



# **Medical image-based diagnostics for cardiovascular diseases using Machine Learning**

By

Alexander M. Martinez Lopez

Carlos F. Gonzalez Rivera

## **RESEARCH FINAL REPORT**

**Advisor:** Edwin Florez Gomez

POLYTECHNIC UNIVERSITY OF PUERTO RICO

UNDERGRADUATE RESEARCH PROGRAM FOR HONOR STUDENTS

HSI STEM GRANT

SAN JUAN, PUERTO RICO

2020



## **Abstract**

With cardiac imaging's important role in the diagnosis of cardiovascular diseases, along with the dawn of big data and machine learning (ML), there are emergent opportunities to build artificial intelligence (AI) tools that will directly assist physicians in heart failure (HF) diagnostics. An important application in biomedical engineering, as HF is very difficult to diagnose because of its complex symptoms, circumstances, and comorbidities. This study aims to: (1) perform accurate and precise cardiac segmentation and quantification of key left ventricle functional indices from CMR; and (2) build a ML tool using decision trees for image-based HF diagnosis. The quantified indices left ventricular end-diastolic, end-systolic volumes and ejection fraction were achieved using Heron's formula and the area-length method. One-sample T tests revealed there were no statistical significance between the obtained mean values and the comparative mean values in each quantified variable. The obtained strong p-values show the quantified values closely resemble those established in the Sunnybrook Cardiac Data. Finally, a Machine Learning tool using decision trees for image-based heart failure diagnosis was successfully built, as every tested patient was classified correctly using the trained ML model, achieving a 100% accuracy.

## Table of Contents

1	Introduction.....	1
2	Literature Review.....	3
2.1	Heart Failure .....	3
2.1.1	Heart failure Diagnosis .....	4
2.2	Cardiovascular Medical Images.....	6
2.3	Cardiac Segmentation .....	6
2.3.1	Active Contour.....	7
2.4	Indices of Cardiac Function.....	8
2.5	Machine Learning .....	9
2.5.1	Heart Failure Diagnosis using Machine Learning .....	10
2.5.2	Inputs, Outputs and Other Features .....	11
2.5.3	ML Technique.....	11
2.5.3.1	Decision Trees .....	12
3	Research Program .....	14
3.1	Objectives and Scope.....	14
3.2	Methodology .....	15
3.2.1	Image Data Acquisition .....	15
3.2.2	Metadata Acquisition & Image Processing.....	15
3.2.3	Cardiac Segmentation .....	16
3.2.4	Quantification of Functional Indices .....	17
3.2.5	Machine Learning Model & Diagnostic Predictions .....	20
3.3	Itinerary .....	20
4	Data.....	21
4.1	Segmentation Data .....	21
4.2	Quantification Data.....	28
4.3	Classification Data .....	30
5	Analysis and Results Interpretation .....	31
6	Conclusions, Recommendations and Future Work.....	35
6.1	Conclusions.....	35
6.2	Recommendations.....	35
6.3	Future Work .....	36
7	References.....	37

## List of Figures

Figure 1: Heart Failure Statistics .....	3
Figure 2: Common Symptoms and Treatments for Heart Failure .....	4
Figure 3: Cardiac Image Segmentation Tasks (Chen et al., 2020) .....	7
Figure 4: Broad Classification of Machine Learning Techniques. (Suryakanthi, 2020) ....	9
Figure 5: Machine Learning Techniques (Martin-Isla et al., 2020).....	12
Figure 6: Decision Tree Algorithm Model .....	13
Figure 7: SAX Basal slices (left) & LAX 4-Chamber View (right) .....	15
Figure 8: Histogram Equalization on SAX Basal slices .....	16
Figure 9: Endocardial Segmentation.....	17
Figure 10: Usage of Heron's Formula.....	18
Figure 11: LV Length Measurement.....	18
Figure 12: Volume Quantification Diagram & Equation .....	19
Figure 13: Endocardial Segmentation for Patients 1 & 2 .....	21
Figure 14: Endocardial Segmentation for Patients 3 & 4 .....	21
Figure 15: Endocardial Segmentation for Patients 5 & 7 .....	22
Figure 16: Endocardial Segmentation for Patients 8 & 9 .....	22
Figure 17: Endocardial Segmentation for Patients 10 & 11 .....	22
Figure 18: Endocardial Segmentation for Patients 12 & 13 .....	23
Figure 19: Endocardial Segmentation for Patients 14 & 15 .....	23
Figure 20: Endocardial Segmentation for Patients 16 & 17 .....	23
Figure 21: Endocardial Segmentation for Patients 18 & 19 .....	24
Figure 22: Endocardial Segmentation for Patients 21 & 22 .....	24
Figure 23: Endocardial Segmentation for Patients 23 & 24 .....	24
Figure 24: Endocardial Segmentation for Patients 25 & 26 .....	25
Figure 25: Endocardial Segmentation for Patients 27 & 28 .....	25
Figure 26: Endocardial Segmentation for Patients 29 & 30 .....	25
Figure 27: Endocardial Segmentation for Patients 31 & 32 .....	26
Figure 28: Endocardial Segmentation for Patients 33 & 35 .....	26
Figure 29: Endocardial Segmentation for Patients 36 & 37 .....	26
Figure 30: Endocardial Segmentation for Patients 38 & 39 .....	27
Figure 31: Endocardial Segmentation for Patients 40 & 41 .....	27
Figure 32: Endocardial Segmentation for Patients 42, 43 & 44 .....	27
Figure 33: End-Diastolic Volume Mean Values by Group with Standard Deviation .....	29
Figure 34: End-Systolic Volume Mean Values by Group with Standard Deviation.....	29
Figure 35: Ejection Fraction Mean Values by Group with Standard Deviation.....	29
Figure 36: Machine Learning Decision Tree Model.....	30

## List of Tables

Table 1: New York Heart Association Functional Classification of Heart Failure .....	5
Table 2: The predicted values for normal LV parameters in males by age. ....	8
Table 3: The predicted values for normal LV parameters in females by age. ....	9
Table 4: SDC Group Statistics.....	19
Table 5: Gantt Chart Itinerary for the Scientific Research .....	20
Table 6: Functional Indices (EDV, ESV and EF) Quantification Data .....	28
Table 7: Functional Indices (EDV, ESV and EF) Quantification Statistical Data .....	28
Table 8: Classification Test Metrics .....	30

# 1 Introduction

Heart failure (HF) currently affects nearly five million Americans and contributes to approximately 287,000 deaths a year [1]. HF is a chronic, progressive, and serious cardiovascular condition that requires immediate medical care. This condition's pathophysiology is defined by the heart muscle being unable to pump enough blood to meet the body's needs for blood and oxygen [2]. Despite being considered a major public health burden and the many recent advances in cardiovascular diagnosis and treatment, misdiagnosis is still a common concern for patients at risk of heart disease. HF is specially one of the most frequently erroneously diagnosed conditions seen by primary care doctors [3,4].

Nowadays modern medical imaging techniques, such as magnetic resonance imaging (MRI), computed tomography (CT) and ultrasound are widely used to assess qualitatively and quantitatively cardiac anatomical structures and functions [5]. Such non-invasive assessments provide support for diagnosis, disease monitoring and treatment planning. Even with the incredible assessments medical images provide, HF patients are very difficult for physicians to diagnose correctly because they possess a variety of complex symptoms, circumstances, and comorbidities.

Emerging disciplines like big data and machine learning (ML) have the potential to facilitate HF diagnosis with the use of artificial intelligence (AI) tools. ML methods can be utilized to address a wide range of data-related problems, from simple analytic queries in heart failure symptomatology existing data to the more complex challenges involving in analyzing raw cardiac MRI images [6]. Such AI image-based diagnostic tools have the

potential to substantially alleviate the burden of HF through by providing fast and accurate diagnostic decision making. This scientific research focuses on developing and validating an AI image-based diagnostic tool for HF using estimated conventional imaging indices such as ventricular volumes in end diastole/systole and ventricular ejection fraction, acquired by semi-automatically contouring the left ventricle's endocardial boundaries from raw MRI images. The intent behind such computational scientific research is to add the body of knowledge, in hopes of reducing cardiovascular disease prevalence.

## 2 Literature Review

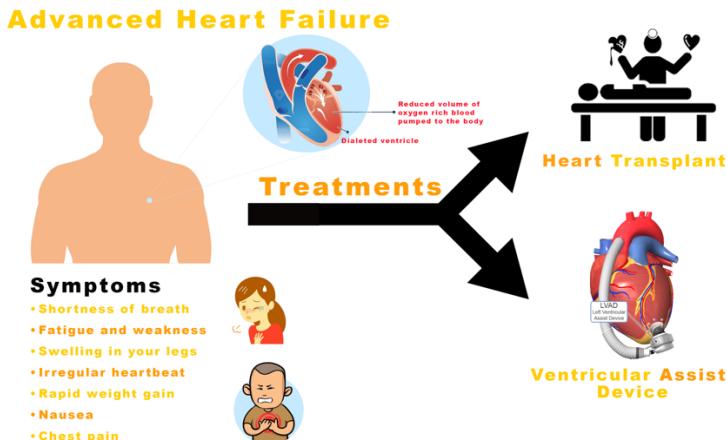
### 2.1 Heart Failure

Heart failure (HF) currently affects nearly five million Americans and contributes to approximately 287,000 deaths a year [1]. HF doesn't mean that your heart has stopped or is about to stop working. Rather it is a chronic, progressive condition in which the heart muscle is unable to pump enough blood to meet the body's needs for blood and oxygen [2]. HF has an estimated overall prevalence of 2.6 percent and is responsible for 9.4% of deaths attributed to cardiovascular disease in the U.S [1,7]. This severe condition is often particularized by having a high end-diastolic and end-systolic volumes, and a low ejection fraction. High end-diastolic volume is attributed to the heart muscle being overly stretched, becoming larger and floppy due to damage. Overall, it is a serious condition that requires medical care, **Figure 1** illustrates several eye-opening statistics that show the high prevalence of HF.



Figure 1: Heart Failure Statistics

Overall, heart failure is a serious condition that requires medical care. Most treatments for people who suffer from heart failure involves a balance of medications and, if needed, use devices that help the heart pump properly. **Figure 2** illustrates some of the common symptoms and treatments of HF.



**Figure 2: Common Symptoms and Treatments for Heart Failure**

### 2.1.1 Heart failure Diagnosis

The American Heart Association and American College of Cardiology define HF as “a complex clinical syndrome that can result from any structural or functional cardiac disorder that impairs the ability of the ventricle to fill with or eject blood [2]. Diagnosing such complex disorder requires information from clinical signs and symptoms, biomarkers, and imaging modalities. Requiring such a large amount of information to correctly diagnose HF makes it extremely hard for physicians to systematically identify the cause of said complex disorder. This leads to HF being one of the most frequently erroneously diagnosed conditions [3,4]. Further evidence of misdiagnosis was found on an observational study of 190 cases of diagnostic errors which found that one of the most encountered diseases was congestive heart failure presenting 5.7% of diagnostic errors [4].

