° (3ch), 45°, 60° (4ch), 90°, 120° (2ch), and 135°	3: Base, mid, and apex
Standard	3: 0° (3ch), 60° (4ch), and 120° (2ch)	3: Base, mid, and apex
Long and base	3: 0° (3ch), 60° (4ch), and 120° (2ch)	1: Base
Long-axis only	6: 0° (3ch), 45°, 60° (4ch), 90°, 120° (2ch), and 135°	0: -

were converted into 3D shape using reconstruction algorithm. 3D reconstruction algorithm was evaluated by comparing reference and reconstructed LV geometries.

The sensitivity to the number of projections used for the reconstruction was verified by comparing the combination of projections which were obtained from same LV. Table2.2 represents the different combinations which used from 4 to 9 projections. Four different combinations of projections were reconstructed and compared. The reference LV geometry was generated from six long-axis and three short-axis.

2.3.4. 3D reconstruction algorithm improvement

In this study, previous 3D reconstruction algorithm was improved. Previous algorithm could not represent the long-axis 2- and 4-chamber projections' shape well on

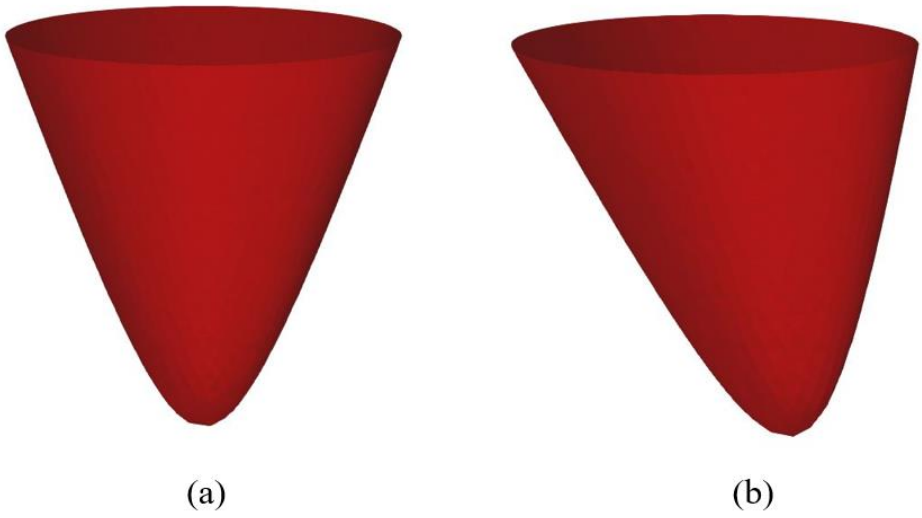

Figure 2.5 Different apex location geometries. (a) Centered: apex is located near the center of the short-axis projections. (b) Tilt: apex is not located near the center.

the reconstructed geometries. Especially, when the apex point was not located near the center of the short-axis projections (*Figure 2.5 (b)*), this problem worsened.

The method of assigning the apex point on the short-axis projection images caused this problem. Previous algorithm assumed the average of the x and y components of all short-axis boundary points denotes the apex point. This assumption was only reasonable when the apex point was located near the center of short-axis (*Figure 2.5 (a)*). However, most of the LV have tilted geometries (*Figure 2.5 (b)*). Long-axis projections of *Figure 2.1* shows that the apex point was not located near the center of the short-axis projections.

The apex point can be estimated from the three long-axis projections. The apex point is the highest point of the each long-axis LV boundary. Six boundary points and apex point were obtained from the three long-axis and compared with the short-axis

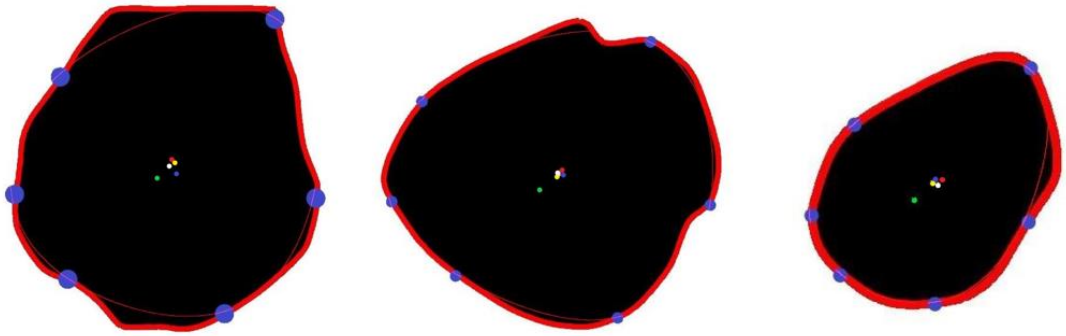

Figure 2.6 Estimation of the apex point. Comparison between short-axis projection and points which were obtained from the three long-axis. Red line denotes the short-axis projection boundaries. Blue points on the boundaries represents the three long-axis. Green point represents the average of six blue points. Blue point inside the boundary shows the average of boundary points which were obtained from the interpolated curve using six points. Yellow point represents the geometric center of the interpolated curve. Red point shows the average of the short-axis boundary. Finally, white point denotes the geometric center of the short-axis boundary.

projections. To estimate the apex point on the short-axis projection, several points were calculated and compared. *Figure 2.6* shows the comparison between points which obtained from short-axis and long-axis projections. Green, blue, and yellow points were calculated from the three long-axis projections. Red and white points were obtained from the short-axis projection. Average of boundary points was calculated as:

$$\mathrm{X_{avg}} = \frac{\sum_{i=1}^{n} X_i}{n} \; and \; \mathrm{Y_{avg}} = \frac{\sum_{i=1}^{n} Y_i}{n} \tag{2.5}$$

Where n denotes the total number of points which consist the boundary and X_i and $\mathrm{Y_i}$ represent each point's X and Y values. Geometric center was calculated as:

$$\mathrm{X_{geo}} = \frac{\sum_{i=1}^{n} X_{geo,i} \times A_i}{\sum_{i=1}^{n} A_i} \; and \; \mathrm{Y}_{geo} = \frac{\sum_{i=1}^{n} Y_{geo,i} \times A_i}{\sum_{i=1}^{n} A_i} \tag{2.6}$$

Where $\mathrm{A_i}$ denotes the area of divided sections and $\mathrm{X_{geo,i}}$ and $\mathrm{Y_{geo,i}}$ represent each section's X and Y values of the geometric center. *Figure 2.6* shows that the white and yellow points, which were calculated as geometric centers, were the closest points. From this result, the difference between geometric center of the interpolated curve using six points and apex point was transferred to the short-axis. Apex point was estimated by adding this difference to the geometric center of short-axis boundary.

2.3.5. Evaluation of reconstruction geometries

LV physiological parameters were calculated to compare the 3D geometries reconstructed from ground truth and segmented images. In this study, reconstructed LV volume with ground truth images is the benchmark one. 3D reconstruction algorithm calculates the LV volume on every time instance over whole cardiac cycle. Physiological parameters, ejection fraction (EF), stroke volume (SV), and cardiac output (CO), were calculated using these volume information. Three parameters have been widely used in cardiac research and performed as the index of the LV assessment [55].

Ejection fraction (EF) is the fraction of blood leaving heart each time when LV contracts. It is defined as:

$$EF = \frac{E_{DV} - E_{SV}}{E_{DV}} \tag{2.7}$$

Where E_{DV} and E_{SV} are end-diastolic and end-systolic volumes [ml] over the given cardiac cycle. Stroke volume (SV) is the difference between the maximum and minimum volumes during the cardiac cycle. It has [ml] dimension. Finally, cardiac output (CO) is the volume of the blood, which is pumped by the LV, per unit time and has [ml/sec] dimension. It is defined as:

$$CO = HR \times SV \tag{2.8}$$

Where HR is heart rate and SV is stroke volume.

3. RESULTS

3.1. Segmentation result

3.1.1. CAMUS dataset segmentation

Before conducting the segmentation task on the in-house dataset, U-net and segAN were trained and tested based on the CAMUS dataset to verify the methods. CAMUS dataset only consists of 2- and 4-chamber projection. Both models were trained and tested without separation of projections. Also, the post-process, which used morphological transformation, was evaluated quantitatively. Among 450 patients' dataset, 400 patients' images were used to train two CNN models and remained 50 patient's images were used to test the trained models. For the comparison, p-value was calculated to verify the statistically different.

