



**Politecnico
di Torino**

Politecnico di Torino

**Corso di Laurea Magistrale
in INGEGNERIA BIOMEDICA**

Tesi di Laurea Magistrale

**ARTIFICIAL INTELLIGENCE APPLIED TO 2D
ECHOCARDIOGRAPHY FOR THE EARLY
DETECTION OF MYOCARDIAL INFARCTION**

Relatore

Prof. Filippo Molinari

Candidata

Carolina Gomez

Tutor aziendale

Annalisa Letizia
Vincenza Tufano

Anno Accademico 2022-2023

Il Gruppo Teoresi, nato a Torino nel 1987, è oggi una realtà internazionale che conta circa mille dipendenti in Europa e negli Stati Uniti. Opera come partner qualificato supportando i clienti nello sviluppo di prodotto e di processo, attraverso tecnologie innovative e un'esperienza globale in ambito ingegneristico. Il Gruppo Teoresi è attivo in una vasta gamma di settori di mercato e rappresenta una combinazione di competenze tecniche e know-how applicativo in grado di coprire l'intero ciclo di sviluppo del prodotto, dalla fase di progettazione concettuale sino al supporto after-market.

Acknowledgements

First and foremost, I would like to thank my mentors from Teoresi, Annalisa and Vincenza, who were always available and constantly followed my progress and offered assistance during the development of this thesis project. I would also like to express my gratitude to my supervisor Prof. Filippo Molinari for being the first person to spark my interest into Deep Learning and processing of medical images, as well as for all of his insightful comments and suggestions during the project.

Finally, I would like to thank everyone who influenced my coming to Italy and studying at Politecnico di Torino and all who supported me during this journey. To my family in Colombia and my friends in Torino, who became family as well, this is as much yours as it is mine. Italia, ti porterò per sempre nel mio cuore.

Sommario

L'**infarto del miocardio** (MI) si verifica in seguito ad una interruzione del flusso sanguigno al muscolo cardiaco, di solito a causa di una sottostante malattia coronarica (CAD). Nel 2019, le CADs di tipo ischemico sono state responsabili di 1 decesso su 6, diventando la principale causa di decesso nel mondo. Allo stato attuale, la diagnosi di malattie cardiache, e in particolare di MI, si definisce a partire da specifici valutazioni di biomarcatori ed elettrocardiografia. Tuttavia, poiché i risultati di questi metodi sono talvolta inconcludenti e possono portare a diagnosi errate, confondendosi con altre sindromi, i cardiologi si affidano frequentemente a tecniche di diagnostica per immagini. L'ecocardiografia è la tecnica più utilizzata, poiché consente la visualizzazione del cuore in modo semplice, in tempo reale ed economico. Da questo esame, i cardiologi studiano il movimento delle pareti del ventricolo sinistro (LV) alla ricerca di contrazioni anomale che compaiono contestualmente all'ischemia. Per valutare ulteriormente la funzionalità cardiaca, dall'ecocardiogramma vengono calcolati anche parametri come la frazione di eiezione del sangue. Al fine di raggiungere questi task, può rendersi necessaria la segmentazione del LV.

La segmentazione manuale è un processo dispendioso, tanto più data la quantità di esami che i cardiologi eseguono quotidianamente a causa dell'elevato numero di pazienti con malattie cardiache. Inoltre, a causa del rumore e della dipendenza dall'operatore, la valutazione ecocardiografica presenta un'elevata variabilità intra e inter-osservatore. Per questi motivi, diversi approcci diagnostici basati sull'intelligenza artificiale (AI) sono stati proposti per la valutazione automatica della funzionalità cardiaca. Questi algoritmi possono ridurre il carico di lavoro dei cardiologi, assistendoli nell'interpretazione degli ecocardiogrammi in modo più rapido, robusto e accurato.

Lo scopo di questa tesi è lo sviluppo di un algoritmo di AI completamente automatico per la diagnosi precoce di MI dall'**ecocardiografia 2-D**. Ciò comporta prima la **segmentazione del LV** e poi, attraverso la valutazione di alcuni parametri, l'**identificazione di MI**. Nello specifico, per la prima fase

viene utilizzato un modello di **Deep Learning** (U-Net), sfruttando la disponibilità di ecocardiogrammi con segmentazioni manuali. Per la seconda fase vengono testati diversi algoritmi di **Machine Learning** supervisionati, unitamente a tecniche di Data Augmentation, ottenendo alla fine i migliori risultati con il modello Random Forest.

Rispetto alla letteratura, l'approccio di questa tesi identifica i segmenti specifici del LV che presentano infarto, utilizzando, per la costruzione del modello, parametri clinici esistenti. Inoltre, il modello dimostra prestazioni e generalizzabilità superiori rispetto ad altri metodi. Infine, le segmentazioni generate ed i parametri calcolati vengono presentati al cardiologo consentendo la verifica umana della diagnosi prodotta. Il successo di questo algoritmo incoraggia la possibilità di future applicazioni nel contesto clinico; tuttavia, questo richiede prima la convalida con più dati.

Abstract

Myocardial infarction (MI) occurs when blood flow to the heart muscle is blocked, usually as the result of underlying coronary artery disease (CAD). In 2019, ischemic CAD was responsible for 1 in 6 deaths globally, making it the principal cause of death worldwide. At present, diagnosis of heart disease, and specifically MI, requires particular biomarker and electrocardiography findings. However, since the findings of these methods are sometimes inconclusive and can result in misdiagnosis with other syndromes, cardiologists recurrently rely on imaging techniques to establish a final judgment. Echocardiography is the technique chosen most often since it allows for visualization of the heart in a simple, real-time and cost-effective manner. From this exam, cardiologists study the movement of the walls of the left ventricle (LV) seeking for abnormal contractions which appear immediately after the onset of ischemia. To further evaluate cardiac functionality, parameters such as blood ejection fraction are also calculated from the echocardiogram. In order to achieve these tasks, segmentation of the LV may be necessary.

Manual segmentation is a tedious and time consuming process, more so given the amount of echo exams that cardiologists perform daily due to the high prevalence of cardiac disease. In addition, due to the noise and operator-dependency inherent to ultrasound imaging, **LV segmentation** and **MI diagnosis** present high intra and inter-observer variability. For these reasons, diagnostic approaches based on artificial intelligence are widely investigated to obtain automatic evaluation of heart functionality from echocardiography exams. These algorithms can help reduce cardiologist's workload by assisting in the interpretation of echocardiograms in a faster, robust and accurate manner.

The purpose of this thesis project is the development of a fully automatic artificial intelligence algorithm for the early detection of MI from **2-D echocardiography**. This involves first the segmentation of the LV and then, through the assessment of certain parameters, the identification of MI. Specifically, for the first step a **Deep Learning** model is used (U-Net), profiting from the availability of echocardiograms with manual segmentations. For the second step, different

supervised **Machine Learning** algorithms are tested together with Data Augmentation techniques, obtaining at the end the best results with the Random Forest model.

Compared to current literature, this thesis's approach identifies the specific segments of the LV that present infarction, using, for the construction of the model, existing clinical parameters. Additionally, the model demonstrates higher performance and generalizability in comparison to other papers. Finally, the generated segmentations and calculated parameters are intended to be presented to the cardiologist allowing for human verification of the produced diagnosis. The success of this algorithm encourages the possibility of future application in the clinical context; however, this first requires validation of the model with further data.

Contents

List of Figures	IV
List of Tables	VII
Acronyms.....	XI
1. Myocardial Infarction and its diagnosis	
1.1 Heart anatomy and physiology.....	1
1.2 Myocardial Infarction.....	4
1.3 Diagnosis of Myocardial infarction	5
1.3.1 Cardiac biomarkers	5
1.3.2 ECG.....	7
1.3.3 Echocardiography	8
2. Artificial Intelligence for MI detection	
2.1 Introduction to Artificial Intelligence	16
2.2 AI in left ventricle segmentation.....	17
2.3 AI in myocardial infarction classification.....	23
3. Materials and Methods	
3.1 Datasets	28
3.1.1 CAMUS	28
3.1.2 HMC-QU	30
3.1.3 Data division	32
3.2 Architecture of solution.....	34
3.3 Preprocessing	35
3.4 Segmentation.....	37
3.4.1 Convolutional Neural Networks	37
3.5 Post-processing	45

3.5.1	Morphological operations	45
3.5.2	Segment division.....	47
3.6	Classification.....	48
3.6.1	Feature extraction.....	49
3.6.2	Machine learning models.....	54
3.6.3	Data augmentation	58
3.6.4	Validation Metrics.....	59
4. Results		
4.1	Pre-processing	63
4.2	Segmentation.....	66
4.3	Post-processing	68
4.3.1	Morphological closing	68
4.3.2	Segment division.....	69
4.4	Classification.....	70
4.4.1	Feature extraction.....	70
4.4.2	KNN	81
4.4.3	SVM	82
4.4.4	RF.....	83
4.4.5	Data augmentation	84
4.5	Cascade approach.....	89
4.6	Processing time	90
5. Conclusions		
5.1	Open points	94
Appendix		
A.1	KNN classification results	96
A.2	SVM.....	99
A.3	RF	103
A.4	RF + DA.....	107

A.5 Cascade approach.....	111
Bibliography	116

List of Figures

Figure 1.1 Structure of the heart. Blood flow through the chambers and heart valves	2
Figure 1.2 System of coronary arteries with corresponding circulation sites..	4
Figure 1.3 Spectrum of myocardial injury, ranging from no injury to myocardial infarction.....	6
Figure 1.4 Pathological changes in ECG signal.....	8
Figure 1.5 Echocardiography modes	10
Figure 1.6 Apical four-chamber view	11
Figure 1.7 Apical four-chamber view structures	11
Figure 1.8 AutoStrain Philips application.....	14
Figure 2.1 Active polynomials approach for LV wall segmentation.....	19
Figure 2.2 Active shape model generation	20
Figure 2.3 U-Net for LV wall segmentation.....	23
Figure 2.4 Manual threshold method for MI detection.....	24
Figure 2.5 Machine learning approach for MI detection by a conventional classifier.....	25
Figure 3.1 Example echocardiography image and segmentation mask for a CAMUS patient (Patient 442).	29
Figure 3.2 Endocardial boundary segment division and division ratios.....	30
Figure 3.3 Example echocardiography image and segmentation mask for a HMC-QU patient (Patient 51).	32
Figure 3.4 Pipeline for proposed solution.....	34
Figure 3.5 Histogram comparison before (old) and after (new) image equalization.....	37
Figure 3.6 General architecture of a Convolutional Neural Network.....	38
Figure 3.7 Convolution layer in a CNN.....	40
Figure 3.8 Pooling operations	41
Figure 3.9 Transformation of a CNN into a FCNN with resulting heatmap ..	42

Figure 3.10 Comparison between ground truth and output segmentation of a net with no skip connections (FCN-32s), skip connection between final layer and a previous pooling layer (FCN-16s) and use of two skip connections (FCN-8s).....	43
Figure 3.11 U-Net architecture	44
Figure 3.12 Overview of morphological transformations.....	47
Figure 3.13 17-segment model of the left ventricle.....	48
Figure 3.14 Biplane Simpson Method using the end diastolic and end systolic apical 4- and 2- chamber views for estimation of LV volume and calculation of the ejection the fraction.....	50
Figure 3.15 Computation of the normalized maximum displacement of the 6 segments of the A4C view echo with the ground-truth labels (normal = 1, infarcted = 2)	53
Figure 3.16 KNN. Selected neighbors to a new point when k = 5	55
Figure 3.17 Division between classes using SVM	56
Figure 3.18 General architecture of a Radom Forest model.....	58
Figure 3.19 Confusion matrix with corresponding metrics.	60
Figure 3.20 ROC curve with model comparison.	61
Figure 4.1 Sample CAMUS and HMC-QU images and corresponding masks after initial formatting.....	64
Figure 4.2 CAMUS sample image and histogram before and after equalization.	65
Figure 4.3 HMC-QU sample image and histogram before and after equalization.....	66
Figure 4.4 Comparison of manual annotation and automatic mask before and after post-processing for an example patient.....	69
Figure 4.5 Segment division	70
Figure 4.6 Box plots of real and estimated LVEF for CAMUS images.....	71
Figure 4.7 Box plot of estimated LVEF for HMC-QU images.	72
Figure 4.8 LV axis length measurement.	73
Figure 4.9 Box plot of estimated GLS for CAMUS images.....	73
Figure 4.10 Box plot of estimated GLS for HMC-QU images.....	73
Figure 4.11 Box plots of estimated LS for all segments of CAMUS images.	74
Figure 4.12 Box plots of estimated LS for all segments of HMC-QU images.	75

Figure 4.13 Box plots of estimated maximum displacement for all segments of CAMUS images.....	76
Figure 4.14 Box plots of estimated maximum displacement for all segments of HMC-QU images.....	77
Figure 4.15 Box plots of estimated dynamic wall thickness for all segments of CAMUS images.....	79
Figure 4.16 Box plots of estimated dynamic wall thickness for all segments of HMC-QU images.....	80
Figure 4.17 ROC curves general classification: Threshold method vs RF+DA	86
Figure 4.18 ROC curves segment classification: Threshold method vs RF+DA	86
Figure 4.19 ROC curves Seg1 classification: Threshold method vs RF+DA	87
Figure 4.20 ROC curves Seg2 classification: Threshold method vs RF+DA	87
Figure 4.21 ROC curves Seg3 classification: Threshold method vs RF+DA	87
Figure 4.22 ROC curves Seg5 classification: Threshold method vs RF+DA	88
Figure 4.23 ROC curves Seg6 classification: Threshold method vs RF+DA	88
Figure 4.24 ROC curves Seg7 classification: Threshold method vs RF+DA	88

