공학석사 학위논문

REAL-TIME ESTIMATION OF THE LEFT VENTRICULAR VOLUME FROM ECHOCARDIOGRAM DURING CARDIOPULMONARY RESUSCITATION USING CONVOLUTIONAL NEURAL NETWORK

합성곱 신경망을 이용한 심폐소생술 중

심초음파 영상에서 실시간 좌심실 부피 추정

2017 년 8 월

서울대학교 대학원

협동과정 바이오엔지니어링 전공

이병탁

합성곱 신경망을 이용한 심폐소생술 중 심초음파 영상에서 실시간 좌심실 부피 추정

지도교수 이 정 찬

이 논문을 공학석사 학위논문으로 제출함

2017 년 7월

서울대학교 대학원

협동과정 바이오엔지니어링 전공

이 병 탁

이병탁의 공학석사 학위논문을 인준함

2017 년 7월

위 원 장 김 희 찬 (인)

부 위 원 장 이 정 찬 (인)

위 원 윤 형 진 (인)

MASTER THESIS

REAL-TIME ESTIMATION OF THE LEFT VENTRICULAR VOLUME FROM ECHOCARDIOGRAM DURING CARDIOPULMONARY RESUSCITATION USING CONVOLUTIONAL NEURAL NETWORK

BY

LEE BYEONG TAK

JULY 2017

INTERDISCIPLINARY PROGRAM IN BIOENGINEERING

THE GRADUATE SCHOOL

SEOUL NATIONAL UNIVERSITY

Abstract

Real-time estimation of the left ventricular volume from echocardiogram during cardiopulmonary resuscitation using convolutional neural network

Lee Byeong Tak

Interdisciplinary program in bioengineering

The graduate school

Seoul National University

This thesis describes the method for real-time segmentation based on echocardiography and three-dimensional transformation model for the left ventricular volume estimation during cardiopulmonary resuscitation (CPR). Because all people have a different structure of thoracic and the position of the heart, it has been required to optimize CPR by a person. As one of the improved methods, bio-signal feedback using echocardiography CPR is carried out. Echocardiography shows how the heart is compressed by chest compression, which directly shows cardiac output. There are two steps in estimating the cardiac output in echocardiography. The left ventricular segmentation from the echocardiography is needed to be segmented. After that, the three-dimensional volume is required to be estimated with two-dimensional segmented images. However, echocardiography during CPR is difficult due to the instability of contact between the transducer and the chest. Moreover, the previous models that map the segmented two-dimensional

image to the left ventricular volume assume the heart is contracted isometrically, which is different from the condition of the heart during CPR. To solve these problems, the method for segmentation of the left ventricle stable during CPR and the model that can be applied to CPR conditions is suggested in this dissertation. The convolutional neural network is adopted to the left ventricular segmentation problem. Based on the structure of 'SegNet' that is a fully convolutional network for real-time segmentation, skip connection and dice coefficient are applied to adapt the model to echocardiography domain. The former one helps the network to preserve the information of original images, and the latter one is used for stable segmentation. Moreover, Gated recurrent unit that is used for time series data analysis is applied to reflect the previous frames. The network achieves robust and accurate segmentation by referencing the previous frames in the segmentation of current frame. Comparing to Geodesic Active Contour method that shows the best performance in echocardiography, the proposed algorithm accomplishes higher accuracy and robust to unclear images. The left ventricular model is derived with applying constraints during CPR for modeling problem. The heart during CPR is not contracted. Thus, the assumption of the same surface between the diastolic heart and compressed heart is used. Moreover, the single ellipsoid model with the same length in the minor and intermediate axes is adopted. In comparison experiment to ETCO2 that affects the cardiac output during CPR, the proposed model show much greater correlation than the previous model.

Keywords: Cardiopulmonary resuscitation, Echocardiography, Segmentation, Convolutional neural network, Gated recurrent unit, Left ventricular model

Student Number: 2015-22888

CONTENTS

List of Tables

List of Figures

1. Introduction

1.1. Problems

CPR is a first aid that provides oxygen to the body when the heart is stopped. According to the CPR guidelines, it is recommended to compress the inter-nipple line for heart-arrested patients. However, it is pointed out that inter-nipple line is not the optimal position for all patients because of the different structure of the thoracic and the position of the heart by the patient[1-3]. Therefore, studies to find the optimal position for the chest compression is conducted with the cardiac output obtained from echocardiography[4]. To estimate cardiac output by echocardiography, the left ventricular region should be segmented based on echocardiography, and the segmented two-dimensional region should be transformed into a three-dimensional volume using the left ventricular model. There are problems in this process. First, analyzing of the echocardiography during CPR is difficult. Ultrasound images have much speckle noise. Moreover, several frames do not have enough information to be segmented because of poor contact between the chest and the transducer. Second, the previous left ventricular model cannot be applied to the arrested heart. The previous model assumes the heart contracts isometric. However, the arrested-heart

does not contract[5]. Therefore, the model considering the specific circumstance of the heart during CPR is required.

1.2. Aims

The objective of the thesis is to solve two problems. The first is to develop the robust left ventricle segmentation system in real-time based on echocardiography during CPR. The second is to design appropriate model for estimating the left ventricle volume based on the segmented region during CPR.

1.3. Related work

1.3.1. Left ventricle segmentation in echocardiography

There are many automated left ventricular segmentation techniques based on echocardiogram. They can be classified into the following classes: explicit method, pattern recognition based method, and deep learning based method. The explicit method uses the rule devised by the programmer, it including bottom-up threshold approaches and edge detection method[6, 7]. In the pattern recognition based method, instead of setting up the whole process, the programmer designate specific objective function for segmentation. This function can involve high-order features. An example of these is Active contour model method and Active shape method [8, 9]. Especially, active contour method is widely used in echocardiography processing because of the robustness against speckle noise. The convolutional neural network is a representative example of deep learning in image analysis[10]. Even though they are just started to be applied to the medical image, the result of them is better than the previous method.

1.3.2. Left ventricle modeling

Various Left ventricle models have been proposed to estimate the volume in three dimensions from 2D region obtained from echocardiography. These are classified into three classes: single-figure based model, multi-figures based model, Simpson's method[11]. Correlation experiment showed that Simpson's method had a higher correlation than the others. On the other hand, as the model uses many figures, more images of the left ventricular layers are required to calculate the model.

1.4. Proposed solution

For image segmentation task, the neural network is implemented to apply to the left ventricle segmentation. It is designed to run in real-time, robust to the speckle noise, and stable in analyzing the image when the contact between the chest and transducer is unstable. Regarding the structure of the neural network, SegNet with skip connection and dice coefficient is combined with stack GRU that is located between encoding path and decoding path[12, 13]. Left ventricle model that can apply in CPR is suggested. With a constraint of non-contractability, the mathematical model of the left ventricle is suggested. Moreover, the single-figure based model is used for the real-time process with a single layer of the left ventricle.

2. Literature review

2.1. Image segmentation method

Image segmentation methods are widely used in computer vision area including medical image analysis. They are roughly classified into three groups: explicit methods, pattern recognition based methods, and deep learning based methods. Explicit methods segment the object with the simple criteria set by the researcher such as the histogram of the image. Thus, they cannot capture the complicated features of images as well as the performance varies by the quality of images. Pattern recognition based methods search the features of images selected by the researchers. They can segment the object with complex features compared to the explicit method. For example, a curvature of the border line of the objects or the size and the circumference of the object can be the features. Although it shows better performance compared to explicit methods, it cannot capture features that are not selected by the investigator. Deep learning based method overcomes this problem. It learns the features of the image by itself, and observe complex features that researchers can not find. The result from deep learning outdo the other methods these days. Even though there have been many studies to segment the left ventricle from echocardiography, automatic segmentation of the structure of the heart is still a

challenge because of acoustic interferences of the echocardiography[14]. Also, CPR makes it harder to analyze the echocardiography due to the severe artifact. This chapter introduces the left ventricle segmentation algorithm applied to echocardiograms and compares their performance. Because there is no case applied to the CPR, all cases described in this chapter are followed to the normal state heart.

	Explicit method	Pattern recognition	Deep learning
Examples	Threshold[7], Edge detection[6]	Active contour[8], Active shape[9]	Deep convolutional neural network[10]
Advantage	Easy to implement	Can capture high-order features	Learning the features by itself
Disadvantage	Can not capture high-order features	Can not capture the features that are not signified	Computationally expensive in training

Table 2.1. Compare of image segmentation methods. Higher-order features mean that more criteria are required to obtain the object. Deep learning of pattern recognized based method requires more criteria than explicit method. For example, threshold method uses a single criterion called image intensity. In contrast, active contour method needs three criteria including the image gradient and shape of the contour.

2.1.1. Explicit methods

Explicit method is straightforward and intuitive. Among explicit methods, threshold and edge detection are often used in segmentation problem in echocardiography. They searched the image intensity map and returned the pixels that satisfy the criteria. The former relies on the intensity of the pixels in the image and the latter relies on

the gradient of the image intensity map.

2.1.1.1. Threshold

Threshold method explores every pixel of the image and assigns suitable threshold value to the pixels which have a certain value of the intensity. It is simple, easy to implement and suitable for medical imaging because the difference in intensity of the pixel is clear. At the same time, it has a limitation in processing under noise or in the case of an image in which the light to be illuminated varies depending on the time. Several thresholding methods applied in echocardiography have been proposed. W. Ohyama segmented the left ventricle segmentation using ternary thresholding method[15]. It uses two steps of segmentation: one for the ternary thresholding at the first step, the other for binary segmentation with the mask of initial ternary segmentation. The average distance from the centroid to endocardium is used to measure the accuracy, and the correlation is 0.942 in their dataset. However, it has problem in vulnerable to noise. J.B. Santos applied window adaptive thresholds to segment the left ventricle. It applied adaptive window that divides the images into blocks and performing threshold using the Otsu's method[16]. Adaptive window thresholding overcomes a problem of threshold method: not robust to the noise.

2.1.1.2. Edge detection

Edge is significant local changes of intensity in an image. That is, edge detection extracts the high gradient line from the image intensity map. The criterion of the gradient is determined by the researcher. Steve M. constructed ventricular cavity area estimation system with edge detection[17]. After background subtraction, sharpening and smoothing, four types of an edge detection algorithm including Sobel,

Laplacian based, spatial difference and gray level thresholding are applied. Sobel edge detection method shows the highest correlation among them: 0.92.

2.1.2. Pattern recognition methods

Unlike an explicit method, pattern recognition based methods segment an object with higher-order features that are devised by the researcher. Those features are represented in a loss function, which is the distance of features between the current model and the ground truth. The model is trained by reducing the loss fnction. Active contour and Active shape are representative models for segmentation in echocardiography. Active contour adopts the shape of the contour with image gradient as features. As another example, Active shape uses the distribution of specific points as features.

2.1.1.1. Active contour

Curves are moved to find object boundaries following the objective function in Active contour method. That is, the purpose of the active contour models is to find a proper contour for an object. The contour is initially set in the range of a particular area, and it is changed until it minimizes the loss function: the contour energy function, which becomes zero when the contour is exactly fitted to an object. The energy function is consist of the internal energy that depends on the shape of the contour itself and the external energy that depends on the image gradient. Internal energy comprises of continuity term and curved term. They represent the curvature and the derivative of the curvature, respectively. External energy reflects the gradient of the image. Thus, energy function converges in the direction where the change of

the curvature and curvature, and the change of image gradient is small.

$$\mathrm{E} = \int \alpha E_{cont} + \beta E_{curv} + \gamma E_{image} ds$$

Chalana suggested a multiple active contour model for cardiac boundary detection, which uses temporal data as well as spatial data[18]. The model is similar to the three-dimensional active contour model proposed by Cohen[19]. An algorithm was evaluated with randomly selected echocardiographic images from 44 patients, and the correlation to the manual segmentation is 0.91. Weiming Wang proposed a combinative active contour with Gaussian pyramid[20]. It divides the raw image into several images of the image pyramid, and segment the object region from the lower resolution, which is used as the prior for segmentation in higher resolution. With image pyramid, it achieves high accuracy and robust to speckle noise at the same time. Experiment result showed excellent correlation with the ground truth: 0.95.

2.1.2.2. Active shape

The Active shape is an algorithm that detects objects that change within a limited range and is often used to detect changes of objects. In training process, several points that represent the object and the line connecting a pair of marked points are trained in a training set. The trained model has mean value and variation of each point, which is the model of the object. When test images are input into the model, the active shape model estimates corresponding points in the model to the input image. Gary Jacob applied active shape model in tracking myocardial borders[21]. With, simple shape model, Kalman filter is adapted for tracking each point. The correlation to manual segmentation is 0.94. Gregg Belous segmented the left

ventricle using random forest with active shape model[22]. The shape model is applied at the initial detection landmark. Following, each landmark of the shape model is directed as a result of random forest classifier. The evaluation was performed on 35 patient ultrasound images. Global overlap coefficient for the active shape model is 0.84, on the other hand, the correlation was 0.9 in the case of Active shape method with random forest.

2.1.3. Deep learning

Deep learning based segmentation is based on convolutional neural network structure. Compared to machine learning based method, deep learning based methods explore the features that are needed for segmentation by itself. In other words, the network adjusts the parameters to extract the proper feature while the neural network is being trained. The convolutional neural network has been applied in the medical image area with the development of deep learning[23]. U-net is the one of the most popular segmentation model used in the biomedical domain[24]. It is designed for segmentation of electron microscopic stacks and overcomes the difficulties of biomedical images: lack of the amount of the images. Also, several methods have been developed for echocardiography. Zhao segmented epicardium and endocardium with localization segmentation method[25]. Classification of the patch is used to segment the region. If the patch matches the region of interest, a value of 1 is assigned to the middle of the patch, and a value of 0 otherwise. Li Yu tried to segment the left ventricle with the image pyramid[26]. Additionally, a matching approach is used for separate the connection region between LV and LA. From their dataset, comparison among active contour method, fixed CNN, their

method experiments. The correlation to the manual segmentation is 0.86, 0.88, 0.95, and the time cost is 8.6, 11, 25.4(s), respectively.

