Automatic Quantification of Cardiac MRI for Hypertrophic Cardiomyopathy

A Thesis

Presented to

the faculty of the School of Engineering and Applied Science

University of Virginia

in partial fulfillment
of the requirements for the degree

Master of Science

by

Anudeep Konda

May 2018

APPROVAL SHEET

This Thesis
is submitted in partial fulfillment of the requirements
for the degree of
Master of Science

Author Signature: Anudeep.k

This Thesis has been read and approved by the examining committee:

Advisor: Dr. Vicente Ordonez Roman

Committee Member: Dr. Craig H Meyer

Committee Member: Dr. Nada Basit

Committee Member: Dr. Xue Feng

Committee Member: _____

Committee Member: _____

Accepted for the School of Engineering and Applied Science:

Craig H. Benson, School of Engineering and Applied Science

May 2018

谨以此纪念你为我展示的世界

纪念山顶上汇成河流的风

纪念我们伸出的手

数不清的凌晨无人的街

声声不绝的铁轨

目击万物在这一刻间苍老

夏日不去解释它的意义

只有到来

- *Harper*

Abstract

Hypertrophic cardiomyopathy (HCM) is the most common monogenic heart disease, characterized by unexplained left ventricular hypertrophy, myofibrillar disarray and myocardial fibrosis. Left and right ventricular mass, ejection fraction and myocardium wall thickness at different segments measured from cardiac cine MRI based on LV, RV segmentation and mean myocardial T1 measured from LV segmentation on native T1 maps are critical biomarkers for diagnosis and prognosis of HCM patients. Deep convolutional neural networks (DCNNs) have shown great promise in many medical image segmentation tasks, including cardiac MRI. However, due to the greatly increased variability in shape and size of heart chambers and often reduced image contrast, the segmentation for HCM is more challenging than healthy and other patient populations and the model trained on generic cardiac MRI is very likely to fail on HCM. In this study, we developed a cascaded deep convolutional neural network to automatically segment the epi and endocardium at end-diastole and end-systole phases from cine and native T1 images to calculate all variables of interest based on a database with 100 HCM patients. Ejection fraction, LV and RV mass, mean myocardial T1 and regional wall thickness at 6 automatically localized segments per slice were calculated with promising results. The model greatly reduces the post-processing time and inter/intra-observer variability in biomarker quantifications for HCM patients.

Table of Contents

List of Figures

Acknowledgements

I would like to thank Dr. Craig Meyer for letting me work on this fascinating project. I am thankful to Dr. Xue Feng for his excellent mentorship through every step in the process. I am also thankful to Dr. Vicente Ordonez Roman for undertaking the duty of thesis advisor and Dr. Nada Basit for her support in weeks leading up to the defense. I would like to thank them all again for serving on my committee and making the defense a wonderful experience.

Thanks mum and dad for being who you are and making me what I am.

Anudeep Konda

Chapter 1

Introduction

Hypertrophic cardiomyopathy (HCM) is the most common monogenic heart disease and the number one cause of heart related death in young adults. Currently known clinical risk markers are only modestly effective at identifying highest risk individuals. There is a pressing need for rapid and robust quantification of biomarkers from cardiac magnetic resonance images (CMR) to find ways of predicting the risk of developing complications associated with HCM, so that the most appropriate advice and treatment can be given to patients. In order to completely understand the techniques we developed for automated biomarker quantification, it will be very helpful to know the anatomy of the heart and basics of CMR images. These introductory sections are followed by the motivation, related work and the research contributions of this study.

1.1 Anatomy of Human Heart

The heart is a muscular pump with four chambers and an equal number of valves. The two chambers at the top of the heart are known as the atria, a right and a left. The two bottom chambers are the ventricles. The atria receive blood that returns from the different parts of the body, while the ventricles pump that blood back to all body tissues. The heart wall consists of three layers: the outer epicardium, the middle myocardium, and the inner endocardium. The epicardium and endocardium are thin layers. The myocardium forms the main bulk of the heart and is made up of cardiac muscle fibers. Vessels that carry blood away from the heart to the body are called arteries; vessels that bring it back are called veins. The largest artery is named the aorta. It arises from the left ventricle. The events that take place in the heart between successive heartbeats constitute the cardiac cycle. Such events include the opening and closing of valves and contraction and relaxation of chambers. The cardiac cycle is split into two phases: systole and diastole. During systole, the ventricles contract and push blood into the arteries. During diastole, the ventricles relax and receive blood from the atria. Figure 1 shows a longitudinal cross-section of heart where all four chambers can be seen.

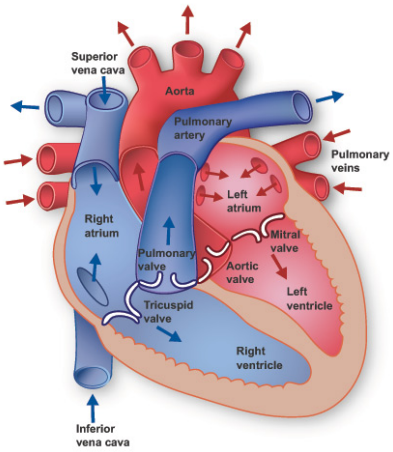


Figure 1: Cross-section of human heart showing all four chambers and all four valves
**Source: https://www.texasheart.org/heart-health/heart-information-center/topics/heart-anatomy*

1.2 Hypertrophic Cardiomyopathy

HCM is a cardiovascular disease that affects the heart muscle, also known as the myocardium. It is characterized by the thickening of heart muscle, especially in the ventricles (or lower heart chambers). Thickening of the myocardium occurs most commonly at the septum, which is the muscular wall that separates the left and right side of the heart. The thickened septum may cause a narrowing that can block or reduce the blood flow from the left ventricle to the aorta - a condition called "outflow tract obstruction." The ventricles must pump harder to overcome the narrowing or blockage. HCM also may cause thickening in other parts of the heart muscle, such as the bottom of the heart called the apex, right ventricle, or throughout the entire left ventricle. The degree and location of thickening is fairly random. Myocardial thickening also causes cellular changes which lead to stiffening of the tissue. This restricts the normal relaxation of the ventricle disabling it from fully filling with blood. Since there is less blood at the end of filling, there is less oxygen-rich blood pumped to the organs and muscles, which leads to ischemia.

Figure 2 shows a pictorial representation of the longitudinal cross-section of normal and HCM hearts. The thickening of the septum which greatly reduces the left ventricular volume can be observed in the HCM heart. In addition, the figure also shows the blockage of aortic valve due to hypertrophy in the ventricular septum.

HCM is very common and can affect people of any age. It affects men and women equally. It is a common cause of sudden cardiac arrest in young people, including young athletes. HCM usually is inherited. It is caused by a change in some of the genes in heart muscle proteins. However, the actual cause of the disease is unknown.

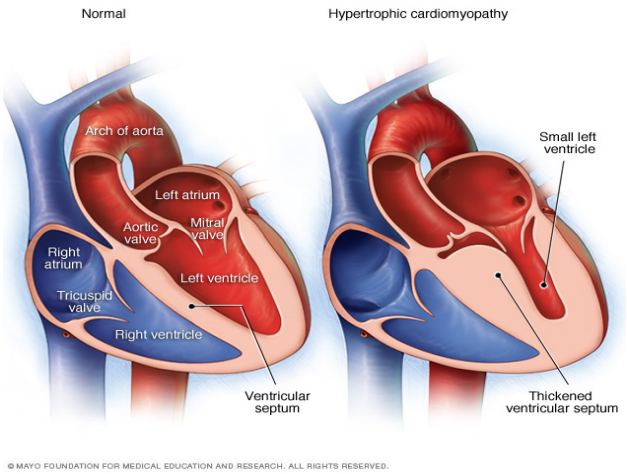


Figure 2: Cross section showing all four chambers in Normal(left) and HCM(right) heart
**Source*: https://www.mayoclinic.org/diseases-conditions/hypertrophic-cardiomyopathy/symptoms-causes/syc-20350198

HCM is particularly interesting because it is asymptomatic and goes undiagnosed in some patients until it manifests itself, causing a cardiac event (heart failure, cardiac arrest, etc). It can be diagnosed by family screening which includes genetic testing and echo cardiogram tests with an accuracy of 80%. CMR, which is considered the gold standard for determining the physical properties of the left ventricular wall, can serve as a more efficient screening tool for several reasons. Echo images are not conclusive and cannot be helpful in determining some cases like segmental lateral ventricular hypertrophy. Moreover, the hypertrophy maybe absent in children because their hearts are not fully developed. It fails to detect HCM as no physical attributes representing it exists. Researchers, however, have studied asymptomatic carriers of an HCM-causing mutation through the use of CMR and have been able to identify crypts in the interventricular septal tissue in these people. Biomarkers extracted from CMR can be very helpful in identifying the patients at highest risk. The following section introduces CMR and discusses different types of CMR images used in this study.

1.3 Cardiac MRI

Cardiac MRI is a medical technology for non-invasive assessment of the function and structure of the cardiovascular system. CMR uses the same basic principles of image acquisition and reconstruction as other MRI techniques which rely on positively charged hydrogen protons in the body, most of which are located in water molecules. A powerful, uniform, external magnetic field is employed to align the protons that are normally randomly oriented within the water nuclei of the tissue being examined. This alignment (or magnetization) is next perturbed or disrupted by introduction of an external Radio Frequency (RF) energy. The nuclei return to their resting alignment through various relaxation processes and therefore emit RF energy. After a certain period following the initial RF, the emitted signals are measured. Fourier transformation is used to convert the frequency information contained in the signal from each location in the imaged plane to corresponding intensity levels, which are then displayed as shades of gray in a matrix arrangement of pixels. By varying the sequence of RF pulses applied and collected, different types of images are created. In this study, we use cine cardiac images in short-axis view and native-T1 maps. Details regarding the same are presented in following sub sections.

1.3.1 Image Views

There are two imaging planes in which the CMR images are acquired, the short axis and the long axis. The long axis of the heart is the axis that aligns the base of the heart and the apex. Short axis is the one perpendicular to it. Short axis images give an excellent cross-sectional view of the left and right ventricles. It is ideal for calculating the ventricular volumes and the wall thickness. Figure 3 shows the long and short axis of the heart along with one slice in short-axis view.