The New York Heart Association (NYHA) Functional Classification provides a simple way of classifying the extent of heart failure [10], as seen on **Table 1**. It classifies patients in one of four categories based on their limitations during physical activity, breathing and varying degrees in shortness of breath and or angina pain.

**Table 1: New York Heart Association Functional Classification of Heart Failure**

New York Heart Association Functional Classification of Heart Failure	
Class	Patient Symptoms
I	No limitation of physical activity. Ordinary physical activity does not cause undue fatigue, palpitation, dyspnea (shortness of breath).
II	Slight limitation of physical activity. Comfortable at rest. Ordinary physical activity results in fatigue, palpitation, dyspnea (shortness of breath).
III	Marked limitation of physical activity. Comfortable at rest. Less than ordinary activity causes fatigue, palpitation, or dyspnea.
IV	Unable to carry on any physical activity without discomfort. Symptoms of heart failure at rest. If any physical activity is undertaken, discomfort increases.
Information from references [10]	

HF can be described by the ejection fraction (EF) or the percentage of blood leaving the heart each time it contracts. According to this measurement, HF can be categorized into two main phenotypes: HF with reduced ejection fraction (HFrEF) and HF with preserved EF (HFpEF) [11,12]. HFrEF or systolic dysfunction occurs when the heart pumps less oxygen-rich blood to the body because the muscle does not contract effectively. In HFrEF, the EF less than or equal to 40%. On the other hand, in HFpEF or diastolic dysfunction the heart muscle contracts normally. However, the heart muscle is overly stretched and floppy due to damage which causes higher ventricular filling. In HFpEF, the EF more than or equal to 41%.

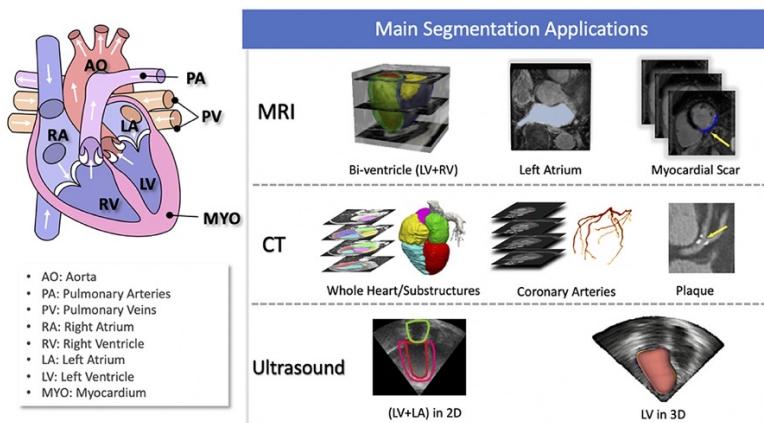
## **2.2 Cardiovascular Medical Images**

Noninvasive cardiac imaging is a combination of methods that can be used to obtain images related to the structure and function of the heart [13]. These tests are easy to perform, safe, and can be used to detect various heart conditions, such as abnormalities that impair the ability of the heart to pump blood. These imaging modalities hold the potential for progress in HF screening, risk stratification and monitoring. A wide variety of imaging tests are now available, including echocardiography (ECHO) and Cardiovascular magnetic resonance (CMR), where all of which together provide invaluable information on cardiac anatomy, physiology, and hemodynamics. Advancements in ECHO and CMR have allowed for improved tissue characterizations, cardiac motion analysis, and cardiac performance analysis under stress. However, CMR provides the gold standard for evaluating cardiac function and anatomy in 3-D-quality moving images. CMR offers a comprehensive assessment of heart failure patients and is now the gold standard imaging technique to assess myocardial anatomy, regional and global function, and viability [14]. Considering HF is one of the most frequently erroneously diagnosed conditions because of the systematic difficulty for physicians to identify the cause, imaging modalities and the emergence of big data and machine learning can potentially assist physicians in diagnosis and care.

## **2.3 Cardiac Segmentation**

Cardiac image segmentation is a method for partitioning the image into a number anatomically meaningful regions, based on which structural and functional indices are desired to be extracted, such as the myocardial mass, wall thickness, left ventricle (LV)

and right ventricle (RV) volume as well as EF. It is an important first step as calculating these CMR structural and functional indices need a delineation of the anatomical structures of interest. Regions of interest (ROI) in cardiac image segmentation include the LV, RV, left atrium (LA), right atrium (RA), and coronary arteries. Error! Reference source not found. gives an overview of task related to cardiac image segmentation based on imaging modalities, as described by Chen et al., 2020 [5].



**Figure 3: Cardiac Image Segmentation Tasks (Chen et al., 2020)**

### 2.3.1 Active Contour

Cardiac segmentation using an image-driven approach, use of no anatomical models, utilize techniques such as thresholding, clustering, pixel or voxel classification and active contour [15]. Active contour is the techniques particularly important for this scientific research. Active contour is one of the most utilized techniques in medical imaging applications, especially in heart segmentation [15]. This type of segmentation technique can be defined as use of energy forces and constraints for segregation of the pixels of interest from the image for further processing and analysis [16]. A particular active contour technique is called snake model. A snake is an energy-minimizing spline guided by external constraint forces and influenced by image forces that pull it toward features such

as lines and edges [17]. Essentially It uses a certain amount of prior knowledge about the target and locks onto nearby edges, localizing the ROI accurately.

## 2.4 Indices of Cardiac Function

Conventional imaging indices include measures commonly used in routine clinical image analysis such as end-diastolic/systolic ventricular volumes and ventricular EF. Estimation of these clinical indices is of great importance for cardiac structural and functional analysis from CMR. The LV is the ROI for this scientific research like many others, as it is the most investigated chamber in cardiac segmentation and quantification. Left Ventricle End-diastolic and End-systolic Volumes (LVEDV and LVESV) are measurements of the amount of blood in the chamber, encompassed by the myocardial tissue, when the heart muscle is relaxed (LVEDV) or contracted (LVESV) [15]. EF quantifies the percentage of blood leaving the heart each time it contracts. EF, LVEDV and LVESV are of great importance when addressing HF diagnosis, normal ranges for these functional indices are gender- and age- dependent. **Table 2** and show normal values for EF, LVEDV and LVESV parameters by age for males and females, respectively.

**Table 2: The predicted values for normal LV parameters in males by age.**

Age (y)	EDV (ml)			ESV (mL)			EF (%)		
	Lower	Mean	Upper	Lower	Mean	Upper	Lower	Mean	Upper
<b>Males</b>									
11 – 20	84	138	192	12	45	79	52	67	82
21 – 30	115	167	219	32	64	96	51	66	81
31 – 40	113	165	217	35	67	99	51	65	80
41 – 50	105	156	208	33	65	97	50	65	79
51 – 60	94	145	197	29	61	93	50	64	79
61 – 70	80	133	185	23	55	88	49	64	78

71 – 80	65	120	174	15	49	83	49	63	78
Information from references [18]									

Table 3: The predicted values for normal LV parameters in females by age.

Age (y)	EDV (ml)			ESV (mL)			EF (%)		
	Lower	Mean	Upper	Lower	Mean	Upper	Lower	Mean	Upper
<b>Females</b>									
11 – 20	77	120	162	13	37	60	55	71	86
21 – 30	76	119	161	17	40	63	55	70	86
31 – 40	75	118	160	19	42	64	54	70	85
41 – 50	74	116	158	20	43	65	54	69	85
51 – 60	73	115	158	21	44	67	53	69	84
61 – 70	72	114	157	22	45	68	53	69	84
71 – 80	69	1113	158	22	46	70	53	68	83
Information from references [18]									

## 2.5 Machine Learning

Machine learning (ML) is a branch of artificial intelligence, statistics and computer sciences that focuses on using data and algorithms to make accurate predictions. It arises at the intersection of statistics, which seeks to learn relationships from data, and computer science, with its emphasis on efficient computing algorithms [19]. The types of ML are subclassified into categories such as supervised learning and unsupervised learning [20].

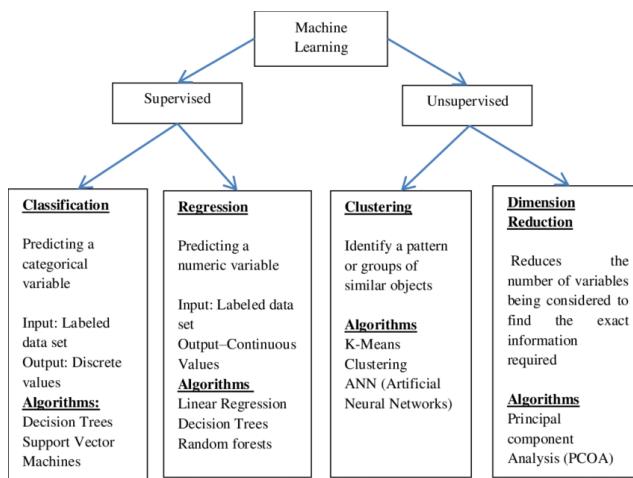


Figure 4: Broad Classification of Machine Learning Techniques. (Suryakanthi, 2020)

Supervised machine learning algorithms apply information that has been learned in the past to new data using labeled parameters to predict future outcomes. This methodology's goal is to predict a known output or target. A known training dataset is provided to the learning algorithm, it analyzes it and produces an inferred function to make predictions about the output values. The learning algorithm compares its output with the correct, intended output to find errors and modify the model accordingly. On the other hand, the goal for unsupervised learning is to find naturally occurring patterns or groupings within the data. This computer learning methodology explores the data to infer associations from the dataset to describe hidden structures from unlabeled data.

### **2.5.1 Heart Failure Diagnosis using Machine Learning**

ML offers new approaches to leveraging the growing volume of imaging data available for analyses of cardiovascular diseases. Such AI image-based diagnostic tools have the potential to substantially alleviate the burden of HF through by providing fast and accurate diagnostic decision making. ML methods can be utilized to address a wide range of data-related problems, from simple analytic queries in heart failure symptomatology existing data to the more complex challenges involving in analyzing raw cardiac MRI images [6].

Image-based cardiac diagnosis using a ML model have three main requirements for the accurate construction of the ML model. Firstly, it requires (1) input imaging datasets from which suitable imaging predictors can be extracted [21]. Secondly, (2) accurate output diagnosis labels, and finally (3) a suitable ML technique chosen depending on the

application to predict the cardiac diagnosis (output) based on the imaging predictors (input) [21].