Table 3.1 reports the dice metric, precision, sensitivity, and Hausdorff distance of two CNN models. To evaluate the effect of the post-process, each model was calculated twice, before and after applying post-process. The Dice metric of the LV cavity is 0.920 and of the LV myocardium 0.860 for the U-net model. For the segAN, the dice metric of the LV cavity is 0.917 and of the LV myocardium is 0.859. Both models show higher performance on LV cavity than the LV myocardium (p-value < 0.05). The precision and sensitivity are calculated similar to the Dice metric. The Hausdorff distance is 4.92 mm for LV cavity and 6.23 mm for LV myocardium on the U-net model. For the segAN, 5.14 mm for LV cavity and 6.18 mm for LV myocardium were calculated.

Table 3.1 The comparison of metrics. The comparison of metrics: Dice metric, precision, sensitivity, and Hausdorff distance (HD) between the before and after the post-process for both U-net and segAN on CAMUS dataset (4-chamber and 2-chamber projections). P-P denotes the post-process

Method	Section	P-P	Dice	Precision	Sensitivity	HD (mm)
U-net	LV	None	0.915 ± 0.066	0.920 ± 0.105	0.919 ± 0.059	5.17 ± 1.39
	cavity	Done	0.920 ± 0.058	0.929 ± 0.093	0.921 ± 0.060	4.92 ± 1.27
	LV	None	0.858 ± 0.066	0.852 ± 0.089	0.867 ± 0.077	6.43 ± 1.40
	myocardium	Done	0.860 ± 0.063	0.861 ± 0.063	0.871 ± 0.077	6.23 ± 1.25
segAN	LV	None	0.912 ± 0.079	0.907 ± 0.123	0.930 ± 0.053	5.25 ± 1.90
	cavity	Done	0.917 ± 0.071	0.916 ± 0.112	0.932 ± 0.053	5.14 ± 1.71
	LV	None	0.855 ± 0.066	0.852 ± 0.092	0.862 ± 0.078	6.28 ± 1.19
	myocardium	Done	0.859 ± 0.064	0.862 ± 0.086	0.866 ± 0.077	6.18 ± 1.17

Table 3.1 shows the effect of the post-process. Dice metric increases from 0.915 to 0.920 for the LV cavity and increases from 0.858 to 0.860 for the LV myocardium on the U-net model (p-value < 0.05). Also, dice metric of segAN increases from 0.912 to 0.917 for the LV cavity and from 0.855 to 0.859 for the LV myocardium (p-value < 0.05). The Hausdorff distance, for which small values indicate better results, decreases from 5.17 mm to 4.92 mm for the LV cavity and from 6.43 mm to 6.23 mm for the LV myocardium on the U-net model (p-value < 0.05). Also, Hausdorff distance of segAN decreases from 5.25 mm to 5.14 mm for the LV cavity and from 6.28 mm to 6.18 mm for the LV myocardium (p-value < 0.05). The effect of the post-process was remarkable on the Hausdorff distance. Three other metrics were improved less than 1%. However, Hausdorff distance decreases from 1.4% to 4.7% for each class.

3.1.2. In-house dataset segmentation

Six projections were separately trained and tested based on the in-house dataset using U-net and segAN. Also, segmentation results using level-set method were reported to comparison. Figure 3.1 represents the segmentation results. Original 2D echo images (Gray), predicted segmentation images using the U-net and segAN models (U-net and segAN), ground truth images (Ground Truth), and level-set segmentation images (Level-set) on six standard projections consist each row. Green and red regions of the U-net, segAN, and Ground Truth rows' images indicate LV cavity and LV myocardium respectively. Red line of the Level-set row's images denotes the LV cavity. Level-set method can only segment the LV cavity.

Figure 3.1 shows that the machine learning models achieve high performance on LV segmentation task. Segmented images agreed with ground truth images, which were converted from expert's delineation for the both long-axis (*Figure 3.1* (a), (b), and (c)) and short-axis (*Figure 3.1* (c), (d), and (e)) projections. Segmentation difference between U-net and segAN is small and hard to evaluate from the images.

Level-set method generated reasonable segmentation results for the long-axis projections. The wall was delineated correctly by the level-set method. However, the existence of valves, which has similar brightness of the LV wall and has intensity difference with LV cavity, made a problematic segmentation near the base. If the valve

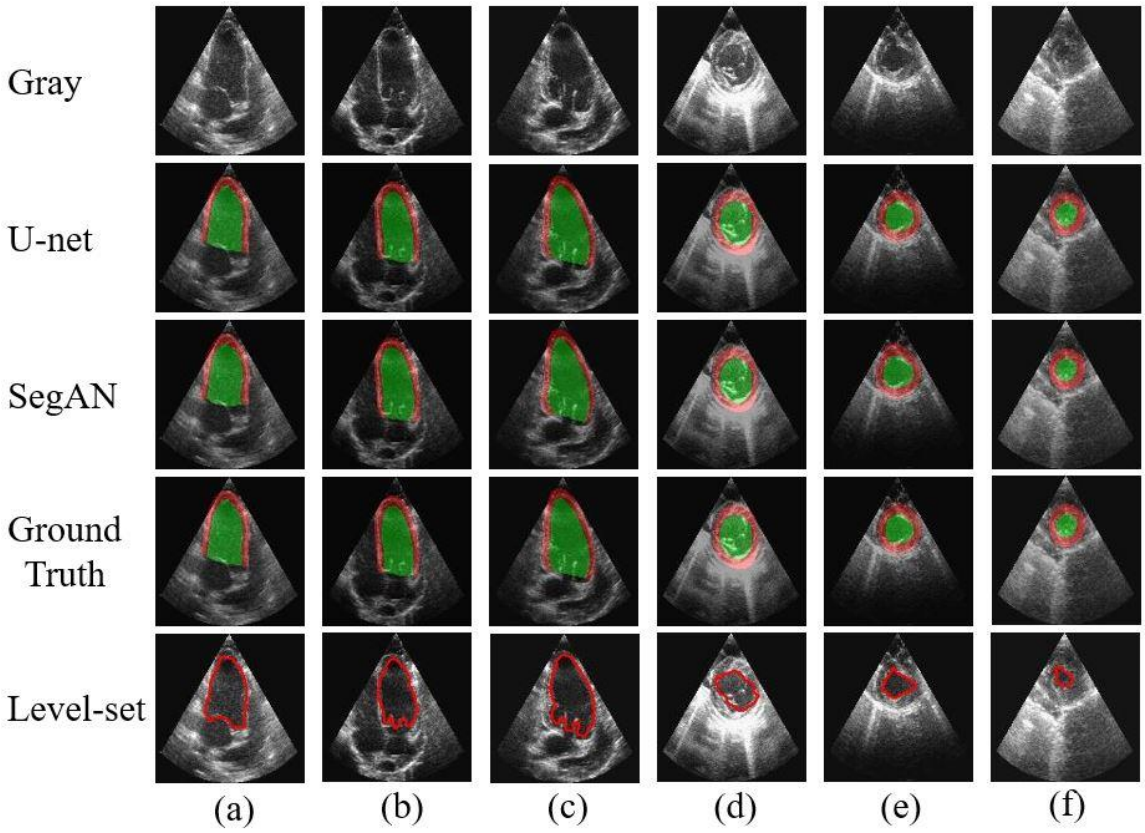

Figure 3.1 Illustration of the segmentation results for six standard projections. The first row shows original 2D echo images (Gray), second and third rows show predicted segmented result from U-net and segAN models (U-net and segAN), fourth row shows the ground truth images (Ground Truth), and the last row represents the segmentation results using level-set method (Level-set). Red region represents LV myocardium and green region represents LV cavity. Red line in the Level-set images denotes the LV cavity. Images are the one time instant of each projection's cardiac cycle. Each column shows long-axis (a) 3-chamber, (b) 4-chamber and (c) 2-chamber and short-axis (d) base, (e) mid and (f) apex projections.

was closed, level-set method could generate a closed curve. However, in case of an open valve, the iteration of the level-set method had to controlled to divide the LV cavity and left atrium cavity. For the short-axis projections, presence of the valve orifice, trabeculae, and papillary muscles disturbed the LV cavity segmentation. Table 3.2, 3.3, and 3.4 report

Table 3.2 U-net model. U-net model: The comparison of metrics: Dice metric, precision, sensitivity and Hausdorff distance (HD)