		Method	Accuracy
Explicit method	Threshold	Window adaptive threshold with Otsu's method[15]	-
		Ternary thresholding[16]	0.942
	Edge detection	Proper processing with Sobel edge detection[17]	0.92
Pattern recognition	Active contour	Combining temporal data with spatial data[18]	0.91
		Gaussian pyramid with geodesic active contour[20]	0.95
	Active shape	Simple shape model[21]	0.94
		Combining with random forest[22]	0.9
Deep learning		Dynamic neural network[25]	0.95
		Localization with segmentation[26]	-

Table 2.2. Accuracy and Process time of the left ventricle segmentation in echocardiography. All of them are applied to a different dataset. Thus, a quantitative comparison is not appropriate. The number in the table is just for reference.

2.2. Left ventricle modeling

The efficiency of the chest compression during cardiopulmonary resuscitation is evaluated by the ratio of the volume perfused to the body from the heart, which is called the ejection fraction of the left ventricle. An ejection fraction is a ratio between the volume of the blood ejected from the left ventricle and the volume of the left ventricle at diastole. The real volume of the left ventricle is impossible to obtain via echocardiography. Thus several left ventricle models have been suggested to estimate the volume of the left ventricle based on the echocardiography. Those models are classified into three groups: a single solid figure model, a combination of several solid figures and Simpson's rule.

2.1.1. A single solid figure model

A single solid figure model assumes the shape of the left ventricle as a single figure, such as a sphere or ellipsoid that estimates the left ventricle as a sphere and an ellipsoid. A sphere model is rarely used because there are differences to the actual anatomical structure of the left ventricle. Instead, the ellipsoid model is often used in single figure models. W. yatt assumed several models of the ellipsoid and compared the correlation between real volume and the estimated volume using ellipsoid model[27]. Among ellipsoid model, there was a correlation of 0.969 in single short-axis are-length method and 0.956 in short-axis diameter cube method. Chaudry, K suggested a simple ellipsoid model that has different three axes[28]. The correlation to the real volume was 0.91 for systole volume and 0.86 for diastole volume. Folland, E compared geometry models including ellipsoid-biplane method, ellipsoid-single plane method, and modified the ellipsoid method for left ventricle

ejection fraction in real-time[29]. Each model has correlation of 0.78, 0.76, and 0.55, respectively.

2.1.2. Combination of solid figures model

A combination of several figures expresses the left ventricle with two or three figures of a circular cone, circular cylinder, and hemiellipsoid. For example, the cylinder hemiellipse model represents the left ventricle with a combination of semi-ellipsoidal and cylindrical. Each of them is divided along with major axis with an equal surface area. A cylinder truncated cone-cone is consist of three different figures including cylindrical, frustum and cone. The more figures are combined, the more accurate in representing the left ventricle. Pascal Gueret studied the correlation of the left ventricle stroke volume determined by echocardiography with a bullet like a cylinder hemiellipse model[30]. The correlation is calculated as 0.9. Folland compared ejection fraction of the cylinder-truncated con-cone model to the left ventricle and obtained a correlation of 0.78[29].

2.1.3. Simpson's rule

Simpson's rule approximates the volume with an integration of cylinders along with major axis. Each cylinder's radius and height are the length of the minor axis and the major axis divided by the number of cylinders. As selecting the proper height of each cylinder, the height converges to infinity, and the model becomes more accurate with higher computation complexity. It is considered as the golden standard in estimating the left ventricle volume. The correlation study to the ejection fraction and the volume was conducted by many researchers including Schiller, Silverman, Wyatt,

and Bommer[27, 31-33].

	Reference Model	Correlation
Single figure	Single short-axis area-length[27]	0.969 (EF)
	Simple ellipsoid that has different three axes[28]	0.91 for systole 0.86 for diastole
	Ellipsoid-biplane, ellipsoid-single plane, modified ellipsoid[29]	0.78, 0.76, 0.55 (EF)
Multi-figure	Bullet like a cylinder hemiellipse[30]	0.9 (EF)
	Cylinder-truncated cone-cone model[29]	0.78 (EF)
Simpson's rule	Three figures based Simpson's method[31]	0.98 (EF)
	Twenty figures based Simpson's method[32]	0.94 for diastole 0.91 for systole
	Computer aided[33]	0.84 (EF)

Table 2.3. Comparison of the left ventricle model.

3. Basic theory

3.1 Deep learning

A similar picture has similar features. Traditional image segmentation method uses the image intensity as the main feature of the segmentation and Pattern recognition based method use high-order feature in segmentation. Both methods have in common that the user specifies the features needed for image analysis. However, deep learning uses another method. That is the main idea of deep learning. A computer learns the features of the objective by trial and errors until the model of deep learning finds the proper features. That is, a computer finds the criteria to distinguish the object from the others by learning hidden regularity of the object.

3.1.1 Artificial Neural Networks

The cortex is made up of billions of neurons. An artificial neural network is a simplified representation of such a biological neural network and was represented by Warren McCulloh[34].

3.1.1.1 Structure of Networks

The structure of the artificial neural network is similar to a real neural network. In

the neural network, synapses convey the signal from dendrites to the axon via cell body. Later, the axon branches out and connects to dendrites of other neurons, which leads cascade activation of following synapses' group. Similarly, in the artificial neural networks, input data are transferred from the input channels to the output channels through the activation function[35]. Subsequently, the output data enter the following input channels. Input channels, output channels, and activation function correspond to dendrites, axon, and cell body, respectively. The activation function is located between the input channels and the output channels, and it decides whether the summation of the input signal is activated or not. A single unit of these elements is called Perceptron. As the neural network does, a set of single Perceptron makes the complicated multi-layer perceptrons. Perceptron is referred to as the input layers, the hidden layers, and the output layers, depending on which part they are located.

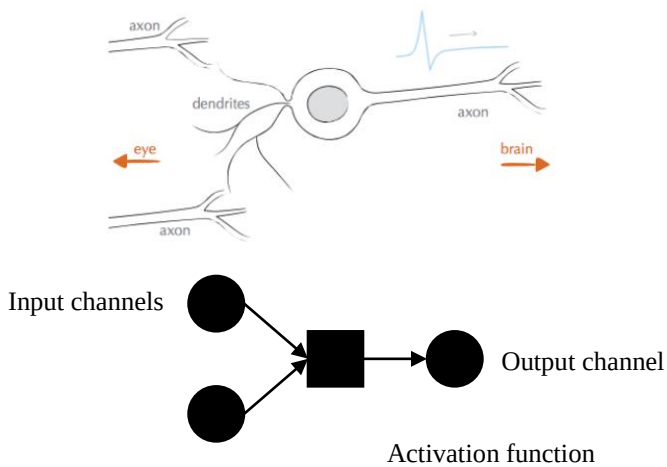


Figure 3.1. The structure of a neuron and a single-perceptron[36]. Top is the structure of the real neural and bottom is the struecture of the artificial neuron.

3.1.1.2 Activation function

The activation function is a mathematical representation that biological neurons are activated when the input stimulus is above a specific level. There are several commonly used activation functions including Sigmoid, tanh, and Relu[37]. The sigmoid function maps the input values to a value range from 0 to 1. It has problems in practice. For example, when an input exceeds a certain range, the differential value of the corresponding point, which is used for updating parameters, is 0. Thus, no update occurs during training. Additionally, it is computationally expensive because of not zero-centered characteristics and complexity of the function. Tanh is introduced to solve these problems. Tanh is a function mapping input signal to range from -1 to 1. It does not solve gradient vanishing problem but solves the not zero-centered problem of the sigmoid. Relu has been used most recently as it solves the problem of gradient vanishing. The gradient does not become zero like sigmoid or tanh. However, it has another problem called dying relu, meaning that most of the cells in the network is deactivated because of negative input.

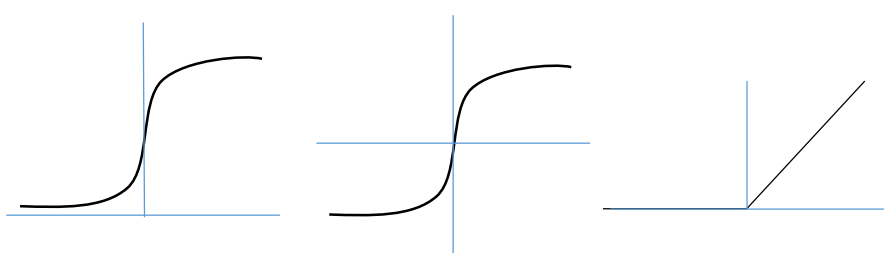


Figure 3.2. Sigmoid, tanh, and Relu activation function

3.1.2 Convolutional Neural Network

The convolutional neural network is similar to the artificial neural network but slightly different. When the information moves from one of the hidden layers in multi-perceptron to following hidden layer, all the Perceptrons between those two layers are connected. This structure has limitations in that it does no express the relation between Perceptron in one layer. However, convolutional neural networks can express the relation within a layer by using convolution layers instead of fully connected layers[37]. Therefore, there is an advantage in representing the spatial information like images. Convolutional neural networks are made up of convolution layers and pooling layers. Each of them plays a different role in the network.

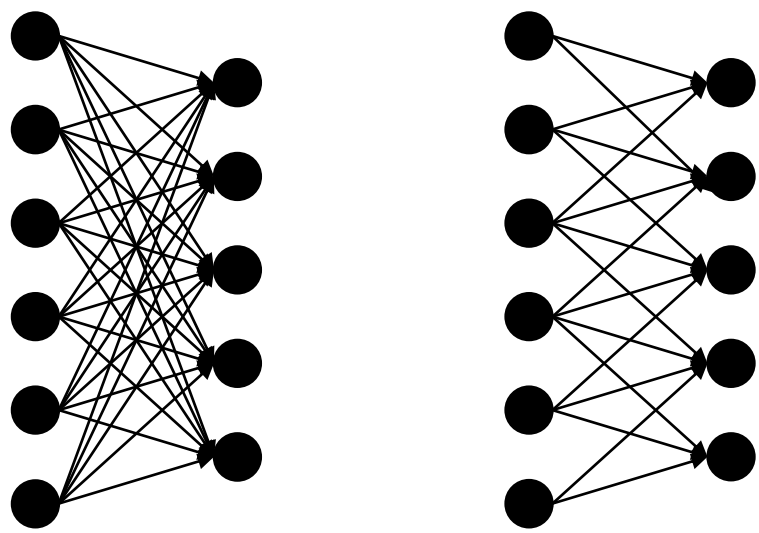


Figure 3.3. Comparison between Artificial neural network and Convolutional neural network (Left: Artifcial neural network, Right: Convolutional neural network)

3.1.2.1 Convolution layers

In image processing, several types of the kernel are applied to the image to acquire the smoothed, edged, or equalized images. This process occurs in convolution layer in the convolutional neural network as well. The difference is that the parameters of a kernel are determined by a programmer in normal image processing, but they are continuously updated in convolution layers by training. There are three hyperparameters to identify in convolution layers: kernel size, strides, and padding[38]. The size of kernel decides how many pixels' information affects the next convolution layers. As the size of kernel bigger, the more parameters are required, which is computationally expensive. In practical, the 3x3 kernel is often used because it can reflect the full information even with smaller kernels as the network becomes deeper. The convolutional operation moves the window of the kernel at regular intervals. The amount of moving interval is called stride. The size of the output differs from the scale of the input by performing a convolution operation described above. Padding is used to adjust the output size.

3.1.2.2 Pooling layers

The purpose of pooling layers is to reduce the number of parameters and to control overfitting problem by lowering the spatial dimension[38]. For example, 2x2 max pooling selects the maximum value in a 2x2 space with a regular interval, resulting in the reduction of the space dimension in half. Similarly, the average pooling returns the average values of given space. Unlike other layers in convolutional neural networks, it does not have parameters to learn.

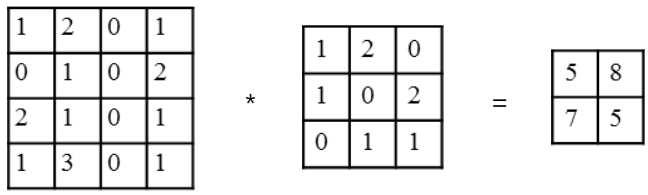


Figure 3.4. Convolution layers

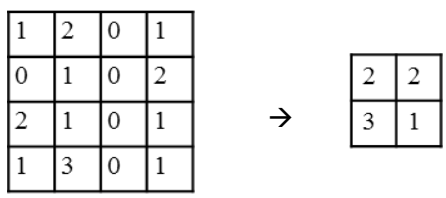


Figure 3.5. Pooling layers

3.1.3 Training Neural Networks

Training the neural network means exploring the optimized parameters for acquiring features. With optimized parameters, it is possible to obtain the proper features, and achieve the desired output from the networks.

3.1.3.1 Objective of training

The neural network explores the parameters based on the criterion that is called loss function. Loss function signifies the distance between the current output and the

desired output. As the output of the network approaches close to the desired output during training, the loss function decreases. Thus, the objective of training is minimizing the loss function. Mean squared error and cross entropy error are the most used loss function. The former is the average of a square of the distance between true label and output of the current network. It is used in the same sense to standard deviation. The latter is to find the distance of the probability distribution between the desired output and the output from the current network. It reduces the distance until the probability of the output of the current network converges to the probability of desired output.