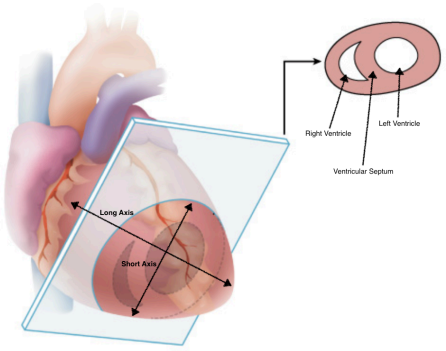


Figure 3: Long and short axis of the heart
**Source: http://fqz301.brinkster.net/Braunwald%20Heart%20Disease%207th-new/heart%207th%20by%20FengQZ%202005-07-02/html/124[1].html*

1.3.2 Cine Cardiac MRI

Cine CMR images (will be referred to as cine images from now on) are short movies that show the heart motion throughout the cardiac cycle, in short-axis. Cine images are obtained with ECG triggered segmented imaging. Segmented acquisition is the process of dividing the cardiac cycle into multiple segments to produce a series of images that can be displayed as a movie. The cardiac cycle begins with the R wave of the ECG, ends with the subsequent R wave and is typically divided into 10 to 20 segments, depending on the heart rate, Figure 4 shows a visual representation of this. Each image in the cine is typically composed of information gathered over several heart beats allowing for a movie to be acquired with a breath hold of 10 to 20 seconds depending on the sequence. As in any movie, the final cine is a sequence of individual frames. These images can be very helpful in studying cardiac function, valvular function, and the movement of blood through the heart.

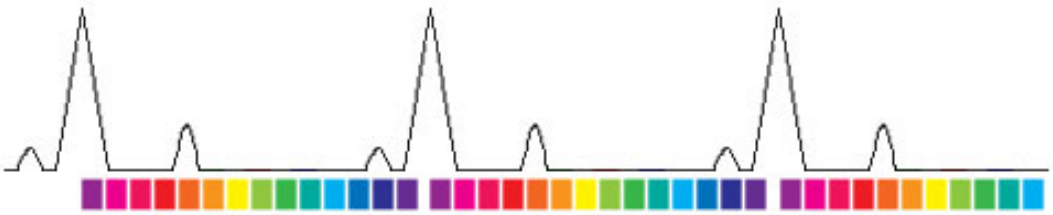


Figure 4: ECG tracing with the colored boxes representing different frames
**Source: https://www.med-ed.virginia.edu/courses/rad/cardiacmr/Techniques/Cine.html*

In the cardiac cycle which is comprised of multiple phases represented by different frames in the cine, we are particularly interested in two phases namely end-diastole and end-systole. At end-diastole, the myocardium is completely relaxed and is fully filled with blood that will be pumped in the following systole phase. At end-systole, the myocardium is completely contracted and has pumped all the blood it can, out of the ventricle. Figure 5 shows the cine on one slice; the cine includes end-systole and end-diastole which are marked correspondingly.

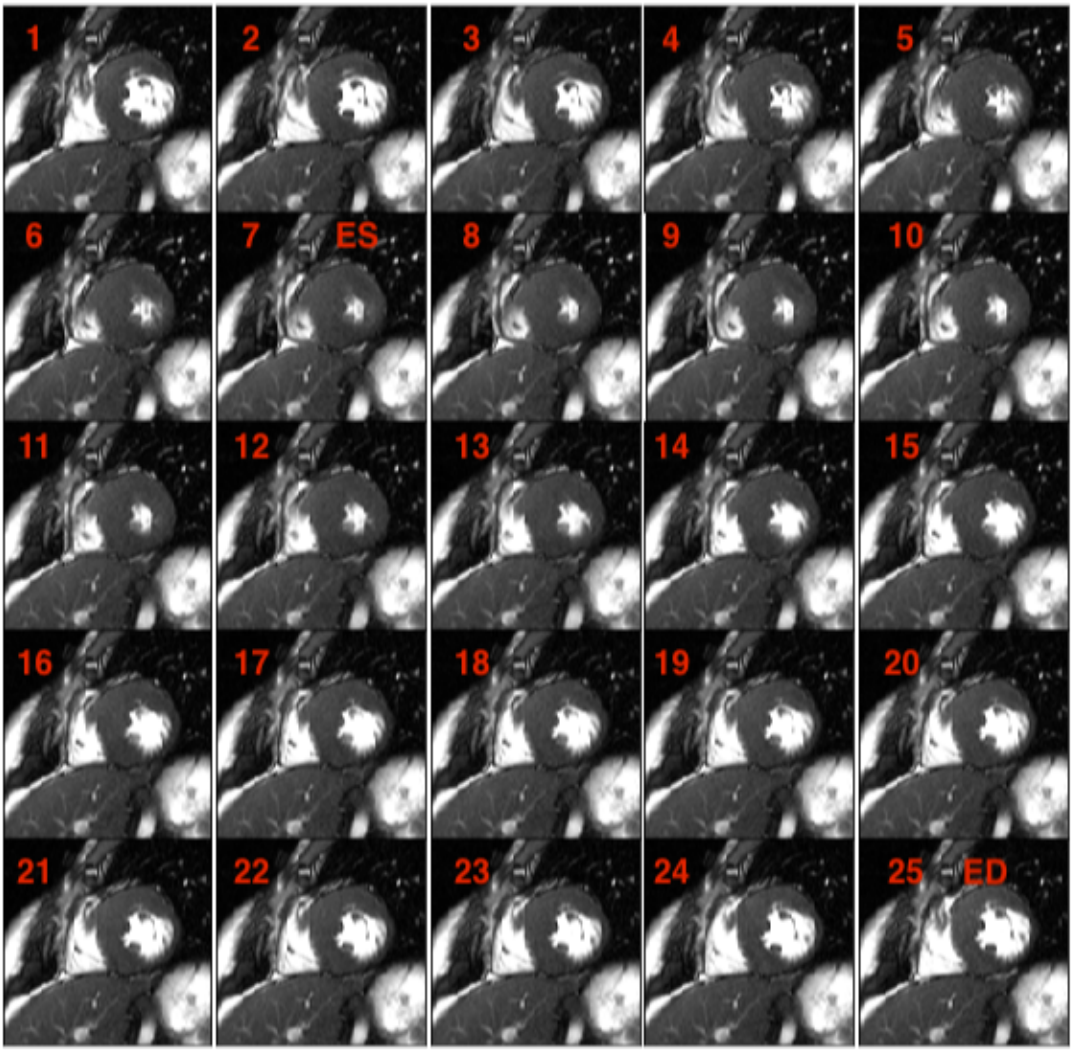


Figure 5: A cine with twenty-five cardiac phases/frames in sequence on a mid-cavity slice. Phase 7 is end-systole end phase 25 is end-diastole

1.3.3 T1 Maps

T1 (longitudinal relaxation time) is the time constant which determines the rate at which excited protons return to equilibrium. It is a measure of the time taken for spinning protons to realign with the external magnetic field. T1 of the myocardium is altered in various disease states due to increased water content or other changes to the local molecular environment. Changes in both native T1 and T1 following

administration of gadolinium (Gd) based contrast agents are considered important biomarkers. In this study, we only use native T1 maps, which represents the absolute T1 value of the tissue.

1.4 Motivation

The goal of this work is to automate the rapid and robust quantification of biomarkers including the left and the right ventricular mass, ejection fraction and myocardial wall thickness, from cine images, and mean myocardial T1 from native T1 maps for identification of HCM patients at risk. To achieve accurate measurements of these variables, segmentation of heart chambers and myocardium regions is required on cine images across all short-axis slices and for at least end-diastole and end-systole phases, and on native T1 maps. Currently the segmentation is performed manually by experienced cardiologists, so it is time-consuming and suffers from inter-observer variability with reduced biomarker quantification accuracy and robustness, especially in a multi-site study. Automating the biomarker quantification involves automating the heart chamber segmentation task.

There are several challenges in automating the heart chamber segmentation task - the heterogeneities in the brightness of LV cavity due to blood flow, presence of papillary muscles with signal intensities similar to myocardium makes it harder to delineate the endocardial wall. Tissues surrounding the epicardium (fat, lung), which have different intensity profiles and show poor contrast with the myocardium, make the segmentation of the epicardium difficult. Segmentation complexity also depends on the slice level of the image. Apical and basal slice images are more difficult to segment than mid-ventricular images. Indeed, MRI resolution is not high enough to resolve size of small structures at the apex and ventricle shapes are strongly modified close to the base of the heart, because of the vicinity of the atria[1]. Dynamic motion of heart causes inhomogeneity of intensity and high variations in contrasts. Irregular crescent shape of the right ventricle makes it much harder to segment in comparison with the left ventricle. Moreover, data from HCM population has a much higher variability in the shape and size of the heart chambers because of the randomness of hypertrophy in comparison with normal and other pathologies. Figure 6 shows basal, mid ventricular and apical slices from HCM hearts (a, b) and a normal heart (c). The myocardial hypertrophy is pointed to by white arrows. It can be observed that in a normal heart the wall thickness is consistent throughout the LV, whereas in HCM hearts the consistency highly varies.

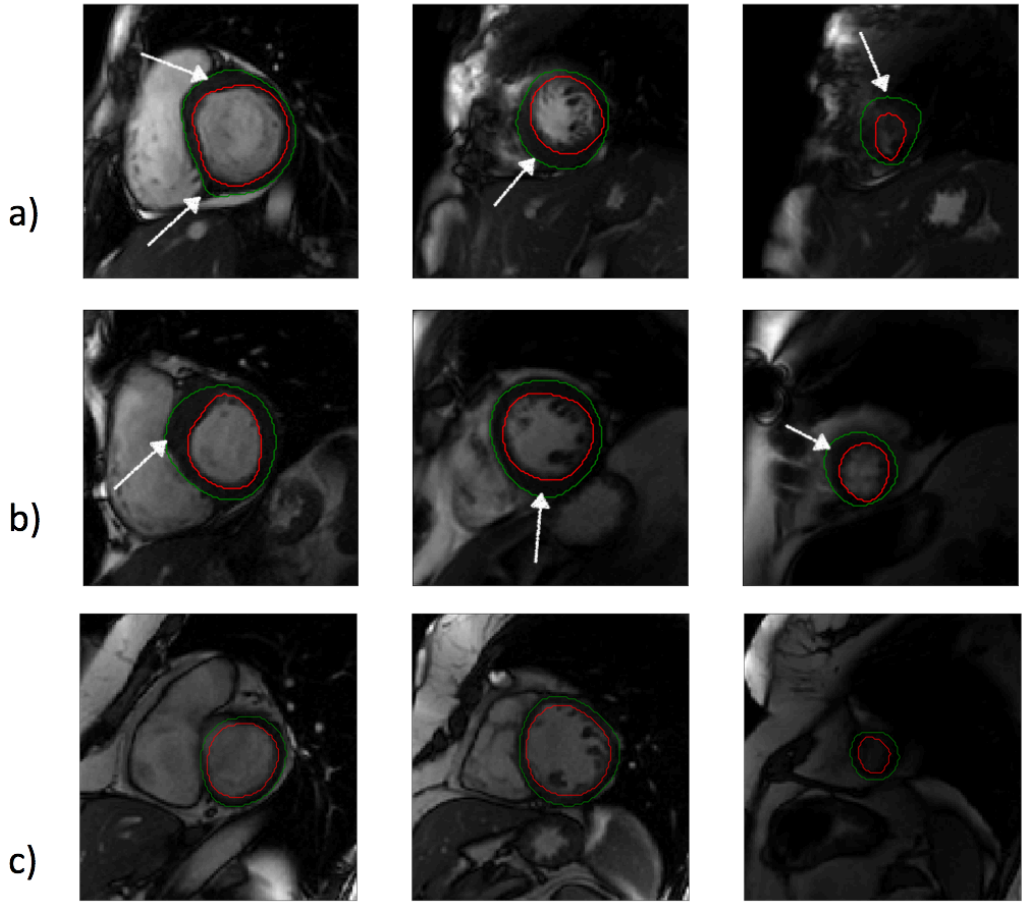


Figure 6: Comparison of HCM (a, b) with normal heart (c). White arrows point to the hypertrophy in myocardium