### **2.5.2 Inputs, Outputs and Other Features**

As previously mentioned building an AI image-based diagnostic tool for HF has three requirements: Input, Output and ML technique. The latter will be discussed later. Suitable imaging inputs are needed to build a ML model for image-based diagnosis estimation. These inputs can be raw imaging data, conventional cardiac indices or radiomics features extracted from the image [21]. Important for this scientific research are conventional cardiac indices, as this is the input that will be utilized. Conventional imaging indices include measures commonly used in routine clinical image analysis such as end-diastolic/systolic ventricular volumes and ventricular EF.

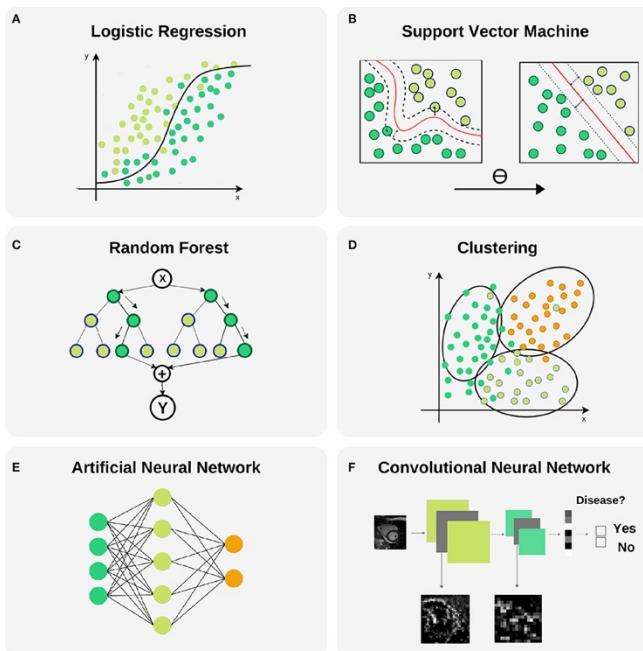
Supervised learning requires accurately labeled training examples to provide, in the simplest form, an output that is a binary variable which takes a value of 1 for a diseased individual and 0 for a control healthy subject [21]. To obtain an accurate ML model, it is important to use a balanced training sample, comprising a similar number of healthy and diseased subjects.

Additional non-imaging factors are also included into the ML model for accurate performance. Some of these factors can be electrocardiogram data, genetic data, sex, or age.

### **2.5.3 ML Technique**

Successful design, implementation, and validation of a new ML tools for image-based HF diagnosis, requires a suitable ML technique. It is important for the right technique

to be chosen for the application for HF diagnosis; however, it must also accordingly use the inputs given. **Figure 5** presents the most used ML techniques in the field of cardiac imaging and diagnosis. However, for the purpose of this research decision trees will be the technique furthered discussed due to their supervised machine learning nature and ability to accurately classify conventional indices of heart anatomy and hemodynamics.



**Figure 5: Machine Learning Techniques (Martin-Isla et al., 2020)**

### 2.5.3.1 Decision Trees

Decision Trees (DTs) are one of the most utilized supervised ML methods in medical diagnosis [22]. This technique continuously splits the data according to a certain parameter. The DT is a set of rules based on the input features values optimized for accurately classifying all elements of the training set. The tree can be explained by two entities, namely decision nodes and leaves. Decision nodes are where the data is split, it represents a choice that will result in the subdivision of all records into two or more mutually exclusive subsets [22]. The leaves are the decisions or the final outcomes.

In the literature, Random Forest or DT have been used frequently and were selected as the best performing model in some works. A DT model along with its decision and leaf nodes is represented in **Figure 6**.

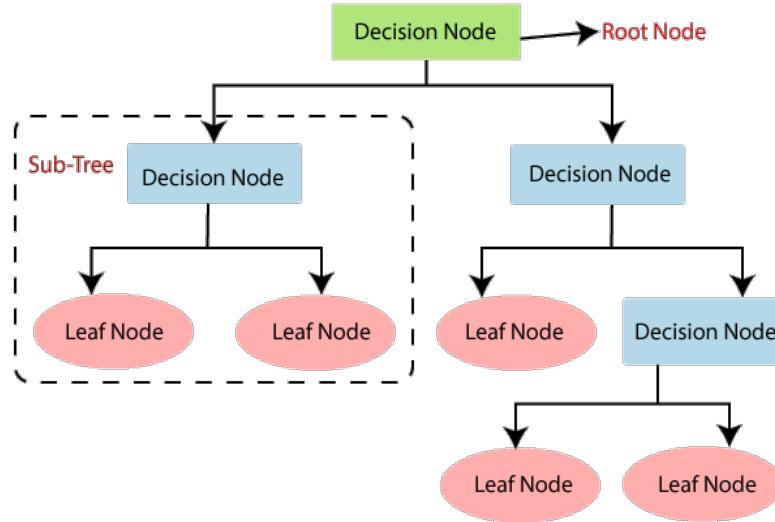


Figure 6: Decision Tree Algorithm Model

### **3 Research Program**

This chapter focuses on defining the objectives along with establishing a procedure and schedule to follow, to successfully achieve the stated and desired objectives for the proposed research.

#### **3.1 Objectives and Scope**

This scientific research looks to accomplish several core objectives. Firstly, the work looks to accomplish accurate and precise cardiac segmentation and quantification of key left ventricle functional indices from CMR. The functional indices of interest are ejection fractions and left ventricular end-diastolic, end-systolic volumes. Following segmentation and quantification, the work looks to successfully build a Machine Learning tool using decision trees for image-based heart failure diagnosis.

The research will be directed using CMR images, from the Sunnybrook Cardiac Data (SCD) [23], in DICOM format as input image data. Python libraries will be instrumental for the computational portion of this research. Scikit-image will be utilized to perform semi-automatic segmentation and quantification of the functional indices of the LV. As for the Machine Learning model the Scikit-learn library will be utilized. The output will provide a classified answer on if the patient possesses the heart failure pathology or it doesn't, deemed as healthy. It is important to mention along with the main libraries, others such as Pydicom, Numpy and Pandas will be utilized in conjunction.

## 3.2 Methodology

### 3.2.1 Image Data Acquisition

The CMR image data utilized for cardiac segmentation and quantification of functional LV indices were from the SCD [23], which consists of 45 cine-MRI images from a mix of patients and pathologies such as: healthy, hypertrophy and heart failure in DICOM format. Two different sets of images were used from the dataset. Short-Axis (SAX) Basal slices were utilized for endocardial LV area segmentation and quantification. In addition to the SAX Basal slices, Long-Axis (LAX) 4-Chamber views, were utilized to measure the length of the LV. The ventricular length, for the research's purpose, is determined from the apex to the mitral valve. The CMR image views can be seen on **Figure 7**.

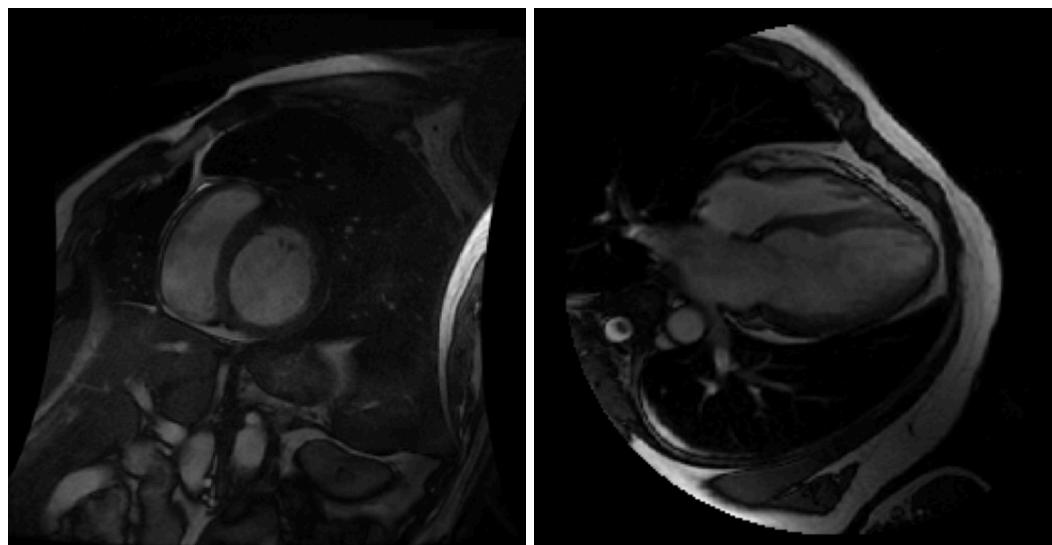


Figure 7: SAX Basal slices (left) & LAX 4-Chamber View (right)

### 3.2.2 Metadata Acquisition & Image Processing

Important information from the image must be extracted from the DICOM files. Relevant information of the analyzed images includes the modality, size and image spacing

parameters. As for metadata that provide information of patient data, the patient's ID, heart rate, and gender was extracted. Furthermore, image processing was needed for executing accurate and reliable segmentation. The focus was to enhance the contrast of the image's local features, as well as the edges and boundaries. The imaging processing utilized to accomplish these enchantments was Histogram Equalization, which adjusts the contrast of an image by using its histogram. The Python library utilized for this processing technique was Skimage.



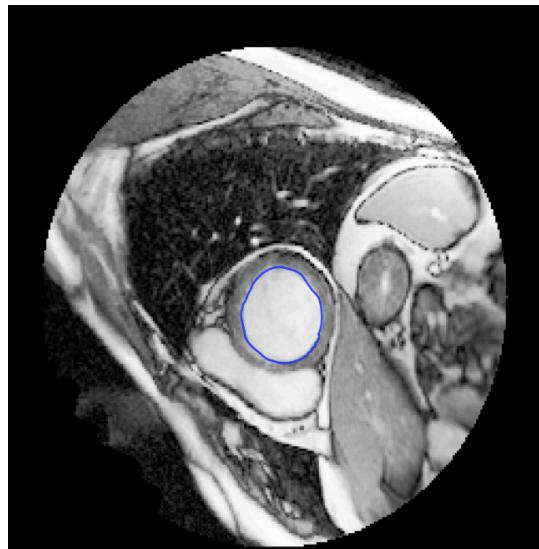
**Figure 8: Histogram Equalization on SAX Basal slices**

The contrast of the image was effectively accomplished by spreading out the most frequent intensity values of the image. This methodology increased the global contrast of images when the image data was represented by close contrast values. The equalization allowed for areas of lower local contrast to gain a higher contrast.