List of Tables

Table 1.1 Normal values for 2D echocardiographic volumes according to gender	13
Table 3.1 Clinical information available for each CAMUS patient	30
Table 3.2 HMC-QU patients with corresponding MI and Non-MI classification for each LV wall segment.....	31
Table 3.3 Data division for deep learning segmentation model	33
Table 3.4 Data division (patients) for machine learning classification model	34
Table 3.5 Data division (segments) for machine learning classification model	34
Table 4.1 U-Net model training hyper-parameters	67
Table 4.2 Network performance over complete dataset.	67
Table 4.3 Network performance over images belonging to the HMC-QU dataset.	68
Table 4.4 Network performance over images belonging to the CAMUS dataset.	68
Table 4.5 Network performance over complete dataset after post-processing.	69
Table 4.6 LVEF approximation methods with corresponding relative and absolute errors.	71
Table 4.7 Statistical values for real and estimated LVEF for CAMUS images.	71
Table 4.8 Statistical values for estimated LVEF for HMC-QU images.	72
Table 4.9 Statistical values for estimated GLS for CAMUS images.....	73
Table 4.10 Statistical values for estimated GLS for HMC-QU images.....	73
Table 4.11 Statistical values of estimated LS for all segments of CAMUS images.....	74
Table 4.12 Statistical values of estimated LS for all segments of HMC-QU images.....	75
Table 4.13 Statistical values of estimated maximum displacement for all segments of CAMUS images.	77

Table 4.14 Statistical values of estimated maximum displacement for all segments of HMC-QU images.	78
Table 4.15 Statistical values of estimated dynamic wall thickness for all segments of CAMUS images.	79
Table 4.16 Statistical values of estimated dynamic wall thickness for all segments of HMC-QU images.	80
Table 4.17 Cross-correlation between MI label and features extracted.	81
Table 4.18 Results for KNN Model 1	81
Table 4.19 Results for KNN Model 2.....	82
Table 4.20 Results for KNN Model 3.....	82
Table 4.21 Results for SVM Model 1	83
Table 4.22 Results for SVM Model 2.....	83
Table 4.23 Results for SVM Model 3.....	83
Table 4.24 Results for RF Model 1.....	84
Table 4.25 Results for RF Model 2.....	84
Table 4.26 Results for RF Model 3.....	84
Table 4.27 Results for RF+DA Model 1.....	85
Table 4.28 Results for RF+DA Model 2.....	85
Table 4.29 Results for RF+DA Model 3.....	85
Table 4.30 General results for segment classification with cascade approach	89
Table 4.31 Results for segment classification with cascade approach evaluating test segments.....	90
Table 4.32 Processing time with and without GPU for a sample echo video	91
Table A.1 KNN model 1: Training set confusion matrix.	96
Table A.2 KNN model 1: Test set confusion matrix.	96
Table A.3 KNN model 3: Seg1 Test set confusion matrix	96
Table A.4 KNN model 3: Seg2 Test set confusion matrix	97
Table A.5 KNN model 3: Seg3 Test set confusion matrix	97
Table A.6 KNN model 3: Seg5 Test set confusion matrix	98
Table A.7 KNN model 3: Seg6 Test set confusion matrix	98
Table A.8 KNN model 3: Seg7 Test set confusion matrix	98
Table A.9 KNN model 2: Training set confusion matrix.	99
Table A.10 KNN model 2: Test set confusion matrix.	99
Table A.11 SVM model 1: Training set confusion matrix.	100

Table A.12 SVM model 1: Test set confusion matrix.	100
Table A.13 SVM model 2: Training set confusion matrix.	100
Table A.14 SVM model 2: Test set confusion matrix.	101
Table A.15 SVM model 3: Seg1 Test set confusion matrix	101
Table A.16 SVM model 3: Seg2 Test set confusion matrix	101
Table A.17 SVM model 3: Seg3 Test set confusion matrix	102
Table A.18 SVM model 3: Seg5 Test set confusion matrix	102
Table A.19 SVM model 3: Seg6 Test set confusion matrix	103
Table A.20 SVM model 3: Seg7 Test set confusion matrix	103
Table A.21 RF model 1: Training set confusion matrix.	103
Table A.22 RF model 1: Test set confusion matrix.	104
Table A.23 RF model 2: Training set confusion matrix.	104
Table A.24 RF model 2: Test set confusion matrix.	104
Table A.25 RF model 3: Seg1 Test set confusion matrix	105
Table A.26 RF model 3: Seg2 Test set confusion matrix	105
Table A.27 RF model 3: Seg3 Test set confusion matrix	106
Table A.28 RF model 3: Seg5 Test set confusion matrix	106
Table A.29 RF model 3: Seg6 Test set confusion matrix	106
Table A.30 RF model 3: Seg7 Test set confusion matrix	107
Table A.31 RF+DA model 1: Training set confusion matrix.	107
Table A.32 RF+DA model 1: Test set confusion matrix.	107
Table A.33 RF+DA model 2: Training set confusion matrix.	108
Table A.34 RF+DA model 2: Test set confusion matrix.	108
Table A.35 RF+DA model 3: Seg1 Test set confusion matrix	109
Table A.36 RF+DA model 3: Seg2 Test set confusion matrix	109
Table A.37 RF+DA model 3: Seg3 Test set confusion matrix	109
Table A.38 RF+DA model 3: Seg5 Test set confusion matrix	110
Table A.39 RF+DA model 3: Seg6 Test set confusion matrix	110
Table A.40 RF+DA model 3: Seg7 Test set confusion matrix	110
Table A.41 General results for Seg1 classification with cascade approach .	111
Table A.42 General results for Seg2 classification with cascade approach .	111
Table A.43 General results for Seg3 classification with cascade approach .	112
Table A.44 General results for Seg5 classification with cascade approach .	112
Table A.45 General results for Seg6 classification with cascade approach .	112
Table A.46 General results for Seg7 classification with cascade approach .	113

Table A.47 Results for Seg1 test segments classification with cascade approach	113
Table A.48 Results for Seg2 test segments classification with cascade approach	114
Table A.49 Results for Seg3 test segments classification with cascade approach	114
Table A.50 Results for Seg5 test segments classification with cascade approach	114
Table A.51 Results for Seg6 test segments classification with cascade approach	115
Table A.52 Results for Seg7 test segments classification with cascade approach	115

Acronyms

MI

Myocardial Infarction

CAD

Coronary Artery Disease

LV

Left Ventricle

EF

Ejection Fraction

ES

End-systolic

ED

End-diastolic

A4C

Apical Four-chamber View

A2C

Apical Two-chamber View

RWMA

Regional Wall Motion Abnormalities

STE

Speckle Tracking Echocardiography

GLS

Global Longitudinal Strain

AI

Artificial Intelligence

ML

Machine Learning

DL

Deep Learning

CNN

Convolutional Neural Network

SVM

Support Vector Machine

KNN

K-Nearest Neighbors

RF

Random Forest

CHAPTER 1

1 Myocardial Infarction and its diagnosis

This first chapter briefly presents the anatomy and physiology of the heart in order to introduce the concept of myocardial infarction (MI) and its epidemiological significance. The current methods for diagnosis of MI are also discussed. Special attention is given to the possibility of detection through echocardiography, for which a general explanation of this imaging technique is offered. Afterwards is presented the difficulties still present in the process of diagnosis through ultrasound imaging and how the implementation of artificial intelligence offers a new path for faster detection and diagnosis, premise on which this thesis project stands.

1.1 Heart anatomy and physiology

The heart is the main organ in the cardiovascular system. It is in charge of pumping blood throughout the body and works together with the nervous and endocrine systems to control heart rate and blood pressure [1]. As can be seen in **Figure 1.1**, the heart is divided into four main chambers: right and left atrium and right and left ventricle. The passage of blood throughout the heart involves the following steps [1]:

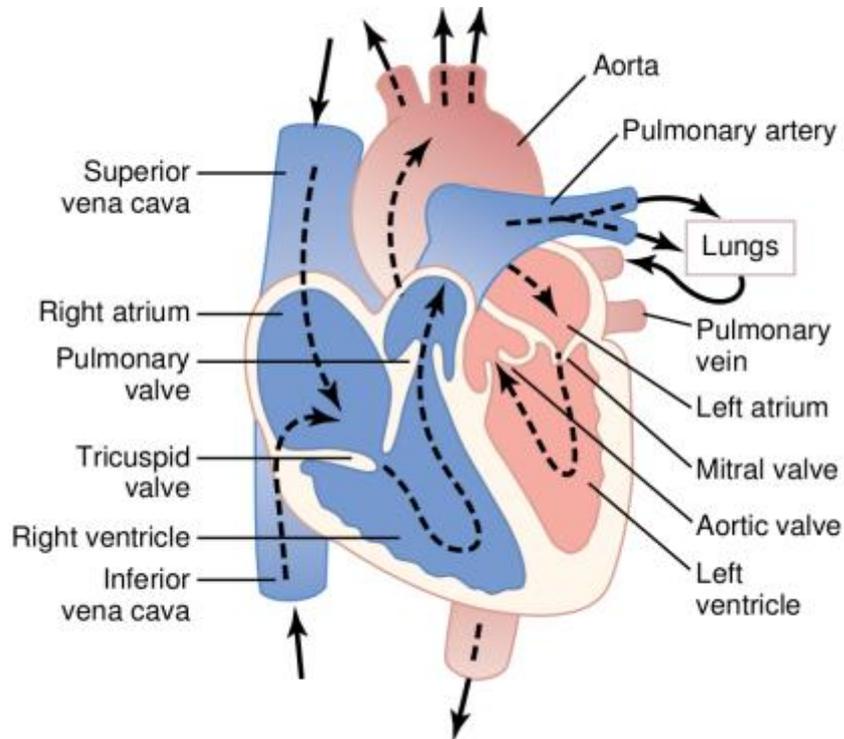


Figure 1.1 Structure of the heart. Blood flow through the chambers and heart valves. [2]

1. The inferior and superior vena cava deliver oxygen-poor blood into the right atrium.
2. The right atrium pumps the oxygen-poor blood into the right ventricle passing through the tricuspid valve.
3. Oxygen-poor blood is pumped from the right atrium to the lungs through the pulmonary artery.
4. Oxygen-filled blood is carried from the lungs to the left atrium through the pulmonary veins.
5. The left atrium pumps the oxygen-rich blood into the left ventricle passing through the mitral valve.
6. The left ventricle pumps the oxygen-rich blood to the rest of the body passing through the aorta.

All the previous steps, occurring between successive heartbeats, make up the cardiac cycle. This cycle is divided into two main phases: diastole and systole. During diastole the heart is relaxed and begins to expand as it receives blood

into both ventricles coming from both atria. Systole occurs when the ventricles contract to eject the blood out of the heart, the right ventricle to the lungs and the left ventricle to the rest of the body systems [3].

The action of pumping is obtained by the contraction of the walls of the heart. The walls are composed of 3 different layers, from inner to outer: endocardium, myocardium and epicardium. The myocardium, middle and thickest layer, is made up of muscle tissue responsible for the contractile function of the heart [4]. Its blood supply comes directly from the system of coronary arteries that runs within the epicardial layer (**Figure 1.2**). Sufficient blood supply is required in order to generate enough force to correctly pump blood, especially in the walls of the left ventricle which are significantly thicker compared to those of the other cavities as a result of the greater contractile force required. This is due to the fact that the LV contracts against the pressure of the whole systemic circulation, facing a great resistance and thus requiring a high oxygen demand. This explains why both left and right coronary arteries supply blood to the left ventricle (**Figure 1.2**). Failure to supply such demand increases the effort made by the myocardium to pump blood which can bring significant implications and consequences for heart muscle functionality [5].