1.5 Related Work

Existing automatic cardiac MRI segmentation techniques can broadly be classified into two groups – those that need no prior information[2–7] and those that need weak or strong prior knowledge[8–10]. The former includes techniques that are primarily image-driven which use the pixel intensity differences to distinguish between different regions of interest. The latter techniques are often model-driven, using statistical information extracted from manually annotated training data that describe the spatial relationships of the LV and RV objects and their surrounding structures, knowledge of the heart biomechanics, or anatomical assumptions about the statistical shapes of the objects. Such assumptions about the objects, through either

weak or strong priors, contributes to the propensity of these methods to overfit on a particular training dataset, thus making them less generalizable.

The approaches that need no prior information includes active contours or snakes techniques, pixel classification using clustering algorithms, region growing algorithms, learning based techniques. Among these approaches, accurate fully automatic segmentation is only achievable using learning based techniques. These include techniques based on random forests[11–13], markov random fields[5,14] (MRF), conditional random fields[15,16] (CRF), restricted boltzman machines[17] (RBM) and deep learning. Methods using random forests rely on image intensity and define the segmentation problem as a classification task. These methods have multiple stages of intensity standardization, estimation and normalization, which are computationally expensive and affect the success of further steps. Moreover, their performance is rather mediocre at basal and apical slices and overall inferior to the state-of-the-art. MRFs and RBMs try to learn the probability of input data. Computing the image probability and parameter estimation in generative models is generally difficult and can lead to reduced performance if oversimplified. Besides, they use the Gibbs sampling algorithm for training, which can be slow, become stuck for correlated inputs, and produce different results each time it is run due to its randomized nature. Alternatively, CRF methods try to model the conditional probability of latent variables, instead of the input data. However, they are still computationally difficult, due to complexity of parameter estimation, and their convergence is not guaranteed. Deep learning techniques, that received much attention over the course of last 5 years, are much more stable and achieve a better performance in comparison with the techniques mentioned earlier.

Deep convolutional neural networks (DCNN) have shown great promise in many medical image segmentation tasks, including cardiac MRI. A majority of these only focus on segmenting the LV for ejection fraction calculation. Yang et al.[18] proposed a fully convolutional architecture to segment the LV myocardium which is relatively shallow, consisting of three convolutional blocks with two of them followed by max pooling and one 4 strided deconvolution to regain the original image dimension. An average of dice score of 0.75 was reported on the CMR dataset from York University. Avendi et al[19]. proposed a hybrid approach that uses deformable models in conjunction with deep learning. A fully convolutional network is used to locate the LV, and a stacked auto encoder is then used to infer the shape of the ventricle which is then used by a deformable model to accurately segment the region of interest. The

main limitation of this method is that it is multi-stage and requires manual offline training along with extensive hyper-parameter tuning, which can be cumbersome and difficult to generalize to multi-site data. Tran et al[20]. proposed a 15-layered architecture that uses 2D data, which achieved the state-of-art dice scores of 0.96 on epicardium and 0.92 on endocardium using the Sunnybrook dataset. Despite being the state-of-art, this technique uses 2D data and ignores the overall shape of the LV which could be crucial in identifying the edge slices that shouldn't be segmented.

Some approaches that perform both LV and RV segmentations treat them as two separate problems, thus ignoring their relative positions and shapes. Avendi et al[21]. also proposed a technique quite similar to the one discussed earlier to segment the RV. A dice score of 0.81 was reported on the endocardium. Tran et al[20]. also suggested their LV segmentation approach can be used, without any changes, to segment the RV. Furthermore, none of these studies focus on HCM populations, and we hypothesize that a model trained on normal and other patient populations is very likely to perform poorly on an HCM dataset due to the notable differences in contrast and shape.

1.6 Contributions

Inspired from the works mentioned earlier, the problem of automatic biomarker quantification is tackled by-

1. Developing a fully automatic cascaded deep learning model to accurately segment both epicardium and endocardium of left and right ventricles from cine images.
2. Developing deep learning models to accurately segment the epicardium and endocardium of LV from native T1 images.
3. Quantifying the following biomarkers
    a. LV wall thickness
    b. LV mass, RV mass
    c. LV ejection fraction, RV ejection fraction
    d. Mean myocardial T1

Chapter 2

Convolutional Neural Networks

We introduce the basic concepts and the lexicon of deep learning models used in this study, which are the convolutional neural networks (CNN), and discuss their history in image segmentation problems. Some common CNN architectures used for image segmentation are briefly described. The very popular 3D-UNet style architecture extensively used in medical image segmentation is described in detail.

2.1 Background

CNNs are multi-layer feed-forward networks specifically designed to recognize features in image data. The architecture of CNNs is inspired by Hubel and Wiesel's study of neurobiological signal processing in cat's visual cortex. A typical application of CNNs consists of recognition of various objects in images. However convolutional networks have been successfully used for various different tasks, too. The neurons in CNNs work by considering a small portion of the image, say patch. The patches are inspected for features that can be recognized by the network. As a simple example, a feature may be a vertical line, an arch, or a circle. These features are then captured by the respective feature maps of the network. A combination of features is then used to classify the image, or in our case, each pixel.

A CNN architecture typically consists of multiple layers. The abstraction, i.e layers, is done based on the functionality of the operations. The series of mathematical operations that perform convolutions on the image using multiple kernels (in this case, filters or feature maps) are termed convolutional layers, ones that reduce the spatial dimension by averaging the local information are called pooling layers, ones that apply a non-linear function to each pixel in the image in order to amplify the effects of a previous convolution are called activation layers, ones that increase the spatial dimension by transposed convolutions or interpolations are respectively called deconvolution or up-sampling layers. A series of many such layers together constitutes a CNN.

2.1.1 Convolutional Layers

Convolutional layer includes a kernel (or filter) whose weights will be 'learned' over time while training. Let K be a kernel with x rows, y columns and depth d. Then the kernel with size Kx * K_y * d works on a receptive field of K_x * K_y on the image. The kernel height and width are smaller than the input image height and width. The kernel slides over (convolves with) the image, producing a feature map (Figure 7). Convolution is the sum of the element-wise multiplication of the kernel and the original image. Note that the depth d of the kernel is equal to the depth of its input. Therefore, it varies within the network. Usually the depth of an image is the number of color channels, the three RGB channels. During training, discussed at the end of this chapter, the values in the convolutional kernels are "learned".

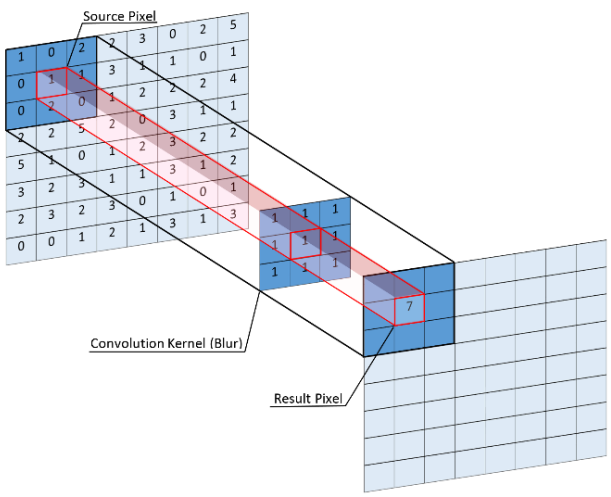


Figure 7: An example of the convolution operation using a kernel whose values are all 1s
**Source: https://datascience.stackexchange.com/questions/23183/why-convolutions-always-use-odd-numbers-as-filter-size*

Although efficient, using the simple convolutional operation might not yield the best feature extraction at times when the regions of interest have complex shapes. Moreover, using larger kernels to increase the effective receptive field (the dimension of the original image viewed by a convolutional layer) size so as to infer more spatial information isn't always efficient as there will be a lot more parameters to train which

not only takes longer, but could also lead to overfitting. To overcome these issues, we tried incorporating dilated convolutions into our CNN.

2.1.1.1 Dilated Convolution

In general, the receptive field of the CNN should be larger the region of interest being segmented so as to acquire enough spatial information. A simple way of increasing the receptive field is to use a larger convolutional kernel. But doing so also increases the number of parameters that are to be trained which not only increases the time to convergence of the gradient descent algorithm but also increases the chance of overfitting on the training data. To overcome this issue the idea of dilated convolutions[22] was put forth.

In simple terms, dilated convolution is just a convolution applied to input with defined gaps. With this definition, given our input is an 2D image, dilation rate k=1 is normal convolution and k=2 means skipping one pixel per input and k=4 means skipping 3 pixels.

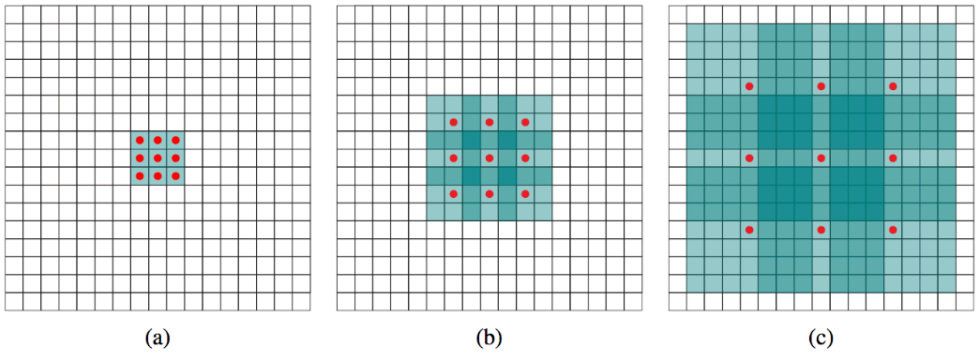


Figure 8: Systematic dilation supports exponential expansion of the receptive field without loss of resolution or coverage. (a) F1 is produced from F0 by a 1-dilated convolution; each element in F1 has a receptive field of 3×3. (b) F2 is produced from F1 by a 2-dilated convolution; each element in F2 has a receptive field of 7×7. (c) F3 is produced from F2 by a 4-dilated convolution; each element in F3 has a receptive field of 15×15. The number of parameters associated with each layer is identical. The receptive field grows exponentially while the number of parameters grows linearly.