### **3.2.3 Cardiac Segmentation**

Endocardial LV area segmentation was carried on SAX Basal slices from processed image data acquired from the SCD. An active contour technique was utilized using

Python's Skimage library, where the function was provided several inputs and parameters other than the image data itself. An initial area or coordinates were given to the algorithm so it could perform the LV segmentation around the manually established area. Due to the manual establishment of the LV coordinates the segmentation was semi-automatic. Other inputs include snake length and smoothness parameters, time stepping parameter and the snake's attraction to brightness and edges. Consistency with all the parameters through all patient segmentation was tried to be maintained; however, to accommodate for noise and ventricle size variance the parameters were adjusted accordingly. **Figure 9** illustrates initial endocardial segmentation success utilizing methodology. Forty-one out of forty-five patients were successfully segmented, exclusions were due to high disturbance or noise.

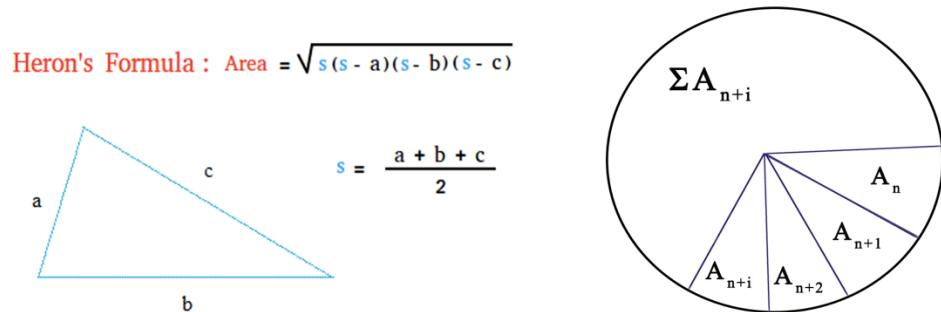


**Figure 9: Endocardial Segmentation**

### 3.2.4 Quantification of Functional Indices

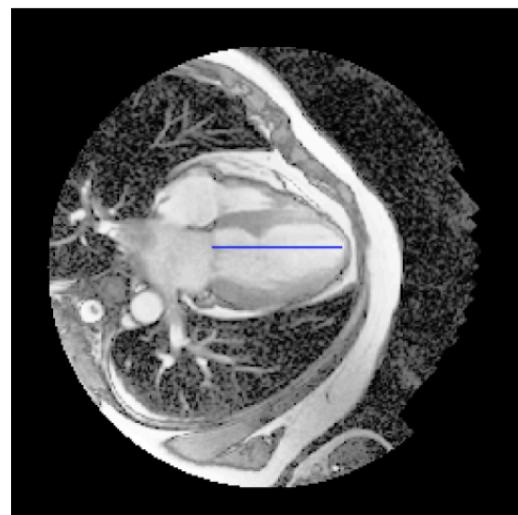
After endocardial segmentation was successfully achieved on SAX Basal slices throughout a complete cardiac cycle, quantification of the area segmented proceeded to take place. Using the set of coordinates formed by the delineated contour, the endocardial

area is calculated using Heron's formula. The ellipsoidal-like SAX Basal slice surface area of the LV is split up into several triangles to reduce geometrical assumptions. The summation of all the triangle areas inside the boundaries make up the SAX Basal slice surface area. The formulas and usage of Heron's formula can be seen on **Figure 10**. It is important to denote pixel spacing parameters acquired from the metadata were utilized for unit conversion purposes.



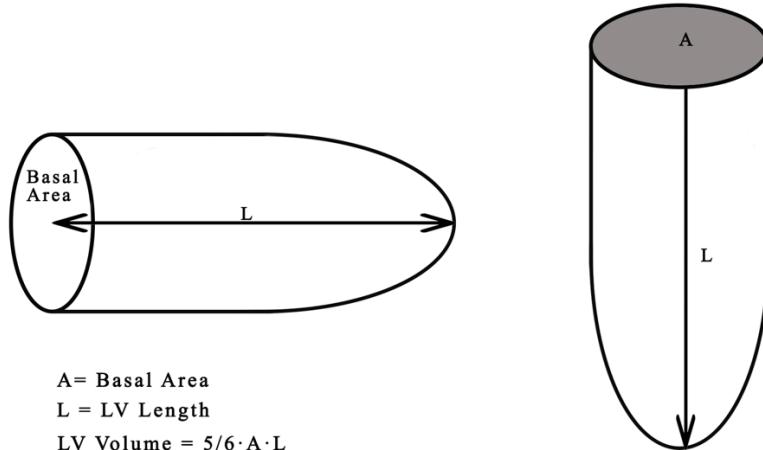
**Figure 10: Usage of Heron's Formula**

LV length was addressed using LAX 4-Chamber views, the ventricular length is determined from the apex to the mitral valve. The highest and lowest calculated lengths were deemed to be in end-diastole and end-systole, respectively. **Figure 11** illustrates the LV length measurement.



**Figure 11: LV Length Measurement**

The area-length method or the “bullet” method was the geometric approximation utilized to calculate the EDV and ESV of the LV. Only SAX and LAX planes were needed for this cylindrical assumption. The cross-sectional area will be calculated from SAX Basal slices and the LV length using LAX 4-Chamber views.



**Figure 12: Volume Quantification Diagram & Equation**

The obtained quantified values were statistically addressed for their accuracy and precision from referenced values of the SCD [23], see **Table 4**.

**Table 4: SDC Group Statistics**

	Normal (n=9)		Hypertrophy (n=12)		HF (n=12)		HF with Infarction (n=12)	
	Mean	STDV	Mean	STDV	Mean	STDV	Mean	STDV
<b>EDV (ml)</b>	115.69	36.89	114.39	50.46	233.67	63.21	244.92	86.02
<b>ESV (ml)</b>	43.10	14.74	43.11	24.50	158.28	56.34	174.34	90.64
<b>EF (%)</b>	62.93	3.65	62.72	9.22	33.09	13.07	32.01	12.27

A one sample T-test was utilized to address whether the obtained quantified values were statistically different from the reference values. The null hypothesis assumes that the difference between the obtained group mean, and the referenced mean value is equal to zero.

### 3.2.5 Machine Learning Model & Diagnostic Predictions

Following the quantification of the LV functional indices, obtained EDV, ESV and EF data was utilized to train and test a decision tree ML model for HF diagnosis. Python's Sklearn library was utilized for the building and evaluation of the decision tree model, where the data was split into a training set (80%) and testing set (20%). The model was set to have a max depth of ten nodes and to use entropy, the measure of impurity, disorder, or uncertainty, to control how the model will split the data and draw boundaries. The entropy equation can be seen on **Equation (1)**.

$$\text{Equation 1: } I_H = \sum_{j=1}^c p_j \log_2(p_j) \quad (1)$$

Once the model was trained it classified the testing data as a binary output which a value of 0 represented a patient with "healthy", and a value of 1 represented a "HF" patient. Several metrics were used to measure the quality of the predictions from the HF classification model. Using measures like precision, recall and F1 score it was addressed how many predictions were True and how many are False.

## 3.3 Itinerary

**Table 5: Gantt Chart Itinerary for the Scientific Research**

Project Planning		Aug.		Sept.			Oct.		Nov.		Dec.		Jan.			Feb.		Mar.			Apr.			May					
Phase	Task	W 3	W 4	W 1	W 2	W 3	W 4	W 1	W 2	W 4	W 3	W 4	W 1	W 2	W 3	W 4	W 1	W 2	W 3	W 1	W 2	W 3	W 4	W 1	W 2	W 3	W 4		
Proposal	Introduction																												
	Literature Review																												
	Objectives & Scope																												
	Methodology																												
Image Dataset	Download Data																												
	Determine SAX																												
	Determine LAX																												
Metadata & Image Processing	Acquire metadata																												
	Histogram Equalization																												
	Optimization																												
Cardiac Segmentation & Quantifying	Segmentation																												
	Length Measurement																												
	Quantification																												
ML Model & Predictions	Decision Tree Design																												
	Optimization																												
	Predictions																												

## 4 Data

This chapter focuses in exhibiting all the data found on this scientific research during the segmentation, quantification, and classification phases.

### 4.1 Segmentation Data

Forty-one out of forty-five patients were successfully segmented, exclusions were due to high disturbance and noise on the SAX images or due to lack of specific SAX and LAX views for the quantification methodology.