3.1.3.2 Training method

Finding the appropriate parameters through the training of the network is a goal of the neural network. In training, how much to improve the parameter based on loss function is determined through gradient descent, and method to apply it to each parameter is called back propagation

3.1.3.2.1 Gradient descent

The parameter space is expansive and complex, making it difficult to find the optimal solution. Thus, it is impossible to find the global minimum analytically. Instead, an approximation method is used as alternatives to find the minimum. Even though there is a concern that the loss function may fall into local minima, it is acceptable if it is close enough to the global minima. Rather, there is an opinion that the probability of falling into the local minima is exceedingly rare[39]. One of the most used methods for minimizing the loss function is stochastic gradient descent method[40].

In stochastic gradient descent method, the loss function is to move a certain distance in the direction of slope at the present position and to make progress in the same approach until it converges to the minimum. Distance in gradient descent is determined by the slope of the position and learning rate that decides how much to learn in a single step. Even though stochastic gradient descent is straightforward and easy to implement, it is inefficient in some cases. Momentum and learning rate decay method is used to improve the inefficiency of gradient descent method[41, 42]. The idea of momentum comes from physics. When a ball rolls down a hill, it is getting faster if a ball keeps going down in the same direction. Momentum method also plays the same role in going down to the minimum point, meaning that the loss converges to the minimum faster. Learning rate decay is method reducing the learning rate as learning progresses. Adagrad is the typical method using learning rate decay. It helps the loss to converge more stable than gradient descent method. Adam method is a fusion of both techniques, converging to the minimum fast and steady[43].

3.1.3.2.2 Backpropagation

With gradient descent method, the information of how much parameters be updated is needed to update each parameter of the network. The process is very demanding and computationally expensive because of the chain of a differential equation. Backpropagation approaches these problems by using recursive application of the chain rule. With chain rule, calculation of back propagation is the linear problem. Thus, transmission of low derivative can lead vanishing gradient problem.

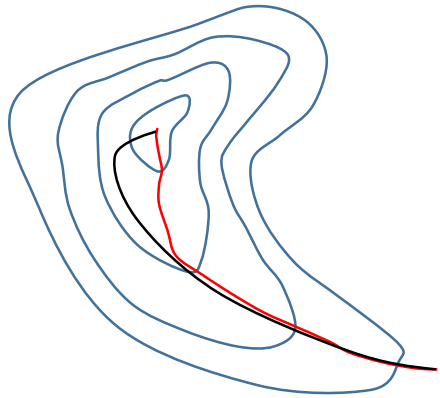


Figure 3.6. Comparison of gradient descent methods. The red line indicates how to find the minimum using stochastic gradient descent (SGD) method, and the black line refers ADAM optimizer. While SGD moves slightly to lower values each time the parameter is updated, ADAM has faster convergence with momentum and reducing learning rate as it approaches convergence point.

3.1.3.3 Improving the training performance

Several techniques improve the efficiency and the accuracy of the Neural network. Problems like overfitting or saddle point stagnation can be solved by selecting and using appropriate technique.

3.1.3.3.1 Regularization

Overfitting means that the model shows good performance only in the training data and not in the test data. It happens when the model has high dimensionality or not enough training data is prepared. When overfitting appears, the gap between training error and test error is large. L2 and L1 regularization method can be applied to

alleviate this problem[44]. These methods penalize the loss function, preventing the model to be fitted only well on training data. Dropout is also commonly used regularization technique[45]. It trains the network with only selected units of the network. The process forces the degree of freedom of neural networks down and makes a neural network to avoid overfitting.

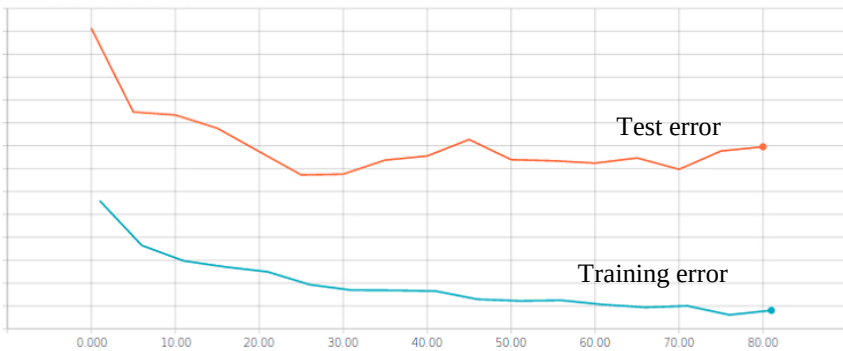


Figure 3.7. Training error and test error in case of overfitting. The gap between test and training error does not decrease during training the network.

3.1.3.3.2 Weight initialization

Assume that all parameters of the weights are initialized with zero. When the network updates the parameters, all the values will be updated with direction and distance. It makes the meaning of having multiple weights disappear. Therefore, to initialize parameter randomly is important. Xavier initialization, which disperses the weight randomly with proper distance is widely used these days[46]. In practical, using transfer learning is one of solution for weight initialization. Transfer learning means to transfer parameters of a pre-trained model trained on another data set[47]. With transfer learning, fine-tuning the parameters in a model can bring desirable

result.

3.1.3.3.3 Batch normalization

Batch normalization makes each layer expand the activation properly, leading the training performance to be increased[48]. There are three advantages in using batch normalization. First, it improves the training time. Second, dependency on weight initialization is lowered. At last, it prevents overfitting. Batch normalization is usually inserted between convolution layers and activation function layers.

3.1.3.3.4 Hyperparameter Optimization

Hyperparameter means the learning rate, the momentum, dropout rate, the number of the neuron every layer and L1 or L2 regularization. If hyperparameters are set poorly, the network may not perform well. There are three types of hyperparameter optimization methods. Grid search means searching the hyperparameter with evenly spaced. Random search explores the hyperparameters by picking random values in given range[49]. Bayesian optimization finds the optimal hyperparameters based on the Gaussian process[50]. Unlike the other tuning methods, Bayesian optimization statistically finds the next hyperparameters to search based on the loss function of the prior hyperparameters.

4. Methods

The process of estimation of ejection fraction consists of 2 parts: segmentation of the left ventricle, mapping the 2D image to 3D image. Segmentation method using the convolutional neural network is described in the first part. With segmented images, 2D images are transformed to 3D, and then ejection fraction is calculated, which is outlined in the second part.

4.1 The left ventricular segmentation

4.1.2 Neural network model

The left ventricle segmentation in echocardiography during chest compression for CPR is a difficult problem. It contains speckle noise, and image disappearing problem often occurs. Thus, three criteria that model must satisfy for stable segmentation of echocardiography during CPR were set, and a neural network was constructed based on these criteria.

4.1.2.1 Criteria for the neural network

The neural network structure is devised based on three criteria to segment the left ventricle in robust to disturbance and in real-time. First, the network must run in real-

time. Segnet is adopted as the base architecture to satisfy the first criterion[12]. Segnet is designed for real-time segmentation problems, so it is very efficient in both terms of memory and computational time. In terms of structure, Segnet consists of 13 layers of compressing path and corresponding 13 layers of an expanding path. Each 13 layers are identical to the vgg 16 network without fully connected layers. Second, the network must be robust to speckle noise that occurs a lot in ultrasound images. As a result of the test, segmentation performance was not good and versatile depending on the quality of the frame when using Segnet alone for the left ventricle segmentation. With using skip connection and dice coefficient, the performance of the segmentation is increased, and the second criterion is satisfied[51]. There are two types of skip connection. Long skip connections connect the layers of contracting path with the layers of expanding path, helping the network to recover spatial information loss. On the other hand, short skip connections connect the pair of layers in contracting path. It is usually used to prevent vanishing gradient problems in fully convolutional networks. As a loss function, dice coefficient is used. Dice coefficient is widely used in biomedical image segmentation instead of cross entropy. It signifies the harmonic of the summation of precision and recall[52]. The advantage of dice is that it does not depend heavily on hyperparameters[23]. Third, the network must estimate the left ventricle when the image quality is poor due to the contact problem between a transducer and the chest. Gated recurrent unit (GRU) that is a type of recurrent neural network is applied[13]. The memory of GRU synthesizes both the information of current frame and the previous frames, making the network to segment the left ventricle based on the change of it. Several studies show a model combining recurrent neural network and convolutional neural network, which

overcomes the limitations of the existing single CNN model.

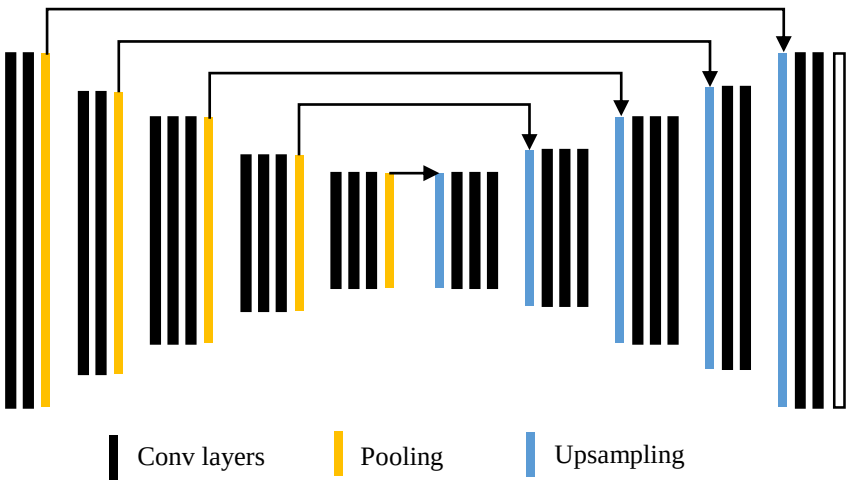


Figure 4.1. The structure of SegNet[12]. The input image enter the left end and come out from left end. The line in the figure indicate the indices memory for upsampling.

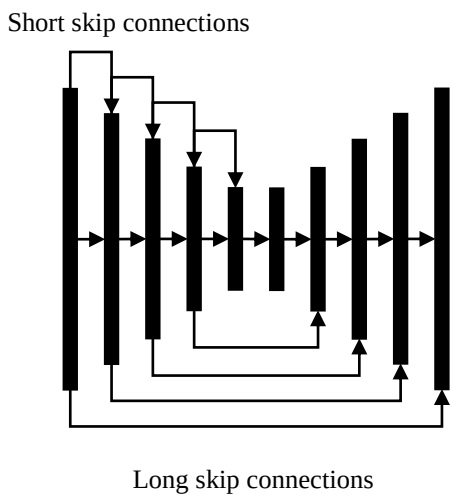


Figure 4.2 Skip connection. The images enter the left end and then come out from the right end.

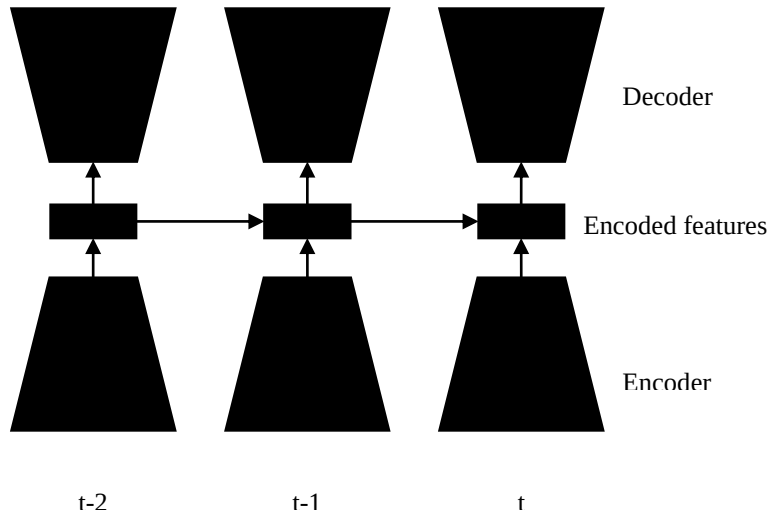


Figure 4.3 CNN combined with RNN. The connection between pairs of encoded features is same as the RNN structure.

4.1.2.2 Proposed neural network

The network is designed based on three criteria described above. It has 13 layers of encoding path, two layers of gated recurrent units, and 13 layers of decoding path.