** Taken from "MULTI-SCALE CONTEXT AGGREGATION BY DILATED CONVOLUTIONS"*

Figure 8 shows the exponential increase in the receptive field as a function of the dilation rate. The red dots are the points at which the kernel values are centered. The green shaded area represents the receptive field of the kernel. Note that the number of red dots, that represents the number of parameters to be trained remain constant.

2.1.2 Pooling Layer

Pooling layers are used to reduce the spatial dimension of the feature maps in order to reduce the number of parameters and the computation in the network. This also helps to control overfitting. Pooling Layer operates independently on every depth slice of the input and resizes it spatially, using either the 'max', 'min' or 'average' operations. The most common form is a pooling layer with filters of size 2x2 applied with a stride of 2 down samples every depth slice in the input by 2 along both width and height, discarding 75% of the activations. If it were using the 'max' operation, every 'max' operation would in this case be taking a 'max' over 4 numbers (little 2x2 region in some depth slice). The depth dimension remains unchanged. This is called max-pooling. Average pooling and minimum pooling are defined similarly. In our work, we mostly use max-pooling. Figure 9 shows an illustration of max-pooling.

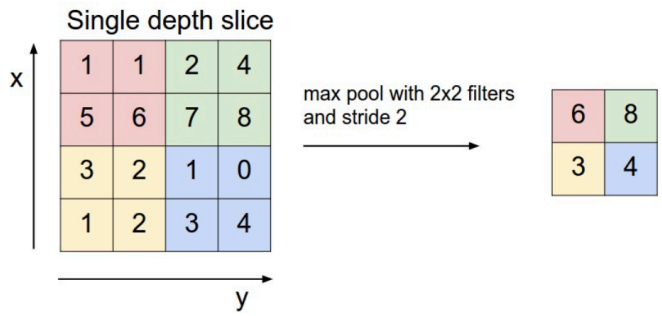


Figure 9: The most common down sampling operation is max, giving rise to max pooling, here shown with a stride of 2. That is, each max is taken over 4 numbers (little 2x2 square)
**Source: http://cs231n.github.io/convolutional-networks/*

2.1.3 Deconvolution Layer

Deconvolutional layers are typically used in CNNs in order to annul the effects of a previous pooling layers and convolutional layers in the network. The job of these layers is to increase the spatial dimension, so that the final output has the same resolution as the input image and direct correlations between the model output and the regions in the output feature map can be obtained. These layers, similar to convolutional layers, have weights that can be 'learned'. Rather than using a fixed function, such as an interpolation kernel, these layers can learn the upsampling operation. These are sometimes also called transposed convolutions. Figure 10 shows an illustration of the deconvolution operation. The 2x2 blue image at the bottom is the input and the 3x3 gray grid is the deconvolutional kernel.

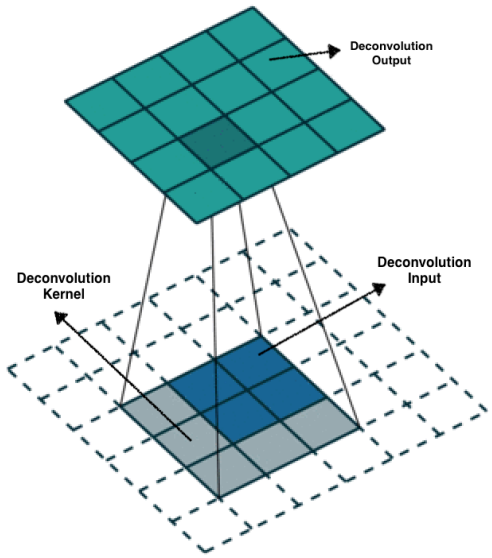


Figure 10: Illustration of deconvolution operation on a 2x2 image using a 3x3 kernel with stride 1
**Source: https://datascience.stackexchange.com/questions/6107/what-are-deconvolutional-layers*

The kernel is moved around with a stride of 1 pixel with no padding. The output size of the deconvolution is given the formula below-

$$S_o = stride(S_i - 1) + S_f - 2 * pad$$

Where S_i is the input size, S_f is the kernel size, pad is the padding and S_o is the output size. Using this formula, the output dimension of our example is 4x4, which is twice the input dimension.

2.1.4 Activation Layer

In general, every convolution layer is immediately followed by an activation layer. Activation layers apply a non-linear function to its inputs. In context of neural networks, activation layers are of much significance as they enable it to estimate a complex non-linear function. Without activations, the neural network would simply be a high order linear equation, which has limited power and does not performs good most of the times. A linear equation is easy to solve but they are limited in their complexity and have less power to learn complex functional mappings from data. One hard requirement that an activation function should satisfy is its differentiability. While 'learning', the back-propagation (described in further sections) algorithm relies on the gradients from these functions to determine the error that needs to be propagated

backwards. For the most part, we use the rectified linear activation (ReLU) and softmax activation in our networks. They are described below.

2.1.4.1 ReLU

ReLU has become very popular in the last few years. It computes the function f(x)=max(0,x). In other words, the activation is simply thresholded at zero. It was found to greatly accelerate the convergence of stochastic gradient descent compared to other complex functions. It is argued that this is due to its linear, non-saturating form. Unfortunately, ReLU can be fragile during training and can "die". For example, a large gradient flowing through a ReLU layer could cause the weights to update in such a way that the neuron will never activate on any datapoint again. If this happens, then the gradient flowing through the unit will forever be zero from that point on. That is, the ReLU units can irreversibly die during training since they can get knocked off the data manifold. It is a major pitfall that should be avoided using ReLU. Despite this technical drawback, in practice, researchers have found that ReLU significantly reduces the training time without any effect on the accuracy. Figure 11 shows a graphical representation of the activation function

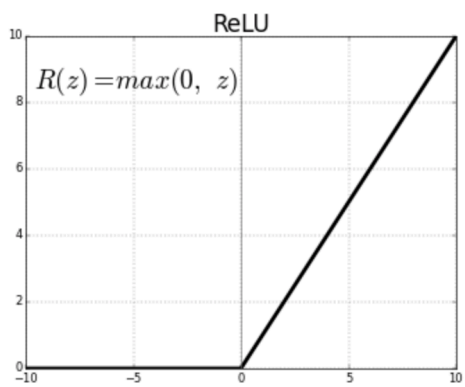


Figure 11: Graphical representation of the ReLU activation

2.1.4.2 Softmax

Softmax function calculates the probabilities distribution of the event over 'n' different events. In a general way of saying, this function will calculate the probabilities of each target class over all possible target classes. Later the calculated probabilities will be helpful for determining the target class for the given inputs. Mathematically, it is defined as –

$$\sigma : \mathbb{R}^K \to (0,1)^K$$
$$\sigma(\mathbf{z})_j = \frac{e^{z_j}}{\sum_{k=1}^{K} e^{z_k}} \quad \text{for } j = 1, \ldots, K.$$

Where K is the dimension of input vector z. Resultant vector σ(Z) has values between (0, 1) whose sum is equals to one. We use softmax activation on the very last layer of our network to get the pseudo-probabilities that represent the confidence of a pixel belonging to a particular class.

2.1.5 Batch Normalization Layer

This layer adds a normalization step (shifting inputs to zero-mean and unit variance) to make the inputs of each trainable layers comparable across features. By doing this it ensures a high learning rate while keeping the network learning. This layer quickly became very popular mostly because it helps to converge faster. Also, it allows activations functions to not get stuck in the saturation mode (e.g. gradient equal to 0).

2.1.6 Training

Training phase of a CNN involves calculation of the loss term and back propagation of the loss through the entire network. The loss term represents the error in prediction made by the CNN on an input. The gradients computed on each layer represent the contribution of that layer to the final loss term. When back propagating, all the trainable parameters are updated according to their gradients. When this is repeated on all training inputs for several epochs, the parameters will be updated in a way that they approximate a non-linear function that models the task at hand.

2.2 Segmentation Using CNNs

Image segmentation using CNNs is a classification task on a pixel level. Fully Convolutional Networks (FCN)(CNNs with all convolutional layers) by Long et al. from Berkeley, popularized CNN architectures for dense predictions. This allowed segmentation maps to be generated for image of any size. Almost all the subsequent state of the art approaches on segmentation adopted this paradigm. One of the main problems with using CNNs for segmentation is pooling layers. Pooling layers increase the field of view and are able to aggregate the context while discarding the 'where' information. However, segmentation requires the exact alignment of class maps and therefore, needs the 'where' information to be preserved. Encoder-

Decoder style network architecture was proposed to tackle this issue. Encoder gradually reduces the spatial dimension with pooling layers and decoder gradually recovers the object details and spatial dimension. There are usually shortcut connections from encoder to decoder to help decoder recover the object details better.

2.2.1 3D UNet

3D UNet was originally proposed by Cicek et al.[23] for automatic segmentation of Xenopus (a highly aquatic frog) kidney. It has an encoder-decoder style architecture with skip connections between corresponding layers in encoding and decoding paths. This architecture is very popular for medical image segmentation. All the deep learning models used in this study have the same architecture, the 3D UNet. 3D in the name indicates that the input to this network is a 3D image. UNet refers to the structure of the network, which resembles the letter 'U'. Figure 12 shows the block representation of 3D UNet architecture.

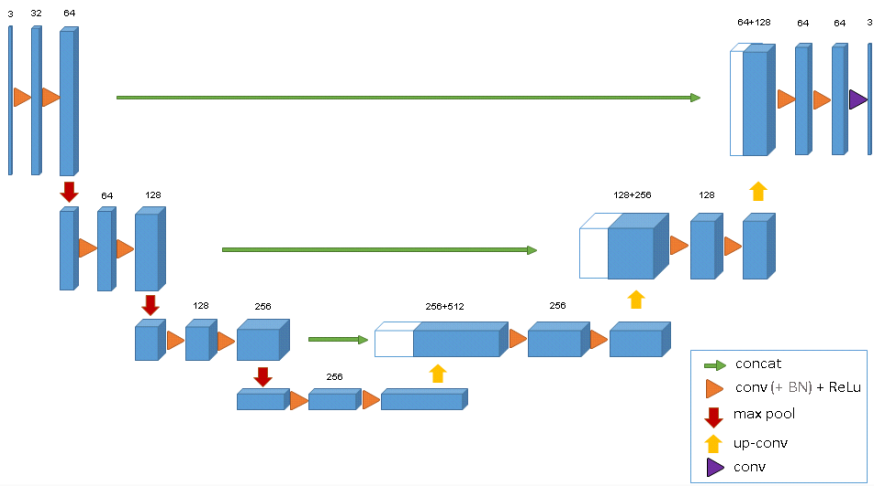


Figure 12: The 3D u-net architecture. Blue boxes represent feature maps. The num- ber of channels is denoted above each feature map.