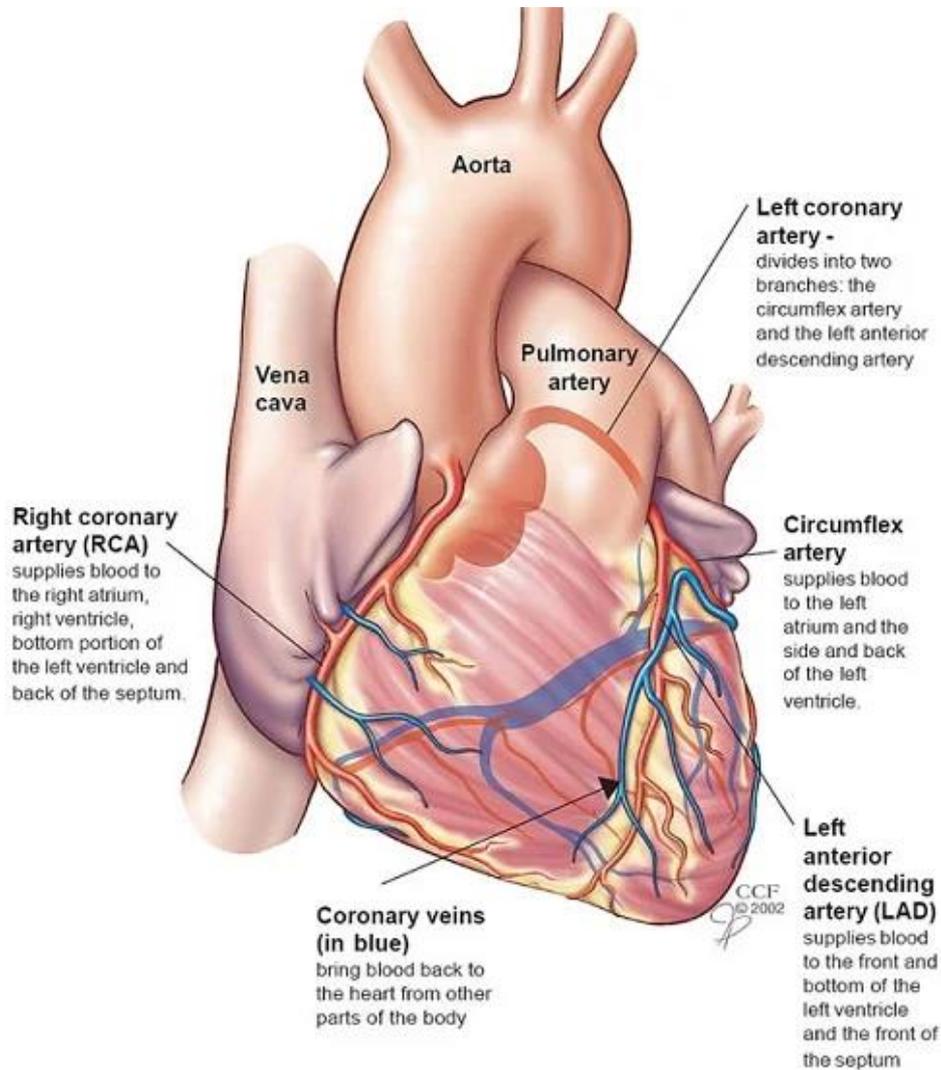


Figure 1.2 System of coronary arteries with corresponding circulation sites. [1]

1.2 Myocardial Infarction

Myocardial infarction (MI) is defined as myocardial cell death due to prolonged ischemia, usually caused by underlying coronary artery disease (CAD) [6]. CAD is the most commonly diagnosed heart disease and the leading cause of death worldwide accounting to 1 in 6 deaths globally [7]. It is caused by plaque buildup in the walls of the arteries from continuous cholesterol deposit which

reduces the effective area for blood circulation. This can set off an inflammatory response which further narrows the inside of the arteries resulting in partial or total block of the blood flow. In turn, not enough oxygen or nutrients reaches the heart muscle which can eventually lead to MI [8].

Presence of MI is generally evident with shortness of breath, pain around the chest, shoulders, back, and arms; however, these symptoms may not occur in the early stages. Due to the blockage of the coronary artery and deprivation of blood supply, there is progressive damage to the affected part of the myocardium with increased risk for heart failure, wall ruptures, arrhythmias, sudden cardiac death, among other complications [4]. An early detection of MI and its localization is critical to limit the extent of damage and prevent death or disability [9].

Particular interest is given to the left ventricle (LV) wall. Not only is right ventricular infarction uncommon but additionally, even if the right ventricle is affected, almost always the left ventricle will also be affected due to the coronary anatomy [5]. Myocardial infarction can be classified by size: microscopic (focal necrosis), small (10% of the LV myocardium), moderate (10–30% of the LV myocardium), and large (30% of the LV myocardium) [10]. Pathologically it can be defined as acute, healing, or healed. Acute MI denotes infarction less than 3-5 days and is characterized by the presence of an inflammatory process [11]. Healing infarction (7 to 28 days) demonstrates reduced inflammation and the start of tissue replacement. Since the myocardium has very little self-renewal capability, healed infarction is manifested as scar tissue without cellular infiltration. The entire process leading to a healed infarction usually takes at least 5–6 weeks [12].

1.3 Diagnosis of Myocardial infarction

1.3.1 Cardiac biomarkers

The clinical definition of MI denotes the presence of acute myocardial injury detected by abnormal cardiac biomarkers in the presence of evidence of acute

Myocardial Infarction and its diagnosis

myocardial ischemia [13]. Cardiac troponin I (cTnI) and T (cTnT) are proteins present in the contractile apparatus of myocardial cells and are expressed almost exclusively in the heart. The release of these markers increases following cardiac injury, which can be detected from a blood sample. If levels rise with at least one value above the 99th percentile upper reference limit, myocardial injury is diagnosed. Increases in cTnI values have been found to occur only following injury to cardiac tissue. In contrast, cTnT is also expressed in the presence of skeletal muscle injury. Still, cardiac troponin increase, either type I or T or both, is clinically used to signal myocardial injury, which encompasses a variety of injuries or diseases among which MI is included (**Figure 1.3**). For MI diagnosis, evidence of ischemia is required and therefore complementary evaluations, such as electrocardiography (ECG) exams, are necessary.

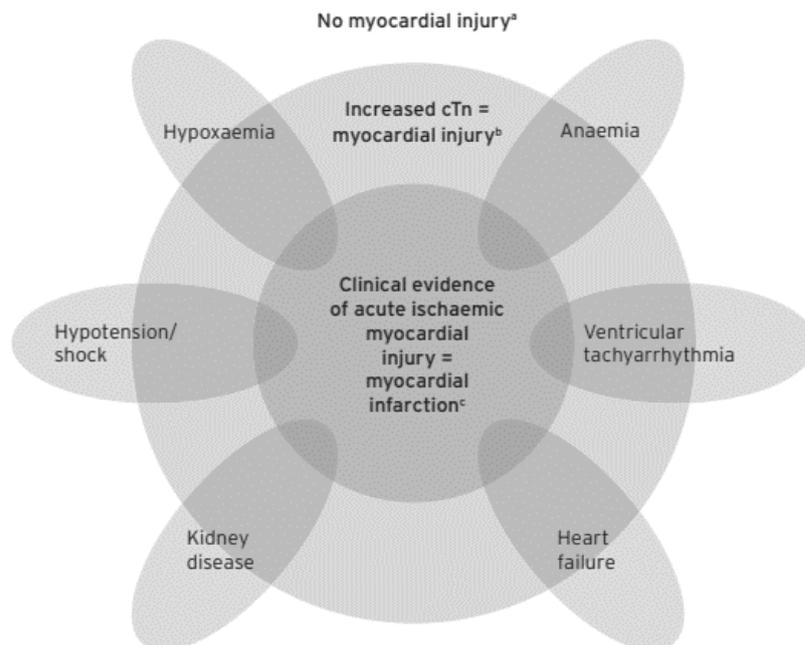


Figure 1.3 Spectrum of myocardial injury, ranging from no injury to myocardial infarction. [13]

1.3.2 ECG

Myocardial ischemia can be detected from ECG changes and, when coupled with a rising pattern of cTn values, acute MI is diagnosed. Acute myocardial ischemia is often associated with dynamic changes in ECG waveform (**Figure 1.4**) like ST-segment elevation or depression (STEMI) or T wave inversions. Some patients with acute coronary occlusion may have an initial ECG without ST-segment elevation and are labeled as NSTEMI. Pathologic Q waves are also sometimes present representing increased prognostic risk (Q-wave MI). The ECG by itself is often insufficient to diagnose acute myocardial ischemia or infarction, since these same ECG abnormalities can be observed in other conditions such as acute pericarditis, LV hypertrophy (LVH), left bundle branch block (LBBB), Brugada syndrome and thrombosis with thrombocytopenia syndrome (TTS). For this reason, it is common for cardiologists to additionally rely on imaging techniques to determine the appropriate diagnosis [13].

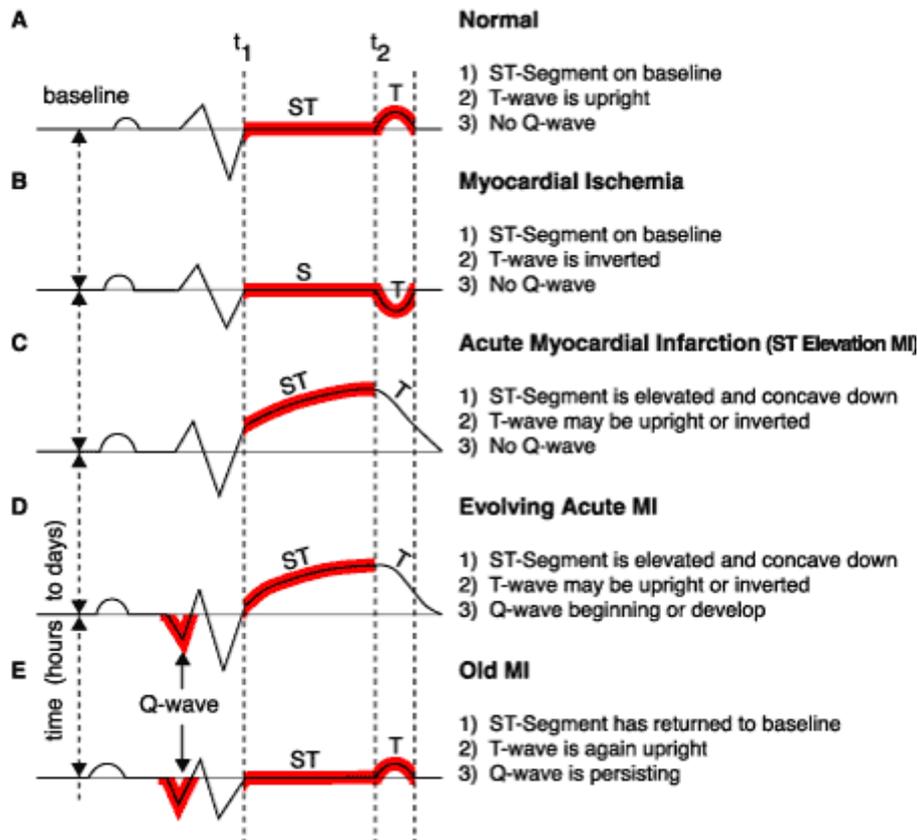


Figure 1.4 Pathological changes in ECG signal. A) Normal ECG B) T-wave inversion C) ST Elevation D) ST Elevation and Q-wave presence E) Q-wave presence. [14]

1.3.3 Echocardiography

Following MI, contractile function of the affected area decreases as a result of cell injury, inflammation, and fibrotic tissue replacement [4]. Reduced wall motion is the first abnormality to set in after the onset of myocardial ischemia, preceding metabolic and electrocardiographic abnormalities [9]. Therefore, visualization of the motion of the heart through medical imaging, proves to be an effective and early-stage technique to detect MI. 2D echocardiography, specifically, is often used to diagnose heart diseases as it is a non-invasive, cheap, and simple procedure [10]. Before describing how MI diagnosis is obtained through an echo analysis, a brief introduction to this technique is required.

1.3.3.1 Basis of echocardiography

Echocardiography is an echography exam of the heart and is thus an application of ultrasound imaging. Ultrasound imaging is highly used in the clinical context as it allows for non-invasive real-time observation of internal body structures with moderate spatial and temporal resolution [15]. For the examination, a transducer (also known as probe) is placed on the skin in the area of interest and ultrasound waves, generated inside the probe, are transmitted into the body. These are usually of frequencies ranging from 1.5 to 7.5 MHz which will travel at a speed of approximately 1540 m/s while in the heart tissue [16]. The waves reflect off of the body structures and are picked up again by the transducer which will generate an electrical impulse proportional to the wave echo amplitude. This signal, in addition to the time it takes for the wave to travel from the transducer and back, provides the information necessary to produce the echocardiography image [15]. The quality of the image is limited by the difficulties inherent to ultrasound imaging, mainly intrinsic noise, low spatial resolution, poor penetration through bone (which difficult imaging of inside-chest structures), and high operator dependency [17].

There are various ultrasound imaging modes: A-mode (amplitude-mode), B-mode, M-mode (motion-mode), color Doppler, power Doppler and spectral Doppler (**Figure 1.5**). The 2-dimensional standard images normally associated with echography exams correspond to the B-Mode. In this mode, a cross sectional view of the structure is obtained in a grayscale image in which each pixel takes on a shade depending on its density and depth. Regarding echocardiography, different views of the heart can be obtained in B-Mode depending on the position and orientation of the transducer [17]. The positions are parasternal, apical, subcostal, suprasternal, and the orientation are long axis, short axis, two chamber and four-chamber [18]. Since this thesis project focuses on the apical four-chamber view (A4C), only this view will be described in detail.