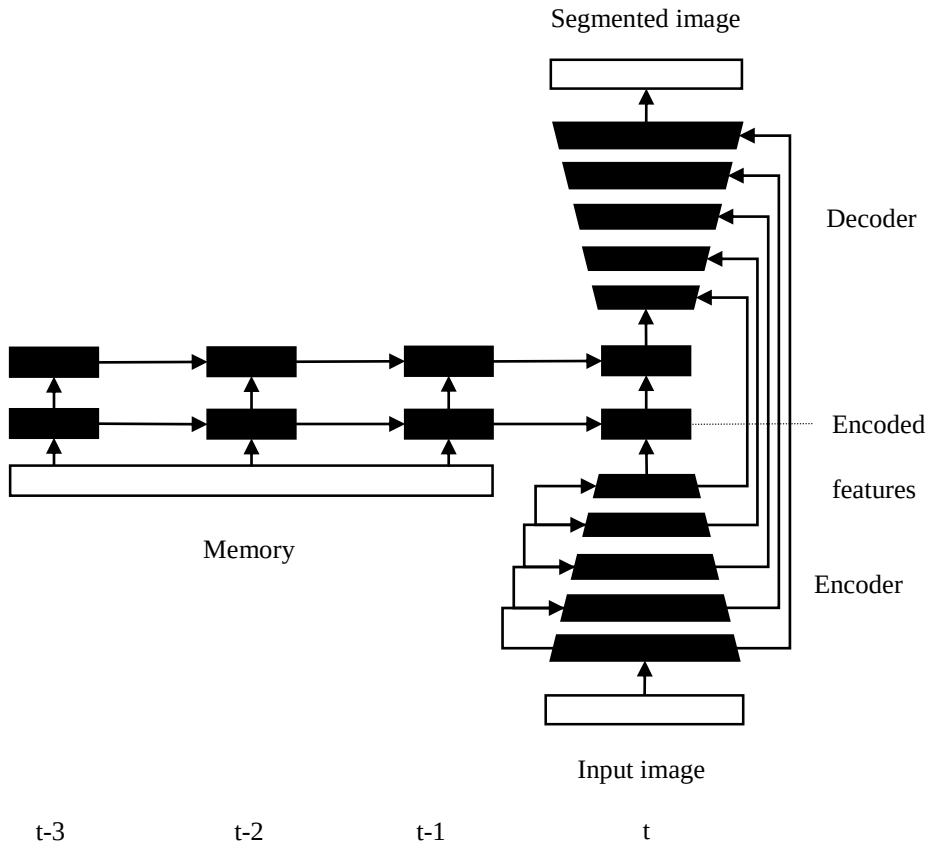


Figure 4.4 Overall structure of the proposed neural network

4.1.2.2.1 Encoding path

Encoding path is made up of 13 convolution layers and five pooling layers, which are a combination of these five sets. Two or three convolution layers and one pooling pair form each set. In convolution layers, 3x3 kernels with zero padding and stride one are applied, and Relu is used for activation function. In pooling layers, 2x2 max-pooling is adopted. At the end of each set, skip connections that are connected from the beginning of each set are placed. The detailed structure of the network is shown below.

-Layer1: a convolutional layer with 3x3x3x64 filters and 64biass

-Layer2: a convolutional layer with 3x3x64x64 filters and 64biass

-Pooling1: a pooling layer with strides 1x2x2x1 and zero-paddings

-Skip connection1: a short skip connection from input

-Layer3: a convolutional layer with 3x3x64x128 filters and 128biass

-Layer4: a convolutional layer with 3x3x128x128 filters and 128biass

-Pooling2: a pooling layer with strides 1x2x2x1 and zero-paddings

-Skip connection2: a short skip connection from pooling1

-Layer5: a convolutional layer with 3x3x128x256 filters and 256biass

-Layer6: a convolutional layer with 3x3x256x256 filters and 256biass

-Layer7: a convolutional layer with 3x3x256x256 filters and 256biass

-Pooling3: a pooling layer with strides 1x2x2x1 and zero-paddings

-Skip connection3: a short skip connection from pooling2

-Layer8: a convolutional layer with 3x3x256x512 filters and 512biass

-Layer9: a convolutional layer with 3x3x512x512 filters and 512biass

-Layer10: a convolutional layer with 3x3x512x512 filters and 512biass

-Pooling4: a pooling layer with strides 1x2x2x1 and zero-paddings

-Skip connection4: a short skip connection from pooling3

-Layer11: a convolutional layer with 3x3x512x512 filters and 512biass

-Layer12: a convolutional layer with 3x3x512x512 filters and 512biass

-Layer13: a convolutional layer with 3x3x512x512 filters and 512biass

-Pooling5: a pooling layer with strides 1x2x2x1 and zero-paddings

Table 4.1. Layers of encoding path.

4.1.2.2.2 Gated recurrent unit

Gated recurrent unit (GRU) learns time-based information through update gate and reset gate[13]. The update gate and the reset gate determine how much memory the past information will be put into and how much it will reflect the new information. GRU with a length of 3 and a double stacked layer is used in the network. The encoded features of the convolutional neural network will be fed into the GRU as input. GRU mixes information from the input and the memory in an appropriate combination and usesthis information as input to the decoding path.

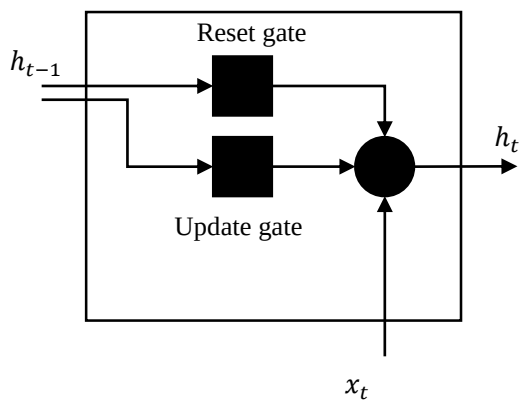


Figure 4.5. Structure of gated recurrent unit[53]

$$updata\ \ gate\colon\ \mathrm{z_t} = \sigma_g(W_z x_t + U_z H_{t-1} + b_z)$$

$$reset\ \ geate\colon\ \mathrm{r_t} = \sigma_g(W_r x_t + U_r H_{t-1} + b_r)$$

$$output\colon\ \mathrm{h_t} = z_t\ \circ\ h_{t-1} + (1 - z_t)\ \circ\ \sigma_h\ \ (W_h x_t + U_h\ \ (r_t\ \circ\ h_{t-1}) + b_h)$$

Table 4.2. Gated recurrent unit equation. z refers to update gate and r refer to reset gate. h is final output that passed through update gate and reset gate.

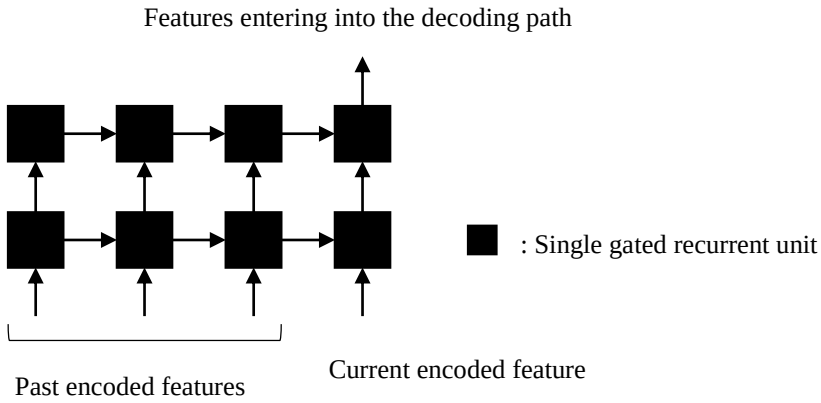


Figure 4.6. Layers of gated recurrent units. Gated recurrent units in the same level layer are identical each other.

4.1.2.2.2 Decoding path

A decoder with last 16 layers constructs a segmentation region from the encoded features. The overall structure the same as reversing the encoding layers. There are five sets of combination of convolution transpose layers and up-sampling layers. All convolution transpose layers adopt 3x3 kernels, and Relu is applied as activation function subsequently. Umsmpling layers, in turn, increase the output resolution to express the segmentation region corresponding to original images. Because the layers are expanded without additional information of the parameters, information loss appears in the up-sampling layer. Long skip connection is introduced to solve the problem by referring to the information of the encoder. Long skip connection is placed between every set of the decoder that is connected from the corresponding location of the encoder. Lastly, the final output of the decoder is fed into the sigmoid function that classifies the label of each pixel; region of interest(1) or otherwise(0).

-Layer1: a convolutional transpose layer with 3x3x512x512 filters and 512biass

-Layer2: a convolutional layer with 3x3x512x512 filters and 512biass

-Layer3: a convolutional layer with 3x3x512x512 filters and 512biass

-Upsampling1: a upsampling layer with strides 1x2x2x1 and zero-paddings

-Skip connection1: a long skip connection from pooling 4

-Layer4: a convolutional layer with 3x3x512x512 filters and 512biass

-Layer5: a convolutional layer with 3x3x512x512 filters and 512biass

-Layer6: a convolutional layer with 3x3x512x256 filters and 256biass

-Upsampling2: a upsampling layer with strides 1x2x2x1 and zero-paddings

-Skip connection2: a long skip connection from pooling 3

-Layer7: a convolutional layer with 3x3x256x256 filters and 256biass

-Layer8: a convolutional layer with 3x3x256x256 filters and 256biass

-Layer9: a convolutional layer with 3x3x256x128 filters and 128biass

-Upsampling3: a upsampling layer with strides 1x2x2x1 and zero-paddings

-Skip connection3: a long skip connection from pooling 2

-Layer10: a convolutional layer with 3x3x128x128 filters and 128biass

-Layer11: a convolutional layer with 3x3x128x64 filters and 64biass

-Upsampling4: a upsampling layer with strides 1x2x2x1 and zero-paddings

-Skip connection4: a long skip connection from pooling 1

-Layer12: a convolutional layer with 3x3x64x64 filters and 64biass

-Layer13: a convolutional layer with 3x3x64x3 filters and 3biass

Table 4.3. Layers of decoding path

4.2 The left ventricle model

There are several models for approximating the volume of the left ventricle of the normal heart as introduced in chapter 3. However, those models cannot be applied to the heart during CPR. The heart during cardiopulmonary resuscitation, unlike the normal heart, is not contracted isometrically but distorted in shape. Therefore, another model for the left ventricle that can be applied to CPR condition is required.

4.2.1 Constraint for model during CPR

The left ventricle volume change should be estimated with an image of one layer of the left ventricle to analyze it in real-time through echocardiography. The larger the number of figures that make up the model, the more images of the left ventricular layers is needed[11]. Therefore, a single-figure model that requires only one layer of the left ventricular image is applied, and an ellipsoid model with the same length of the minor axis and the intermediate axis for the diastole is selected. Then, the surface area of both compressed heart and heart at diastole are same. In the case of a normal heart, the surface area of the heart changes continuously due to the contractility of the heart. However, in the case of the CPR, the surface area is constant because the heart is not capable of contracting.

4.1.3 Training

Initial weights are taken with vgg16 trained in the Image-net dataset[37]. Adam optimizer is used by the optimizer. Each layer is batch normalized. The Dice coefficient is employed as loss function, which is the inverse number of the mean of summation of precision and recall[52]. Thus, the higher the recall and precision, the

smaller dice coefficient. The parameters of Adam optimizer, learning rate, and weight decay are selected as hyper-parameters.

$$DICE = \frac{1}{\frac{1}{Precision} + \frac{1}{Recall}}$$

$$\left(Precision = \frac{TP}{TN + TP}, Recall = \frac{TP}{TP + FP}\right)$$

4.2.2 Mathematical model

The ejection fraction (EF) is calculated as follow. It indicates how much blood is pumped to the body in one cycle of the heart. There are three variables required to obtain the left ventricular volume if it is applied to the single-ellipsoid model. Thus six variables need to know for estimation of EF.

$$EF = \frac{diastolic\ \ voume - systolic\ \ volume}{diastolic\ \ volume} = \frac{V_d - V_s}{V_d} = 1 - \frac{V_s}{V_d}$$

$$1 - \frac{V_s}{V_d} = 1 - \frac{\frac{4}{3}\pi x_s y_s z_s}{\frac{4}{3}\pi x_d y_d z_d} = f(x_s, y_s, z_s, x_d, y_d, z_d)$$

Four variables of the equation can be acquired from the echocardiography at diastole and compressed heart. The application of the ellipsoid model with the same length along with minor and intermediate and the constraint on the preservation of the surface area of the left ventricle reduces the variables that need to know to four.

$Constraints1$:

$$x_d \cong z_d$$

$Constraints2$:

$$4\pi\left(\frac{(x_s y_s)^{1.6} + (y_s z_s)^{1.6} + (z_s x_s)^{1.6}}{3}\right)^{\frac{1}{1.6}} \cong 4\pi\left(\frac{(x_d y_d)^{1.6} + (y_d z_d)^{1.6} + (z_d y_d)^{1.6}}{3}\right)^{\frac{1}{1.6}}$$

$$(x_s y_s)^{1.6} + (y_s z_s)^{1.6} + (z_s x_s)^{1.6} = (x_d y_d)^{1.6} + (y_d x_d)^{1.6} + (x_d)^{3.2}$$

$$Let, x^{1.6} = X, y^{1.6} = Y, z^{1.6} = Z$$

$$X_s Y_s + Z_s(Y_d + X_d) = X_d(X_d + 2Y_d)$$

$$Z_s = \frac{X_d(X_d + 2Y_d) - X_s Y_s}{X_d + Y_d}$$

$$z_s = \left(\frac{x_d^{1.6}(x_d^{1.6} + 2y_d^{1.6}) - x_s^{1.6} y_s^{1.6}}{x_d^{1.6} + y_d^{1.6}}\right)^{\frac{1}{1.6}}$$

$$EF = 1 - \frac{\frac{4}{3}\pi x_s y_s z_s}{\frac{4}{3}\pi x_d y_d z_d} = 1 - \left(\frac{x_s y_s}{x_d^2 y_d}\right)\left(\frac{x_d^{1.6}(x_d^{1.6} + 2y_d^{1.6}) - x_s^{1.6} y_s^{1.6}}{x_d^{1.6} + y_d^{1.6}}\right)^{\frac{1}{1.6}} = f(x_s, y_s, x_d, y_d)$$

5. Experiment result

5.1 Experiment method

5.1.1 Left ventricle segmentation

5.1.1.1 Data acquisition

The training and test images are obtained from 15 pigs which are receiving cardiopulmonary resuscitation (CPR). All pigs are female, and the weight range is from 30kg to 40kg. CPR is performed 100 times per minutes. The transducer is placed below the chest compression position, and the four chamber view is obtained. The size of the images acquired from echocardiography is 636 x 420. Because of the unstable contact between the chest and the transducer, there is a big variation between images. Ground truths for segmentation of each image are labeled manually by a cardiologist for supervised learning. Seg3D which is a software purposing of segmentation of Dicom images is used for manual segmentation works[54]. Additionally, the adjustment of brightness and contrast are applied to segment the left ventricle in case the image is blurry or difficult to segment correctly on the original image.