**Taken from 'Learning Dense Volumetric Segmentation from Sparse Annotation'*

Each convolutional block has two convolutions followed by max pooling. Every convolution is immediately followed by ReLU activation and batch normalization layer. Each deconvolutional block consists of two convolutions followed by a deconvolution to regain spatial dimension. Moreover, there are skip connections from the encoding path to decoding path at corresponding spatial dimensions. These are

shown by green arrows. The very final convolution (shown by a purple arrow) that generates a three-dimensional feature map is followed by a softmax activation in order to obtain a pseudo-random probability distribution at each pixel representing its class membership. All the deep learning models used in this work have the UNet architecture.

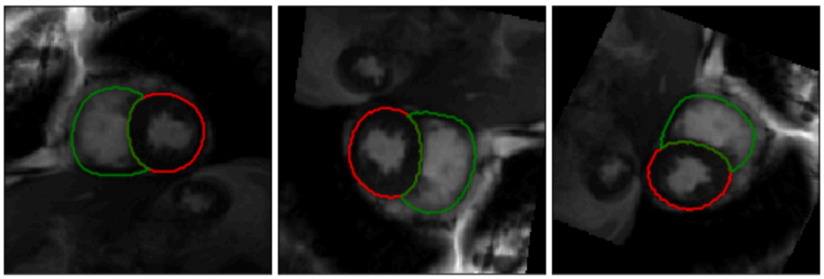
Chapter 3

HCM Registry

The data from HCM Registry[24], a National Heart Lung and Blood Institute-sponsored 2750-patient, 44-site, international registry and natural history study designed to address limitations in extant evidence to improve prognostication in HCM, is used in this study. We only use 100 datasets from this registry of which 60% were used for training and the rest for testing. Images acquired using cine and T1 protocols are used in this study.

3.1 Cine Data

Cine data has images for all cardiac phases on slices covering the entire heart. Ground truth segmentation masks for both epi and endocardial LV and RV are available at end-diastole phase and only endocardial LV and RV are available at end-systole phase. Data preprocessing includes building 4D image matrices from DICOM. 4D images represent a 3D heart in time, i.e. different cardiac phases. Building 4D matrices requires the parsing of information from DICOM meta data to find the appropriate slice location and phase number for each 2D image that it contains. These images are later center cropped from 256x256 to 154x154 to reduce the overall size and increase training speed. As the region of interest is always at the center of the image in CMR, we can be certain that such cropping wouldn't exclude it. Also, resizing the images to 32 slices allows all of them to have uniform size while being fed into the deep learning model. Intensity normalization by subtracting the mean and dividing with standard deviation of the image.

a) 

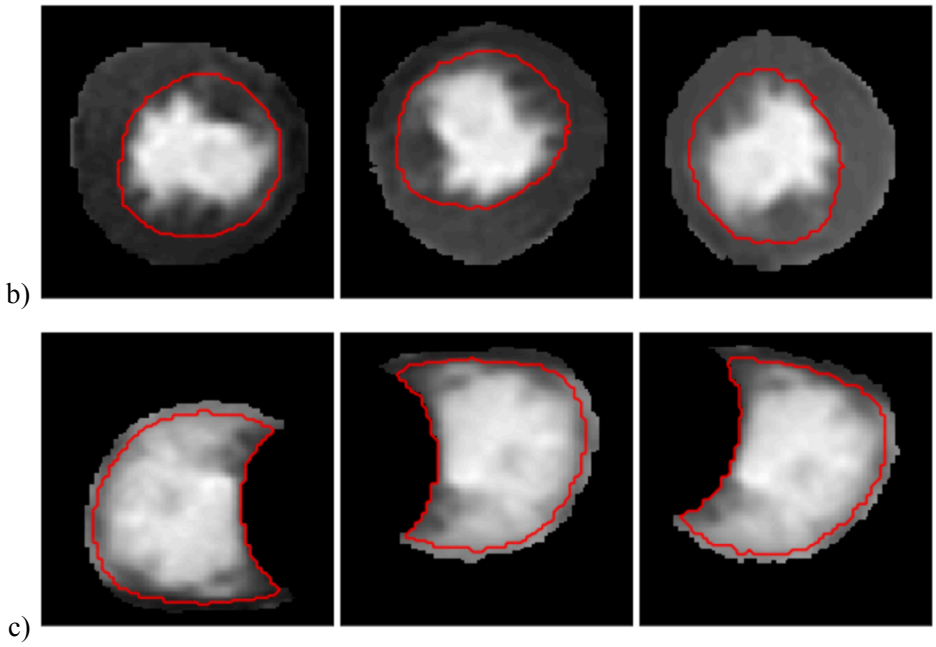


Figure 13: Examples of cine data and its augmentation

During training, random 3D affine transformation including translation, rotation, scaling and shear was used on the fly to augment the dataset. Figure 13a shows some 2D images representing one slice of the heart along with some introduced augmentations for epicardial boundaries of LV and RV, Figure 13b shows the same but for LV endocardium and Figure 13c for RV endocardium. The first column is the source image and columns 2 and 3 are augmentations on the source image.

3.2 T1 Data

Data used for T1 segmentation is also part of the HCM Registry. We use the native T1 maps to segment out the left ventricle. The LV epicardial and endocardial ground truth contours are available for all corresponding images. Data preprocessing is fairly similar to that of cine data. As T1 maps are only acquired at basal, mid-cavity and apical positions rather than on all slices that cover the entire heart, we use 2D versions of the augmentations used for cine data to warp each 2D image. Figure 14 shows native T1 maps at basal, mid-sequence and apical slices along with the epi and endocardial (red and green respectively) contours. The corresponding 2D augmentations are shown in Figure 15.

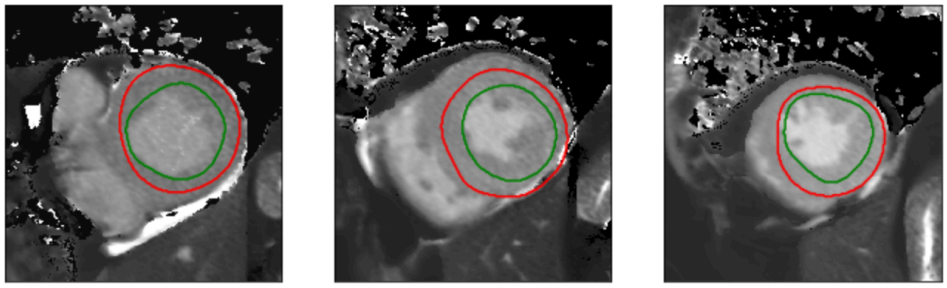


Figure 14: Native T1 maps at basal, mid-sequence and apical slices

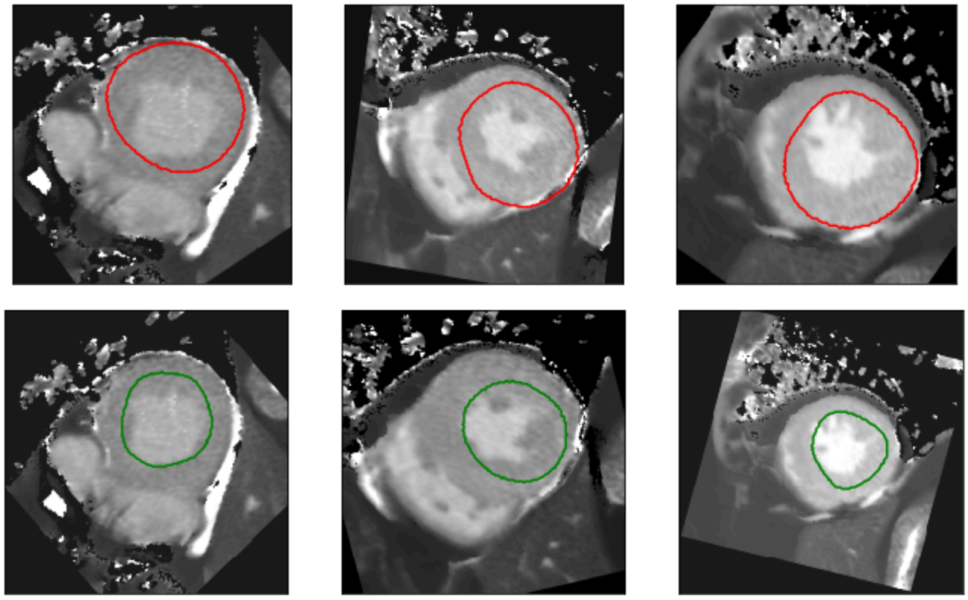


Figure 15: Augmentation results on epi and endo contours shown in Figure 14

Chapter 4

Approach

This chapter describes the fully automatic approach that we developed to achieve accurate segmentations on both cine and native T1 maps. Also, automatic algorithms for biomarker quantification based on the segmentations are detailed along with evaluation metrics that were used to assess the segmentation quality and biomarker quantification accuracy.

4.1 Cardiac Chamber Segmentation

A cascaded deep learning based approach was developed to accurately segment the heart chambers and thereby automate the quantification of HCM biomarkers. First, accurate segmentations for LV and RV epicardium are obtained with one network. Results from the epi segmentation are then used to obtain tight bounding boxes that encloses the LV and RV chamber, respectively. Separate models are trained for endocardium segmentation for each chamber. Input images are also masked by the results from epi segmentation so that the network can focus on the inside of the chamber. Figure 16 shows a workflow of this approach. The following sub sections further discuss each step.

4.1.1 DCNN Architecture and Loss Metric

A 3D-UNet style architecture is used for the segmentation in both steps. Convolutions in the encoding phase of 3D-UNet were replaced with dilated convolutions in order to increase the receptive field without having to train more parameters. In medical images, the ground truth masks are dominated by background leading to an overwhelming class imbalance, especially when there are multiple foreground classes. This can be addressed by applying a weight map to the categorical-cross entropy loss function or by using a dice-similarity metric based loss function. The latter is usually preferred as it does not rely on hyper parameters. With multiple foreground labels, the loss is given as-

$$Loss = 1 - \frac{\sum_{i=1}^{n} Dice_i}{n}$$

where n is the number of classes, excluding background and each individual dice is calculated using Sorenson's coefficient.