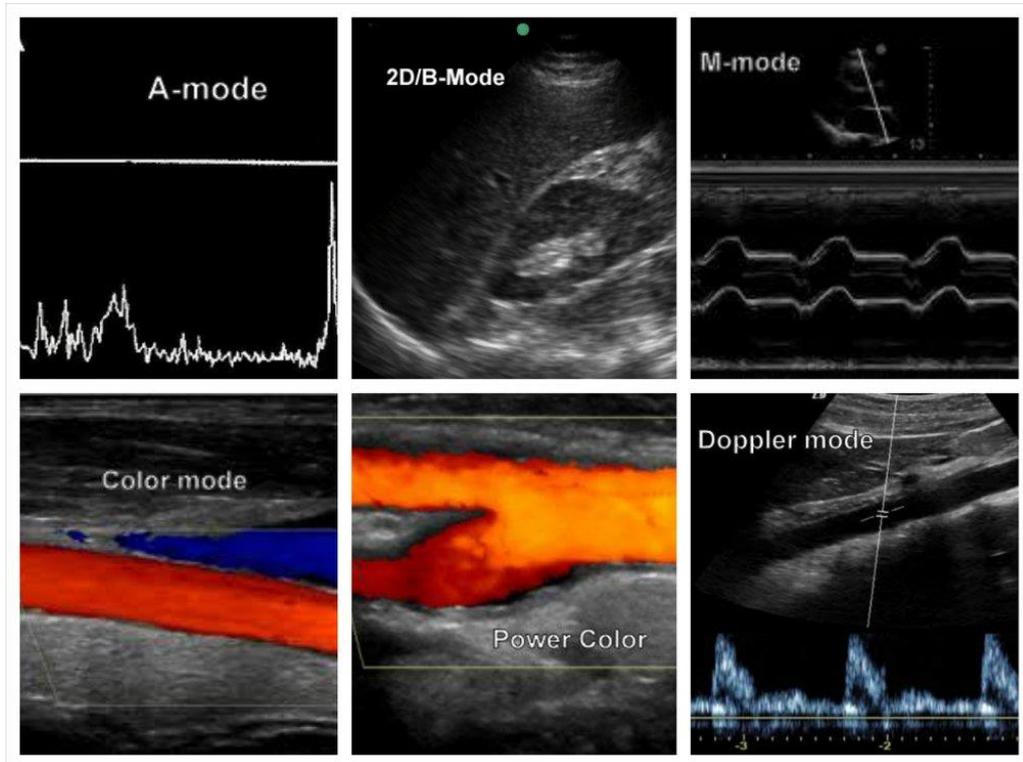


Figure 1.5 Echocardiography modes. [19]

The A4C image is obtained with the patient in left lateral position and by placing the transducer in the 4th or 5th intercostal space oriented towards the patient's left shoulder (**Figure 1.6**) [20]. This produces an image where all four chambers of the heart are seen simultaneously as depicted in **Figure 1.7**. The left ventricle appears as a truncated ellipse, with the interventricular septum, apex, and lateral walls visualized. Both mitral and tricuspid valves can be observed. In the lower half of the image are the left and right atria and the interatrial septum. The left and right superior pulmonary veins are visible at the bottom of the image.

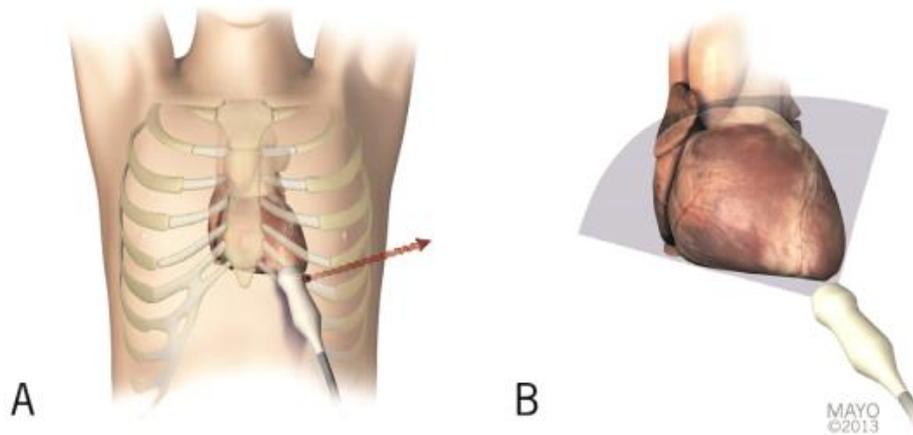


Figure 1.6 Apical four-chamber view. A) Orientation of the probe indicator B) Direction of the ultrasound beam on the heart. [21]

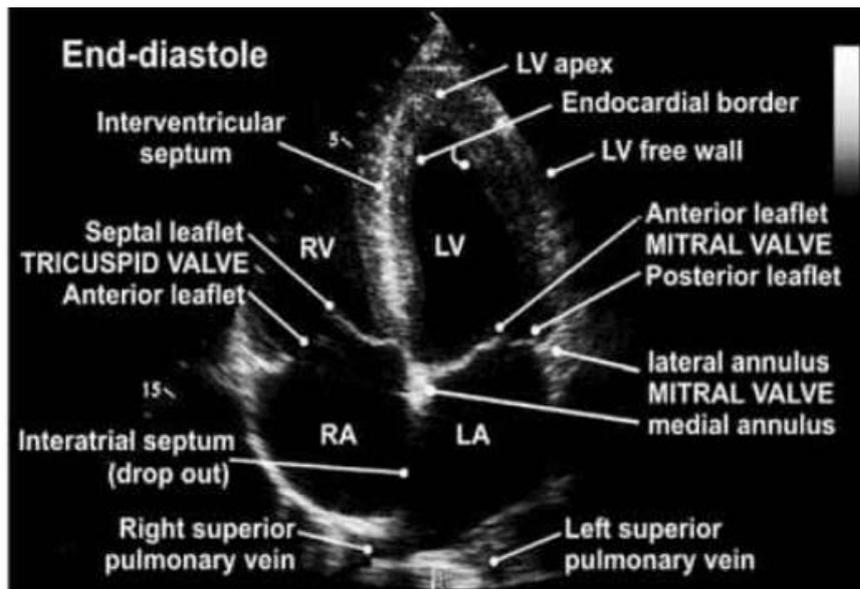


Figure 1.7 Apical four-chamber view structures. [22]

1.3.3.2 Identification of ischemia with echocardiography

A patient's echocardiogram is the primary tool to detect, identify and quantify regional wall motion abnormalities (RWMA) resulting from ischemia and thus

MI [23]. As stated before, particular interest is given to the observation of the motion of the LV wall. Currently, evaluation of RWMA is a manual or semi-automatic process. In both cases, cardiologists first have to segment the left ventricle either in manual mode, by completely sketching the area of interest in each frame of the echo video, or semi-automatic mode, by carefully selecting specific points of the left ventricle in at least one frame [24]. Either way, segmentation is a tedious and time consuming process and more so when the analysis of a complete cardiac cycle is required [25]. In addition, there is no unified standard to evaluate the quality of segmentation results which leads to high intra and inter-observer variability [24].

Evaluation of motion of the LV, after it has been segmented, is in many cases done visually by the cardiologist making interpretation and detection of RWMA highly subjective and thus operator dependent. This analysis can only offer semi-quantitative information by scoring the myocardial segments of the LV wall based on the severity of the wall motion abnormality as follows: 1–normal or hyperkinesia, 2–hypokinesia, 3–akinesia, 4–dyskinesia [26]. Dyskinesia means abnormal movement in which instead of contracting during systole, the myocardium expands during this phase. Akinesia refers to a lack of movement or no contraction while hypokinesia means reduced movement. Lastly, hyperkinesia occurs when the myocardium contracts more than usual and this often appears as a compensatory mechanism in segments which do not present MI [27].

From the segmentation of the LV an important quantitative clinical parameter can be calculated: left ventricle ejection fraction (LVEF). This value corresponds to the fraction of stroke volume (calculated as the difference between end-diastolic volume and end-systolic volume) in relation to the volume of blood in the LV at the end of diastole, making it a measure for assessing LV systolic function [28]. As such, it can aid cardiologists in clinical evaluation of different cardiovascular diseases. For this reason The American Society of Echocardiography and the European Association of Cardiovascular Imaging, have defined normal ranges for 2D echocardiography obtained LVEF

(**Table 1.1**). Even though this is the most commonly used echocardiographic parameter for cardiac evaluation, it offers only a global assessment of function and gives no information in regards to contractility or displacement, for which it cannot be used by itself to assess infarction.

Table 1.1 Normal values for 2D echocardiographic volumes according to gender. Adapted from [29]

Parameter	Male		Female	
	Mean +/- SD	2-SD range	Mean +/- SD	2-SD range
LV EDV (mL)	106 +/- 22	62-150	76 +/-15	46 – 106
LV ESV (mL)	41 +/- 10	21-61	28 +/- 7	14-42
LVEF	62 +/- 5	52-72	64 +/- 5	54-74
LV, left ventricular; EDV, end-diastolic volume; ESV, end-systolic volume; EF, ejection fraction				

In order to obtain a quantitative assessment of motion, displacement and deformation of the LV, cardiologists are now being aided by speckle tracking echocardiography (STE), the most advanced technology currently available in the market. Speckle tracking uses computer vision to track the displacement of “speckle patterns” of the myocardium in 2D echocardiography to then assess not only deformation but also strain during the cardiac cycle (**Figure 1.8**) [30]. Global longitudinal strain (GLS) is the most commonly used strain-based measure for LV systolic function. Similarly to LVEF, GLS is calculated as the fraction of myocardial deformation (difference between myocardial length at end systole and end diastole) in relation to myocardial length at end-diastole [29]. Through assessment of deformation and semi-automatic tracking of the myocardial wall, STE removes part of the subjectivity from the process of visual estimation and diagnosis. However, this technique still presents some drawbacks such as sensitivity to frame rate and image quality. A high cardiac frequency is associated with a lower frame rate and hence low possibility of tracking [31]. Availability only of low-quality echoes with high noise presence

and lack of contrast, which is not uncommon especially in the hospitals of many developing countries, further hinders tracking accuracy [32]. Additionally, differences in algorithms between vendors leads to different performances and results during analysis, demonstrating a lack of standardization [29]. Even under ideal conditions, many studies using the speckle tracking technique in 2D echocardiography, only reach around 80% to 85% sensitivity in the identification of the infarcted segments [23] [33].

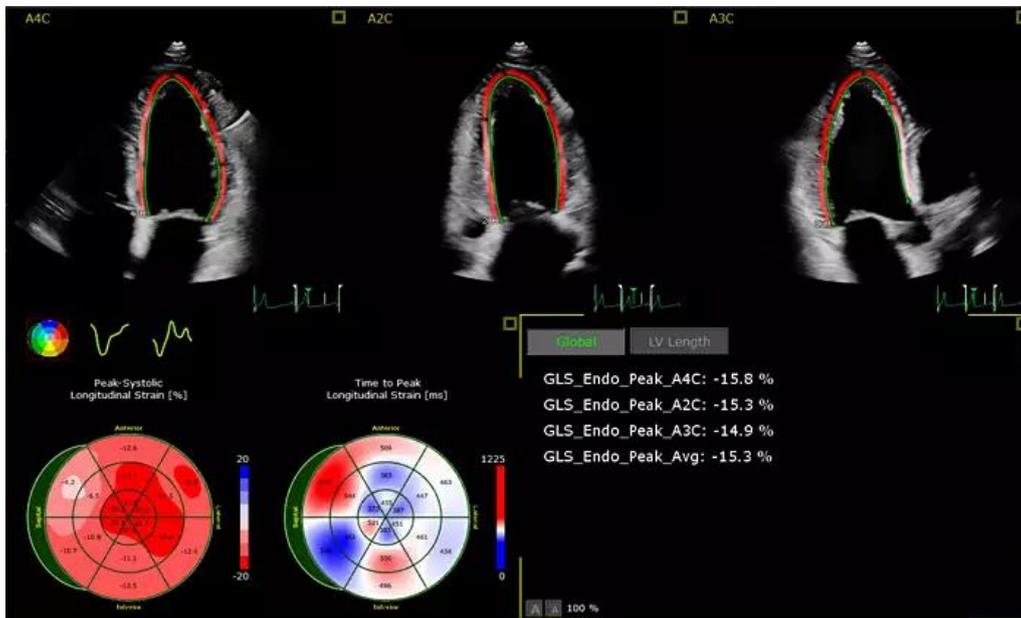


Figure 1.8 AutoStrain Philips application. [34]

STE has evidenced the advantages of computer-aided diagnosis, offering support to doctors by speeding up analysis and calculating parameters that visually could only be estimated. In order to solve the difficulties still present in the diagnosis of MI through medical imaging and to obtain more accurate diagnostic results, the next step in computer-aided diagnosis is the use of more robust artificial intelligence (AI) tools; tools that assist cardiologists both in the step of segmentation of the LV as well as diagnosis of MI. This would facilitate interpretation of echo exams, reduce cardiologists work pressure, and more importantly it could lead to saving lives.

CHAPTER 2

2 Artificial Intelligence for MI detection

Artificial intelligence encompasses many very different techniques. Therefore, an extensive literature review, regarding current AI approaches for LV segmentation in echocardiographic images and MI classification, was required before defining the specific approach that would be selected for this thesis project. The following chapter first introduces the concept of AI and machine learning to then present the most relevant methods and articles found during the literature review conducted.

2.1 Introduction to Artificial Intelligence

Artificial intelligence is broadly defined as “the capability of a machine to imitate intelligent human behavior” [35]. This encompasses many subfields such as natural language processing, robotics, computer vision and machine learning (ML). As far as this thesis project is concerned, the last two categories are the ones of interest. Regarding computer vision, this field of AI trains computers to capture meaningful information from images, videos and other visual data and take decisions or make reports based on that information [36]. Computer vision is not completely unlinked to machine learning, in fact computer vision can be obtained through the use of machine learning techniques. ML encloses a variety of models and algorithms designed to give computers the ability to learn without explicitly being programmed. This requires interpretation of external data, learning and adaptation from such data, and using the knowledge obtained to complete a certain task or goal [37]. Artificial neural networks (ANN’s) are brain-inspired machine learning models. A special type of ANN is a deep artificial neural network: model that gave birth to deep learning. Unlike other ML techniques, deep learning is able

to automatically learn features from the input data eliminating the need of manual feature extraction and enabling the use of larger data sets [38].