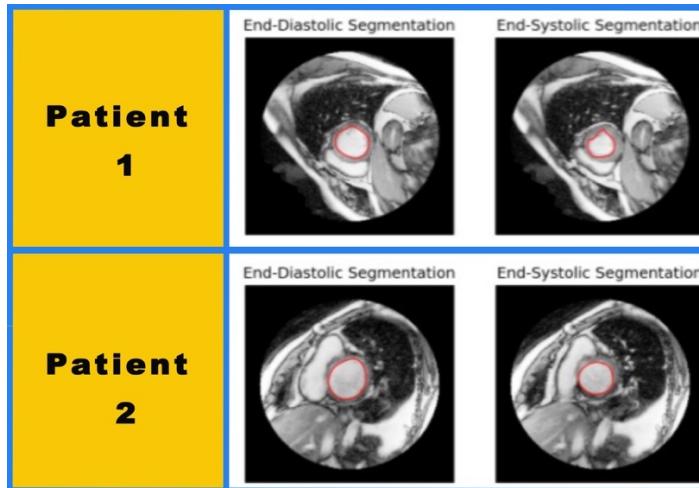


Figure 13: Endocardial Segmentation for Patients 1 & 2

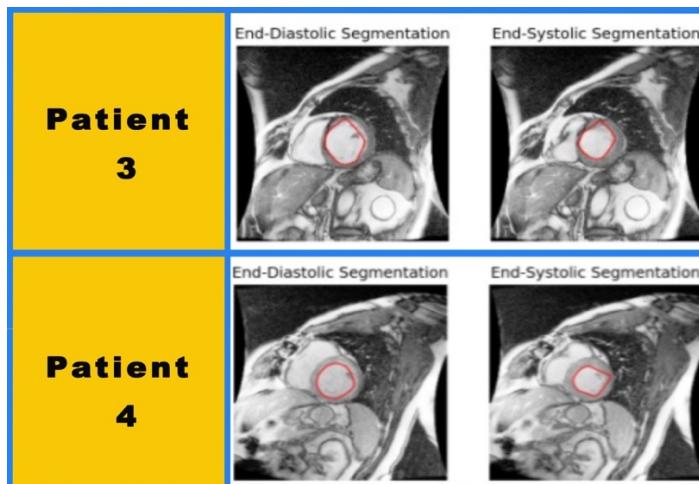
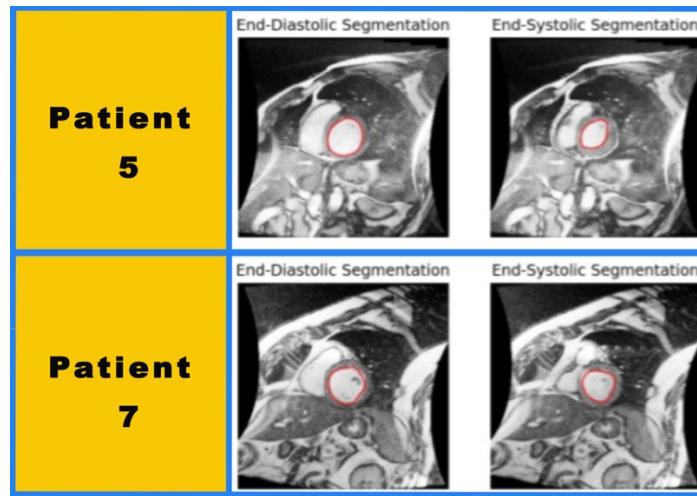
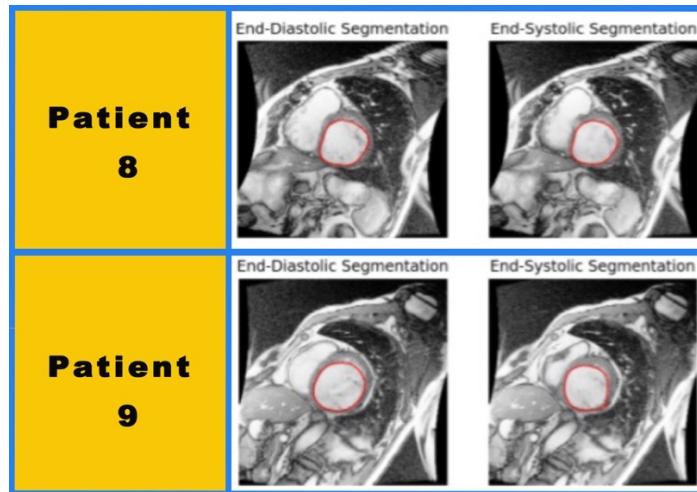


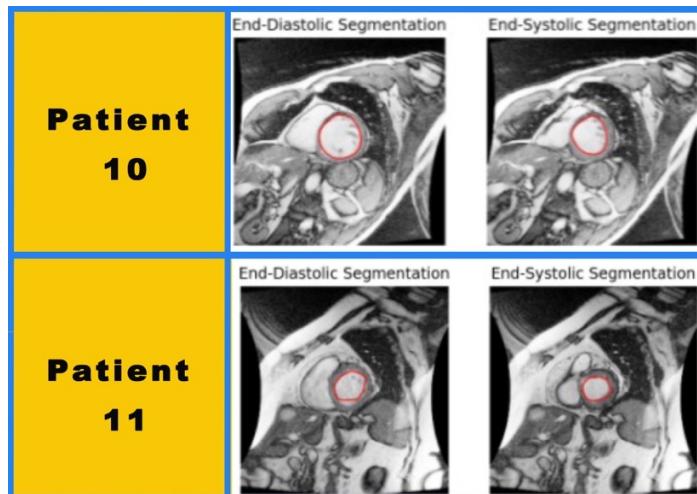
Figure 14: Endocardial Segmentation for Patients 3 & 4



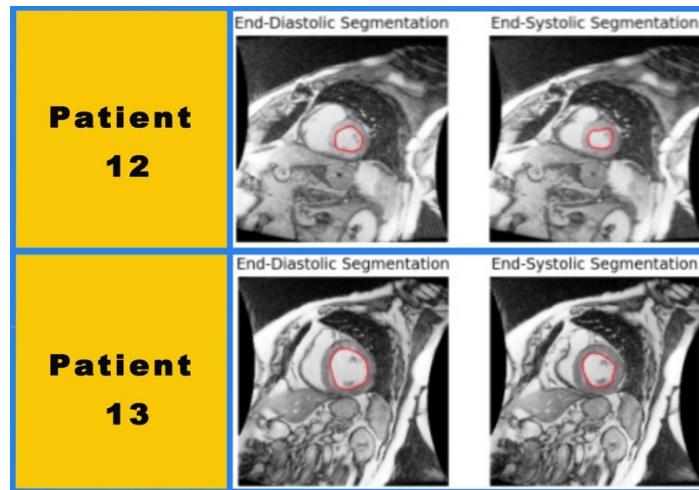
**Figure 15: Endocardial Segmentation for Patients 5 & 7**



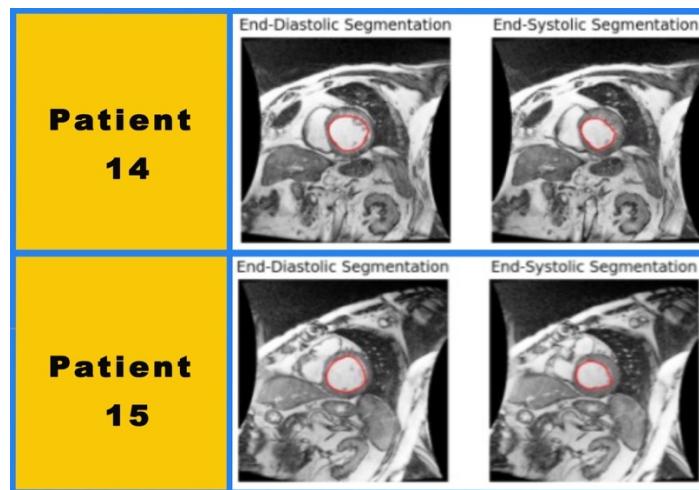
**Figure 16: Endocardial Segmentation for Patients 8 & 9**



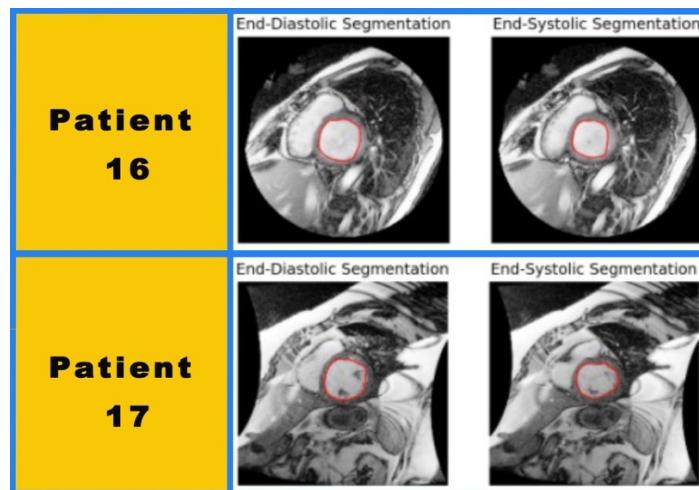
**Figure 17: Endocardial Segmentation for Patients 10 & 11**



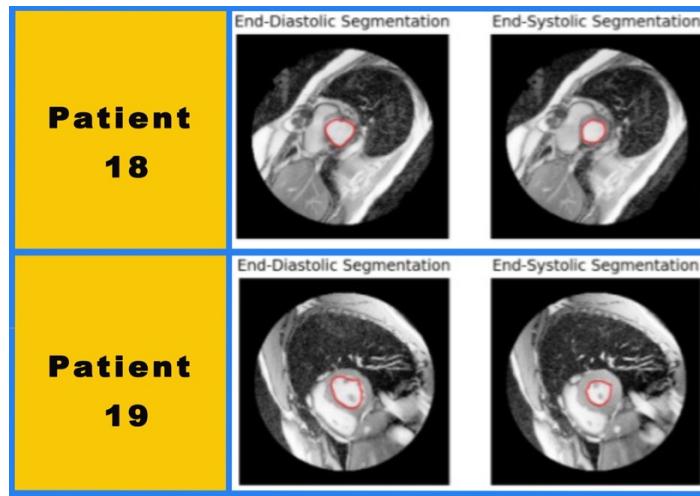
**Figure 18: Endocardial Segmentation for Patients 12 & 13**



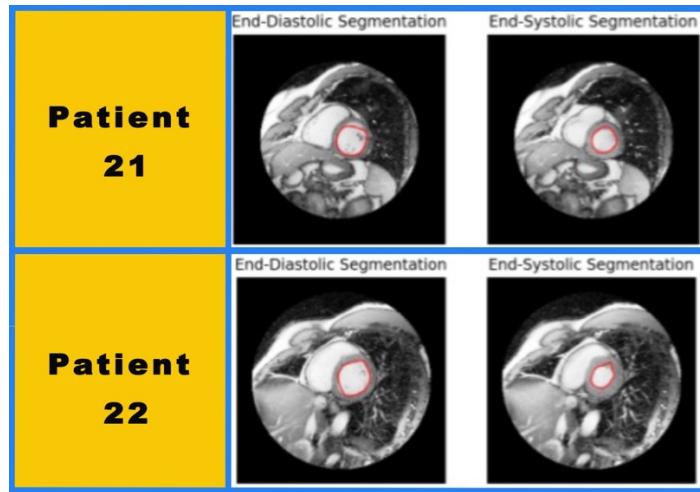
**Figure 19: Endocardial Segmentation for Patients 14 & 15**



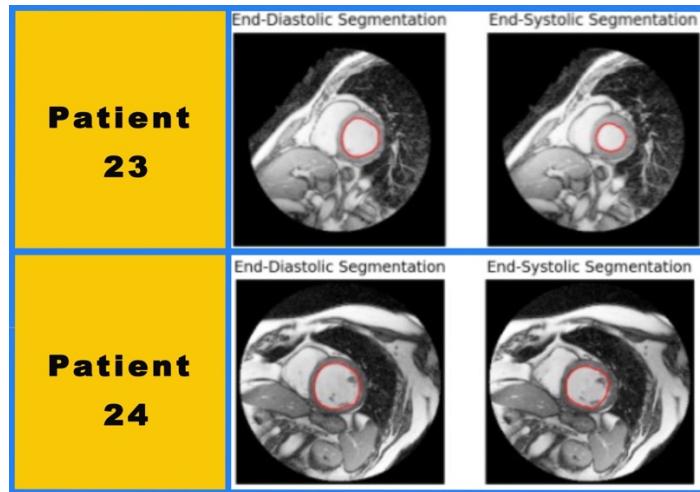
**Figure 20: Endocardial Segmentation for Patients 16 & 17**