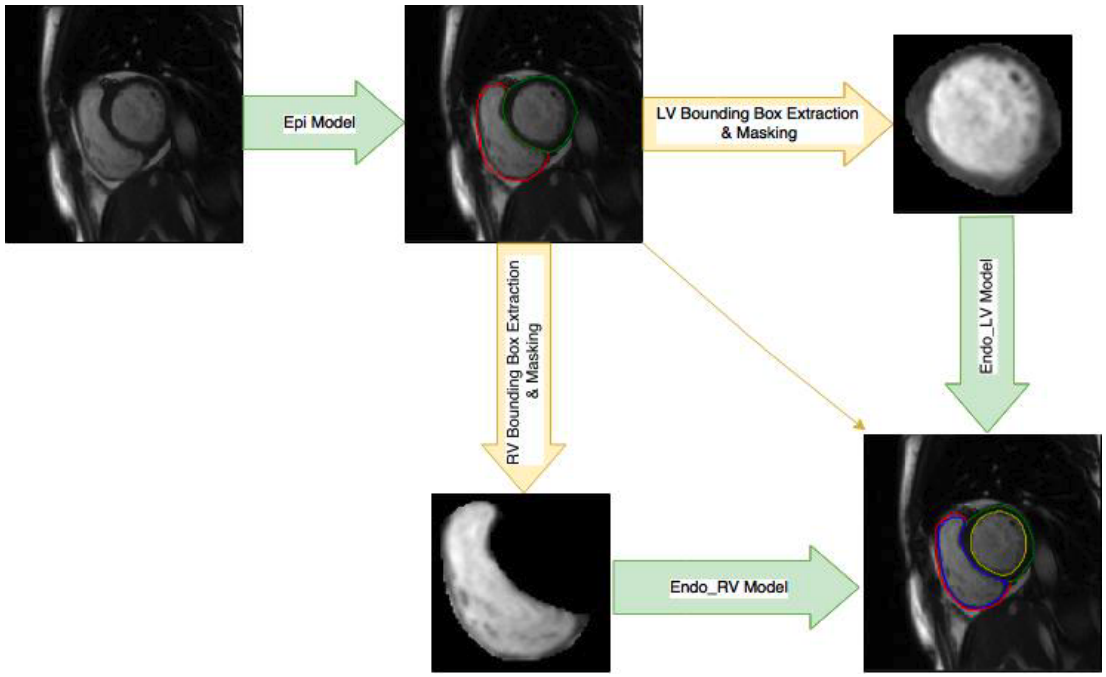


Figure 16: Workflow of the cascaded DCNN approach for end-diastole cardiac chamber

4.1.2 Epicardium Segmentation - Diastole

A single model is used to segment both the LV and RV epicardium with three labels of output: background, LV chamber and RV chamber. During training, random 3D affine transformation including translation, rotation, scaling and shear was used on the fly to augment the dataset. Adam optimizer is used with a learning rate of 5e-4 to minimize the loss. The model was trained for 800 epochs on a TitanX GPU for 8 hrs. For both LV and RV, only one fully connected component is retained during post-processing.

4.1.3 Endocardium Segmentation - Diastole

As the endocardium is always inside the epicardium, to maximize the efficiency of the network by focusing on the myocardium-blood contrast, two separate models are trained for LV and RV endocardial segmentation on masked images. The segmentation task is now a pixel-wise binary classification problem. During training, ground truth epicardium masks are used to obtain a tight bounding box that encloses the contour. The images are then cropped accordingly and pixels outside the epicardium are also masked out. As a comparison, we also trained models without images to check their contribution to the performance. The cropped images are then resized to 128x128x32 dimensions to ensure that all inputs to the model have the same size. Similar affine augmentations is performed as well. During testing, results from the corresponding epi segmentation generated by the model are used for masking and bounding box extraction. Adam optimizer was used with a learning rate of 5e-4 to minimize the loss on both models which were trained for 400 epochs on a TitanX GPU for 4 hrs each.

4.1.4 Endocardium Segmentation – Systole

Although training augmentation of the images at end-diastole phase can mimic end-systole images for general cardiac MRI, we found that the model trained with end-diastole images performs poorly on end-systole images for HCM patients since the shape and contrast is drastically different between the two phases. In some extreme cases where wall thickening is severe, no blood signal is visible, yet the chamber boundary still should be drawn. As ground truth epi masks are not available at systole, the cascaded approach can't be used. Individual models were then used for the endo segmentation of LV and RV on the original images. Adam optimizer was used with a learning rate of 5e-4 to minimize the loss on both models which were trained for 200 epochs on a Titan X GPU for 4 hours in total.

4.1.5 T1 Segmentation

Segmentation of T1 maps is fairly straight forward and highly similar to the strategy we followed for systole endocardium. We train two different models, one to segment LV epicardium and the other to segment LV endocardium. While segmenting the endocardium, we mask the input image with the results of epicardium segmentation. Adam optimizer was used with a learning rate of 5e-4 to minimize the loss on both models which were trained for 200 epochs on a Titan X GPU for 4 hours in total.

4.2 Biomarker Quantification

4.2.1 Wall Thickness

Wall thickness calculations are made following the 17-segment model recommended by the AHA with slight modifications to be able to compare the results with available data. All slices including apical are automatically divided into 6 segments. Figure 2 shows the division of left ventricle into six segments of 60° each on a mid-ventricular slice. Segments 2 and 3 are first found by identifying the septum, which is the overlap of LV and RV epi contours. The exact boundaries of the two segments are symmetrically adjusted to make sure the angle is 120°. Segments 1, 6, 5, 4 can be found by dividing the remaining area into four equal parts. Myocardium thickness is calculated at the beginning of each segment – represented by solid white lines in Figure 17. The same angles are used for all other slices for myocardium division.

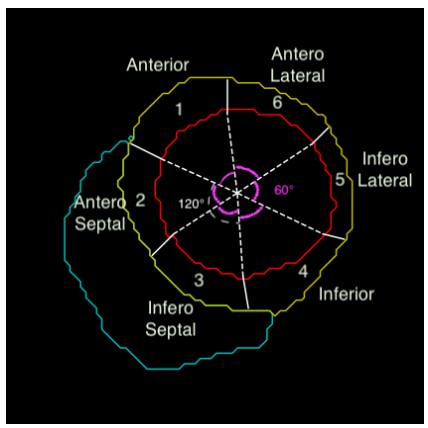


Figure 17: Division of the left ventricle into six equal segments – anterior, antero septal, infero septal, inferior, infero lateral, antero lateral.

4.2.2 Ejection Fraction, LV & RV Mass

Simpson's rule is used to calculate endocardial volumes at end-diastole (EDV) and end-systole (ESV). Ejection fraction (EF) can be found by the below formula. EFs of both LV and RV are calculated.

$$EF = \frac{EDV - ESV}{EDV} * 100$$

Mass calculations require epicardial volumes in addition to endocardial volumes. As epi contours are not available at end-systole, mass is only calculated at end-diastole. Myocardial volume is calculated as the difference of epi and endocardial volumes. Mass is calculated as the product of myocardial volume and density, which was assumed to be constant at 1.05g/cc.

4.2.3 Mean Myocardial T1

Changes in myocardial T1 can be a very helpful biomarker in identifying risk associated with HCM. It is calculated by taking the average of all the pixel values that lie in the myocardium, which is identified from the corresponding segmentation masks on T1 maps.

4.3 Evaluation

Dice scores and average perpendicular distance are calculated to evaluate the segmentation quality. To demonstrate the necessity of different models for HCM patients, we also trained a model for LV epi and endocardium segmentation using the SunnyBrook dataset, which contains normal and other patient populations, and tested the model on HCM patients. Symmetric mean absolute percentage error and root mean squared error values were calculated to evaluate the biomarker quantification results.

4.3.1 Dice Score

The dice score is a statistic used for comparing the similarity of two samples. When applied to boolean data, using the definition of true positive (TP), false positive (FP), and false negative (FN), it can be written as –

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

The value of dice score ranges from 0 to 1 with 0 being complete mismatch and 1 being perfect overlay. Figure 18 shows a visual representation of dice metric.

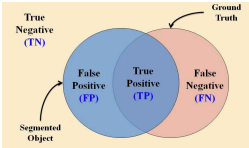


Figure 18: Visual representation of dice score

4.3.2 Average Perpendicular Distance

The average perpendicular distance (APD) measures the distance from the automatically segmented contour to the corresponding manually drawn expert contour, averaged over all contour points. A high value implies that the two contours do not match closely. In general, an APD value less than 5mm is considered a good contour. Figure 19 shows a visual representation of APD.

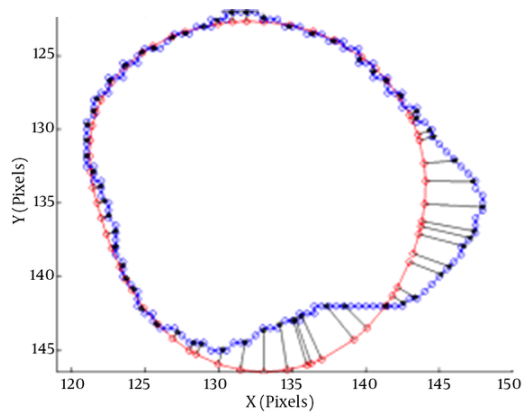


Figure 19: Visual representation of point pairs for APD calculation

Considering a pixel spacing of 1mm, the APD for the above example is 1.426mm, maximum distance is 5.45mm and minimum distance is 0mm.

4.3.3 Symmetric Mean Absolute Percentage Error

Symmetric mean absolute percentage error (sMAPE) is an accuracy measure based on percentage (or relative) errors. It is used to evaluate the quantification of biomarkers. For a set of actual values A and predicted values P, sMAPE is given by - $sMAPE = \frac{100\%}{n} \sum_{t=1}^{n} \frac{|P_t - A_t|}{(|P_t| + |A_t|)}$

4.3.4 Root Mean Squared Error

The root mean squared error (RMSE) represents the sample standard deviation of the differences between predicted values and observed values. It is calculated as the square root of average of squared errors.

Chapter 5

Results & Discussion

5.1 Cardiac Chamber Segmentation

When using the model trained on the Sunnybrook dataset, a dice score of 0.368 on endo_LV and 0.685 on epi LV were achieved while the scores were 0.92 and 0.95 on its own validation dataset. Upon visual inspection of the results, the model failed to locate the endo_LV ROI on many cases, causing the dice score to be 0. Figure 20 shows some results from this experiment.