ML algorithms can be classified according to the type of learning process, the two most recognized being supervised and unsupervised [37]. Supervised learning methods require inputs with their corresponding correct outputs (labels) and the model learns through error to correctly yield the desired output [39]. In this way, the algorithm will always assign new data a specific label from the ones used during training. Unsupervised learning, on the contrary, uses unlabeled data which is analyzed in order to find patterns or data groupings without the need for human intervention. This type of learning method will not be able to label data but will rather cluster it in groups based on their similarities and differences [35]. Supervised ML is the most common type used today.

The remainder of this chapter will explore in greater depth the most relevant, supervised and unsupervised, computer vision and machine learning techniques found in literature related to the subject of this thesis. Such review was conducted previous to the definition of this project's approach in order to select the most viable and appropriate strategy.

2.2 AI in left ventricle segmentation

In literature can be found many different AI techniques applied to the process of image segmentation, however only some are suitable for particular images and applications as is ultrasound imaging and LV segmentation. Regarding computer vision, the most common technique reviewed, aside from speckle tracking, is active contours (AC) also called snakes [40]. This model consists of an energy-minimizing spline which is controlled by external constraint forces and internal image forces that pull it towards features such as lines and edges. Since the snake depends on a minimization function, it can follow moving visual features by simply tracking the same local minimum, proving its active property. In order to converge to the expected shape, this model requires previous knowledge of the desired contour and relies on other

mechanisms (manual or automatic) to initially place the snake near the area of interest.

In the year 2020, Kiranyaz *et al.* [23], proposed a novel approach called Active Polynomials to capture the global motion of the LV wall in a robust and accurate way. Their method builds upon the AC model by fitting a 4th-order smooth polynomial over the converged snake in order to deal with the problem of noise in echocardiography images. The overview of this approach is seen in **Fig. 2.1**. The study constructed a public dataset (HMC-QU) of 160 echocardiography videos on which they tested their model. In the end, their approach proved to correctly track the LV wall motion despite the fact that some echoes were poor quality with high noise levels and low temporal resolution (25 fps), characteristics that would have made speckle tracking unviable. However, their method still required manual location of the initial snake. In their case, a cardiologist has to manually select three specific reference points on the LV wall for the first frame. Such points define the final segmentation obtained, consequently, this study is not able to completely eliminate intra and inter-observer variability currently present in LV segmentation. On top of that, the solution presented in this study takes around 36 seconds to process a single frame, proving that most active contour methods are very time-consuming and inefficient. With a 25 fps rate it would take 900 seconds to process a single second of echo video.

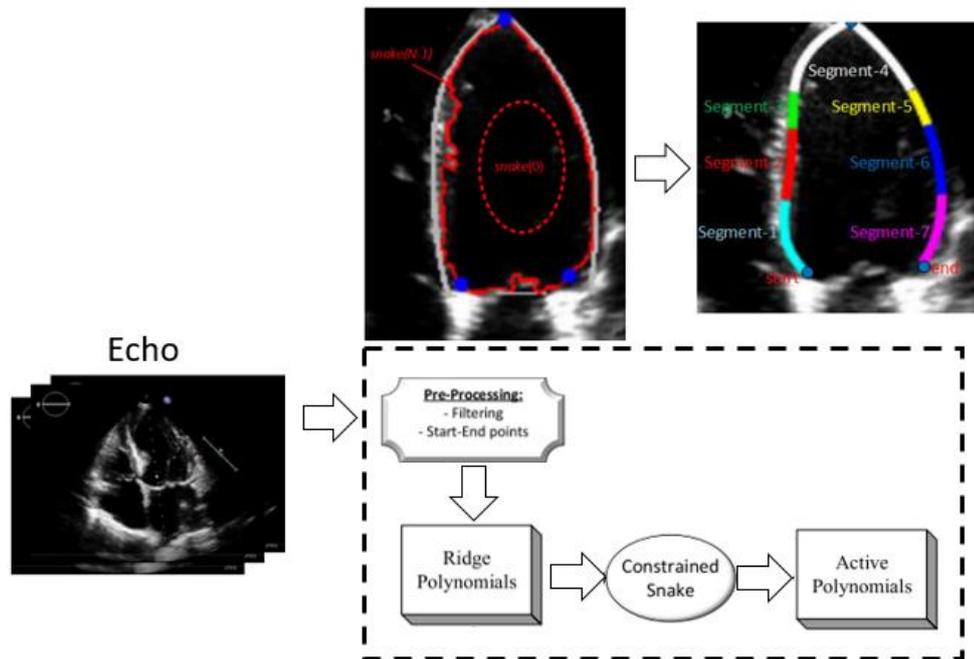


Figure 2.1 Active polynomials approach for LV wall segmentation. Adapted from [23]

Another computer vision technique commonly found in literature for segmentation and tracking of the LV are the active shape (AS) models. AS models integrate prior shape knowledge which is learned from training images with manually drawn contours. Through Principal Component Analysis (PCA), the model defines main variations between the training images. This enables an automatic recognition of possible/acceptable contours and a generation of a starting shape. In addition, AS models contain matrices that describe the texture of the lines perpendicular to control points selected from the drawn contour. This additional information is used for correcting positioning in the search step through modification of the starting shape boundary points [41]. This type of model has been used for ultrasound image segmentation as an alternative computer vision method to AC because it shows more robustness to speckle noise, shadows, and occlusions since its decisions are based on training data [42]. The generation of the ASM for an application for ultrasound prostate image segmentation is presented in **Figure 2.2**.

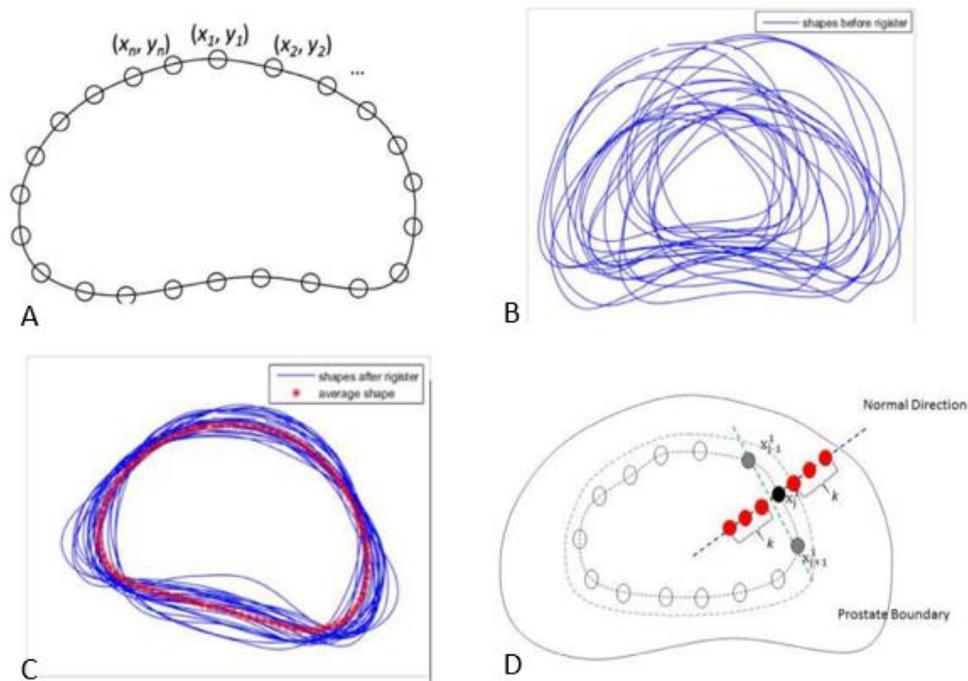


Figure 2.2 Active shape model generation. A) Manual labeling of boundary point B) Contour of all training shaped C) Aligned contours with the average shape (in red) D) Gradient profile (red dots) of the j -th boundary point (black dot). Adapted from [43]

One specific study from 2019 generated a fully annotated dataset for the purpose of echocardiographic assessment [44]. The CAMUS dataset includes manual segmentation of the endocardium and epicardium of the left ventricle and left atrium wall of 500 patients. The study applies different AI techniques to the same dataset for comparison, one such technique being an active shape model. Since ASM take the geometry of the target into consideration, a single model was necessary to obtain the three contours (endocardium and epicardium of left ventricle and left atrium). Regarding the left ventricle, the article presents the following results: for end diastole dice values of 0.920 and 0.917 for LV endocardial and epicardial contour, respectively; for end systole dice values of 0.861 and 0.900 for LV endocardial and epicardial contour, respectively. It is well known that the LV shape in the end systole frame is more difficult to segment than the end diastole which confirms why performance is lower for ES frames. No information about inference time is given in this article regarding the ASM approach.

Although computer vision has proven to obtain acceptable performance for the task of LV echocardiography segmentation, the techniques used are usually at most semi-automatic and very time consuming. With this in mind, currently more studies are placing greater focus on the use of machine learning approaches. Applications of unsupervised clustering ML techniques are not commonly found in literature for ultrasound images for LV segmentation due to the difficulties it presents, mainly the similarity in gray level between myocardial wall and other structures in the left ventricle. This creates a problem since clustering techniques use global image features, such as intensity level, to group pixels together [45]. In consequence, most ML algorithms found in literature for echocardiography segmentation belong to the supervised learning category.

The same article discussed before which used the CAMUS dataset and tested an AS model for LV segmentation, tested as well a random forest (RF) approach. RF is constituted by multiple decision trees which, in this case, take as input a patch of the echo image and in output supply a label patch which is obtained as the mean prediction of the individual trees. For the training process each tree is given a random subset of the training data and corresponding input features from which it will learn a set of split functions. Those functions will group patches with similar intensity levels and segmentation patterns. The images used for testing will be divided in same dimension patches as used for training, each patch will pass through the splitting functions of each tree, and the output segmentation for each specific patch will correspond to the mean label patch computed from the reached leaves [44]. This random forest approach, however, did not obtain better results in comparison to the computer vision AS model method. The dice values obtained for the segmentation were the following: for end diastole frames dice values of 0.895 and 0.914 for LV endocardial and epicardial contour, respectively; for end systole frames dice values of 0.848 and 0.901 for LV endocardial and epicardial contour, respectively.

Not many articles were found that used other classic ML methods to solve the problem of LV segmentation in 2D echocardiography images, possibly because

manual feature extraction from ultrasound images of the heart does not produce sufficient meaningful information to obtain an accurate segmentation. Besides, the success of such methods in delineating fuzzy boundaries of organs in other types of medical images, such as radiological images, has proven to be limited [46]. Instead, articles testing ML deep learning methods are becoming ever more abundant as more data becomes publicly available.

The CAMUS dataset was also used for testing the accuracy of a deep learning model. For this, a fully convolutional neural network was applied with a U-Net architecture. A single network was trained for the automatic segmentation of the three regions of interest: endocardium and epicardium contour of the left ventricle and left atrium wall. Of all the methods tested in this article, deep learning obtained the highest dice score: for end diastole frames dice values of 0.939 and 0.954 for LV endocardial and epicardial contour, respectively; for end systole frames dice values of 0.916 and 0.945 for LV endocardial and epicardial contour, respectively. Furthermore, of all the methods this one presented the shortest inference time, 0.14 +- 0.06 seconds per frame [44].

Another study focused on the segmentation only of the left ventricle wall from low-quality echocardiography by using deep learning [32]. Again a U-Net architecture was employed and in this case trained with 109 echocardiography exams and their corresponding ground-truth segmentations, which were added to the HMC-QU dataset as another contribution made by this article. **Figure 2.3** depicts the LV wall segmentation of each frame in an echo using the trained encoder-decoder convolutional neural network (E-D CNN) model. The proposed approach achieved a sensitivity of 95.72% and specificity of 99.58%, highest values among all the articles reviewed. Again, time efficiency is very high and the designed network takes 2.58 seconds to process a one-cardiac-cycle echo, approximately 0.1 seconds per frame.