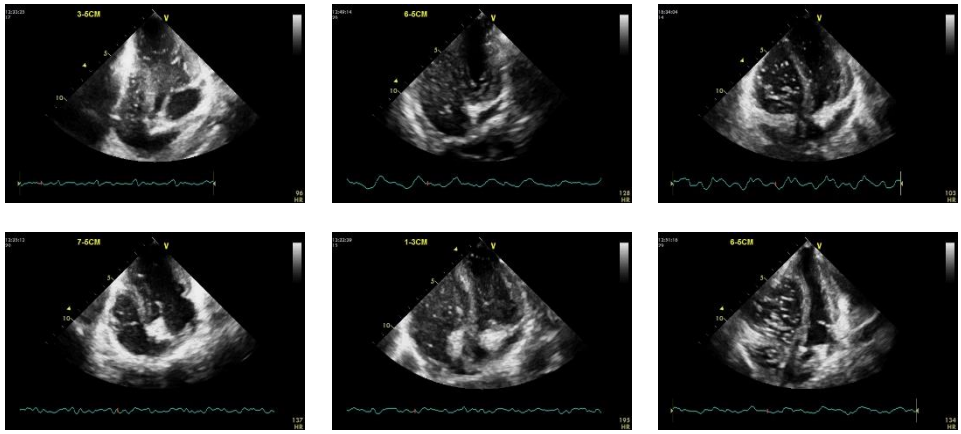


Figure 5.1. Acquired echocardiography from pigs.

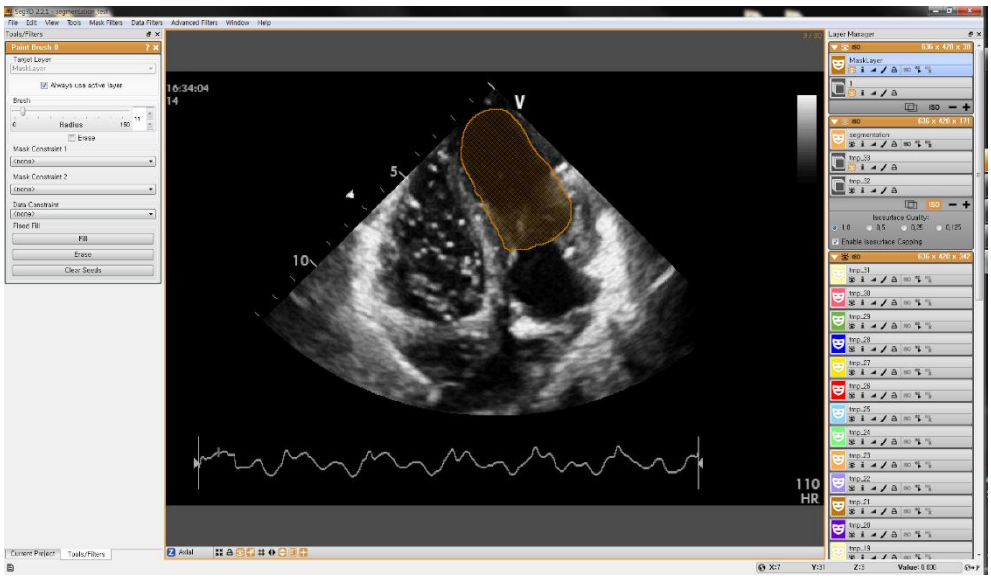


Figure 5.2. Seg3d for annotating ground truth.

5.1.1.2 Pre-processing

Data augmentation with speckle noise is applied to improve accuracy and prevent the overfitting. It allows the network to learn invariance to noise. After that, normalization method is applied to filtered image to make it zero-centered. Images are resized to 224 x 224 to match the input size to the network.

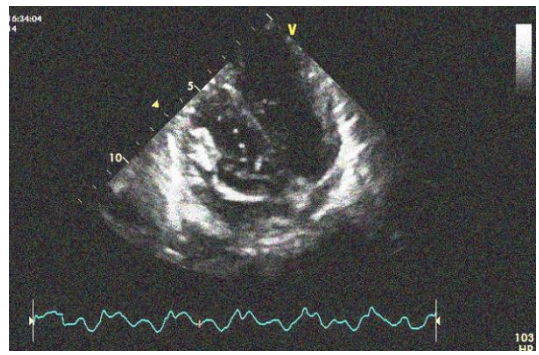


Figure 5.3. Data augmentation

5.1.1.3 Training

All 15 videos are split into training, validation, and test sets at a ratio of 3:1:1. Training set and validation set are used to determine the hyperparameters, and the test set is used to obtain the final accuracy. Models are evaluated with K-fold cross validation[55]. Bayesian optimization is adopted as the hyperparameter search method[50]. Bayesian Optimization is a method of gradually narrowing down the range in which an optimal hyperparameter can exist, taking into consideration the influence of observed hyperparameters on the result. Learning rate, momentum parameters are chosen as the hyperparameter, and the range of each of them is set as follow: learning rate: 0.001-0.0005, beta1 of Adam optimizer (momentum):0.85-0.95, beta2 of Adam optimizer: 0.985-0.995. The other hyperparameters for regularization are not applied because batch normalization is already used. The segmentation based on proposed model is compared to active geodesic contour that is considered as the fastest with high accuracy segmentation method. All programs are run in Intel i5, Titan x Pascal hardware environment, and as a software environment, Tensorflow that is for a framework of designing neural network architecture is used based on python[56].

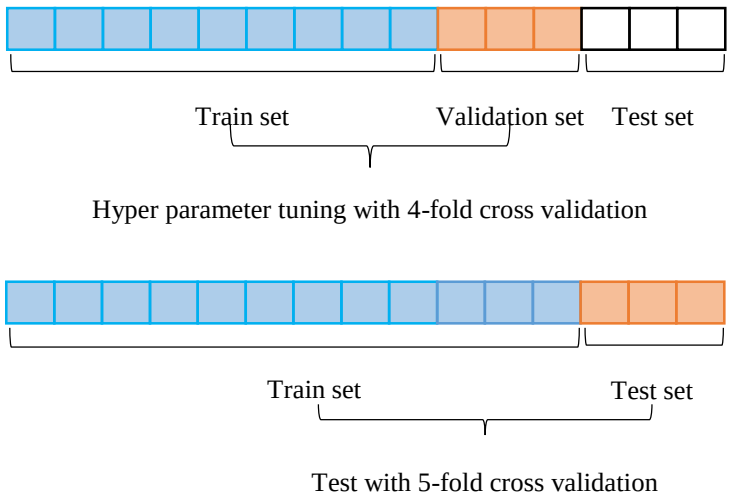


Figure 5.4. Model evaluation method. Hyperparameter is tuned with 4-fold cross validation. After that, the accuracy of model is tested with 5-fold cross validation.

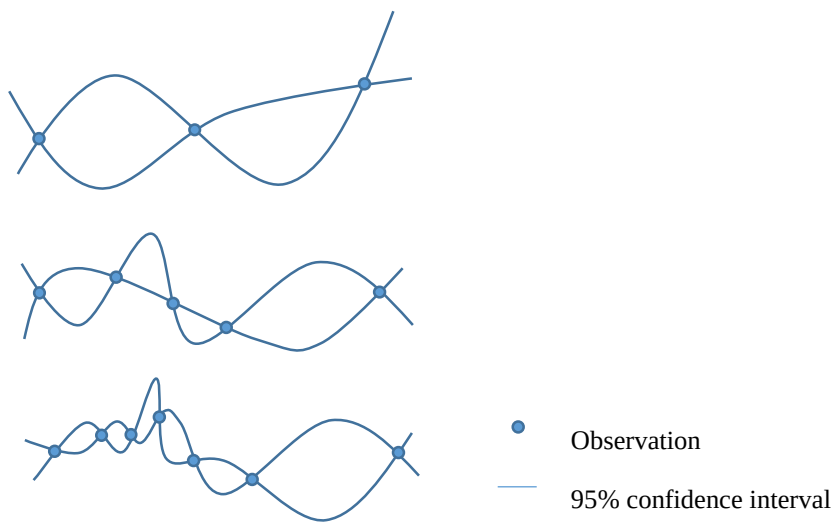


Figure 5.5. Bayesian Optimization. The x-axis is set of hyperparameters, and the y-axis is the loss followed by a set of hyperparameters. From top to bottom, each of them has three, five, and seven observation. Each observation is based on previous observation. As observation increases, the range of confidence interval converges

5.1.2 Left ventricle modeling

The segmented area from the echocardiography is fitted with an ellipse, and then the length of short and long axes of the ellipse is fitted to sine wave. The peak and the valley values of the wave are applied to the left ventricular model to estimate the volume. The estimated volume variation of the left ventricle is compared to end-tidal carbon dioxide tension (ETCO2)[57-60]. ETCO2 refers the partial pressure at the end of an exhaled breath. As the carbon dioxide transfer is affected by the blood flow, it can show how much is cardiac output directly. According to studies on the correlation between cardiac output and ETCO2.

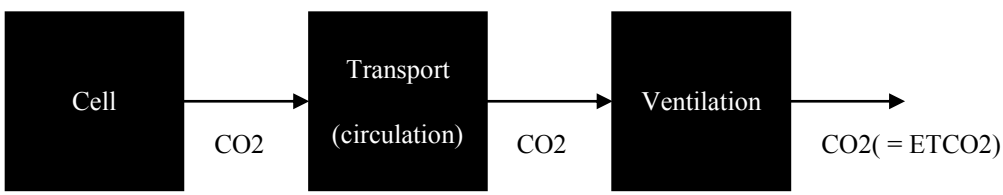


Figure 5.6. Physiological relation between end-tidal CO2 and circulation.

5.2 Result

5.1.1 Left Ventricle Segmentation

5.1.1.1 Hyperparameters

The hyperparameters are tuned to 0.0016131, 0.86210, and 0.99051 using Bayesian optimization. The number of the retrieved points is 20 in the specified range.

5.1.1.2 Correlation comparison

Correlation comparison to geodesic active contour method, which is considered as

the best performing model for real-time left ventricular segmentation is conducted. The left ventricular segmentation based on convolutional neural network based studies are not a real-time purpose, so no comparative study is conducted. The accuracy term is defined as dice.

	Geodesic ACM	Proposed model
Dice	0.726	0.947
Processing speed	0.41s/frame	0.026s/frame

Table 5.1. Correlation of segmentation methods.

5.1.2 LV modeling

5.1.2.1 Correlation to etco2

Correlations with golden standard values show a correlation of 0.392 for the conventional method, while 0.736 for the proposed model. P-value of both of them is 0.13 and 0.04, respectively. From the correlation and p-value result, it is reasonable to reject the existing model and select the current one.

	Previous method	Proposed method
correlation	0.392	0.736
p-value	0.13	0.04

Table 5.2. Correlation of left ventricle models

5.3 Result analysis

5.2.1 Segmentation

5.2.1.1 Hyperparameter Optimization

Every hyperparameter candidate is tested with 4-fold cross validation. With Bayesian

Optimization, 20 candidates of hyperparameters are compared.

	Learning rate	Beta1(Adam)	Beta2(Adam)
1	0.0019222	0.85176	0.98656
2	0.0017196	0.88884	0.99050
3	0.0010605	0.85117	0.99257
4	0.0018753	0.89413	0.98909
5	0.0010697	0.93632	0.99151
6	0.0014236	0.93531	0.98694
7	0.0014597	0.87273	0.99147
8	0.0013527	0.89425	0.98887
9	0.0010129	0.88969	0.99339
10	0.0019930	0.86229	0.99008
11	0.0018632	0.85392	0.98860
12	0.0019693	0.90194	0.98553
13	0.0010033	0.89897	0.98889
14	0.0018282	0.88975	0.99380
15	0.0014012	0.85264	0.98594
16	0.0014126	0.87984	0.98931
17	0.0017680	0.86977	0.99054
18	0.0018231	0.86197	0.99203
19	0.0016504	0.87754	0.98914
20	0.0016131	0.86210	0.99051