Methods	Projections	Section	Dice	Precision	Sensitivity	HD (mm)
U-net	3-chamber	Cavity	0.920 ±0.017	0.902 ±0.044	0.940 ±0.023	2.96 ±0.24
		Myocardium	0.814 ±0.017	0.869 ±0.040	0.769 ±0.046	3.61 ±0.25
	2-chamber	Cavity	0.927 ±0.023	0.930 ±0.065	0.929 ±0.027	2.72 ±0.38
		Myocardium	0.715 ±0.031	0.818 ±0.040	0.640 ±0.056	4.23 ±0.29
	4-chamber	Cavity	0.900 ±0.011	0.889 ±0.028	0.913 ±0.020	3.18 ±0.41
		Myocardium	0.772 ±0.019	0.811 ±0.034	0.738 ±0.020	3.89 ±0.39
	Base	Cavity	0.931 ±0.037	0.917 ±0.083	0.954 ±0.054	2.20 ±0.26
		Myocardium	0.860 ±0.041	0.896 ±0.039	0.827 ±0.049	3.14 ±0.33
	Mid	Cavity	0.888 ±0.022	0.918 ±0.063	0.863 ±0.036	2.92 ±0.24
		Myocardium	0.761 ±0.079	0.751 ±0.151	0.791 ±0.035	4.19 ±0.57
	Apex	Cavity	0.854 ±0.040	0.928 ±0.109	0.811 ±0.093	2.26 ±0.38
		Myocardium	0.804 ±0.059	0.829 ±0.094	0.785 ±0.040	3.05 ±0.29
	Mean	Cavity	0.903	0.914	0.901	2.71
		Myocardium	0.787	0.829	0.758	3.68

the dice metric, precision, sensitivity, and Hausdorff distance (HD) of U-net, segAN, and level-set methods, respectively. All metrics were calculated on the testing dataset. Six projections' results were calculated separately. The mean value of each metric was defined as:

$$M_{Mean} = \frac{M_{3ch} + M_{2ch} + M_{4ch} + M_{Base} + M_{Mid} + M_{Apex}}{6} \tag{3.1}$$

where M is each metric (3-chamber: 3ch, 2- and 4-chamber: 2 and 4ch).

Table 3.3 SegAN model. SegAN model: The comparison of metrics: Dice metric, precision, sensitivity and Hausdorff distance (HD)

Methods	Projections	Section	Dice	Precision	Sensitivity	HD (mm)
segAN	3-chamber	Cavity	0.915 ±0.014	0.903 ±0.033	0.929 ±0.025	3.57 ±0.71
		Myocardium	0.787 ±0.034	0.860 ±0.033	0.728 ±0.056	3.76 ±0.36
	2-chamber	Cavity	0.923 ±0.021	0.912 ±0.051	0.938 ±0.022	3.09 ±0.33
		Myocardium	0.697 ±0.032	0.816 ±0.036	0.611 ±0.048	4.44 ±0.28
	4-chamber	Cavity	0.914 ±0.016	0.894 ±0.040	0.936 ±0.015	3.29 ±0.49
		Myocardium	0.805 ±0.022	0.845 ±0.029	0.769 ±0.029	3.64 ±0.30
	Base	Cavity	0.917 ±0.039	0.891 ±0.088	0.954 ±0.053	2.27 ±0.31
		Myocardium	0.847 ±0.042	0.889 ±0.043	0.810 ±0.046	3.21 ±0.43
	Mid	Cavity	0.903 ±0.038	0.874 ±0.079	0.941 ±0.045	2.46 ±0.27
		Myocardium	0.799 ±0.102	0.758 ±0.162	0.865 ±0.043	3.64 ±0.77
	Apex	Cavity	0.901 ±0.072	0.894 ±0.142	0.930 ±0.060	2.27 ±0.40
		Myocardium	0.864 ±0.101	0.864 ±0.138	0.852 ±0.060	3.03 ±0.54
	Mean	Cavity	0.912	0.895	0.938	2.82
		Myocardium	0.801	0.842	0.772	3.62

From Table 3.2, mean dice metric of U-net model is 0.903 for the LV cavity and 0.787 for the LV myocardium. Base projection has largest dice metric, 0.931 for LV cavity and 0.860 for LV myocardium and smallest Hausdorff distance, 2.20 mm for LV cavity and 3.14 mm for LV myocardium. Apex projection has smallest dice metric, 0.854 for LV cavity and 0.804 for LV myocardium and mid projection has largest Hausdorff distance, 2.92 mm for LV cavity and 4.21 mm for LV myocardium . Long-axis projections' average dice metric is 0.916 for LV cavity and 0.767 for LV myocardium. Short-axis projections' average dice metric is 0.891 for LV cavity and 0.808 for LV myocardium. Long-axis projections' average Hausdorff distance is 2.96 mm for LV cavity and 3.91 mm for LV

Table 3.4 Level-set method. Level-set method: The comparison of metrics: Dice metric, precision, sensitivity and Hausdorff distance (HD)

Methods	Projections	Section	Dice	Precision	Sensitivity	HD (mm)
Level-set	3-chamber	Cavity	0.857 ±0.048	0.775 ±0.090	0.968 ±0.032	3.61 ±0.57
	2-chamber	Cavity	0.896 ±0.022	0.907 ±0.049	0.890 ±0.056	4.02 ±0.54
	4-chamber	Cavity	0.885 ±0.012	0.549 ±0.057	0.931 ±0.053	3.88 ±0.41
	Base	Cavity	0.815 ±0.050	0.788 ±0.103	0.856 ±0.040	4.37 ±0.36
	Mid	Cavity	0.784 ±0.067	0.892 ±0.120	0.704 ±0.080	3.26 ±0.46
	Apex	Cavity	0.695 ±0.081	0.866 ±0.057	0.555 ±0.114	2.68 ±0.29
	Mean	Cavity	0.821	0.846	0.817	3.64

myocardium. Short-axis projections' average Hausdorff distance is 2.46 mm for LV cavity and 3.45 for LV myocardium.

From Table 3.3, mean dice metric of segAN is 0.912 for the LV cavity and 0.801 for the LV myocardium. 2-chamber projection has largest dice metric value, 0.923 for LV cavity. Apex projection has smallest Hausdorff distance, 2.27 mm for LV cavity and 3.03 mm for LV myocardium. Long-axis projections' average dice metric is 0.917 for LV cavity and 0.763 for LV myocardium. Short-axis projections' average dice metric is 0.907 for LV cavity and 0.837 for LV myocardium. Long-axis projections' average Hausdorff distance is 3.31 mm for LV cavity and 3.94 mm for LV myocardium. Short-axis projections' average Hausdorff distance is 2.33 mm for LV cavity and 3.26 for LV myocardium.

Finally, Table 3.4 reports the results of level-set method. The mean dice metric is 0.821 for the LV cavity and mean Hausdorff distance is 3.64 mm. The highest dice metric is 0.896 for the 2-chamber projection and the lowest Hausdorff distance is 2.68mm for the

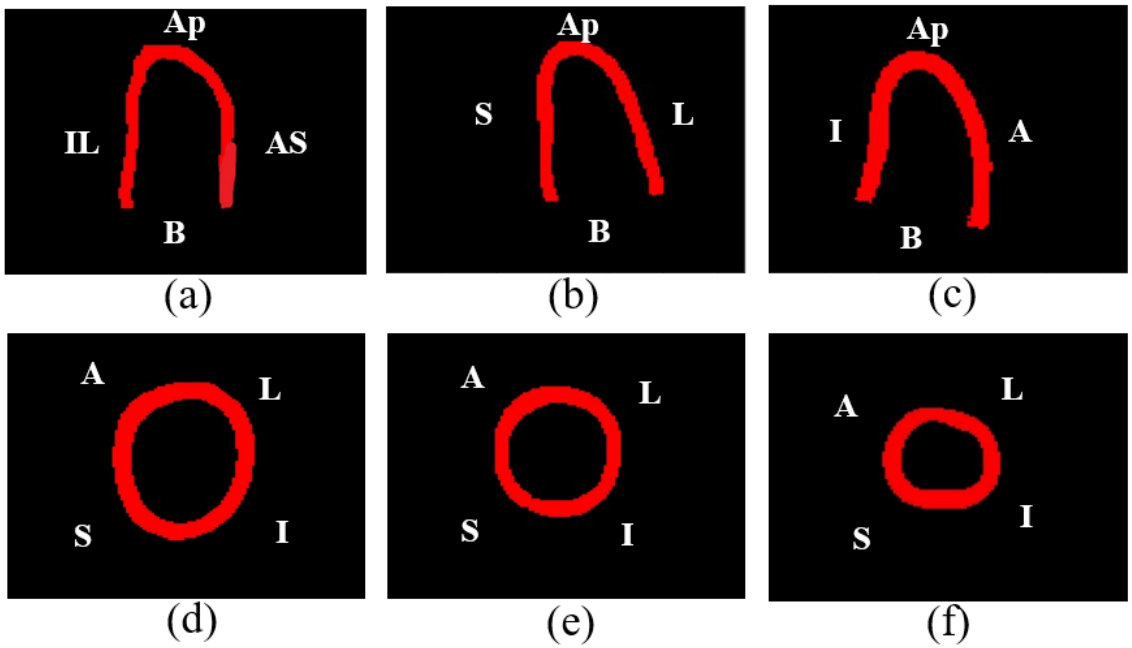

Figure 3.2 One time instant of LV segmented input videos. One time instant of LV segmented input videos using U-net for the 3D reconstruction algorithms. Long-axis (a) 3-chamber, (b) 4-chamber, and (c) 2-chamber and short-axis (d) base, (e) mid, and (f) apex projections. A: Anterior, AS: Anteroseptal, Ap: Apical, B: Basal, I: Interior, IL: Inferolateral, L: Lateral, and S: Septal.

apex. Long-axis projections' average dice metric is 0.879 for LV cavity. Short-axis projections' average dice metric is 0.764 for LV cavity.