**Figure 21: Endocardial Segmentation for Patients 18 & 19**



**Figure 22: Endocardial Segmentation for Patients 21 & 22**



**Figure 23: Endocardial Segmentation for Patients 23 & 24**

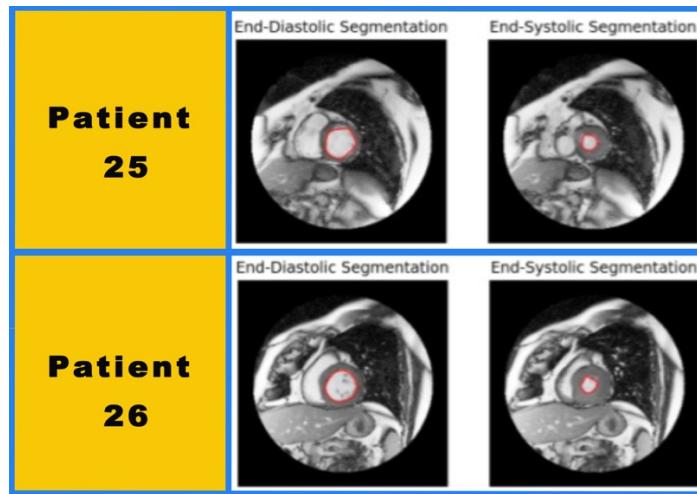


Figure 24: Endocardial Segmentation for Patients 25 & 26

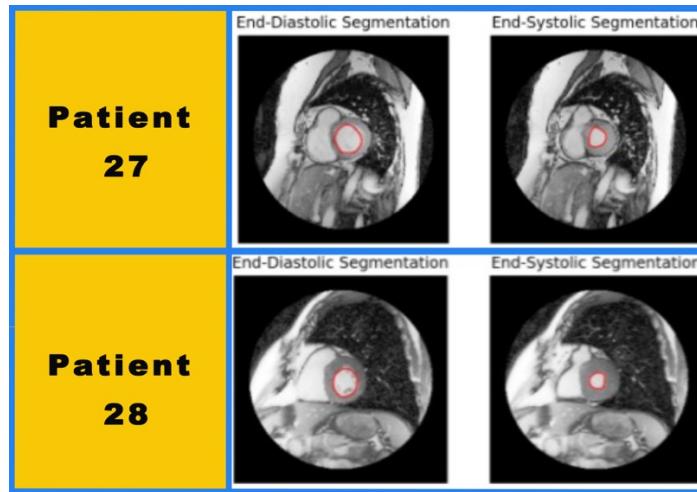


Figure 25: Endocardial Segmentation for Patients 27 & 28

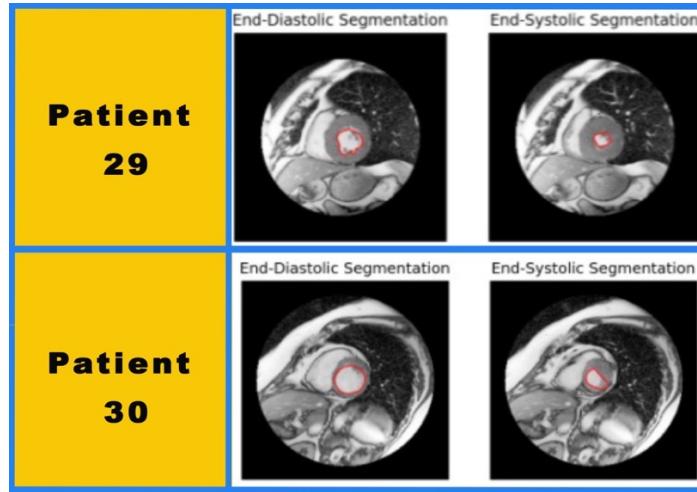


Figure 26: Endocardial Segmentation for Patients 29 & 30

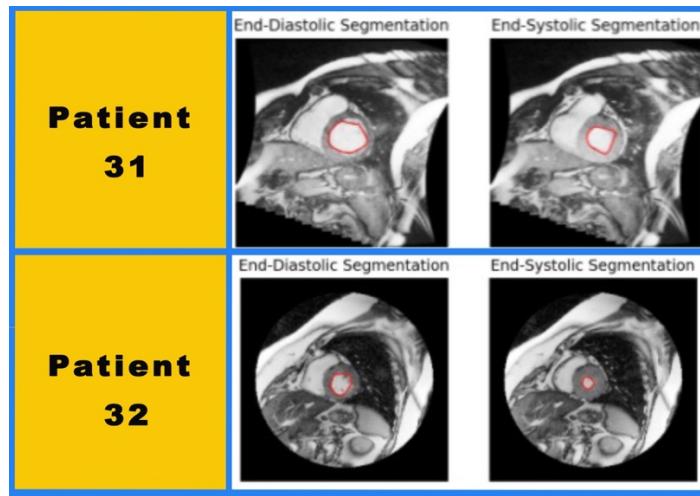


Figure 27: Endocardial Segmentation for Patients 31 & 32

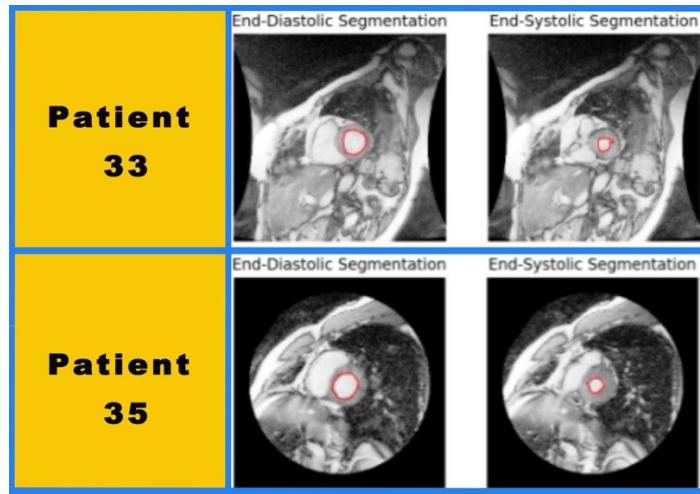


Figure 28: Endocardial Segmentation for Patients 33 & 35

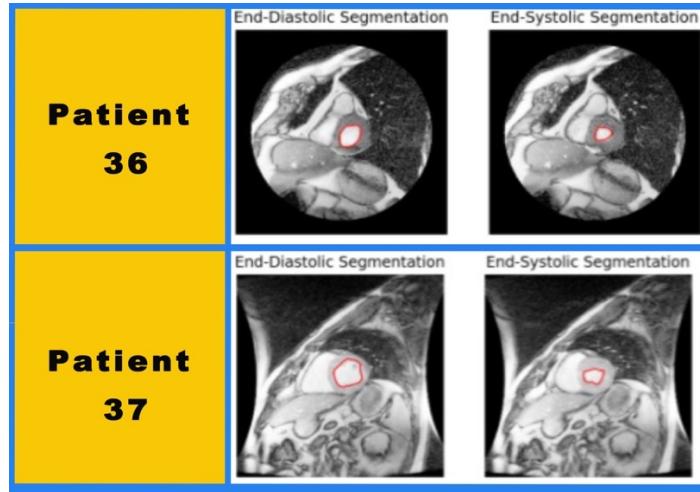
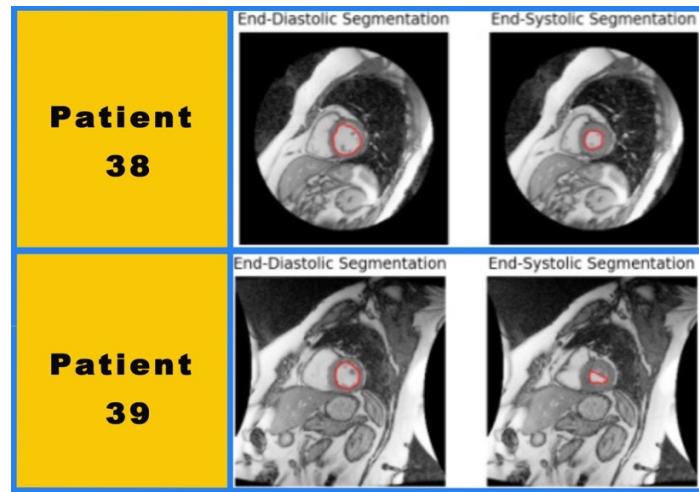
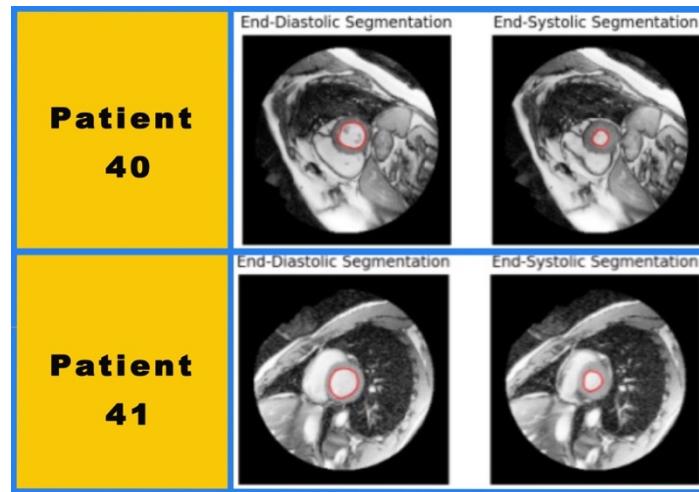


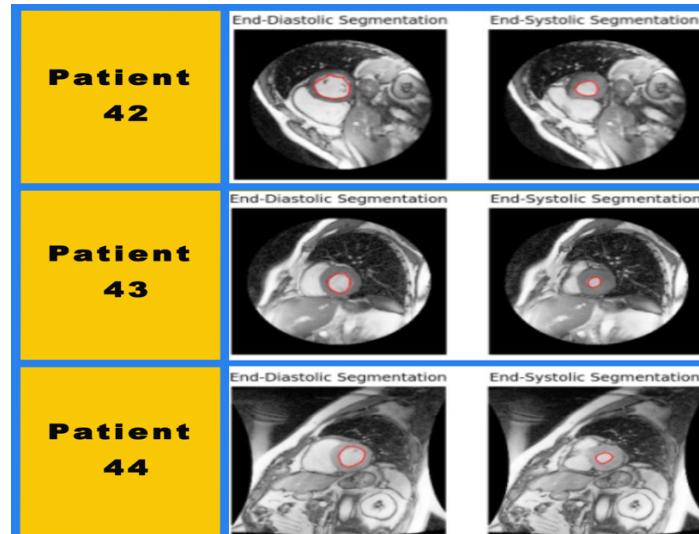
Figure 29: Endocardial Segmentation for Patients 36 & 37