Table 5.3. Hyperparameter candidates selected by Bayesian Optimization

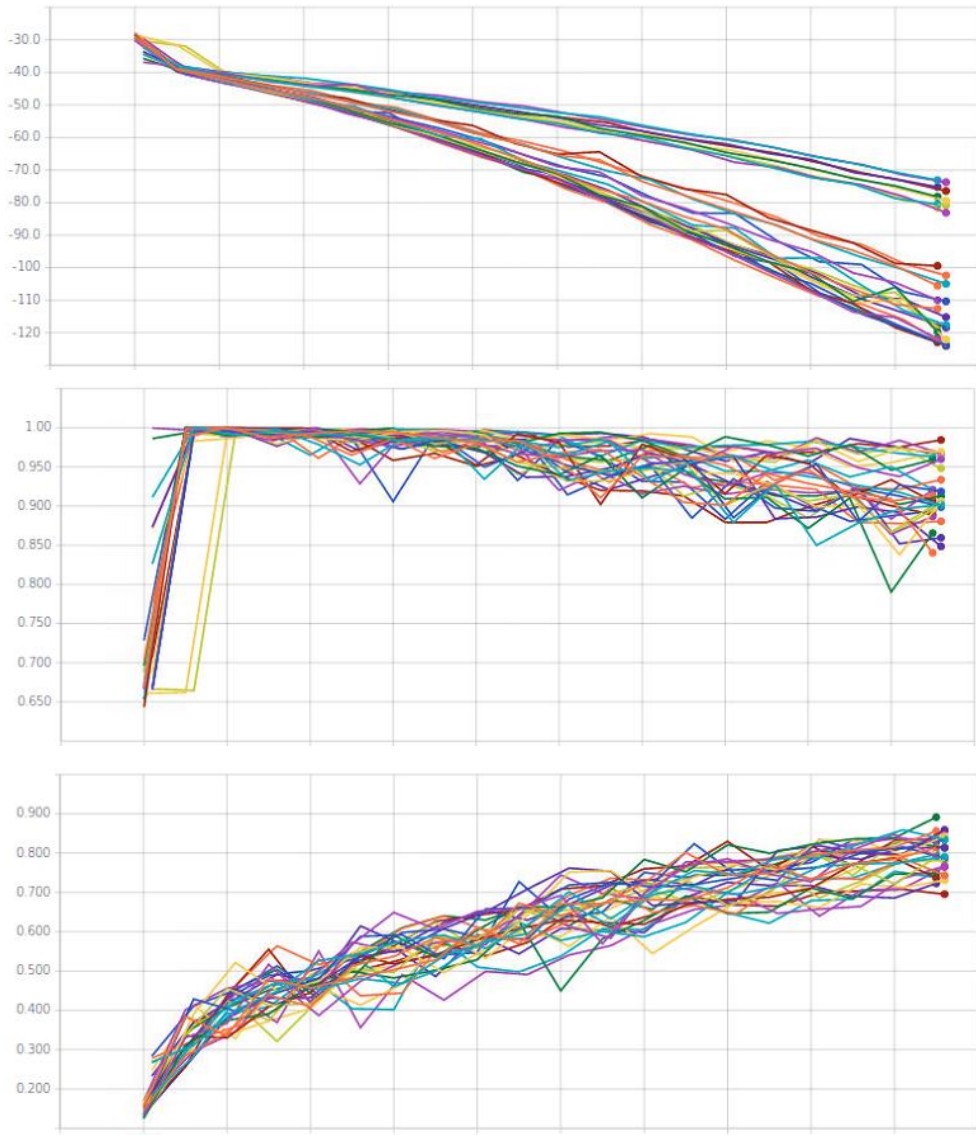


Figure 5.7. Loss, precision and recall change by the hyperparameters during training. The x-axis of the graphs is training epoch and the y-axis of the graphs loss, precision and recall from top to bottom. Those show how the network is trained depending on different hyperparameter set.

5.2.1.2 Train-test comparison

In the cross-validation test, the dice of train set are 0.978, 0.983, 0.965, 0.971 and 0.988, respectively, and the dice of the test set are 0.943, 0.959, 0.939, 0.944, and 0.950, respectively. Training seems to be well done without overfitting because the gap between the train set and the test set is small. On the other hand, the dice of the active geodesic contour is below 0.80 in most case and falls to 0.6 in the case of the low-quality image caused by poor contact between the chest and the transducer. It is much lower than the results of the reference literature because echocardiography during CPR appears to be poor quality and image acquisition instability compared with echocardiography in the absence of CPR.

	1	2	3	4	5	Average
Train	0.978	0.983	0.965	0.971	0.988	0.977
Test	0.943	0.959	0.939	0.944	0.950	0.947

Table 5.4. Five-fold cross validation for test

	1	2	3	4	5	6	7	8
dice	0.817	0.742	0.712	0.693	0.571	0.752	0.705	0.581
	9	10	11	12	13	14	15	Average
dice	0.782	0.775	0.794	0.824	0.732	0.743	0.664	0.726

Table 5.5. Dice of geodesic active contour for 15 videos. Each dice is the mean value of 30 frames of a video

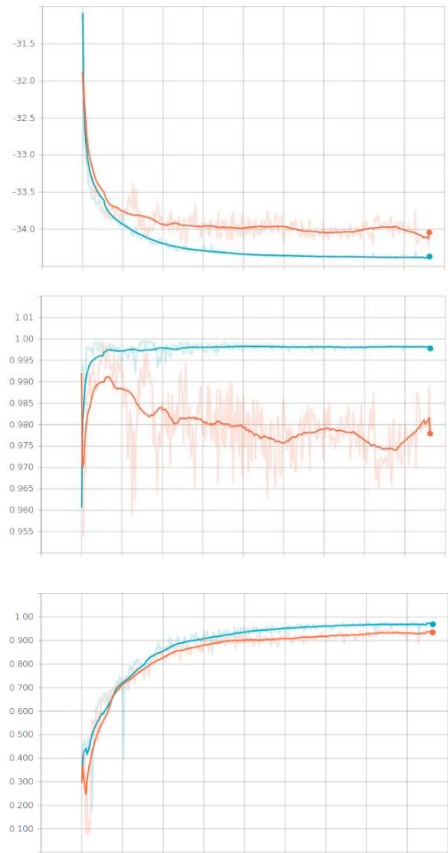


Figure 5.8. Loss, precision, and recall by epoch. They are a training graph of the hyperparameter-tuned neural network. Each graph is about the loss, precision, and recall from top to bottom. The Orange line indicates the training data and blue line means the test data.

5.2.1.3 Processing speed

As a result of the processing speed experiment, the processing time per frame is between 0.02(s) and 0.03(s), and 0.026(s) in average. It means the network can process over 30frames in a second, which is enough to perform in real-time. By contrast, active geodesic contour took 0.41(s) to analysis one frame, meaning that it can process only two frames per second.

5.2.1.3 Image comparison

Proposed model tend to segment the left ventricle closer to the ground truth compared to geodesic active contour method. Geodesic ACM segmented well in clear echocardiography, however, when the inner wall is captured in the echocardiography, it could not distinguish inner wall. Moreover, when the echocardiography becomes blur due to unstable contact, geodesic ACM failed segmentation properly.

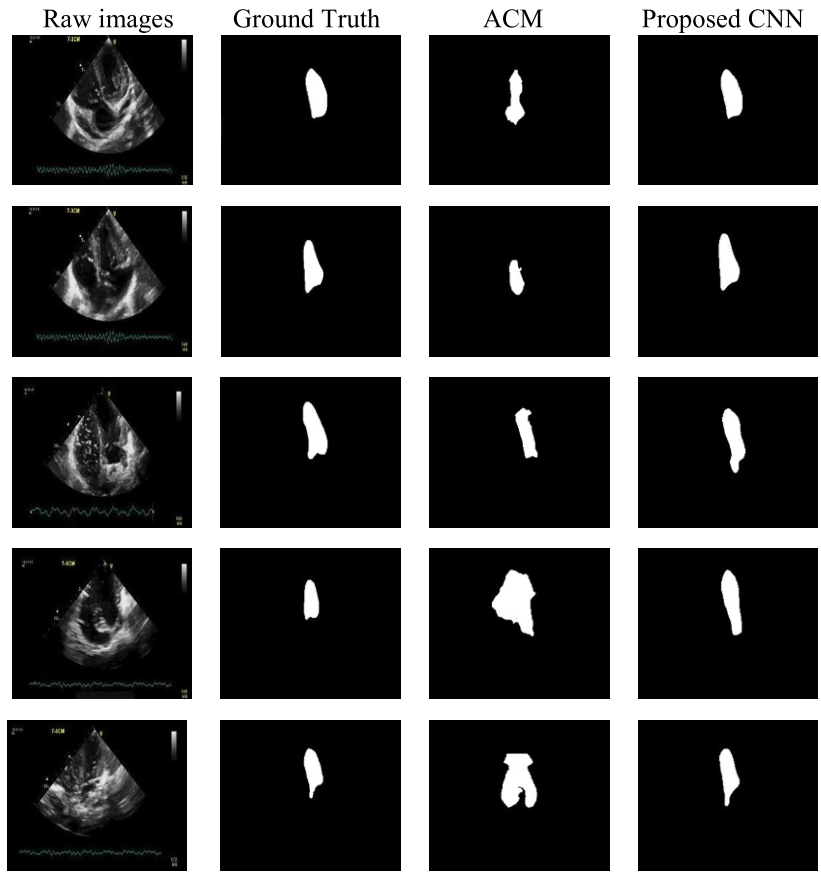
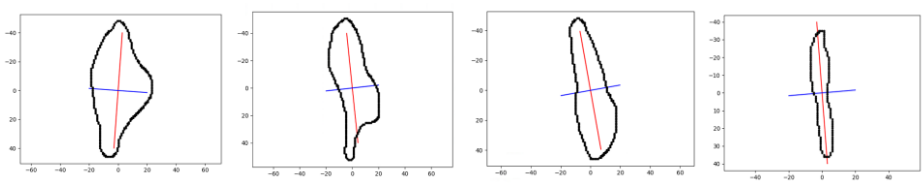


Figure 5.9. Segmented image comparison to the ground truth. From the left to right column: raw image, ground truth, geodesic active contour method(ACM), proposed convolutional neural network(CNN).

5.2.2 LV modeling

5.2.2.1 Ellipse fitting

After the border of the segmented area is extracted, the fitting to the ellipse is made by principle component analysis (PCA)[61]. The mean value of the boundary becomes the center of the ellipse. Following, the major axis and the minor axis of the PCA is same as the long and short axis of the ellipse. The root square of eigenvalues of PCA is mapped to the length of axes. After that, the length of long and short axes of the ellipse is fitted to a sine wave to find the end diastole point and the end compressed point. Both of point is put into the left ventricular model to estimate the volume variation.



Short axis: 19.07	Short axis: 15.5	Short axis: 14.6	Short axis: 6.4
Long axis: 38.9	Long axis: 42	Long axis: 37.5	Long axis: 29

Figure 5.10. Examples of ellipse fitting

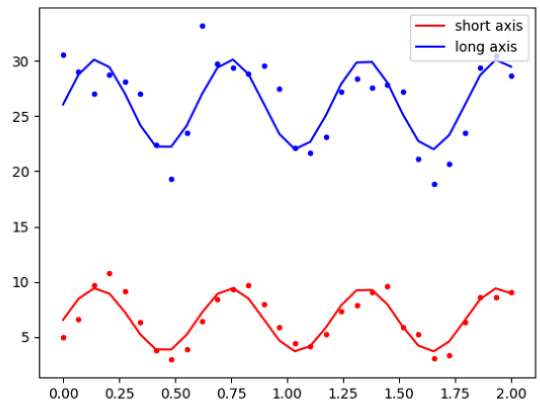


Figure 5.11. Sine wave of the length of short and long axes. The x-axis is the time(s), and the y-axis is the relative length of short and long axes of left ventricle segmented from echocardiography.

5.2.2.2 Three-dimensional volume estimation

Since no machine could measure the cardiac output during CPR, the correlation is indirectly compared using ETCO2, which is frequently used bio-signal during CPR[62, 63]. According to research conducted by Joseph P Ornato, there is a logarithmic correlation between etCO2 and cardiac output[60].

$$\mathrm{ETCO}_2(\%) = 2.49 * \log(CO) + 3.23$$

$$\mathrm{ETCO}_2(mmhg) = 18.675 * \log(EF * HR * EDV) + 24.225 \;\; at \;\; atm$$

Thus, the estimated volume variation of the left ventricle using the model is compared to ETCO2 to evaluate the model. The correlation ratio of proposed model is 0.736, while the previous model shows the correlation ratio of 0.392.

	Etco2(mmhg)	EF(previous)	EF(suggested)
0	37.87	0.5211	0.2128
1	37.45	0.5361	0.2436
2	32.33	0.4550	0.1759
3	31.36	0.6416	0.3317
4	26.69	0.4243	0.1004
5	34.01	0.4557	0.2111
6	32.09	0.4518	0.1590
7	26.88	0.3905	0.1437
8	28.31	0.4649	0.1497
9	26.50	0.3416	0.1356
10	15.26	0.2064	0.0719
11	18.57	0.3060	0.1016
12	36.45	0.5237	0.2304
13	42.68	0.6809	0.3554
14	42.99	0.894	0.607

Table 5.6. End-tial CO2(ETCO2) and ejection fraction(EF) calculated by models of each video. ETCO2 is measured, and EF is estimated from the echocardiography during cardiopulmonary resuscitation.

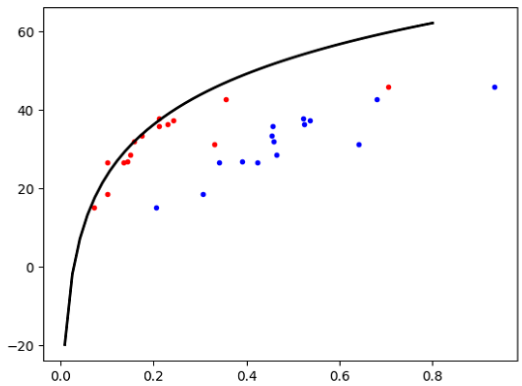


Figure 5.12. Relation between EF from the models and ETCO2. The x-axis is ejection fraction, and the y-axis refers to ETCO2. There is a strong relationship between EF calculated by the suggested model and equation devised by Joseph P Ornato. On the other hand, there is a less correlation for the previous model.