From all metrics, U-net and segAN generated better results the level-set method (p-value < 0.05). Especially, level-set method made lower performance on the short-axis projections. Both U-net and segAN models generated better results on LV cavity than LV myocardium (p-value < 0.05). *Figure 3.2* shows the final segmentation results. After the resizing process, the LV in *Figure 3.2* has the original height-width ratio and pixel-mm ratio. The videos generated by these images became the input of the 3D reconstruction

algorithm. Even the LV cavity made better segmentation results than LV myocardium, the reconstruction algorithm only uses the inner boundary of LV. In this reason, this study used the LV myocardium images as the input of the reconstruction algorithm.

3.2. 3D reconstruction

3.2.1. Algorithm improvement result and validation

The previous reconstruction algorithm failed to represent the 2- and 4- chamber projections because of a not reasonable assumption on apex location for the short-axis projection method (section 2.3.4). This problem was improved using information from the long-axis projections and matching geometric centers method as explained in section 2.3.4. *Figure 3.3* shows the three long-axis (upper row) and short-axis (middle row) projections, which were extracted from the benchmark geometry and used to reconstruction, and the reconstructed geometries using different algorithms (lower). Yellow shape, which was made by original algorithm, had a simple symmetrical cone shape and failed to represent the 4-chamber projection's biased apex and curvature of the right side boundary. However, the red geometry, which was reconstructed by the improved algorithm, had an eccentric apex position and succeed to represent the bump configuration near the apex, which was came from 4-chamber projection. Yellow model's apex was located near the center of the short-axis projections. However, red geometry's apex was not fixed near the center.

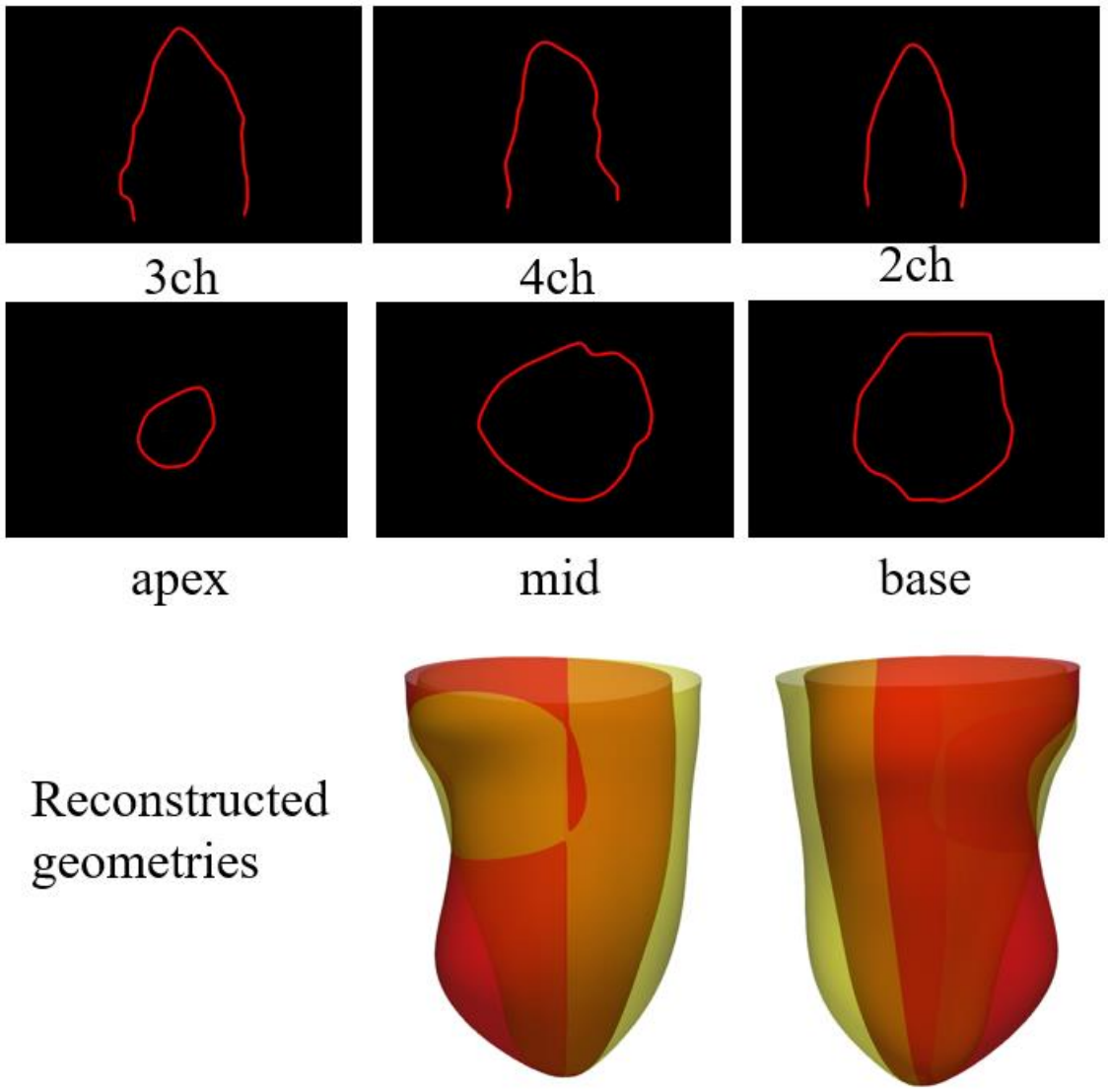

Figure 3.3 Comparison of reconstructed geometries. Long-axis projections (upper row), shot-axis projections (middle row) and comparison of reconstructed geometries (bottom row). Among the reconstruction geometries, red shape is reconstructed by improved algorithm and yellow shape is made by original algorithm. Ch denotes the -chamber projection.

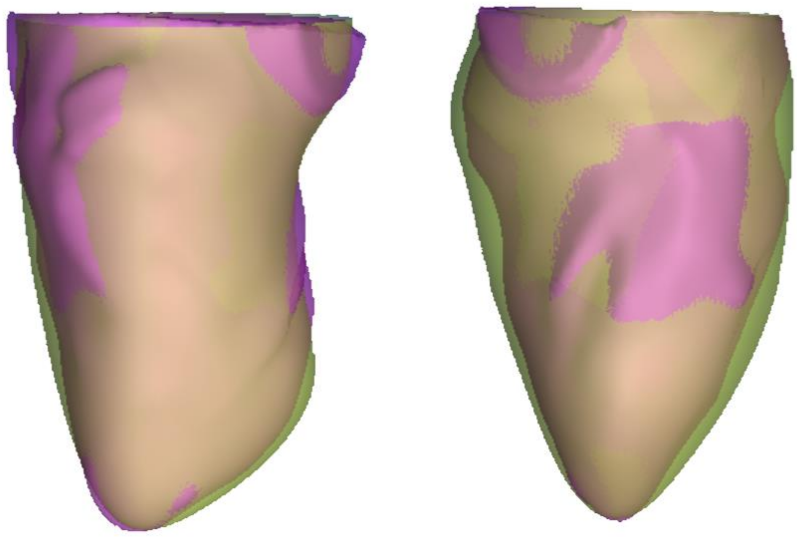

Figure 3.4 Comparison of LV geometries. Purple shape is the original LV geometry, which was obtained from MRI. Yellow shape is reconstructed geometry using improved algorithm.

The improved algorithm was validated by comparing the reconstructed geometry with original shape. *Figure 3.4* shows the comparison between original and reconstructed LV geometries. Purple shape was obtained from the MRI and perform as the reference LV geometry. Yellow shape was reconstructed geometry by improved algorithm using three long- and three short-axis projections, which were extracted from reference geometry (*Figure 3.3* upper and middle row). The reconstructed model succeeded to detect tilted apex point and overall curvature of surface. The smoothing section of the reconstruction algorithm generated the smooth surface, which could not represent the detailed coarse surface of original surface. Also, the surface which was not located near the three long-axis projections made difference with reference geometry.