**Figure 30: Endocardial Segmentation for Patients 38 & 39**



**Figure 31: Endocardial Segmentation for Patients 40 & 41**



**Figure 32: Endocardial Segmentation for Patients 42, 43 & 44**

## 4.2 Quantification Data

**Table 6: Functional Indices (EDV, ESV and EF) Quantification Data**

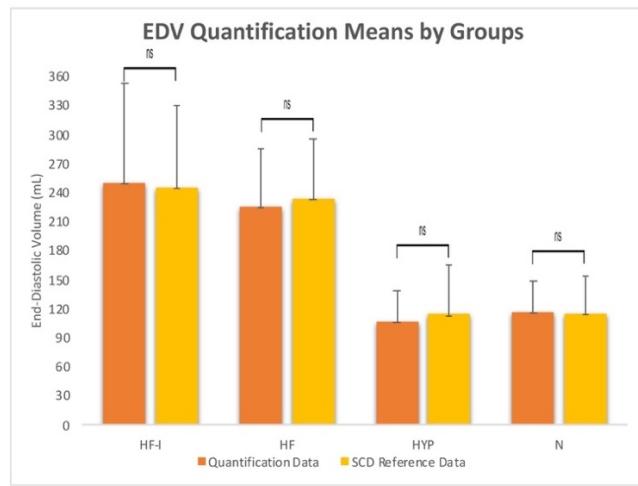
Patient	Gender	Age	Heart Rate	EDV (mL)	ESV (mL)	EF (%)	Pathology
1	Male	53	79	180.69	108.71	39.84	HF-I
2	Male	48	87	239.56	180.46	24.66	HF-I
3	Male	79	65	271.43	167	38.47	HF-I
4	Male	45	54	199.57	136.55	31.58	HF-I
5	Male	60	88	219.27	110.83	49.45	HF-I
7	Male	46	78	198.14	102.04	48.5	HF-I
8	Male	57	90	428.47	380.61	11.17	HF-I
9	Male	69	75	407.15	322.68	20.75	HF-I
10	Male	55	72	343.21	216.89	36.81	HF-I
11	Female	71	62	154.23	97.96	36.48	HF-I
12	Male	77	75	117.15	72.34	38.26	HF-I
13	Female	81	53	249.55	159.87	35.94	HF
14	Male	77	51	170.17	100.09	41.19	HF
15	Male	81	80	224.86	138.92	38.22	HF
16	Male	71	74	285.07	212.57	25.43	HF
17	Female	63	63	314.81	226.16	28.16	HF
18	Female	68	74	128.1	79.57	37.88	HF
19	Male	82	46	164.24	96.22	41.41	HF
21	Male	52	52	205.11	124.36	39.37	HF
22	Male	47	62	194.79	81.88	57.97	HF
23	Male	52	64	248.46	130.98	47.28	HF
24	Female	77	67	295.94	221.2	25.26	HF
25	Male	83	64	114.19	26.95	76.4	HYP
26	Male	42	73	119.38	26.11	78.13	HYP
27	Female	46	88	97.01	35.84	63.05	HYP
28	Male	61	67	122.53	43.67	64.36	HYP
29	Male	62	55	117.83	37.16	68.46	HYP
30	Female	48	71	146	51.43	64.77	HYP
31	Male	47	59	160.23	81.22	49.31	HYP
32	Female	53	66	57.25	15.66	72.65	HYP
33	Female	76	61	74.98	18.26	75.64	HYP
35	Male	38	79	94.02	29.54	68.58	HYP
36	Male	68	82	66.34	27.86	58	HYP
37	Male	63	62	109.77	38.48	64.94	N
38	Female	53	56	111.98	40.15	64.14	N
39	Female	77	73	80.45	21.06	73.82	N
40	Male	70	76	84.72	25.09	70.38	N
41	Male	23	55	143.31	60.99	57.44	N
42	Male	51	56	175.53	65.52	62.67	N
43	Male	61	65	94.44	24.31	74.25	N
44	Male	60	67	131.2	31.07	76.32	N

**Table 7: Functional Indices (EDV, ESV and EF) Quantification Statistical Data**

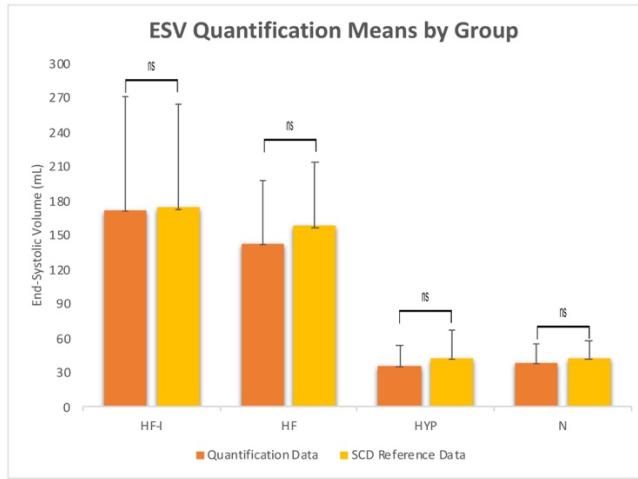
	Normal (n=8)			Hypertrophy (n=11)			HF (n=11)			HF with Infarction (n=11)		
	Mean	STDV	P - value	Mean	STDV	P - value	Mean	STDV	P - value	Mean	STDV	P - value
<b>EDV (mL)</b>	116.43	32.25	0.9504	106.34	32.15	0.4257	225.55	59.37	0.66	250.81	101.78	0.8517
<b>ESV (mL)</b>	38.33	16.81	0.449	35.79	18.33	0.2148	142.89	55.18	0.3768	172.37	98.82	0.9486
<b>EF (%)</b>	67.99	6.68	0.069	67.21	8.63	0.1148	38.01	9.64	0.1215	34.18	11.51	0.5459

\*p-values are obtained from one sample T-test.

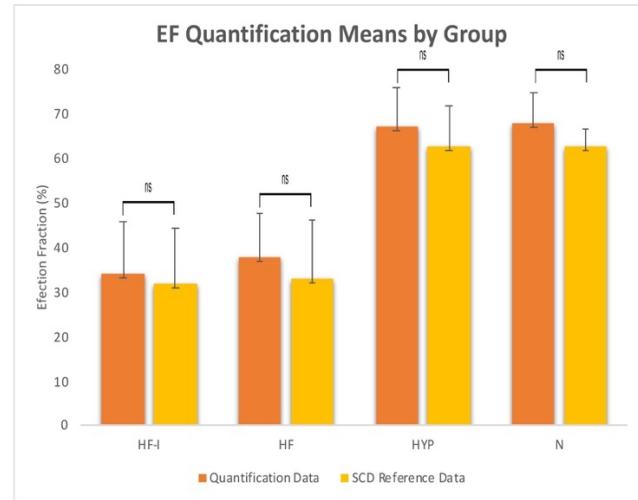
\*\*No statistical differences were found.



**Figure 33: End-Diastolic Volume Mean Values by Group with Standard Deviation**



**Figure 34: End-Systolic Volume Mean Values by Group with Standard Deviation**



**Figure 35: Ejection Fraction Mean Values by Group with Standard Deviation**

## 4.3 Classification Data

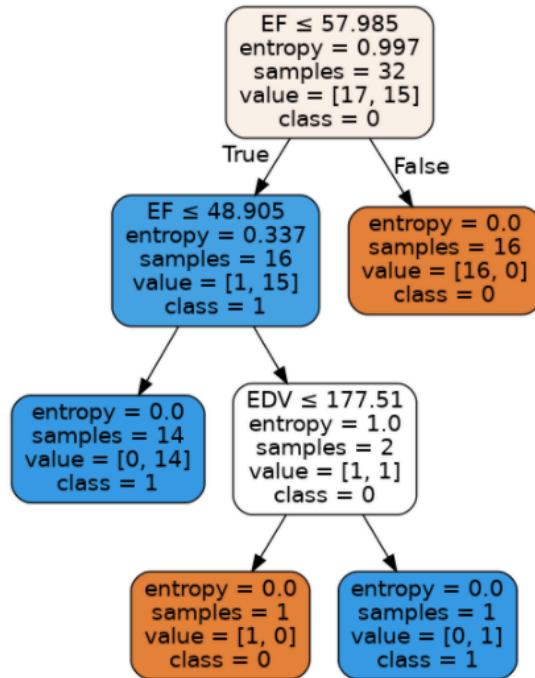


Figure 36: Machine Learning Decision Tree Model

Table 8: Classification Test Metrics

Classification	Precision	Recall	F1 Score	Number of Occurrences
Healthy (0)	1.0	1.0	1.0	2
HF (1)	1.0	1.0	1.0	7
		Total Accuracy	100 %	9

## 5 Analysis and Results Interpretation

Cardiovascular diagnostics using computational methods require a plethora of procedures from obtaining the image data to identifying a disease in a particular patient. The methodology utilized throughout this research, focused on accurately performing each procedure to achieve the research's objectives. Firstly, the work looked to accomplish accurate and precise cardiac segmentation and quantification of key left ventricle functional indices from CMR. The functional indices of interest were the ejection fraction and left ventricular end-diastolic, end-systolic volumes. Following segmentation and quantification, the work looked to successfully build and evaluate a Machine Learning tool using decision trees for image-based heart failure diagnosis.

Endocardial segmentation of the LV on SAX Basal slices was achieved through a semi-automatic procedure. Where initial coordinates and diverse parameters were provided, to an active contour technique from Python's Skimage library, to identify the endocardial boundaries. Unfortunately, four out of forty-five patients were unable to be segmented due to high disturbance and noise on the SAX images or due to lack of specific SAX and LAX views for the quantification methodology. However, those that were segmented were visually successful as they correctly highlight the endocardial boundaries of the LV, see from **Figure 13** to **Figure 32**. Some segmentations, particularly in end-systole, held difficulties in identifying the outer or bottom edges of the LV due to the contrast between the endocardial boundary and epicardial region being very low. These difficulties led to segmenting a slightly bigger or smaller area than desired. Even though most segmentations were visually successful, such qualitative data could be misinterpreted

due to its subjective nature. Therefore, the quantification data is used to validate the accuracy and precision of the segmentation and quantification procedures.

The endocardial segmented area was quantified using Heron's formula, where the surface area of the LV was split up into several triangles to reduce geometrical assumptions, the summation of all the triangle inside the boundary represented the overall surface area value. This quantified cross-sectional area along with the calculated LV lengths were utilized to calculate the EDV and ESV of the LV, using the area-length method. The obtained values, see **Table 7**, were addressed for their precision and accuracy compared to the SDC referenced values, see **Table 4**, using a one-sample T test. The results for the one-sample T tests revealed there were no statistical significance between the obtained mean values and the referenced mean values in each quantified variable. End-diastolic volumes possessed strong p-values, meaning the obtained values closely resemble those established in the SDC. On the other hand, the end-systolic volume data resembled lower accuracy than the end-diastolic data, due to their p-values being lower. The difficulties segmenting the end-systolic areas are clearly shown in these statistics, as the problems identifying the outer or bottom edges of the LV derailed a consistent accurate segmentation and quantification. Furthermore, the HF-I group did remain consistent with both end-systolic and end-diastolic data, achieving p-values of 0.95 and 0.85, respectively. As the heart muscle expands due to both HF and infarction, the end-systolic and end-diastolic areas remained very similar, resulting in accurate segmentation and quantification regardless. Even though further optimization of the segmentation algorithm might resolve these inconsistencies, the p-values still represent the quantified ESV data with no significance from the referenced SDC values and as accurate. As for the calculated EF data,

the p-values were even lower. Since the EF is a calculated value from both EDV and ESV, it was expected to be the variable with the lowest p-value. However, the HF-I group did obtain a high p-value of 0.55 due to a consistent end-systolic and end-diastolic data. All groups held a higher EF mean than their respective comparative EF means, this was attributed to the quantified ESV being lower than established in the SDC. The p-values still represent the quantified EF data with no statistical significance from the referenced SDC values and as fairly accurate.