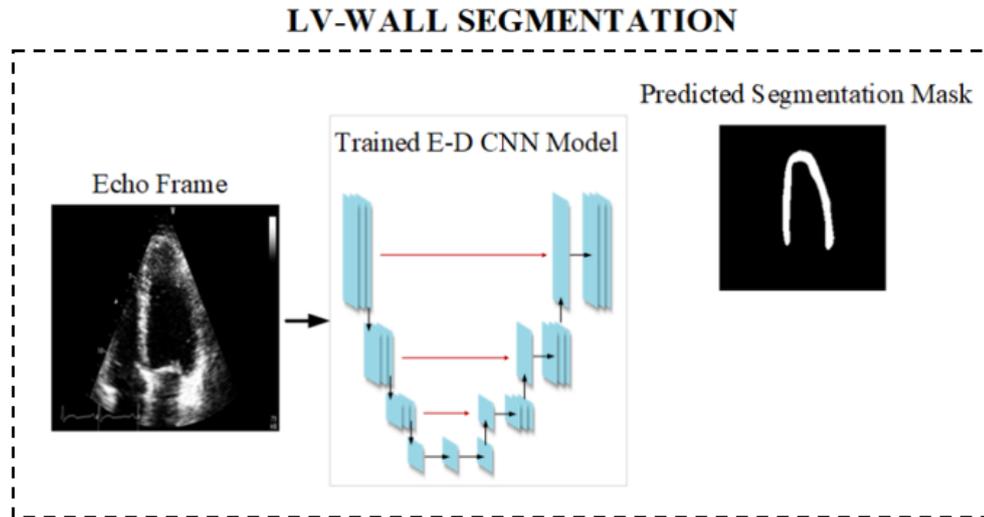


Figure 2.3 U-Net for LV wall segmentation. [32]

2.3 AI in myocardial infarction classification

Not many studies have been conducted regarding the classification of MI from ultrasound images, due to the lack of publicly available datasets. Released in 2020, HMC-QU is the first public echocardiographic dataset for MI detection. Two different approaches have been tested by the creators of this dataset. The first applies a threshold based method in which a fixed value defines the separation between the two possible classification categories (MI and non-MI) [23]. The article first obtains the LV segmentation through Active Polynomials, end diastole and end systole frames are defined and left ventricle ejection fraction (LVEF) is estimated. If LVEF is between 15% and 55%, the maximum displacement of each segment of the myocardium is calculated using the segmentation obtained. The manual threshold is applied to this value: a maximum displacement less than 19% classifies a segment as infarcted. An overview of the classification method after obtaining the LV wall segmentation is presented in **Figure 2.4**. This approach obtains the following results: 87.94% accuracy, 92.86% sensitivity and 87.64% precision.

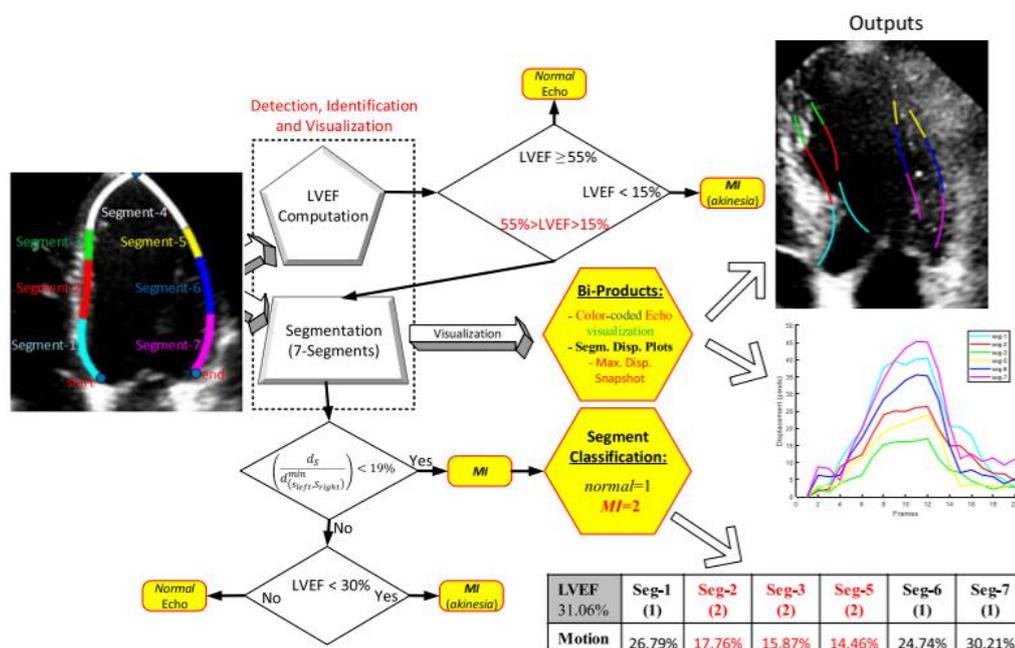


Figure 2.4 Manual threshold method for MI detection. [23]

One year later, their second article was released and for this occasion a machine learning approach was tested. The scheme of this new approach is illustrated in **Figure 2.4**. Four different supervised learning techniques were evaluated: Discriminant Analysis (LDA), Decision Tree (DT), Random Forest (RF), and Support Vector Machine (SVM). From the obtained LV segmentation, three different features were extracted for each segment: two motion features corresponding to maximum displacement of the endocardial boundary point and maximum displacement of each segment's center of mass point, and one area feature corresponding to the minimum area intersection of segments. These features were used to train each of the four algorithms. In the end, RF and SVM obtained the highest values for the evaluation parameters: RF obtained a specificity value of 71.81% and precision of 85.99%, SVM a sensitivity of 85.97% and an accuracy of 80.24%. Inference time for all methods was very low, 0.2ms for SVM and 7ms for RF, for the classification of one echo. It is worth mentioning that the previous feature engineering stage is executed in around 391ms.

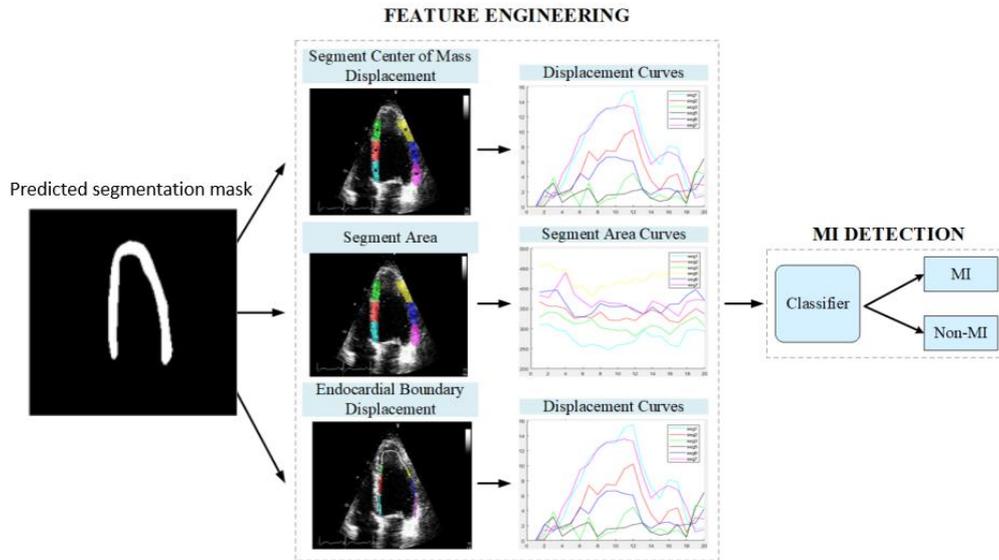


Figure 2.5 Machine learning approach for MI detection by a conventional classifier. [32]

The first approach tested by the authors of these articles obtained high sensitivity even though it used the simplest classification strategy. However, the threshold used is not a clinically defined value but was defined specifically for the HMC-QU dataset, making it valid for those images in particular. In reality, different studies have found that the extent of RWMA is not fixed among MI patients, even patients with similar medical history present significantly different wall motion abnormalities which can be due to differences in metabolic characteristics and ischemic tolerance of the involved myocardium [47]. Additionally, in a single patient these motion abnormalities evolve throughout the whole progression of MI [48]. Such characteristics of RWMA make binary threshold a technique that is inevitably constrained to certain conditions of infarction. Thus, it makes sense that in their successive article the authors would move towards a machine learning approach with more features. Nonetheless, the features selected once again focus mainly on displacement and do not correspond to real life clinical parameters used in echocardiography evaluation. What is more, the sensitivity value they reached with ML was lower than with the threshold method. To obtain better performance and accurate clinical diagnosis more parameters need to be included for classification, particularly clinical parameters which the cardiologist can understand. LVEF is currently the most used parameter in

echocardiography but, once more, a threshold method would not work with this parameter as it is not constant among patients. In fact, many studies have found that around one-third of MI patients present preserved LVEF [49] [50]. GLS and wall thickness are additional parameters that take into account physiological and morphological characteristics and could offer more information for ischemia identification [51] . Again, these variables cannot generate a classification using a threshold method since no consensus on cut off value exists [51] [52]. Due to the number of parameters that are affected with the presence of ischemia and the heterogeneity of parameter combinations that can be presented, a ML approach is the most appropriate path for classification.

After conducting the literature review of the state of the art methods for MI detection, it is clear that machine learning techniques are the most used and most adapted for this thesis specific task. In particular, for segmentation of the left ventricle deep learning and for classification of MI more classical techniques with manual feature extraction.

CHAPTER 3

3 Materials and Methods

The following chapter delves into the materials and methods used for the realization of the algorithm proposed in this thesis project. First will be introduced the datasets used for training and testing of the algorithm. Next is presented the complete architecture of the proposed solution which is composed of four main stages:

1. Pre-processing
2. Segmentation
3. Post-processing
4. Classification

Each stage is explained separately with a thorough description of the applied techniques.

3.1 Datasets

Two different publicly available datasets were used for the realization of this thesis project: CAMUS and HMC-QU. Both were obtained from acquisitions of apical-four-chamber (A4C) and apical-two-chamber (A2C) views in 2D echocardiography.

3.1.1 CAMUS

The Cardiac Acquisitions for Multi-structure Ultrasound Segmentation (CAMUS) dataset contains A4C and A2C echocardiography clinical exams from 500 patients, acquired at the University Hospital of St Etienne (France) [44]. The acquisitions were optimized for left ventricle ejection fraction (LVEF) measurements. For each patient the echo frames corresponding to end systole and end diastole are given (two images for each type of view) alongside

the following information: sex, age, image quality, LV volume at end systole and end diastole and LVEF. However, at the time of realization of this thesis project, of the 500 patients only 450 include the manual annotations for the left ventricle endocardium, the myocardium and the left atrium (LA), based on the analysis of one clinical expert. These three classes are differentiated in the manual annotations by pixel intensity (range 0-255), where pixels corresponding to the LV chamber take value 169, pixels corresponding to the left atrium take value 172 and those corresponding to the LV wall a value of 170 or 171. Though manual threshold, all pixels were set to 0 except for those corresponding to the area of interest (LV wall) which were set to 1. It is important to note that for many of the patients, parts of the wall were not visible in the images. Such patients were not included in this thesis project, therefore of a total of 450 patients containing manual annotations, only 180 were used. Additionally, only the A4C view was used.

Of those 180 patients, 44 have a LVEF lower than 45%, thus being considered at pathological risk. Regarding image quality, of the 180 patients 63 rank “good”, 80 rank “medium” and 37 rank “poor”. **Figure 3.1** and **Table 3.1** present an example patient from the CAMUS dataset, both ES and ED images with the corresponding annotations along with clinical information are shown.

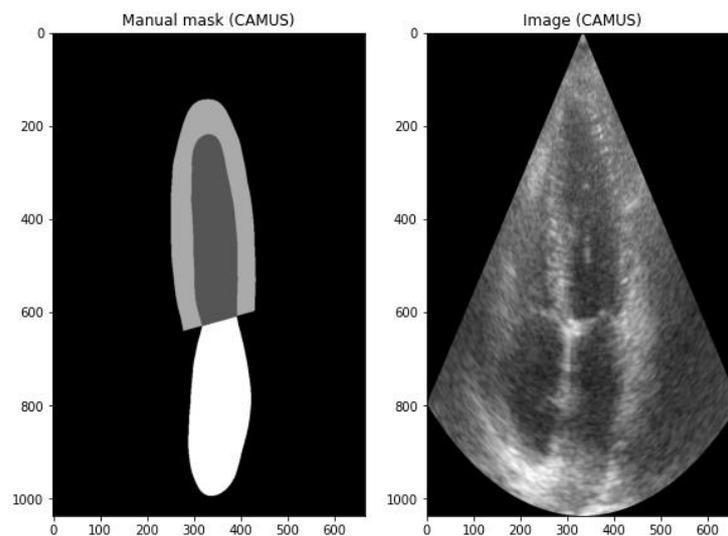


Figure 3.1 Example echocardiography image and segmentation mask for a CAMUS patient (Patient 442).

Table 3.1 Clinical information available for each CAMUS patient

patient	ED Frame	ES Frame	Sex	Age	Image Quality	LVedv	LVesv	LVEF
Patient 442	1	20	F	79	Medium	114.3	42.6	62.70

3.1.2 HMC-QU

Built in collaboration between Hamad Medical Corporation Heart Hospital (HMC) in Doha, Qatar & Qatar University (QU) & Tampere University, the HMC-QU dataset includes a collection of A4C and A2C view 2D echocardiography recordings obtained during the years 2018 and 2019 [9]. The spatial resolution of the echocardiography recordings varies from 422×636 to 768×1024 , and the temporal resolution is 25 fps. Regarding the A4C view, the dataset consists of 162 recordings belonging to 93 MI patients (all first-time and acute MI) and 69 non-MI subjects. The MI term indicates any sign of RWMA, whereas subjects without RWMA are labeled as non-MI in the dataset. These ground-truth labels are provided for each of the six myocardial segments (segment 1, 2, 3, 5, 6 and 7) which can be identified in **Figure 3.2**.