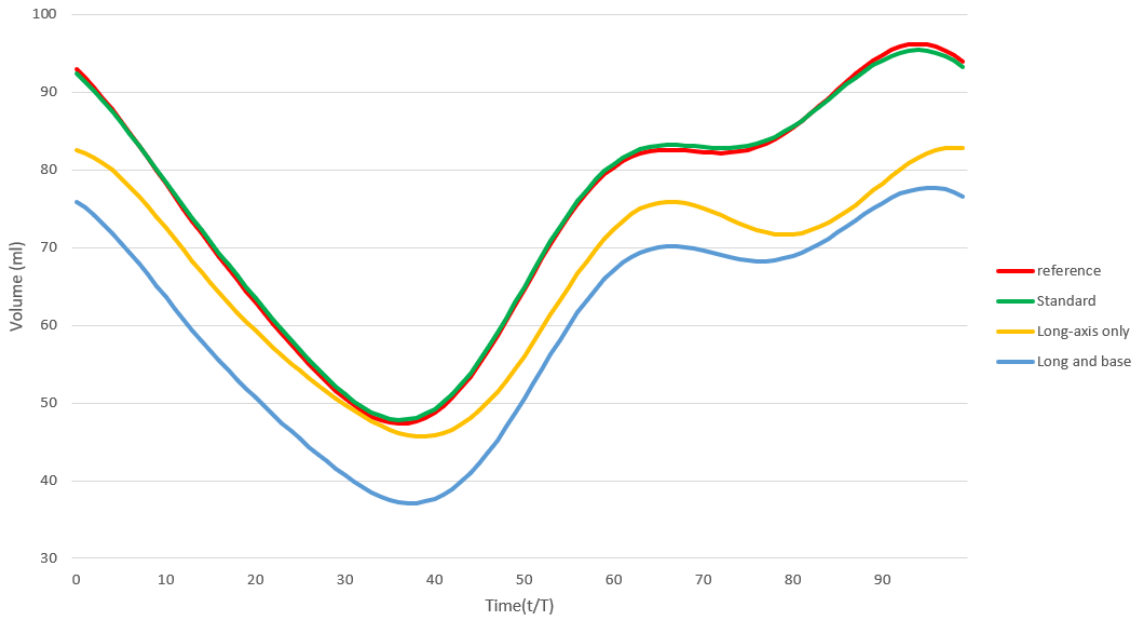

Figure 3.5 Volume-time graphs of one cardiac cycle. Red, green, yellow, and blue lines denote the reference, standard, long-axis only, and long and base projections combinations (Table 2.2), respectively.

3.2.2. Projection validation

In this study, six standard projections were used to reconstruct the LV. To verify the usage of standard projections, four different combinations of projections (Table 2.2) were tested and compared. *Figure 3.5* shows the four volume-time graphs using different projection combinations. The reference combination, which was reconstructed with nine projections, conducted the benchmark for the comparison. Standard projections, three long-axis and three short axis, showed similar volume change trend during the one cardiac cycle with reference case. Also, the absolute volume had close value with benchmark. The largest volume difference was 2.3% near the end-systolic time instance. Long-axis only

Table 3.5 Physiological parameter comparison. GT denotes the geometry which was reconstructed from ground truth images. GT geometries performed as reference. EF: ejection pressure, SV: stroke volume, and CO: cardiac output.

Porcine	parameter	GT	U-net	Error (%)	segAN	Error (%)
Case1	EF	0.37	0.33	10.0	0.36	2.3
	SV [ml]	29.5	29.8	1.2	32.4	9.9
	CO [L/min]	2.18	2.21	1.3	2.39	9.8
Case2	EF	0.40	0.33	18.1	0.34	14.4
	SV [ml]	26.3	23.2	11.7	23.6	10.4
	CO [L/min]	1.94	1.72	11.6	1.74	10.5

projection combination, includes six long-axis, generated reconstructed geometries which had similar absolute volume value with the benchmark. However, it failed to detect the volume change trend. In contrast, long and base projection combination, three long-axis and one short-axis, succeeded to detect the volume change trend besides the lowest volume time instance. However, the entire volumes were 15% smaller than the benchmark volumes. This error is similar to the previous methods using 2D echo compared to the gold standard MRI [56-58].

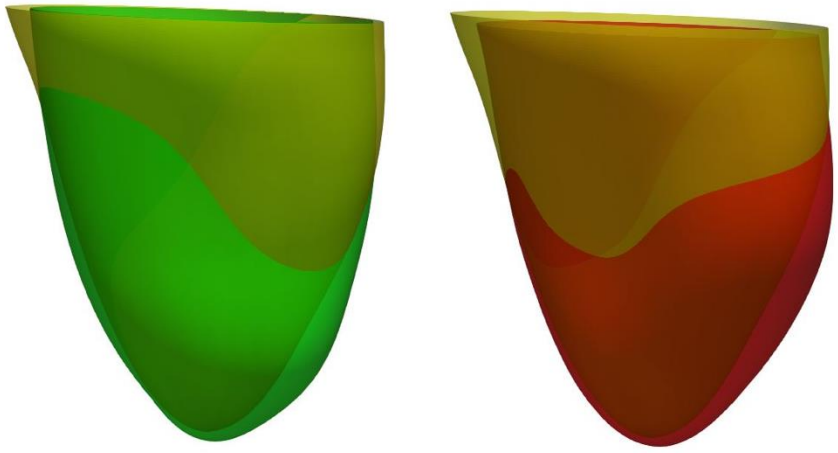

Figure 3.6 LV 3D reconstructed geometry comparison. Yellow LV shape is reconstructed with ground truth images. (a) Green LV shape is reconstructed with segAN segmented images. (b) Red LV shape is reconstructed with segAN segmented images.

3.2.3. Reconstruction using segmented results

Six projections' images segmented by machine learning technique (*Figure 3.2*) were reconstructed by the improved algorithm. In this study, the benchmark to evaluate the reconstructed geometries' physiological parameters was the model reconstructed from ground truth images, which were delineated by experts. Two porcine LV geometries were reconstructed using test dataset images. Table 3.5 reports the physiological parameters of reconstructed geometries from the ground truth, U-net segmented, and segAN segmented images. The average error of ejection pressure using U-net model was 14.0%. SegAN model made 8.4% average error of ejection pressure. For the stroke volume and cardiac output, U-net model made average 6.4% error and segAN model made 10.2% error. These

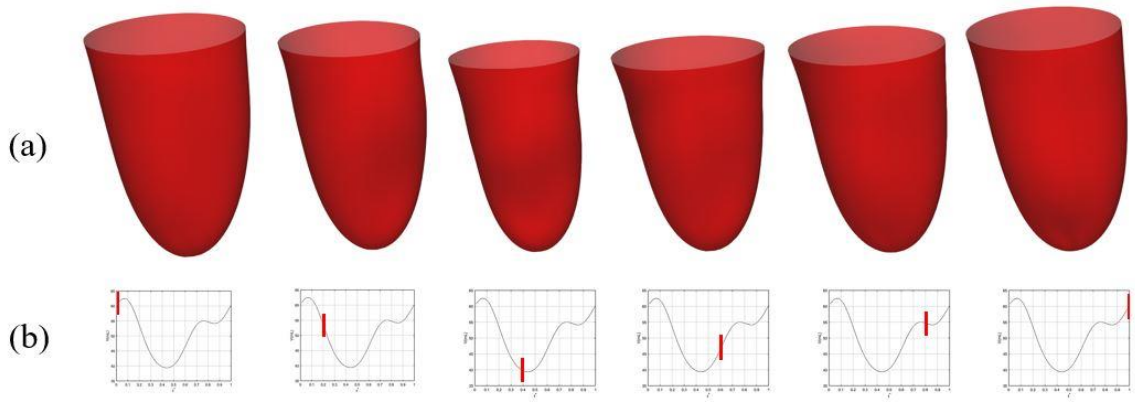

Figure 3.7 One cardiac cycle LV 3D reconstruction results. (a) Reconstruction geometries using segAN and (b) Volume-time graphs over one cardiac cycle with red line which represents corresponding time instance.

results showed that the U-net model generated reconstructed volumes which represented the absolute volume value better than segAN model. However, segAN model made LV shapes which represented the volume trend better than U-net model. *Figure 3.6* shows the comparison between LV geometries. Yellow LV shape was reconstructed with ground truth images. Red and green LV shapes were reconstructed with segAN and U-net segmented images, respectively. *Figure 3.7* illustrates the 3D reconstruction results over one cardiac cycle generated from segAN segmentation results. Each shape represents the specific time instances with the same time intervals during one cardiac cycle using a fixed apex position and the same point of view. Starting at the end-diastole, the model contracts during the systole and expands during the diastole.