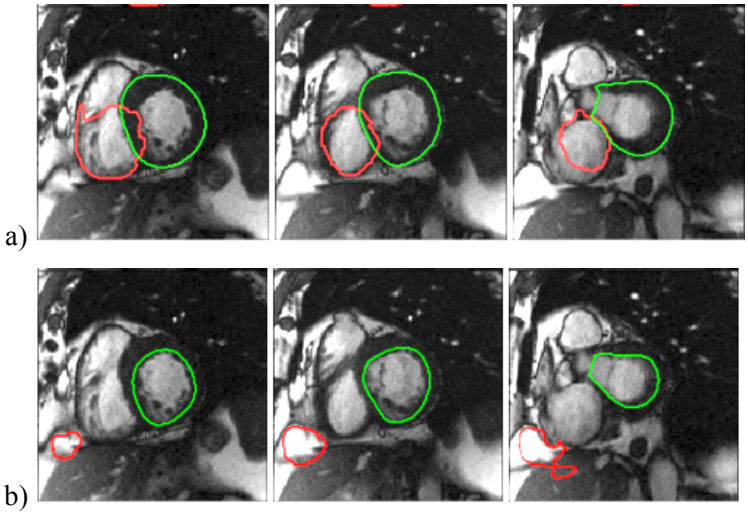


Figure 20: Prediction errors when a model trained on sunnybrook data is deployed on HCM data, Green contours are ground truth and red are predictions. a) result on epicardium, Dice: 0.04; b) result on endocardium, Dice: 0.

When using the cascaded approach trained on end-diastole images to test on end-systole images, the dice scores of 0.45 on endo_LV and 0.36 on endo_RV were obtained, justifying the need to train separate networks for end-systole segmentation. When masking on the images for endo segmentation was not used, the dice score was 0.78 on endo_RV and 0.90 on endo_LV. With masked image, the dice score of RV

increased to 0.81. Table 1 summarizes the final results from our automatic segmentation models for cine data. In general, an APD less than 5mm is considered good segmentation for large objects.

Table 1. Results from the cascaded approach used at end-diastole and separate networks used for endocardium segmentation at end-systole.

ROI	***Dice***	***APD (mm)***
Endo_LV – Dia	0.902 ± *0.025*	2.25 ± *0.58*
Epi_LV – Dia	0.927 ± *0.025*	2.21 ± *0.73*
Endo_RV – Dia	0.813 ± *0.045*	2.98 ± *0.89*
Epi_RV – Dia	0.820 ± *0.040*	3.08 ± *0.88*
Endo_LV – Sys	0.752 ± *0.071*	3.96 ± *1.82*
Endo_RV – Sys	0.613 ± *0.197*	6.23 ± *4.88*

In an attempt to understand the model failures, we identified slices with poor segmentation (dice score < 0.80). The bar chart in Figure 21 shows the percentage of bad slices at basal, mid ventricular and apical locations for End-Diastole. These locations were identified by choosing the first, middle and the last slices from the 3D ground truth mask containing contours.

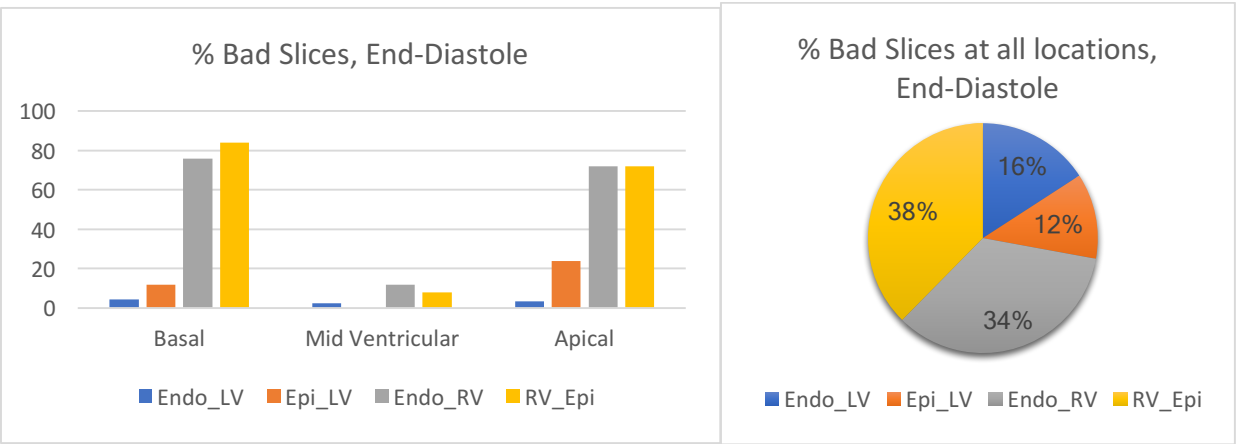


Figure 21: Percentage of bad slices at basal, mid ventricular and apical locations is shown in the bar chart. RV_Epi and RV_Endo have relative more issues at basal and apical locations. The Pie chart shows the contribution of each ROI to the total number of bad slices among all slice locations

The RV segmentation is performing poorly at basal and apical locations and is significantly better at the mid ventricular location. As the endocardial segmentation is contingent on the accuracy of epicardial segmentation, the former is expected to perform poorly when the latter fails, which explains the relatively

high percentage of RV_Endo bad. LV segmentation is following the same trends but with better performance. Among slices at all locations, approximately, 25% of them were bad. The pie chart in Figure 21 shows the contribution of each ROI to the total number of bad slices identified. This represents the overall performance of the model in segmenting these ROIs, irrespective of their location. The model is doing a better job at segmenting the LV relative to RV. This is expected, given the complex crescent shape of the RV.

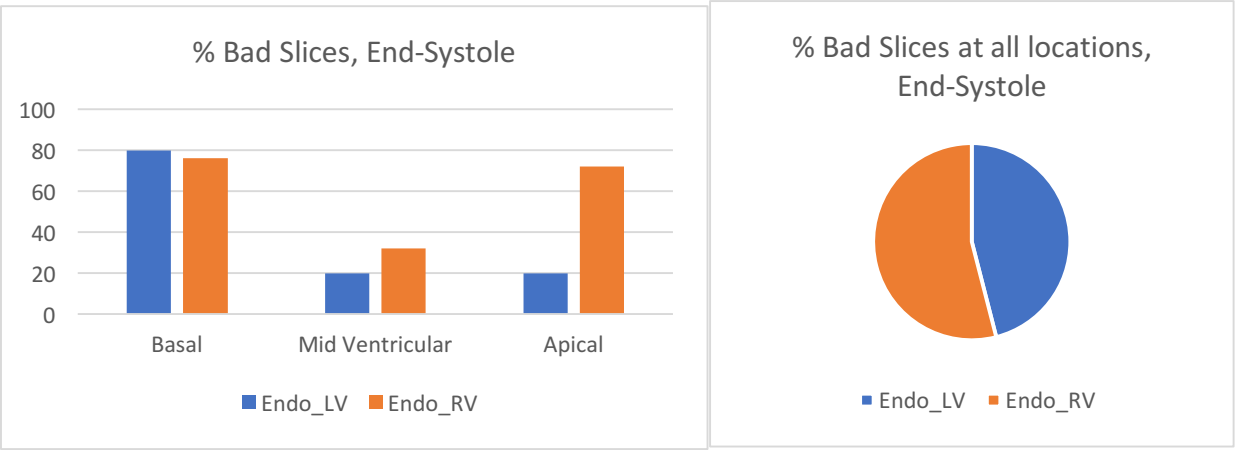
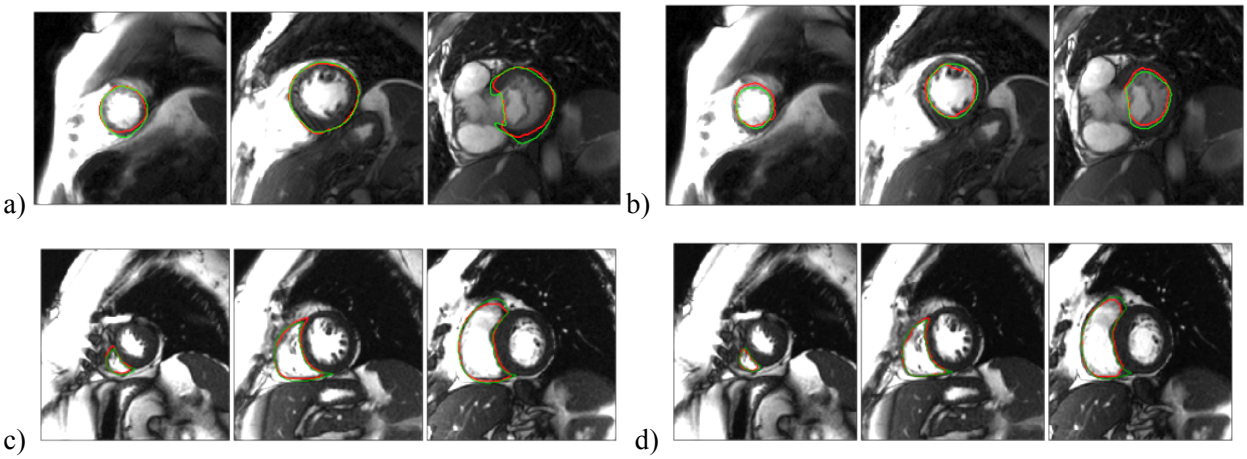


Figure 22: Percentage of bad slices at basal, mid ventricular and apical locations is shown in the bar chart. The Pie chart shows the contribution of Endo_LV and Endo_RVto the total number of bad slices among all slice locations.

Figure 22 shows the same information for End-Systole phase. The model is performing rather poorly in identifying the basal and apical slices for Endo_RV. The percentage of bad slices at end-systole was 42%. The segmentation results for different ROIs are shown in Figure 23. It can be observed that a good epicardial segmentation allows for a good endocardial segmentation.