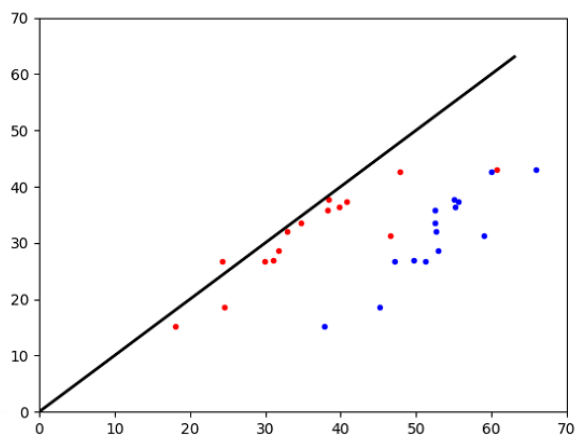


Figure 5.13. Correlation between acquired ETCO2 and estimation from the model. The x-axis is estimated ETCO2 from the ejection fraction acquired from the echocardiography. The y-axis is measured ETCO2. Correlation between red dots and the black line is 0.73 with a p-value of 0.04. In another hand, the blue dots show a correlation of 0.392 and p-value of 0.13.

6. Discussion

6.1 Left ventricle segmentation

With the suggested algorithm, the real-time stable segmentation based on echocardiography is achieved. The suggested algorithm shows higher accuracy as well as fast processing time compared to geodesic ACM that is considered to have the highest performance in real-time segmentation based on echocardiography. However, there are several technical and methodological limitations in suggested model. First, the region of the interest in the raw images is a tiny part. The crop of the ROI or background removal is applied to increase the accuracy in case of this. If attention algorithm or background removal is applied ahead of the segmentation neural network, the accuracy of the network probably increases. Second, the data augmentation may not be enough. Even though there is no exact standard for data augmentation, there is a need to compare the results by applying more data augmentation techniques. Third, the number of layers inserted between encoding path and decoding path is not optimized. The accuracy of the segmentation may be different by the depth and the time series layers of GRU. Thus, the number of the nodes of the neural network is also considered as the hyper-parameters. To find the optimal size of the GRU, the depth and the time series length should have included

in hyper-parameters. Fourth, the iteration of Bayesian optimization is not enough to search the optimal hyperparameters. In usual, the iteration is determined if the updated hyper-parameters are converged toward specific value. Fifth, because all the data in the test set and training set are recorded by same echocardiography machine, the variety of the data may not be enough to learn the unique characteristic. It is expected to obtain improved result with solving the problems described above.

6.2 Left ventricle model

The left ventricular model applicable to CPR condition for the condition of CPR is suggested with the single figure model. With the constraints for the condition of CPR, suggested model shows a higher correlation to the ETCO2 which has a specific relation to cardiac output compared to the previous model. According to the reference in chapter two, many types of the single ellipsoid is adopted. However, the ellipsoid that has the same length along with the intermediate and minor axis is used as the dilated left ventricle model in this thesis. If the different ratio between intermediate and minor axes are tested, the ratio with higher correlation can be found.

6.3 Combining segmentation and 3D transformation

6.3.1 Comparison method

In segmentation problem, Active contour method and Neural network based method are compared. In the left ventricular three-dimensional conversion problem, a conventional left ventricular model that assumes the left ventricle contracts isometric and CPR applied left ventricular model are compared. From now on, the ejection fraction of the left ventricle derived from the echocardiography is compared with the

result this thesis. It contains two steps. First, the left ventricle area in echocardiography is segmented with two segmentation models: ACM method, CNN method. Second, the left ventricular volume variation is estimated by applying the segmented area to two types of three-dimensional conversion model.

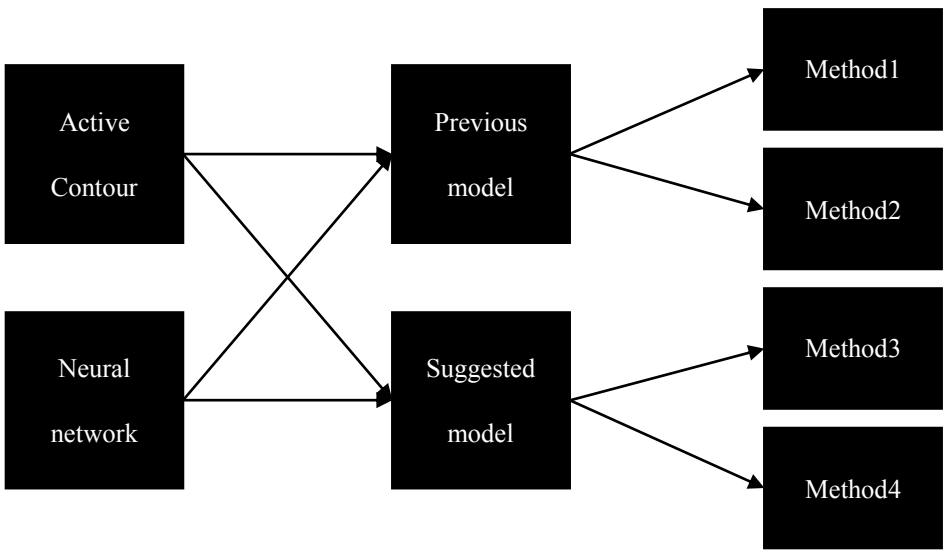


Figure 6.1. Four methods combining the segmentation and left ventricular model

6.3.1 Comparison result

Method 1 shows that correlation of -0.00786 and p-value of 0.98. It means two terms do not have specific relation and it does not fit the model. Method 2 shows that correlation of -0.00673 and p-value of 0.98, which is similar to Method1. The correlation and p-value of method 3 are 0.370 and 0.15, respectively. It confirms that there is little correlation between the measured ETCO2 and estimated ETCO2 from the echocardiography. The result of method 4 shows higher correlation and lower

significance compared to other methods. The correlation is 0.727, and the p-value is 0.05. Thus, method four can be said to explain the model enough.

Segmentation	ACM		Suggested CNN		ETCO2
LV model	Original	Proposed	Original	Proposed	(mmHg)
0	0.863	0.569	0.497	0.216	37.87
1	0.578	0.249	0.550	0.279	37.45
2	0.834	0.606	0.387	0.158	32.33
3	0.813	0.559	0.498	0.237	31.36
4	0.471	0.254	0.466	0.171	26.69
5	0.329	0.117	0.426	0.205	34.01
6	0.439	0.212	0.223	0.050	32.09
7	0.300	0.110	0.322	0.112	26.88
8	0.497	0.262	0.468	0.194	28.31
9	0.548	0.310	0.415	0.161	26.50
10	0.541	0.298	0.369	0.122	15.26
11	0.591	0.330	0.194	0.079	18.57
12	0.764	0.505	0.493	0.248	36.45
13	0.547	0.305	0.674	0.386	42.68
14	0.262	0.106	0.833	0.618	42.99

Table 6.1. The correlation of the result of four methods. ACM refers Active contour method and CNN indicates convolutional neural network.

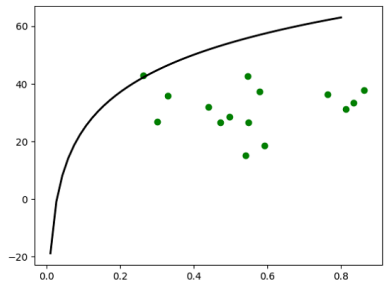
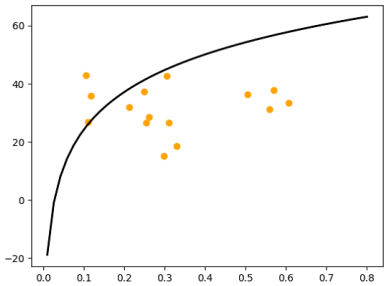
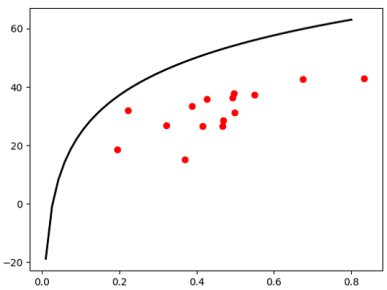
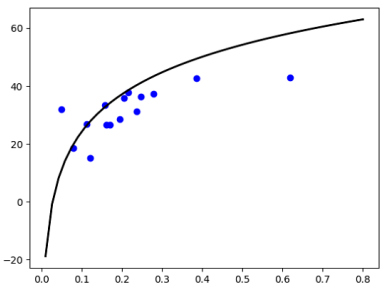
LV model / Segmentation	Previous method	Suggested method
ACM	Method1	Method2
		
	correlation: -0.00786 p-value: 0.98	Correlation: -0.00673 p-value: 0.98
CNN	Method3	Method4
		
	correlation: 0.370 p-value: 0.15	Correlation: 0.727 p-value: 0.05

Table 6.2. Graph of the four types of method. The x-axis of each graph is ejection fraction, and the y-axis is end-tidal CO2. Method 4 that use suggested neural network and suggested left ventricular model shows higher correlation and lower significance compared to other methods.

7. Conclusion

Two tasks for estimation of volume variation in real-time during cardiopulmonary resuscitation are performed. First, a segmentation of the left ventricle of the heart in 4 chamber echocardiography is conducted in real-time using the convolutional neural network. Moreover, the model for three-dimensional mapping based on the two-dimensional segmented region during cardiopulmonary resuscitation is carried out. The neural network consists of 13 layers of encoding path, 13 layers of decoding path, and gated recurrent unit that is located between encoding layers and decoding layers. Overall network is build based on SegNet. Skip connection and dice coefficient are applied to improve the accuracy of echocardiography. Gated recurrent unit is used to reflect the time base information. Instead of initializing the weight randomly, transfer learning from the model trained on image-net is adopted. Second, with Bayesian optimization, hyperparameters including initial learning rate and parameters of Adam optimizer are tuned. The model shows a correlation of 0.97 to the ground truth, which is much higher than the correlation of the previous one: 0.75. Based on segmented area, the left ventricle model for estimating the volume variation during cardiopulmonary resuscitation is suggested. Considering the arrest-heart conditions, constraints of the unchanged size and same length along with

intermediate and short axes are applied. In the correlation test with ETCO2 that reflects the cardiac output indirectly, there is a higher correlation than the existing model: 0.95 for CPR adapted model, 0.5 for the traditional model.

References

[1] J. Shin, J. E. Rhee, and K. Kim, "Is the inter-nipple line the correct hand position for effective chest compression in adult cardiopulmonary resuscitation?," *Resuscitation,* vol. 75, no. 2, pp. 305-310, 2007.

[2] K. C. Cha *et al.*, "Optimal position for external chest compression during cardiopulmonary resuscitation: an analysis based on chest CT in patients resuscitated from cardiac arrest," *Emergency Medicine Journal,* pp. emermed-2012-201556, 2012.

[3] A. Pickard, M. Darby, and J. Soar, "Radiological assessment of the adult chest: implications for chest compressions," *Resuscitation,* vol. 71, no. 3, pp. 387-390, 2006.

[4] J. Jung *et al.*, "Application of Robot Manipulator for Cardiopulmonary Resuscitation," in *International Symposium on Experimental Robotics*, 2016, pp. 266-274: Springer.

[5] S. M. Pogwizd, K. Schlotthauer, L. Li, W. Yuan, and D. M. Bers, "Arrhythmogenesis and contractile dysfunction in heart failure," *Circulation research,* vol. 88, no. 11, pp. 1159-1167, 2001.

[6] D. Marr and E. Hildreth, "Theory of edge detection," *Proceedings of the Royal Society of London B: Biological Sciences,* vol. 207, no. 1167, pp. 187-217, 1980.

[7] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE*

transactions on systems, man, and cybernetics, vol. 9, no. 1, pp. 62-66, 1979.

[8] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models," *International journal of computer vision,* vol. 1, no. 4, pp. 321-331, 1988.

[9] T. F. Cootes, C. J. Taylor, D. H. Cooper, and J. Graham, "Active shape models-their training and application," *Computer vision and image understanding,* vol. 61, no. 1, pp. 38-59, 1995.

[10] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, "Semantic image segmentation with deep convolutional nets and fully connected crfs," *arXiv preprint arXiv:1412.7062,* 2014.

[11] A. E. Weyman, *Cross-sectional echocardiography*. Lea & Febiger, 1982.

[12] V. Badrinarayanan, A. Kendall, and R. Cipolla, "Segnet: A deep convolutional encoder-decoder architecture for image segmentation," *arXiv preprint arXiv:1511.00561,* 2015.

[13] J. Chung, C. Gulcehre, K. Cho, and Y. Bengio, "Gated feedback recurrent neural networks," in *International Conference on Machine Learning*, 2015, pp. 2067-2075.

[14] S. Mazaheri, R. Wirza, P. S. Sulaiman, M. Z. Dimon, F. Khalid, and R. M. Tayebi, "Segmentation methods of echocardiography images for left ventricle boundary detection," *Journal of Computer Science,* vol. 11, no. 9, p. 957, 2015.

[15] W. Ohyama, T. Wakabayashi, F. Kimura, S. Tsuruoka, and K. Sekioka, "Automatic left ventricular endocardium detection in echocardiograms based on ternary thresholding method," in *Pattern Recognition, 2000. Proceedings. 15th International Conference on*, 2000, vol. 4, pp. 320-323:

IEEE.

[16] J. Santos, D. Celorico, J. Varandas, and J. Dias, "Automatic segmentation of echocardiographic Left Ventricular images by windows adaptive Thresholds," in *Proceedings of the International Congress on Ultrasonics, Vienna, April*, 2007, pp. 9-13.

[17] S. M. Collins *et al.*, "Computer-assisted edge detection in two-dimensional echocardiography: comparison with anatomic data," *The American journal of cardiology,* vol. 53, no. 9, pp. 1380-1387, 1984.

[18] N. Lin, W. Yu, and J. S. Duncan, "Combinative multi-scale level set framework for echocardiographic image segmentation," *Medical image analysis,* vol. 7, no. 4, pp. 529-537, 2003.