3.3. Calculation time

U-net and segAN models generated segmentation images using NVIDIA GeForce GTX 1050 GPU. Level-set method segmented the LV boundary on the Intel® Core™ i7-7700HQ CPU. U-net took 53±2 seconds for segmenting the one cardiac cycle images, consists of between 40 to 42 images. It spent average 1.3 seconds per one images. SegAN required 104±8 seconds to segment the one cardiac cycle images. Average 2.51 seconds were taking per one images. Level-set method took 75±2 seconds for segmenting one image. 3D reconstruction algorithm generated one cardiac cycle's LV geometries on the Intel Xeon E5-2670 v2 CPU. It required average 49 seconds to reconstruct the geometry of one cardiac cycle. Total average 367 seconds were required to segment the six projection images over one cardiac cycle and 3D reconstruct using U-net and reconstruction algorithm. SegAN segmentation and reconstruction algorithms took average 673 seconds.

4. DISCUSSION

4.1. Segmentation

In-house dataset, which was used to train CNN models, was different with previous datasets about LV segmentation task. Previous datasets consisted of one or two images, e.g., end-diastolic and end-systolic images, from one case. These datasets had 60 to 450 different cases to acquire diversity [10, 34, 35]. However, in-house dataset consisted of 30 to 40 images from one case and had 10 different cases. The segmentation results showed that, if the number of images were enough, this new kinds of dataset also achieve high performance on the segmentation task. However, due to small number of cases, current segmentation model only worked well on the similar 2D echo images, which was obtained from same transducer. This problem can be solved by increasing the number of cases.

For a LV segmentation tasks, reasonable results can be obtained from a relatively small number of images comparing to other image segmentation tasks. This is because of two features of LV images. First, there are clear standard projections for acquiring the 2D echo images, e.g., six standard projections. Therefore, each projection's images have similar arrangement of structures and shape of each structure. Second, the shape of LV is quite simple. For the long-axis projections, the LV has half-ellipse shape. Also, short-axis projections represent the circular shape of LV. These features facilitate successful segmentation using small amount of data.

CNN models' segmented results showed some unexpected island-like structures in the background. The post-process was followed segmentation task to remove these

island-like structures. The island-like structures caused the higher Hausdorff distance because they were not appeared near LV. However, the dice metric, precision, and sensitivity did not decrease, because they occupied small number of pixels comparing with LV cavity and myocardium. Table 3.1 represents the effect of the post-process. Comparing the metrics before and after conducting the post-process, the Hausdorff distance was decreased from 1.4% to 4.7% and other metrics increased less than 1%. This results show the post-process succeeded in removing the island-like structures.

However, the metrices of LV myocardium were still worse than LV cavity's metrices after post-process. There are two main reasons that generated worse metric values. First, the total number of pixels were inherently different between LV cavity and myocardium. LV cavity had a larger number of pixels which is advantageous for calculating the metrics. Second, island-like structures which had larger diameter than the LV myocardium thickness were not removed using current post-process and these were included on the LV myocardium label. Different background removing method, such as region of interest (ROI) technique, should be developed to increase the segmentation accuracy.

Segmentation results show the usage of 2D echo image can be extended with CNN models by removing segmentation restriction. Currently, CNN models are actively used in the segmentation of MRI and CT cardiac images [26]. Most of the LV MRI and CT datasets consist of dense short-axis images. CNN models achieve more than 0.9 on the dice loss and precision for these image segmentation task [28-30]. This study shows that

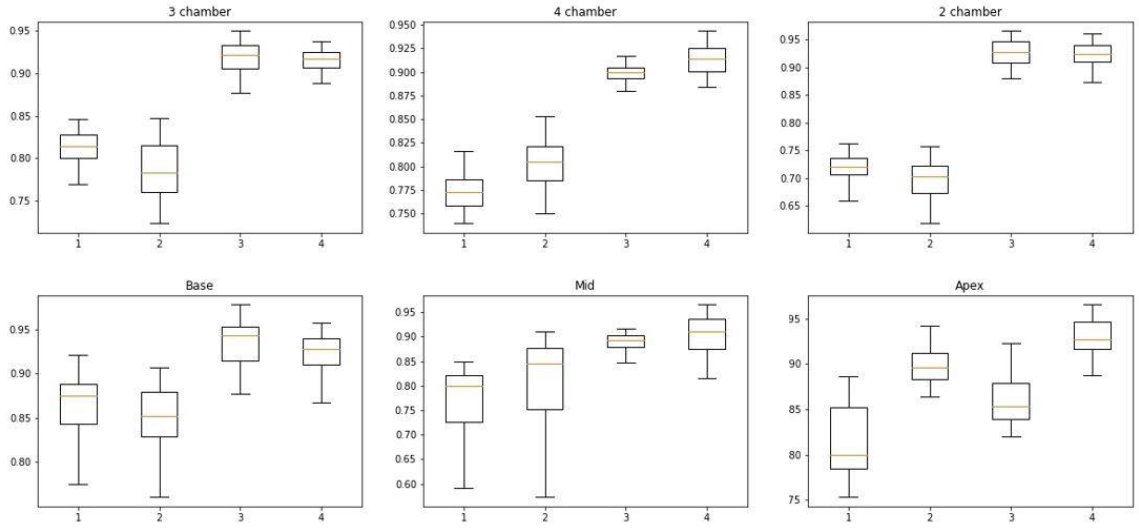

Figure 4.1 Tukey box plots computed from the dice metric. Each plot's numbers denote dice metric of 1: LV myocardium using U-net, 2: LV myocardium using segAN, 3: LV cavity using U-net, and 4: LV cavity using segAN.

the CNN models also can achieve a similar accuracy on 2D echo projections. This suggests that if there is a large dataset for the segmentation, the data types, e.g., MRI, CT and echo, would not significantly affect the performance of the CNN models.

Figure 4.1 shows the comparison between U-net and segAN models using Tukey box calculated from the dice metric. From these plots, U-net outperformed the segAN for the 2-chanmber and base projections. In contrast, for the apex projection, segAN performed better than the U-net. For the other three projections, the superiority of one model is not clear.

Figure 3.1 and Table 3.2, 3.3, and 3.4 show that the U-net and segAN models outperformed the level-set method on the LV segmentation task for the 2D echo images. Level-set method caused several specific problems. For the long-axis, the Aortic and

Mitral valves which were located end of the LV disturbed the exact LV cavity segmentation. The seriousness of this problem changed over the cardiac cycle with opening and closing the valves. The worst case was occurred when the valve was open. Level-set method could not distinguish between LV and the left atrium. To get reasonable breakpoint between two structures, the total number of iterations should be carefully controlled by experts for the level-set, i.e., not fully automatic. For the short-axis, the existence of trabeculae and papillary muscle, which had similar brightness on 2D echo images, caused the distorted segmentation results. However, CNN models generated segmentation results regardless of these problems on both long- and short-axis. Furthermore, CNN techniques were fully automatic. In contrast, the level-set method needed manual starting box which should be chose delicately. Finally, level-set method spent more calculation time than the machine learning models. Level-set method took almost 50 minutes to segment the images for one cardiac cycle.

To overcome the limitation of insufficient dataset, transfer learning and augmentation techniques were employed. However, the different quality of the 2D echo images, e.g., sparse border, variation in echo, and image noise, restricted the application of the trained models. When the model tries to segment different quality images, it generates non-robust segmentation results with incomplete segmentation lines and large island-like structures. Also, some 2D echo images depict cavities other than the LV. These are the common limitations of the machine learning segmentation models and could be improved by training with a larger and diverse dataset.

For the future work, automatic delineation method needs to be extended to segment the valves and superior walls, which is however, non-trivial task because in the standard 2D echo projection image, only small numbers of pixels represent the valves thus making them hard to distinguish from LV ends in long-axis and tendineae in short-axis projections.

4.2. 3D reconstruction

During validating the usage of six standard projections for the reconstruction, the role of the different kinds of projections for the reconstruction was estimated. *Figure 3.5*. showed the comparison between long-axis only, six long-axis projections, and long and base, three long-axis and base projections. Long-axis only combination succeed with capturing the volume change trend over the cardiac cycle. Long-axis and base combination obtained more accurate absolute LV volume than the long-axis only combination, even the small number of projections. This result shows that additional number of long-axis projections helps detecting the absolute volume value and additional short-axis projection supports to detect the volume change trend.