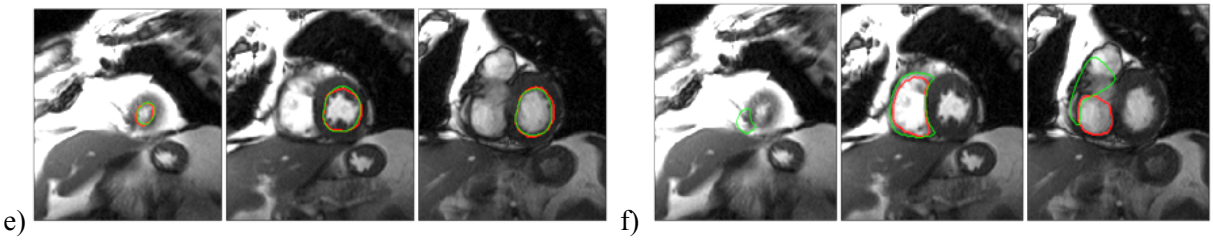


Figure 23: Segmentation results on apical, mid ventricular and basal slices. Green contour is the ground truth and red is the prediction a) Epi_LV, diastole, Dice:0.95; b)Endo_LV, diastole, for images in 'a' , Dice:0.92; c)Epi_RV, diastole,Dice:0.91; d)Endo_RV, diastole, for images in 'c', Dice:0.91; e)Endo_LV, systole, Dice:0.93; f)Endo_RV, systole, Dice:0.73

Dice scores of 0.91 and 0.84 were achieved on epicardial and endocardial regions of native T1 maps. As the native T1 contrasts are very poor, it is often difficult for a human segmenter to accurately delineate the endocardium, which leads to higher variability. Figure 24 shows some segmentation results on native T1 maps.

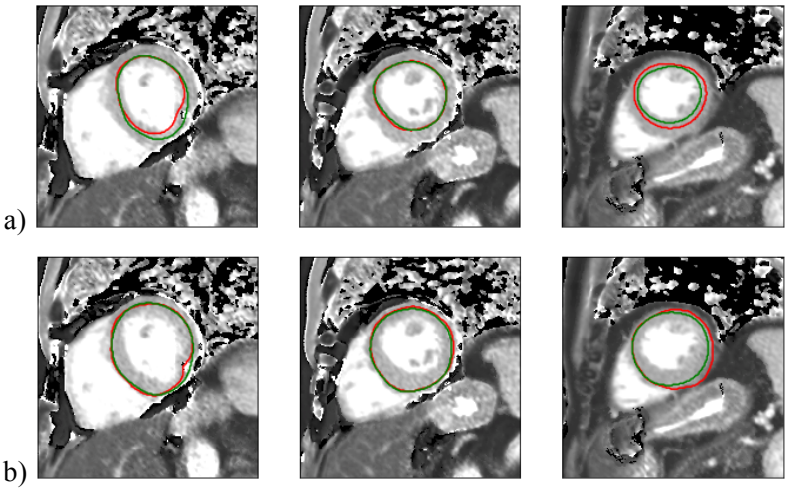


Figure 24: Native T1 map segmentation results. Ground truth contour is in green and prediction in red. a)Endocardium, Dice:0.92; b)Epicardium, Dice:0.95

5.2 Quantification of Biomarkers

Table 2 summarizes the results of biomarker quantification. The 'Mean_Value' column shows the average value of the quantified biomarker on test data. For comparison, root mean square error (RMSE) values from model predictions and an inter-observer study on generic cardiac MRI data are reported [15]. Wall thickness measurements are only done at end-diastole as no epi contours are available for end-systole.

Table 2. Quantification results for LV, RV Mass, Ejection Fraction, Mean myocardial T1 and the End-Diastole wall thickness.

Biomarker	***Mean_Value***	***sMAPE in %***	***Model_RMSE***	***InterObserver_RMSE (Generic Cardiac MRI)***
LV Mass	158.8 gm	13.7 ± 8.9	52.4 gm	17.5 gm
RV Mass	23.9 gm	32.2 ± 18.9	16.4 gm	N/A
LV Ejection Fraction	64.5 %	5.8 ± 4.5	9.5%	4.2%
RV Ejection Fraction	58.0 %	9.2 ± 5.9	21.6%	N/A
Wall Thickness	6.2 mm	20.8 ± 8.1	2.97 mm	N/A
Mean Myocardial T1	916 msec	2.91 ± 2.96	54.5 msec	N/A

The inter observer RMSE values reported in Table 2 are from a population that is representative of healthy patients[25]. RMSE values on HCM population are expected to be higher given the increased variability in the size and shape of heart chambers. Higher errors in RV related values are a result of poor segmentation in comparison with LV. Moreover, the poor performance in basal slice segmentation contributes significantly to the errors in mass and ejection fraction calculations. For myocardial T1, in general the values are around 1000msecs. RMSE of 54.5 msec and sMAPE of 2.9% indicates a robust quantification. Time taken for automatic quantification using the methods we developed is approximately 10sec for each subject.

Conclusion

We have successfully developed a cascaded DCNN approach for cardiac segmentation and quantification of biomarkers in HCM. The underlying aim of the cascaded approach was to achieve improved results on endo when we have an accurate segmentation for epi, since epicardium segmentation usually performs better due to more distinct contrast. Our results concur with this aim; endo APD and dice scores are very close to epi values at end-diastole. Upon visual inspection, it was found that the cascaded approach works perfectly on the mid-cavity slices, but has trouble when segmenting the very apical and very basal slices where the ROI vanishes suddenly. The accurate delineation of LV and RV is very challenging and ambiguous even for cardiologists, especially in RV, which explains the relatively poor RV mass calculations. Even though the inter-observer variability values on HCM data are likely to be higher than on a generic cardiac MRI data, the large differences of the latter in comparison with model prediction signifies the need for improvement. In the future, we will continue to improve the model focusing on basal and apical slices and use more data for training. In addition, we will build models to automatically quantify biomarkers from post contrast T1 and LGE images, which are also acquired in the study protocol to capture more information for HCM diagnosis and prognosis.

Bibliography

1. Petitjean C, Dacher J-N. A review of segmentation methods in short axis cardiac MR images. *Med Image Anal*. 2011;15(2):169-184.
2. Kedenburg G, Cocosco CA, Köthe U, Niessen WJ, Vonken EPA, Viergever MA. Automatic cardiac MRI myocardium segmentation using graphcut. In: *Medical Imaging 2006: Image Processing*. Vol 6144. ; 2006:61440A.
3. Cocosco CA, Niessen WJ, Netsch T, et al. Automatic image-driven segmentation of the ventricles in cardiac cine MRI. *J Magn Reson Imaging*. 2008;28(2):366-374.
4. Jolly M. Fully automatic left ventricle segmentation in cardiac cine MR images using registration and minimum surfaces. *MIDAS Journal-Cardiac MR Left Vent Segmentation Chall*. 2009;4:59.
5. Huang S, Liu J, Lee LC, et al. An image-based comprehensive approach for automatic segmentation of left ventricle from cardiac short axis cine mr images. *J Digit Imaging*. 2011;24(4):598-608.
6. Billet F, Sermesant M, Delingette H, Ayache N. Cardiac motion recovery and boundary conditions estimation by coupling an electromechanical model and cine-MRI data. In: *International Conference on Functional Imaging and Modeling of the Heart*. ; 2009:376-385.
7. Pluempitiwiriyawej C, Moura JMF, Wu Y-JL, Ho C. STACS: New active contour scheme for cardiac MR image segmentation. *IEEE Trans Med Imaging*. 2005;24(5):593-603.
8. Zhang H, Wahle A, Johnson RK, Scholz TD, Sonka M. 4-D cardiac MR image analysis: left and right ventricular morphology and function. *IEEE Trans Med Imaging*. 2010;29(2):350-364.
9. Van Assen HC, Danilouchkine MG, Frangi AF, et al. SPASM: a 3D-ASM for segmentation of sparse and arbitrarily oriented cardiac MRI data. *Med Image Anal*. 2006;10(2):286-303.
10. Zhuang X, Hawkes DJ, Crum WR, et al. Robust registration between cardiac MRI images and atlas for segmentation propagation. In: *Medical Imaging 2008: Image Processing*. Vol 6914. ; 2008:691408.
11. Margeta J, Geremia E, Criminisi A, Ayache N. Layered spatio-temporal forests for left ventricle segmentation from 4D cardiac MRI data. In: *International Workshop on Statistical Atlases and Computational Models of the Heart*. ; 2011:109-119.
12. Lempitsky V, Verhoek M, Noble JA, Blake A. Random forest classification for automatic delineation of myocardium in real-time 3D echocardiography. In: *International Conference on Functional Imaging and Modeling of the Heart*. ; 2009:447-456.

13. Geremia E, Clatz O, Menze BH, Konukoglu E, Criminisi A, Ayache N. Spatial decision forests for MS lesion segmentation in multi-channel magnetic resonance images. *Neuroimage*. 2011;57(2):378-390.
14. Cordero-Grande L, Vegas-Sánchez-Ferrero G, Casaseca-de-la-Higuera P, et al. Unsupervised 4D myocardium segmentation with a Markov Random Field based deformable model. *Med Image Anal*. 2011;15(3):283-301.
15. Cobzas D, Schmidt M. Increased discrimination in level set methods with embedded conditional random fields. In: *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference On*. ; 2009:328-335.
16. Dreijer JF, Herbst BM, Du Preez JA. Left ventricular segmentation from MRI datasets with edge modelling conditional random fields. *BMC Med Imaging*. 2013;13(1):24.
17. Ngo TA, Carneiro G. Fully Automated Non-rigid Segmentation with Distance Regularized Level Set Evolution Initialized and Constrained by Deep-Structured Inference. In: *CVPR*. ; 2014:3118-3125.
18. Yang XL, Gobeawan L, Yeo SY, Tang WT, Wu ZZ, Su Y. Automatic segmentation of left ventricular myocardium by deep convolutional and de-convolutional neural networks. In: *Computing in Cardiology Conference (CinC), 2016*. ; 2016:81-84.
19. Avendi MR, Kheradvar A, Jafarkhani H. A combined deep-learning and deformable-model approach to fully automatic segmentation of the left ventricle in cardiac MRI. *Med Image Anal*. 2016;30:108-119.
20. Tran PV. A fully convolutional neural network for cardiac segmentation in short-axis MRI. *arXiv Prepr arXiv160400494*. 2016.
21. Avendi MR, Kheradvar A, Jafarkhani H. Automatic segmentation of the right ventricle from cardiac MRI using a learning-based approach. *Magn Reson Med*. 2017;78(6):2439-2448.
22. Yu F, Koltun V. Multi-scale context aggregation by dilated convolutions. *arXiv Prepr arXiv151107122*. 2015.
23. Bui TD, Shin J, Moon T. 3d densely convolution networks for volumetric segmentation. *arXiv Prepr arXiv170903199*. 2017.
24. Kramer CM, Appelbaum E, Desai MY, et al. Hypertrophic Cardiomyopathy Registry: The rationale and design of an international, observational study of hypertrophic cardiomyopathy. *Am Heart J*.

2015;170(2):223-230.

25. Suinesiaputra A, Bluemke DA, Cowan BR, et al. Quantification of LV function and mass by cardiovascular magnetic resonance: multi-center variability and consensus contours. *J Cardiovasc Magn Reson*. 2015;17(1):63.