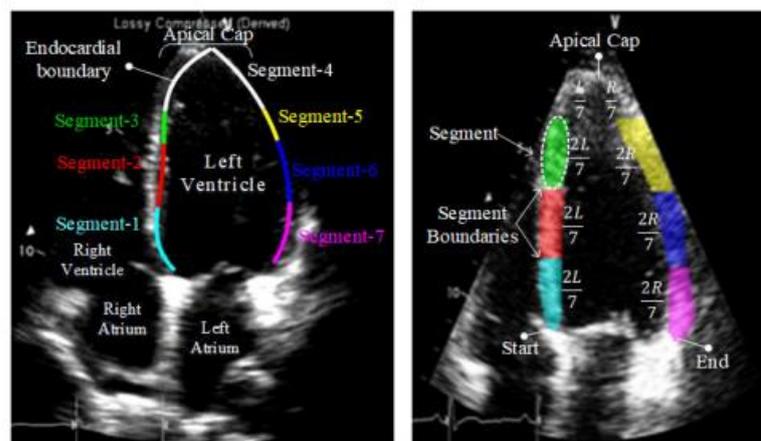


Figure 3.2 Endocardial boundary segment division and division ratios. [32]

As with the CAMUS dataset, some recordings included frames in which a portion of the wall of the LV disappears from the area covered by the ultrasound probe. These recordings were not included in this thesis. Therefore, of the 162 available echo videos, only 148 were used for the step of MI classification. These 148 videos correspond to 89 MI and 59 non-MI patients for which the ground-truth labels of each segment are divided as can be seen in **Table 3.2**. It is important to note that there is a clear imbalance among the number of MI segments, which introduces greater difficulty in the process of classification.

Table 3.2 HMC-QU patients with corresponding MI and Non-MI classification for each LV wall segment

LV wall segment	# MI patients	# Non-MI patients
Segment 1	29	119
Segment 2	54	94
Segment 3	71	77
Segment 5	53	95
Segment 6	30	118
Segment 7	18	130
TOTAL	255	633

Furthermore, the dataset includes a subset of 109 A4C view 2D echocardiography recordings with a corresponding ground-truth segmentation mask for the whole LV wall for each frame in one cardiac cycle. Of the 109 recordings, 72 correspond to MI patients and 37 to non-MI. The ground-truth LV segmentation masks for all the frames in the 109 videos were obtained through a pseudo labeling technique using only a few masks actually segmented by cardiologists. These masks were used to train an E-D CNN which was then used to segment the frames of the other echoes that had no ground-

truth masks. The correct masks generated by the network were added to the previous training set and this new enriched set was used to train the next E-D CNN. This process was repeated iteratively until all frames had a corresponding LV segmentation mask visually acceptable to the cardiologists. However, through manual inspection of the ground-truth masks included in this dataset for the use of this thesis, it was noted that some presented defects not attributable to the LV wall and, once again, some segmentations stretched outside the area covered by the ultrasound probe. Such masks and corresponding A4C frames were removed and not used for this thesis project. In the end, of the 109 recordings (corresponding to 2349 frames) only 95 were used (2034 frames) for the step of LV segmentation. **Figure 3.3** presents an example frame extracted from the echocardiography recording of a patient from the HMC-QU dataset with the corresponding annotated mask.

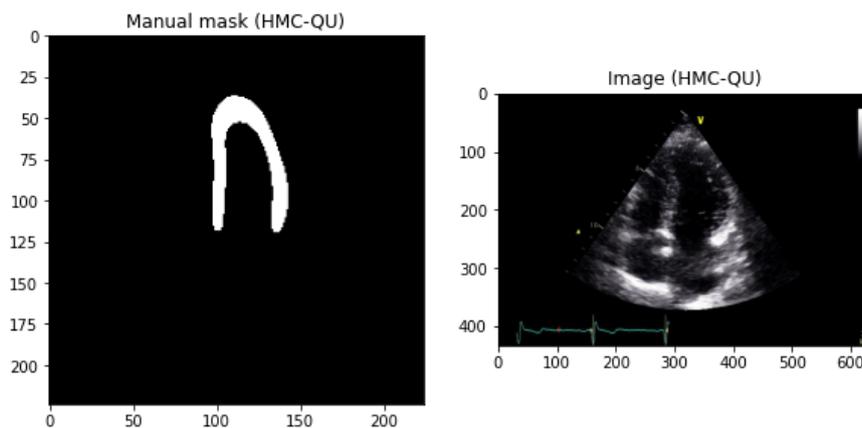


Figure 3.3 Example echocardiography image and segmentation mask for a HMC-QU patient (Patient 51).

3.1.3 Data division

To summarize, from the available datasets the following data was used for the two main steps of this thesis project:

Segmentation of the LV: 180 CAMUS patients which corresponds to 362 echo frames and 95 HMC-QU patients which correspond to 2034 frames. In total 2396 images from 275 patients.

Classification of MI: 148 HMC-QU patients with labels for each of the 6 myocardial segments, accounting to 888 segment labels in total.

In both cases, machine learning models were trained. In order to obtain valid results, the dataset needs to be further divided in order to use part of the samples and labels to train the model and another part to test and evaluate results with unseen data [53]. For the deep learning model used in segmentation, the dataset is divided in three subsets: training, validation and testing. 70% of the data is used for generating the model (training set), 10% is used to evaluate the model's skill while tuning hyper-parameters (validation set) and the remaining 20% is used to obtain an unbiased estimate of the final performance of the completed model (**Table 3.3**). For the machine learning model used for MI classification, 80% of patients were used for training and 20% for testing, which results in the following division (**Table 3.4** and **Table 3.5**).

Table 3.3 Data division for deep learning segmentation model

	Total (frames)	Training	Validation	Test
CAMUS	362	253	37	72
HMC-QU	2034	1424	203	407
Total	2396	1677	240	479

Table 3.4 Data division (patients) for machine learning classification model

	Total	Training	Test
MI	89	71	18
Non-MI	59	47	12
Total	148	118	30

Table 3.5 Data division (segments) for machine learning classification model

	Total	Training	Test
MI	255	204	51
Non-MI	633	506	127
Total	888	710	178

3.2 Architecture of solution

The following diagram (Errore. L'origine riferimento non è stata trovata.) displays the four main stages of this thesis project as well as the techniques that compose each. The algorithm was constructed using Python version 3.9, an open source, object-oriented programming language, which facilitates coding through the use of pre-existing libraries and functions.

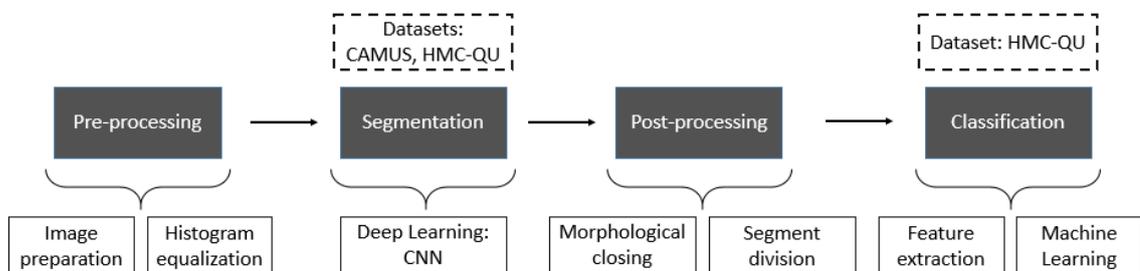


Figure 3.4 Pipeline for proposed solution

The algorithm foresees an initial preprocessing step where the echocardiography video is split into frames and resized and LV myocardial wall is enhanced through histogram equalization. The images alongside the manual segmentation masks constitute the dataset used as input for training and testing of the neural network. Post processing of the masks obtained as output from the network allows the correction of small defects that may appear from the automatic segmentation. Finally, a procedure of feature extraction is performed on the segmentation masks and the obtained features are used to train different machine learning methods. This final step of classification uses only the HMC-QU dataset.

3.3 Preprocessing

Before implementing AI techniques, digital imaging processing makes up the initial stage of this project's algorithm. This, in order to facilitate model training and increase accuracy. The first step regards image extraction and formatting which are necessary to have a homogeneous dataset, especially since two different datasets are implemented. All images used in this project correspond to A4C 2D echoes, however, HMC-QU dataset provides the echo videos while CAMUS directly provides two images corresponding to end systole and end diastole frames. Therefore, frames are initially extracted from the HMC-QU videos and those images are converted to grayscale to match the format of the CAMUS images. Additionally, CAMUS images are already cropped to fit only the actual ultrasound image so the same process is done with HMC-QU. All images are finally resized to the same dimension (512x512). Regarding CAMUS segmentation masks and as explained previously, a threshold is applied to leave only the segmentation corresponding to the LV wall.

Afterward, more advanced preprocessing techniques are used for enhancing image comprehension. This is necessary due to the complexity of the heart and limitations of ultrasound imaging which make LV segmentation a challenging image segmentation task. Achieved accuracy is usually limited by low signal to noise ratio, varying speckle noise and low resolution, all inherent

characteristics of ultrasound imaging [54]. In addition, low contrast between the myocardium and the trabecula and papillary muscles present in the blood pool, shadows cast by dense structures such as muscle and ribs, among other anatomical aspects, further difficult the distinction of the myocardium in the echocardiography image [24].

To address these issues, different filtering techniques were studied and tested, including averaging filters, edge detection filters and speckle noise reduction filters. Although visually these techniques seemed to facilitate myocardial detection, after being tested in the segmentation portion of the project, they proved to reduce model accuracy. Clustering techniques were also evaluated such as k-means and mixture model clustering, but these cannot be generalized to every single echocardiography image and thus required to treat separately both datasets. In the end, the chosen technique which proved to increase segmentation robustness and could be applied to both datasets unitedly, was the histogram equalization method, specifically Contrast limited Adaptive histogram equalization.

Equalization is an image enhancement method that uses the image's histogram for contrast adjustment. The histogram of a digital image is a distribution of its discrete intensity levels, associating each level to the total number of pixels in the image with that specific intensity value [55]. Through equalization, all pixel's intensities in an image are transformed in order to produce a histogram with a uniform distribution. An example image with its corresponding histogram before and after equalization can be seen in **Figure 3.5**. Adaptive histogram equalization (AHE) differs from ordinary equalization in that multiple histograms are computed, each corresponding to a different region of the image [56]. This allows enhancement of local regions that are darker or lighter than most of the image. However, to prevent over amplifying noise, Contrast limited AHE limits the contrast amplification [57].

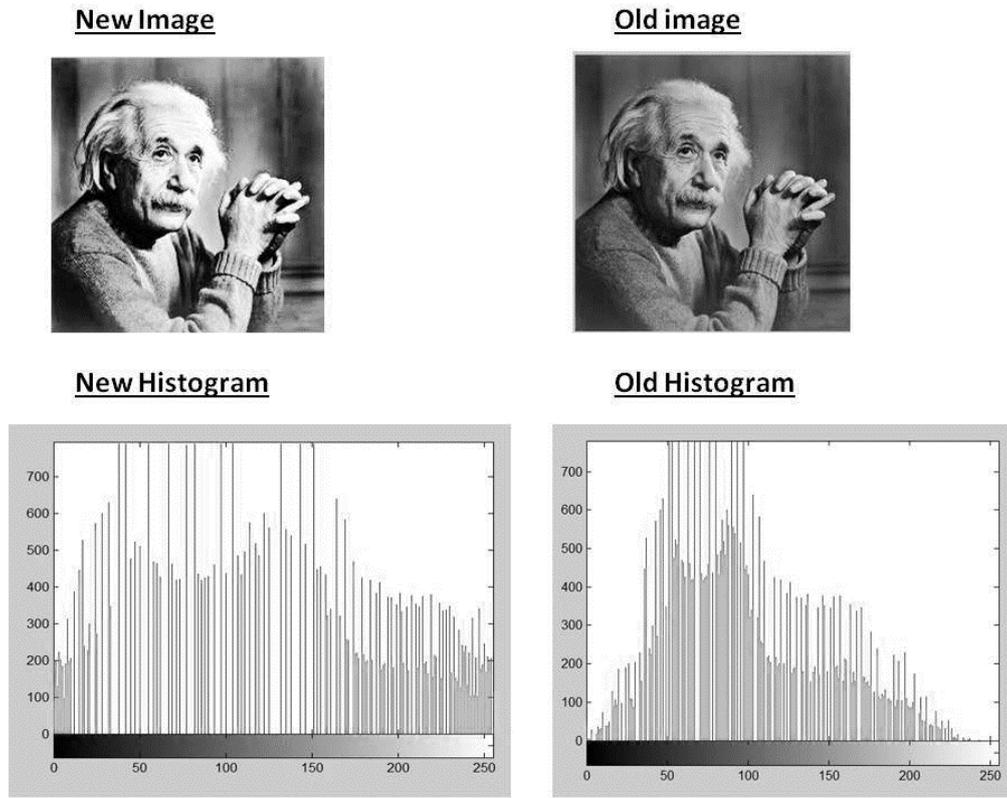


Figure 3.5 Histogram comparison before (old) and after (new) image equalization. [58]

3.4 Segmentation

After preprocessing follows the stage of LV wall segmentation. Based on results found with the current state of the art, it was decided that the best approach would be deep learning, taking into consideration data availability and the requirement for short inference time. In particular, a fully convolutional neural network was implemented with a U-Net architecture. Following, greater detail will be given regarding this type of network structure and how it works to automatically generate the segmentation.

3.4.1 Convolutional Neural Networks

As previously introduced, deep learning (DL) is a machine learning technique based on artificial neural networks. Deep learning has gained a lot of attention

because it overcomes the problem of manual feature selection, necessary in other ML methods, since it automatically extracts the significant features from raw input data for the specific task at hand [59]. These features are obtained through nonlinear processing occurring at the hidden layers of the network which transform the data into different levels of abstraction. Different deep learning models are available such as deep belief networks, recurrent neural networks and convolution neural networks (CNN). The latter is the most common type of deep neural network for image analysis [54].