For the reconstruction using segmented results, inner wall boundary of the myocardium section was used (*Figure 3.2*). The accuracy of the inner wall boundary could be estimated from the LV cavity accuracy of segmentation results. Comparing to the average dice, precision, and sensitivity values, which were near 90%, the physiological parameters had 10 - 20% error. This showed that 10% error of the segmentation increased to 10 – 20% error after reconstruction and this error was coming from both segmentation and reconstruction.

5. CONCLUSIONS

In this study, dynamic 3D LV geometries over one cardiac cycle are automatically generated from six standard 2D echo projections. This algorithm consists of automatic segmentation and reconstruction parts. The machine learning segmentation techniques, especially CNN models, facilitate automatic, rapid and consistent 3D reconstruction from 2D echo images.

Two CNN models, U-net and segAN, have been employed to automatically segment the LV from the 2D echo images. These models were trained with both CAMUS dataset and an in-house dataset. CAMUS dataset was used for the transfer learning to compensate the small amount of training dataset. Also, effect of the post-process was quantified using CAMUS dataset. All metrices were improved by using post-process, which removed the island-like background.

In-house dataset consisted of total 2108 images. Standard projections of 2D echo, three long-axis and three short-axis projections, images over one cardiac cycle were extracted from 10 pigs.

The assessment metrics and segmented images show that both U-net and segAN outperform the level-set method. Especially, CNN models achieve high performance on LV cavity segmentation. For the LV myocardium segmentation, both model generated higher error than the cavity. Relatively small number of pixels which consist the myocardium and background island-like structure caused this high error. U-net achieved average 0.903 and 0.787 dice metrices for the LV cavity and myocardium. SegAN

achieved average 0.912 and 0.801 dice metrices for the LV cavity and myocardium. The performance difference between two models was small. However, segAN took two times more computation time than U-net.

In this study, previous 3D reconstruction algorithm was improved and validated with two perspective. Validity of using six standard sections was verified by comparing with other combinations of different projections. Also, reconstructed geometry was verified with benchmark. During this process, problem of algorithm was confirmed, and the cause of this problem was identified, the method of estimating the apex point inside the short-axis projections. After resolving this problem using information from three long-axis projections and geometric center matching process, algorithm generated the geometry, which represented the all projections and showed the high similarity with reference shape. Segmented images were reconstructed to the 3D LV geometry using improved algorithm. Based on the comparison of physiological parameters, the average predicted error was around 15%, which is similar to the previous methods using 2D echo compared to the gold standard MRI.

Fully automated pipeline, from 2D echo images to a 3D geometric geometry of LV, was generated by combining the machine learning segmentation technique and 3D reconstruction algorithm. This pipeline can help the fast analysis of the vast numbers of echo images in clinical applications to visualize 3D LV wall motion and evaluate global function. Also, this algorithm facilitates the patient-specific LV modeling and simulation without expert's knowledge and effort.

REFERENCE

1. Benjamin EJ, Virani SS, Callaway CW, Chamberlain AM, Chang AR, Cheng S, et al. Heart disease and stroke statistics-2018 update: a report from the American Heart Association. 2018;137(12):e67.

2. WEBER K, KI W, JS J. The heart as a muscle-pump system and the concept of heart failure. 1979.

3. Kono T, Sabbah HN, Rosman H, Alam M, Jafri S, Goldstein SJJotACoC. Left ventricular shape is the primary determinant of functional mitral regurgitation in heart failure. 1992;20(7):1594-8.

4. Van Dantzig JM, Delemarre BJ, Koster RW, Bot H, Visser CAJAhj. Pathogenesis of mitral regurgitation in acute myocardial infarction: importance of changes in left ventricular shape and regional function. 1996;131(5):865-71.

5. Lang RM, Badano LP, Tsang W, Adams DH, Agricola E, Buck T, et al. EAE/ASE recommendations for image acquisition and display using three-dimensional echocardiography. European heart journal cardiovascular Imaging. 2012;13(1):1-46.

6. Yodwut C, Weinert L, Klas B, Lang RM, Mor-Avi V. Effects of frame rate on three-dimensional speckle-tracking-based measurements of myocardial deformation. Journal of the American Society of Echocardiography : official publication of the American Society of Echocardiography. 2012;25(9):978-85.

7. Lang RM, Bierig M, Devereux RB, Flachskampf FA, Foster E, Pellikka PA, et al. Recommendations for chamber quantification. European journal of echocardiography : the journal of the Working Group on Echocardiography of the European Society of Cardiology. 2006;7(2):79-108.

8. Porshnev S, Bobkova A, Zyuzin V, Bobkov VJFr. Method of semi-automatic delineation of the left ventricle of the human heart on echographic images. 2013(8):44-8.

9. Armstrong AC, Ricketts EP, Cox C, Adler P, Arynchyn A, Liu K, et al. Quality control and reproducibility in M-mode, two-dimensional, and speckle Tracking echocardiography acquisition and analysis: the CARDIA Study, year 25 examination experience. 2015;32(8):1233-40.

10. Leclerc S, Smistad E, Pedrosa J, Ostvik A, Cervenansky F, Espinosa F, et al. Deep Learning for Segmentation Using an Open Large-Scale Dataset in 2D Echocardiography. IEEE transactions on medical imaging. 2019;38(9):2198-210.

11. Petitjean C, Dacher JN. A review of segmentation methods in short axis cardiac MR images. Medical image analysis. 2011;15(2):169-84.

12. Cocosco CA, Niessen WJ, Netsch T, Vonken EJ, Lund G, Stork A, et al. Automatic image-driven segmentation of the ventricles in cardiac cine MRI. Journal of magnetic resonance imaging : JMRI. 2008;28(2):366-74.

13. Liu H, Hu H, Xu X, Song E. Automatic left ventricle segmentation in cardiac MRI using topological stable-state thresholding and region restricted dynamic programming. Academic radiology. 2012;19(6):723-31.

14. Billet F, Sermesant M, Delingette H, Ayache N, editors. Cardiac motion recovery and boundary conditions estimation by coupling an electromechanical model and cine-MRI data. International Conference on Functional Imaging and Modeling of the Heart; 2009: Springer.

15. Zhang H, Wahle A, Johnson RK, Scholz TD, Sonka M. 4-D cardiac MR image analysis: left and right ventricular morphology and function. IEEE transactions on medical imaging. 2010;29(2):350-64.

16. Zhuang X, Hawkes DJ, Crum WR, Boubertakh R, Uribe S, Atkinson D, et al., editors. Robust registration between cardiac MRI images and atlas for segmentation propagation. Medical Imaging 2008: Image Processing; 2008: International Society for Optics and Photonics.

17. Litjens G, Kooi T, Bejnordi BE, Setio AAA, Ciompi F, Ghafoorian M, et al. A survey on deep learning in medical image analysis. Medical image analysis. 2017;42:60-88.

18. Menze B, Jakab A, Bauer S, Kalpathy-cramer J, Farahani K, Kirby JJITo. The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS). Medical Imaging. 2014:1-32.

19. Yang J, Veeraraghavan H, Armato III SG, Farahani K, Kirby JS, Kalpathy-Kramer J, et al. Autosegmentation for thoracic radiation treatment planning: A grand challenge at AAPM 2017. 2018;45(10):4568-81.

20. Morra J, Tu Z, Toga A, Thompson PJGCWMSLSC. Automatic segmentation of MS lesions using a contextual model for the MICCAI grand challenge. 2008:1-7.

21. de Vos BD, Wolterink JM, Leiner T, de Jong PA, Lessmann N, Isgum I. Direct Automatic Coronary Calcium Scoring in Cardiac and Chest CT. IEEE transactions on medical imaging. 2019;38(9):2127-38.

22. Ciompi F, de Hoop B, van Riel SJ, Chung K, Scholten ET, Oudkerk M, et al. Automatic classification of pulmonary peri-fissural nodules in computed tomography using an ensemble of 2D views and a convolutional neural network out-of-the-box. Medical image analysis. 2015;26(1):195-202.

23. Wolterink JM, Leiner T, de Vos BD, van Hamersvelt RW, Viergever MA, Isgum I. Automatic coronary artery calcium scoring in cardiac CT angiography using paired convolutional neural networks. Medical image analysis. 2016;34:123-36.

24. Ronneberger O, Fischer P, Brox T, editors. U-net: Convolutional networks for biomedical image segmentation. International Conference on Medical image computing and computer-assisted intervention; 2015: Springer.

25. Xue Y, Xu T, Zhang H, Long LR, Huang X. SegAN: Adversarial Network with Multi-scale L1 Loss for Medical Image Segmentation. Neuroinformatics. 2018;16(3-4):383-92.