[19] L. D. Cohen and I. Cohen, "Finite-element methods for active contour models and balloons for 2-D and 3-D images," *IEEE Transactions on Pattern Analysis and machine intelligence,* vol. 15, no. 11, pp. 1131-1147, 1993.

[20] W. Wang, L. Zhu, J. Qin, Y.-P. Chui, B. N. Li, and P.-A. Heng, "Multiscale geodesic active contours for ultrasound image segmentation using speckle reducing anisotropic diffusion," *Optics and Lasers in Engineering,* vol. 54, pp. 105-116, 2014.

[21] G. Jacob, J. A. Noble, C. Behrenbruch, A. D. Kelion, and A. P. Banning, "A shape-space-based approach to tracking myocardial borders and quantifying regional left-ventricular function applied in echocardiography," *IEEE Transactions on Medical Imaging,* vol. 21, no. 3, pp. 226-238, 2002.

[22] G. Belous, A. Busch, and D. Rowlands, "Segmentation of the Left Ventricle

from Ultrasound Using Random Forest with Active Shape Model," in *Artificial Intelligence, Modelling and Simulation (AIMS), 2013 1st International Conference on*, 2013, pp. 315-319: IEEE.

[23] B. Kayalibay, G. Jensen, and P. van der Smagt, "CNN-based Segmentation of Medical Imaging Data," *arXiv preprint arXiv:1701.03056,* 2017.

[24] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 2015, pp. 234-241: Springer.

[25] Z. Yue, L. Wenqiang, J. Jing, J. Yu, S. Yi, and W. Yan, "Automatic segmentation of the Epicardium and Endocardium using convolutional neural network," in *Signal Processing (ICSP), 2016 IEEE 13th International Conference on*, 2016, pp. 44-48: IEEE.

[26] L. Yu, Y. Guo, Y. Wang, J. Yu, and P. Chen, "Segmentation of Fetal Left Ventricle in Echocardiographic Sequences Based on Dynamic Convolutional Neural Networks," *IEEE Transactions on Biomedical Engineering,* 2016.

[27] H. Wyatt *et al.*, "Cross-sectional echocardiography. II. Analysis of mathematic models for quantifying volume of the formalin-fixed left ventricle," *Circulation,* vol. 61, no. 6, pp. 1119-1125, 1980.

[28] K. R. Chaudry, S. Ogawa, F. J. Pauletto, F. E. Hubbard, and L. S. Dreifus, "Biplane measurements of left and right ventricular volumes using wide angle cross-sectional echocardiography," *The American Journal of Cardiology,* vol. 41, no. 2, p. 391, 1978.

[29] E. D. Folland, A. F. Parisi, P. F. Moynihan, D. R. Jones, C. L. Feldman, and D. E. Tow, "Assessment of left ventricular ejection fraction and volumes by real-time, two-dimensional echocardiography. A comparison of cineangiographic and radionuclide techniques," *Circulation,* vol. 60, no. 4, pp. 760-766, 1979.

[30] J. C. Mercier *et al.*, "Two-dimensional echocardiographic assessment of left ventricular volumes and ejection fraction in children," *Circulation,* vol. 65, no. 5, pp. 962-969, 1982.

[31] N. B. Schiller *et al.*, "Left ventricular volume from paired biplane two-dimensional echocardiography," *Circulation,* vol. 60, no. 3, pp. 547-555, 1979.

[32] N. H. Silverman, T. A. Ports, A. R. Snider, N. B. Schiller, E. Carlsson, and D. C. Heilbron, "Determination of left ventricular volume in children: echocardiographic and angiographic comparisons," *Circulation,* vol. 62, no. 3, pp. 548-557, 1980.

[33] W. Bommer, T. Chun, O. L. Kwan, A. Neumann, D. T. Mason, and A. N. DeMaria, "Biplane apex echocardiography versus biplane cineangiography in the assessment of left ventricular volume and function: validation by direct measurements," *The American Journal of Cardiology,* vol. 45, no. 2, p. 471, 1980.

[34] W. S. McCulloch and W. Pitts, "A logical calculus of the ideas immanent in nervous activity," *The bulletin of mathematical biophysics,* vol. 5, no. 4, pp. 115-133, 1943.

[35] J. J. Hopfield, "Artificial neural networks," *IEEE Circuits and Devices*

Magazine, vol. 4, no. 5, pp. 3-10, 1988.

[36] E. Alcevska, "Segmentation of the Left Ventricle of the Heart in 2D Ultrasound Images using Convolutional Neural Networks," Master, Department of Signals and Systems, CHALMERS UNIVERSITY OF TECHNOLOGY, CHALMERS UNIVERSITY OF TECHNOLOGY, EX093/2016, 2016.

[37] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in neural information processing systems*, 2012, pp. 1097-1105.

[38] F. L. A. Karpathy, and J. Johnson, "n. CS231n Convolutional Neural Network for Visual Recognition," 2016.

[39] Y. N. Dauphin, R. Pascanu, C. Gulcehre, K. Cho, S. Ganguli, and Y. Bengio, "Identifying and attacking the saddle point problem in high-dimensional non-convex optimization," in *Advances in neural information processing systems*, 2014, pp. 2933-2941.

[40] S.-i. Amari, "Backpropagation and stochastic gradient descent method," *Neurocomputing,* vol. 5, no. 4, pp. 185-196, 1993.

[41] I. Sutskever, J. Martens, G. Dahl, and G. Hinton, "On the importance of initialization and momentum in deep learning," in *International conference on machine learning*, 2013, pp. 1139-1147.

[42] M. D. Zeiler, "ADADELTA: an adaptive learning rate method," *arXiv preprint arXiv:1212.5701,* 2012.

[43] D. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980,* 2014.

[44] S. Han, J. Pool, J. Tran, and W. Dally, "Learning both weights and connections for efficient neural network," in *Advances in Neural Information Processing Systems*, 2015, pp. 1135-1143.

[45] S. Wager, S. Wang, and P. S. Liang, "Dropout training as adaptive regularization," in *Advances in neural information processing systems*, 2013, pp. 351-359.

[46] X. Glorot and Y. Bengio, "Understanding the difficulty of training deep feedforward neural networks," in *Aistats*, 2010, vol. 9, pp. 249-256.

[47] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Transactions on knowledge and data engineering,* vol. 22, no. 10, pp. 1345-1359, 2010.

[48] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," *arXiv preprint arXiv:1502.03167,* 2015.

[49] J. Bergstra and Y. Bengio, "Random search for hyper-parameter optimization," *Journal of Machine Learning Research,* vol. 13, no. Feb, pp. 281-305, 2012.

[50] J. Snoek, H. Larochelle, and R. P. Adams, "Practical bayesian optimization of machine learning algorithms," in *Advances in neural information processing systems*, 2012, pp. 2951-2959.

[51] M. Drozdzal, E. Vorontsov, G. Chartrand, S. Kadoury, and C. Pal, "The importance of skip connections in biomedical image segmentation," in *International Workshop on Large-Scale Annotation of Biomedical Data and Expert Label Synthesis*, 2016, pp. 179-187: Springer.

[52] W. R. Crum, O. Camara, and D. L. Hill, "Generalized overlap measures for

evaluation and validation in medical image analysis," *IEEE transactions on medical imaging,* vol. 25, no. 11, pp. 1451-1461, 2006.

[53] J. Chung, C. Gulcehre, K. Cho, and Y. Bengio, "Empirical evaluation of gated recurrent neural networks on sequence modeling," *arXiv preprint arXiv:1412.3555,* 2014.

[54] C. Seg3D, "Volumetric image segmentation and visualization," *Scientific Computing and Imaging Institute (SCI),* 2013.

[55] R. Kohavi, "A study of cross-validation and bootstrap for accuracy estimation and model selection," in *Ijcai*, 1995, vol. 14, no. 2, pp. 1137-1145: Stanford, CA.

[56] M. Abadi *et al.*, "Tensorflow: Large-scale machine learning on heterogeneous distributed systems," *arXiv preprint arXiv:1603.04467,* 2016.

[57] P. K. Chandrasekharan *et al.*, "Continuous End-Tidal Carbon Dioxide Monitoring during Resuscitation of Asphyxiated Term Lambs," *Neonatology,* vol. 109, no. 4, pp. 265-273, 2016.

[58] R. Sehra, K. Underwood, and P. Checchia, "End tidal CO2 is a quantitative measure of cardiac arrest," *Pacing and clinical electrophysiology,* vol. 26, no. 1p2, pp. 515-517, 2003.

[59] D. Steedman and C. Robertson, "Measurement of end-tidal carbon dioxide concentration during cardiopulmonary resuscitation," *Archives of emergency medicine,* vol. 7, no. 3, pp. 129-134, 1990.

[60] J. P. Ornato, A. R. Garnett, and F. L. Glauser, "Relationship between cardiac output and the end-trial carbon dioxide tension," *Annals of emergency medicine,* vol. 19, no. 10, pp. 1104-1106, 1990.

[61] S. N. Wijewickrema and A. P. Papliński, "Principal component analysis for the approximation of an image as an ellipse," 2005.

[62] C. V. Gudipati, M. H. Weil, J. Bisera, H. G. Deshmukh, and E. C. Rackow, "Expired carbon dioxide: a noninvasive monitor of cardiopulmonary resuscitation," *Circulation,* vol. 77, no. 1, pp. 234-239, 1988.

[63] R. P. Trevino, J. Bisera, M. H. Weil, E. C. Rackow, and W. G. Grundler, "End-tidal CO2 as a guide to successful cardiopulmonary resuscitation: a preliminary report," *Critical care medicine,* vol. 13, no. 11, pp. 910-911, 1985.

국문초록

본 논문은 심폐소생술 효과의 실시간 평가를 위한 심초음파영상 분석과 삼차원 변환 모델링에 관한 내용을 담고 있다.

최근 심폐소생술의 효과를 최적화 하기 위해서 생체신호 피드백 심폐소생술에 관한 연구가 많이 수행되고 있으며, 피드백 신호로는 심폐소생술 시 심박출량을 보여줄 수 있는 심초음파가 사용되고 있다. 심초음파를 통해 심장에서 박출되는 혈액량을 추정하기 위해서는 두 단계를 거쳐야 한다. 첫 번째 단계는 심초음파에서 좌심실의 영역을 분할하는 것이며, 두 번째 단계는 분할된 좌심실의 2차원 이미지를 통해 3차원 부피 추정을 수행하는 것이다. 하지만, 심폐소생술 수행 중에는 초음파 탐촉자(Acoustic transducer)와 가슴과의 접촉면을 일정하게 유지시키기 어렵기 때문에 안정적인 심초음파 영상을 획득하기 어렵다는 문제점이 있다. 또한, 좌심실의 2차원 이미지로 3차원 부피를 나타내는 기존의 모델은 등수축하는 정상상태의 심장을 대상으로 만들어졌기 때문에 수축 없이 눌러지는 심폐소생술 중의 심장에 적용하여 심박출량을 추정하는 것이 적절하지 못하다. 따라서, 심폐소생술 중의 불안정한 심초음파에서의 안정적인 실시간 좌심실 분할, 그리고 심폐소생술상황에 적용될 수 있는 좌심실 모델의 개발에 대해서 연구를 수행하였다.

안정적인 실시간 좌심실 분할을 위해서 합성곱 신경망 (Convolution neural network)을 이용한 분할 알고리즘을 개발하였다. 실시간 분할 알고리즘으로 높은 성능을 보여주고 있는 ‘Segnet’ 을 기반으로 심초음파 영상에서의 정확도 향상을 위해 스킵커넥션 (skip conncetion)과 다이스 (dice) 계수를 적용하였다. 스킵커넥션을 통해 풀링과 업샘플링

의 정보손실을 최소화 하였고 다이스 계수를 통해 전체적인 분할의 일치도를 향상시키도록 했다. 마지막으로 게이티드 재귀 유닛 (Gated recurrent unit)을 사용해서 현재의 프레임에 분할하기 위한 충분한 정보가 포함되어 있지 않은 경우 과거의 프레임에서 참조하도록 설계했다. 본 연구에서 제안하는 합성곱 신경망 기반의 모델의 정확도 및 처리 속도를 비교하기 위하여 현재까지 제안된 심초음파 실시간 분할 알고리즘 중 가장 성능이 좋은 것으로 평가되는 측지선 활성 외곽선 (Geodesic active contour) 방법과 비교 실험을 수행해 본 결과 본 연구에서 개발한 모델이 기존의 방법에 비해 더 높은 정확도와 더 처리 빠른 속도를 가지는 결과를 보여주는 것을 확인할 수 있었다.

위의 방법을 통해 분할된 영상을 바탕으로 심폐소생술 중 좌심실의 부피를 추정하기 위한 모델을 유도하였다. 모델을 유도하기 위해 두 가지 가정을 적용하였다. 첫 번째 가정은 심정지 상태의 심장은 수축하지 않기 때문에 압박기와 이완기의 심장의 겉넓이가 일치한다는 것이며, 두 번째는 심박출량을 구하기 위해 필요한 변수의 수를 감소 시키기 위해 이완기의 중축과 단축의 일치한다는 것이다. 이러한 가정이 적용된 본 연구에서 제안하는 부피 추정 모델과 기존에 사용되고 있는 부피 추정 모델을 호기 말 이산화탄소 (EtCO2) 를 이용하여 결과를 비교하였다. 그 결과, 기존 모델 보다 본 연구에서 개발한 부피 추정 모델이 더 높은 상관계수를 가지는 것을 확인할 수 있었다.

핵심어: 심폐소생술, 심초음파, 이미지 분할, 합성곱 인공 신경망, 게이티드 재귀 유닛, 좌심실 모델

학번: 2015-22888