The obtained quantification data of the EDV, ESV and EF of the LV was utilized to train and test a decision tree ML model for HF diagnosis. The model was trained to classify patients or give a binary output, which a value of 0 represented a patient with “HF”, and a value of 1 represented a “healthy” patient. The building and evaluation of the decision tree model was done with Python’s Sklearn library, where the data was split into a training set (80%) and testing set (20%). This data split meant it would utilize thirty-two patients to build and train the model, and nine patients to evaluate it. The trained decision tree model, see **Figure 36**, established three decision nodes to classify patients based on the training data. The root node verifies if the patient in question holds an EF equal or lower to 57.99 %, if true it will move on to the next decision node; however, if false the patient is classified as healthy. The second verifies if the patient in question holds an EF equal or lower to 48.91 %, if true the patient is diagnosed with HF. On the other hand, if false it will move on to the next decision node which verifies if the patient in question holds an EDV equal or lower to 177.51 ml. If true, the patient is classified as healthy; however, if false the patient is diagnosed with HF. The trained ML model was evaluated with nine patients, where two of them were healthy and seven possessed HF. Every tested patient was

classified correctly using the trained ML mod, achieving a 100% accuracy. Considering the data provided to the model for training only utilized thirty-two patients that only held five variables (Age, Gender, EDV, ESV and EF) the built decision tree model is very accurate and holds great potential for providing accurate clinical diagnosis. Unfortunately, cardiovascular diseases, especially HF, are very complex and computational diagnostic approaches require more data to consistently provide accurate clinical diagnosis. Adding additional input data regarding ethnicity, ventricular mass, electrocardiogram data and genetic data will heavily influence making this application a viable assistance for physicians to rely on. However, the findings of this scientific research provide a solid foundation to build on approaches for cardiovascular diagnosis, throughout the whole procedure, starting from raw CMR images to converting them into functional indices and finally use that data to diagnose patients using ML.

## **6 Conclusions, Recommendations and Future Work**

This chapter focuses on establishing the research project's overall conclusions along with any recommendations that hold the potential for the improvement of the research and future areas of research identified during the process that should be explored.

### **6.1 Conclusions**

As a result of the findings, the core objectives of this scientific research were accomplished. Firstly, accurate and precise cardiac segmentation and quantification of key left ventricle functional indices from CMR were accomplished. The quantified indices of ejection fraction, and left ventricular end-diastolic, end-systolic volumes were achieved using Heron's formula and the area-length method. The accurateness of the quantified functional indices was addressed using a one-sample T test. The results for the one-sample T tests revealed there were no statistical significance between the obtained mean values and the referenced mean values in each quantified variable. The obtained strong p-values show the quantified values closely resemble those established in the SDC. Finally, a Machine Learning tool using decision trees for image-based heart failure diagnosis was successfully built, as every tested patient was classified correctly using the trained ML model, achieving a 100% accuracy.

### **6.2 Recommendations**

Based on the research's findings, continual improvement and optimization for the segmentation and quantification algorithm is recommended in order achieve more accurate

results. In addition, more data should be utilized to test the ML algorithm to better asses its viability to assist physicians to diagnose HF.

### **6.3 Future Work**

Future work for this project and research area involves optimizing the segmentation and quantification algorithm to address difficulties in identifying the outer or bottom edges of the LV in end-systole. An automatic segmentation methodology using ML should also be implemented to reduce human inputs. Finally, additional input data regarding ethnicity, ventricular mass, electrocardiogram data and genetic data should be added to the ML model to heavily influence making this application a viable assistance for physicians to rely on.

## 7 References

1. Heart & Vascular: Conditions & Treatments. Retrieved October 25, 2020, from <https://www.emoryhealthcare.org/heart-vascular/wellness/heart-failure-statistics.html>
2. American Heart Association. What is Heart Failure? Retrieved October 25, 2020, from <https://www.heart.org/en/health-topics/heart-failure/what-is-heart-failure#:~:text=Heart%20failure%20is%20a%20chronic,keep%20up%20with%20its%20workload>.
3. Taking the Risks to Heart: Misdiagnosis of Heart Disease. (2017, February 20). Retrieved November 11, 2020, from <https://www.acc.org/membership/join-us/benefits/additional-member-only-benefits/acc-and-the-doctors-company/the-doctors-company-updates/2017/02/20/12/55/taking-the-risks-to-heart-misdiagnosis-of-heart-disease>
4. Singh, H., Schiff, G., Graber, M., Onakpoya, I., & Thompson, M. (2017, June 01). The global burden of diagnostic errors in primary care. Retrieved November 10, 2020, from <https://qualitysafety.bmj.com/content/26/6/484>
5. Chen, C., Qin, C., Qiu, H., Tarroni, G., Duan, J., Bai, W., & Rueckert, D. (2020, February 17). Deep Learning for Cardiac Image Segmentation: A Review. Retrieved November 8, 2020, from <https://www.frontiersin.org/articles/10.3389/fcvm.2020.00025/full>
6. Henglin, Mir & Stein, Gillian & Hushcha, Pavel & Snoek, Jasper & Wiltschko, Alexander & Cheng, Susan. (2017). Machine Learning Approaches in Cardiovascular Imaging. *Circulation: Cardiovascular Imaging*. 10. e005614. 10.1161/CIRCIMAGING.117.005614.

7. Hunt SA, Abraham WT, Chin MH, et al. 2009 focused update incorporated into the ACC/AHA 2005 guidelines for the diagnosis and management of heart failure in adults: a report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines: developed in collaboration with the International Society for Heart and Lung Transplantation [published correction appears in Circulation. 2010;121(12):e258]. Circulation. 2009;119(14):e391–e479.
8. What is heart failure? Retrieved May 15, 2021, from <https://www.heart.org/en/health-topics/heart-failure/what-is-heart-failure>
9. Mitter SS, Shah SJ, Thomas JD. A test in context: E/A and E/e' to assess diastolic dysfunction and LV filling pressure. J Am Coll Cardiol 2017; **69**: 1451– 1464.
10. Criteria Committee, New York Heart Association, Inc. Diseases of the Heart and Blood Vessels. Nomenclature and Criteria for diagnosis, 6th edition Boston, Little, Brown and Co. 1964, p 114.
11. Chen, Y., Wong, L., Liew, O., & Richards, A. (2019, December 16). Heart failure with Reduced ejection Fraction (HFrEF) and Preserved ejection Fraction (HFpEF): The diagnostic value of Circulating MicroRNAs. Retrieved May 15, 2021, from <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6952981/>
12. Hajouli, S. (2021, January 04). Heart failure and ejection fraction. Retrieved May 15, 2021, from <https://www.ncbi.nlm.nih.gov/books/NBK553115/>
13. Blankstein, R. (2012, January 24). Introduction to noninvasive cardiac imaging. Retrieved May 15, 2021, from <https://www.ahajournals.org/doi/full/10.1161/circulationaha.110.017665>

14. Karamitsos, T. D., Francis, J. M., Myerson, S., Selvanayagam, J. B., & Neubauer, S. (2009). The role of cardiovascular magnetic resonance imaging in heart failure. *Journal of the American College of Cardiology*, 54(15), 1407–1424. Retrieved Ma15, 2021, from <https://doi.org/10.1016/j.jacc.2009.04.094>
15. Peng, P., Lekadir, K., Gooya, A. *et al.* A review of heart chamber segmentation for structural and functional analysis using cardiac magnetic resonance imaging. *Magn Reson Mater Phy* 29, 155–195 (2016). Retrieved April 15, 2021, from <https://doi.org/10.1007/s10334-015-0521-4>
16. Hemalatha, R., Thamizhvani, T., Dhivya, A., Joseph, J., Babu, B., & Chandrasekaran, R. (2018, July 04). Active contour Based Segmentation techniques for medical image analysis. Retrieved May 15, 2021, from <https://www.intechopen.com/books/medical-and-biological-image-analysis/active-contour-based-segmentation-techniques-for-medical-image-analysis>
17. Kass, M., Witkin, A. & Terzopoulos, D. Snakes: Active contour models. *Int J Comput Vision* 1, 321–331 (1988). Retrieved May 10, 2021 <https://doi.org/10.1007/BF00133570>
18. Cain, P.A., Ahl, R., Hedstrom, E. *et al.* Age and gender specific normal values of left ventricular mass, volume and function for gradient echo magnetic resonance imaging: a cross sectional study. *BMC Med Imaging* 9, 2 (2009). Retrieved May 2, 2021 <https://doi.org/10.1186/1471-2342-9-2>
19. Deo, R. (2015, November 17). Machine learning in medicine. Retrieved January 11, 2021, from <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5831252/#:~:text=Machine%20learn>

- ing%20is%20the%20scientific,emphasis%20on%20efficient%20computing%20algor  
ithms.
20. Suryakanthi, T.. (2020). Evaluating the Impact of GINI Index and Information Gain  
on Classification using Decision Tree Classifier Algorithm\*. International Journal of  
Advanced Computer Science and Applications. 11. 10.14569/IJACSA.2020.0110277.
21. Martin-Isla, C., Campello, V., Izquierdo, C., Raisi-Estabragh, Z., Baeßler, B.,  
Petersen, S., & Lekadir, K. (2020, January 06). Image-Based Cardiac Diagnosis With  
Machine Learning: A Review. Retrieved November 1, 2020, from  
<https://www.frontiersin.org/articles/10.3389/fcvm.2020.00001/full>
22. Song, Y., & Lu, Y. (2015, April 25). Decision tree methods: Applications for  
classification and prediction. Retrieved May 10, 2021, from  
<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4466856/>
23. Radau P, Lu Y, Connelly K, Paul G, Dick AJ, Wright GA. Evaluation framework for  
algorithms segmenting short axis cardiac MRI. *MIDAS J.* (2009) 49. Retrieved  
November 18, 2020, from <http://www.cardiacatlas.org/studies/sunnybrook-cardiac-data/>