26. Gardner BI, Bingham SE, Allen MR, Blatter DD, Anderson JL. Cardiac magnetic resonance versus transthoracic echocardiography for the assessment of cardiac volumes and regional function after myocardial infarction: an intrasubject comparison using simultaneous intrasubject recordings. Cardiovascular ultrasound. 2009;7:38.

27. Suinesiaputra A, Cowan BR, Al-Agamy AO, Elattar MA, Ayache N, Fahmy AS, et al. A collaborative resource to build consensus for automated left ventricular segmentation of cardiac MR images. Medical image analysis. 2014;18(1):50-62.

28. Bernard O, Lalande A, Zotti C, Cervenansky F, Yang X, Heng PA, et al. Deep Learning Techniques for Automatic MRI Cardiac Multi-Structures Segmentation and Diagnosis: Is the Problem Solved? IEEE transactions on medical imaging. 2018;37(11):2514-25.

29. Bai W, Sinclair M, Tarroni G, Oktay O, Rajchl M, Vaillant G, et al. Automated cardiovascular magnetic resonance image analysis with fully convolutional networks. Journal of cardiovascular magnetic resonance : official journal of the Society for Cardiovascular Magnetic Resonance. 2018;20(1):65.

30. Zreik M, Leiner T, De Vos BD, van Hamersvelt RW, Viergever MA, Išgum I, editors. Automatic segmentation of the left ventricle in cardiac CT angiography using convolutional neural networks. 2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI); 2016: IEEE.

31. Lieman-Sifry J, Le M, Lau F, Sall S, Golden D, editors. FastVentricle: cardiac segmentation with ENet. International Conference on Functional Imaging and Modeling of the Heart; 2017: Springer.

32. Çiçek Ö, Abdulkadir A, Lienkamp SS, Brox T, Ronneberger O, editors. 3D U-Net: learning dense volumetric segmentation from sparse annotation. International conference on medical image computing and computer-assisted intervention; 2016: Springer.

33. Tan LK, Liew YM, Lim E, McLaughlin RA. Convolutional neural network regression for short-axis left ventricle segmentation in cardiac cine MR sequences. Medical image analysis. 2017;39:78-86.

34. Veni G, Moradi M, Bulu H, Narayan G, Syeda-Mahmood T, editors. Echocardiography segmentation based on a shape-guided deformable model driven by a fully convolutional network prior. 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018); 2018: IEEE.

35. Zhang J, Gajjala S, Agrawal P, Tison GH, Hallock LA, Beussink-Nelson L, et al. Fully Automated Echocardiogram Interpretation in Clinical Practice. Circulation. 2018;138(16):1623-35.

36. Schenkel T, Malve M, Reik M, Markl M, Jung B, Oertel HJAobe. MRI-based CFD analysis of flow in a human left ventricle: methodology and application to a healthy heart. 2009;37(3):503-15.

37. Mangual JO, Kraigher-Krainer E, De Luca A, Toncelli L, Shah A, Solomon S, et al. Comparative numerical study on left ventricular fluid dynamics after dilated cardiomyopathy. 2013;46(10):1611-7.

38. Rajan NK, Song Z, Hoffmann KR, Belohlavek M, McMahon EM, Borazjani I. Automated Three-Dimensional Reconstruction of the Left Ventricle From Multiple-Axis Echocardiography. Journal of biomechanical engineering. 2016;138(1).

39. Pourmorteza A, Schuleri KH, Herzka DA, Lardo AC, McVeigh ERJCCI. A new method for cardiac computed tomography regional function assessment: stretch quantifier for endocardial engraved zones (SQUEEZ). 2012;5(2):243-50.

40. Dekker DL, Piziali RL, Dong Jr EJC, research b. A system for ultrasonically imaging the human heart in three dimensions. 1974;7(6):544-53.

41. Sawada H, Fujii J, Kato K, Onoe M, Kuno YJH. Three dimensional reconstruction of the left ventricle from multiple cross sectional echocardiograms. Value for measuring left ventricular volume. 1983;50(5):438-42.

42. Raichlen JS, Trivedi SS, Herman GT, Sutton MGSJ, Reicher NJJotACoC. Dynamic three-dimensional reconstruction of the left ventricle from two-dimensional echocardiograms. 1986;8(2):364-70.

43. Ghosh A, Nanda NC, Maurer GJUiM, Biology. Three-dimensional reconstruction of echocardiographic images using the rotation method. 1982;8(6):659-61.

44. Jakrapanichakul D, Tanabe K, Belohlavek M, Seward JBJJouimojotAIoUiM. Shape reconstruction of the left ventricle: accuracy of limited-plane three-dimensional echocardiography. 2001;20(7):767-74.

45. Gustavsson T, Pascher R, Caidahl KJCmi, graphics. Model based dynamic 3D reconstruction and display of the left ventricle from 2D cross-sectional echocardiograms. 1993;17(4-5):273-8.

46. Crick SJ, Sheppard MN, Ho SY, Gebstein L, Anderson RH. Anatomy of the pig heart: comparisons with normal human cardiac structure. Journal of anatomy. 1998;193 (Pt 1):105-19.

47. Manovel A, Dawson D, Smith B, Nihoyannopoulos P. Assessment of left ventricular function by different speckle-tracking software. European journal of echocardiography : the journal of the Working Group on Echocardiography of the European Society of Cardiology. 2010;11(5):417-21.

48. Bagger T, Sloth E, Jakobsen CJ. Left ventricular longitudinal function assessed by speckle tracking ultrasound from a single apical imaging plane. Critical care research and practice. 2012;2012:361824.

49. Haralick RM, Sternberg SR, Zhuang X. Image analysis using mathematical morphology. IEEE transactions on pattern analysis and machine intelligence. 1987;9(4):532-50.

50. Li C, Xu C, Gui C, Fox MD. Distance regularized level set evolution and its application to image segmentation. IEEE transactions on image processing : a publication of the IEEE Signal Processing Society. 2010;19(12):3243-54.

51. Dietenbeck T, Alessandrini M, Friboulet D, Bernard O, editors. CREASEG: a free software for the evaluation of image segmentation algorithms based on level-set. 2010 IEEE International Conference on Image Processing; 2010: IEEE.

52. Dice LRJE. Measures of the amount of ecologic association between species. Ecology. 1945;26(3):297-302.

53. Powers DM. Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation. 2011.

54. Huttenlocher DP, Klanderman GA, Rucklidge WJJITopa, intelligence m. Comparing images using the Hausdorff distance. 1992 IEEEComputerSociety Conference on Computer Vision and Pattern Recognition. 1993;15(9):850-63.

55. Reiter U, Reiter G, Manninger M, Adelsmayr G, Schipke J, Alogna A, et al. Early-stage heart failure with preserved ejection fraction in the pig: a cardiovascular magnetic resonance study. Journal of cardiovascular magnetic resonance : official journal of the Society for Cardiovascular Magnetic Resonance. 2016;18(1):63.

56. Hoffmann, R., et al., Analysis of Left Ventricular Volumes and Function: A Multicenter Comparison of Cardiac Magnetic Resonance Imaging, Cine Ventriculography, and Unenhanced and Contrast-Enhanced Two-Dimensional and Three-Dimensional Echocardiography. Journal of the American Society of Echocardiography, 2014. 27(3): p. 292-301.

57. Greupner, J., et al., Head-to-Head Comparison of Left Ventricular Function Assessment with 64-Row Computed Tomography, Biplane Left Cineventriculography, and Both 2- and 3-Dimensional Transthoracic EchocardiographyComparison With Magnetic Resonance Imaging as the Reference Standard. Journal of the American College of Cardiology, 2012. 59(21): p. 1897-1907.

58. Hoffmann, R., et al., Assessment of systolic left ventricular function: a multi-centre comparison of cineventriculography, cardiac magnetic resonance imaging, unenhanced and contrast-enhanced echocardiography. European Heart Journal, 2005. 26(6): p. 607-616.

APPENDIX A

IN-HOUSE DATASET: TRAINING AND TESTING DATASET

Training dataset

Projections	Pig identification number	Number of images
3-chamber	15, 6921, 6926, 6927, and 8550	196
2-chamber	6918, 6926, 6933, and 8550	151
4-chamber	6918, 6926, 6927, 6933, and 8550	194
Base	6918, 6921, 6926, 6937, and 8550	189
Mid	15, 6921, 6927, 6933, and 6937	200
Apex	6918, 6927, and 6933	118

Testing dataset

Projections	Pig identification number	Number of images
3-chamber	6933 and 6937	82
2-chamber	6921 and 6937	80
4-chamber	6921 and 6937	80
Base	6927 and 6933	83
Mid	6918 and 6926	69
Apex	6921 and 6937	53