CNNs are composed of three types of layers: convolution layer with activation function, pooling layer, and a final fully connected layer (**Figure 3.6**) [59]. The first two layers serve the purpose of feature extraction and the last layer of classification. The degree of image abstraction is related to the number of convolution layers, earlier layers focus on simpler features such as colors and edges, while progressive layers are able to recognize more complex elements or shapes until the intended object is identified.

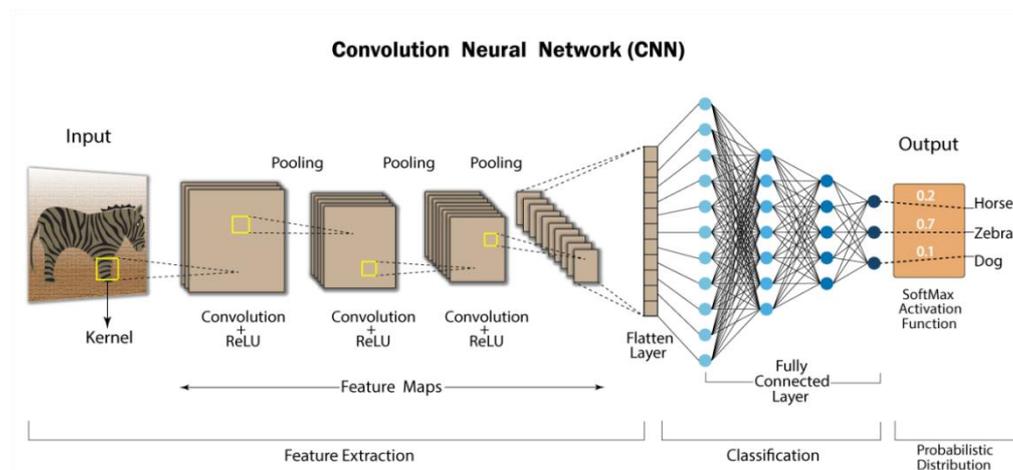


Figure 3.6 General architecture of a Convolutional Neural Network. [60]

Convolution:

In the convolution process, a filter (kernel) working as a feature detector moves across the whole image checking if the defined pattern is present. The kernel is a 2D array of learnable weights, typically with a 3x3 dimension. It is applied

to a section of the image by calculating the dot product between the filter and the image's pixel it covers. The obtained scalar value is fed into an output array to which more values are added as the kernel sweeps across the whole image shifting by a fixed stride [61]. Mathematically, the convolution operation between the image pixel $I(x, y)$ and the filter $F(x, y)$ of $K \times K$ dimension, is defined as follows (3.1).

$$H(x, y) = I(x, y) * F(x, y) = \sum_{i=0}^{K-1} \sum_{j=0}^{K-1} I(i, j)F(x - i, y - i) \quad (3.1)$$

After searching the whole image, the final output matrix is known as a feature map, activation map or heat-map for the searched pattern. Its dimension will be reduced compared to the original image. Reduction of resolution as more convolutional layers are applied throughout the CNN, favors the extraction of more informative features. Output dimensionality reduction after each convolutional layer will be the result of [62]:

- **Number of kernels used:** the number of filters used corresponds to the numbers of researched patterns;
- **Stride:** parameter that defines the number of pixels that the filter moves after each convolution operation;
- **Zero-padding:** technique that allows to preserve the original input size by adding pixels of value 0 at the image border.

The output feature map tensor dimension will be $O \times O \times M$, with M corresponding to the number of filters applied. By noting I the length of the input volume size, F the length of the filter, P the amount of zero padding and S the stride, then the output size O of the feature map along that dimension is given by (3.2):

$$O = \frac{I - F + P_{start} + P_{end}}{S} + 1 \quad (3.2)$$

Afterwards, a Rectified Linear Unit (ReLU) transformation is applied to the resulting feature maps which replaces any negative values with value 0, introducing nonlinearity to the model and speeding up training [63]. An example of the complete convolution process is depicted in **Figure 3.7**.

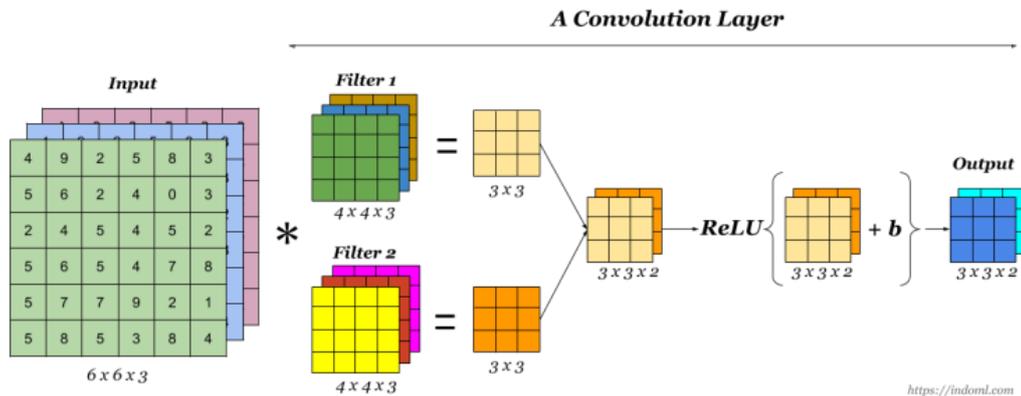


Figure 3.7 Convolution layer in a CNN. [64]

Pooling:

After each convolution layer a pooling operation takes place in order to reduce signal dimensionality without losing robustness of the feature map with respect to translations and deformations [46]. This process consists in taking a small block from the activation map and producing a single output from it. Pooling can be done in different ways such as averaging the values in the selected block or taking only the maximum value; both approaches are demonstrated in **Figure 3.8**. CNNs usually apply max-pooling with a down sampling by a factor of 2 [54].

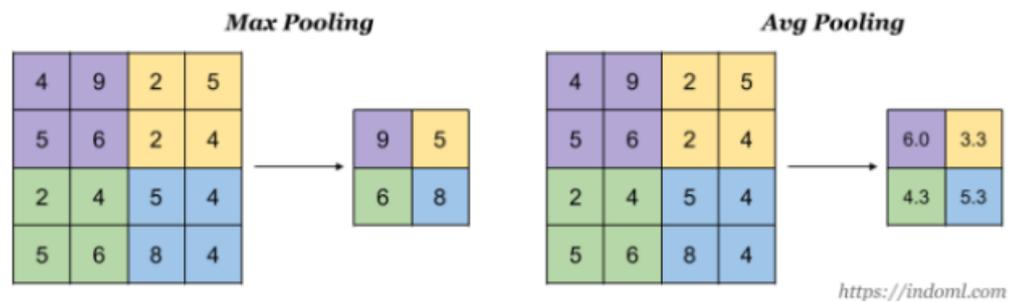


Figure 3.8 Pooling operations. [64]

Fully connected:

If present, the fully connected layer is found after the last pooling layer and is used to obtain the final predicted value. It operates on a single vector input thus, it is required for the output feature maps from the last pooling layer to be flattened before the fully connected layer. In a fully connected layer, each neuron in output connects directly to all nodes in the previous layer and a ReLU activation function is used, with the exception of the last layer. The last layer usually employs a softmax activation function to produce a probability from 0 to 1 of the input belonging to the corresponding target class [61]. The output of the network is a fix-sized vector where each element is a probabilistic score of membership to each category [54]. Since the fully connected layer works with a fixed input vector dimension, the input image for the network needs to have a certain dimension to guarantee that after the last pooling layer the matrix will be of the required size.

3.4.1.1 Fully Convolutional Neural Network

In visual tasks, as is biomedical image segmentation, the designed network should give as output not a single class label for the whole image but rather a single class label for each pixel. In this way, localization is included as output since the network provides a mask of the original image with each pixel corresponding to a single class [65]. To obtain this, Long et al. [66] took a normal CNN and transformed the final fully connected layers to convolutions. Specifically, a 1x1 convolution layer is added which uses a 1x1xD kernel where

D corresponds to the depth of the input feature maps tensor. In this way, activation maps resulting from the feature extraction process are condensed together to generate an activation map representative of the whole process, as can be seen in **Figure 3.9**. This allows to maintain information of location and solves the constraint of CNNs of having a fixed input image dimension since the $1 \times 1 \times D$ kernel can work with any size input feature map.

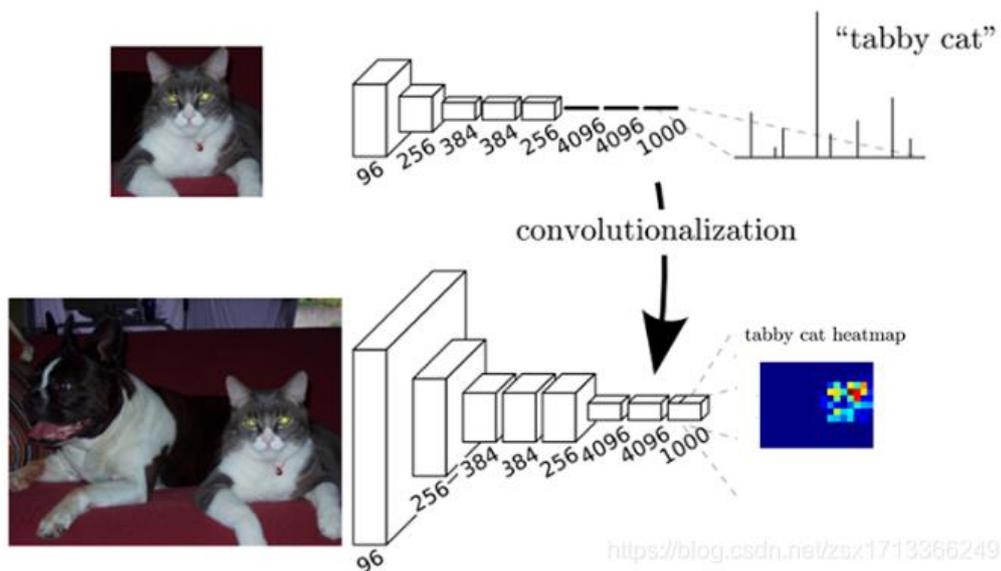


Figure 3.9 Transformation of a CNN into a FCNN with resulting heatmap. [66]

The heatmap obtained as output however, does not retain the original image dimension and lacks detail due to condensation of information throughout the network. In order to solve this issue, the final segmentation mask should not be generated immediately after the feature extraction process (also known as down-sampling or encoding) but instead an up-sampling process should occur first. Up-sampling layers implement the inverse operations of pooling and convolution. Unpooling maps the content of the input tensor to a larger one generating new activation maps which present more distributed activation. In order to compact these dispersed activations, deconvolution (or backwards convolution) layers are introduced. In these deconvolutions, activations specific for the class of interest are amplified while signals attributed to noise are suppressed. Through the use of both unpooling and deconvolution layers, the network is able to scale the produced feature maps back to the original

image dimension before applying the 1x1 convolution. In the end, a network with this configuration is capable of generating a more accurate segmentation map.

However, since the upsampling enlarges information coming from a small spatial dimension, the finer details of the image are inevitably lost and the output mask is still a coarse representation of the desired segmentation. Long et al. [66] overcomes this by proposing the addition of skip connections. Skip connections combine deep, coarse, semantic information with shallow, fine, appearance information so the model can make local predictions that respect global structure. **Figure 3.10** compares the final segmentations obtained when no skip-connections are used, when combining information from one previous layer and when using two previous layers. Improving over this final design, in 2015 Ronneberger et al. [65] proposed the U-Net architecture, a FCNN whose main feature is the usage of the same number of convolutional layers in upsampling and downsampling and the usage of a larger number of feature channels. Additionally, skip connections are present between each upsampling layer and the symmetrical downsampling layer. By increasing the number of layers and connections, the U-Net is able to integrate more context information and obtain greater segmentation accuracy.

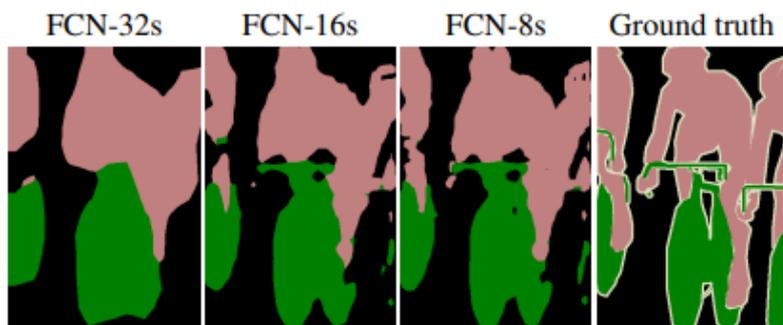


Figure 3.10 Comparison between ground truth and output segmentation of a net with no skip connections (FCN-32s), skip connection between final layer and a previous pooling layer (FCN-16s) and use of two skip connections (FCN-8s). [66]

3.4.1.2 U-Net

The FCNN designed by Ronneberger et al. [65] consists of a contracting path (encoder) and an expansive path (decoder) which are more or less symmetrical producing the u-shaped architecture that inspired this network's name (**Figure 3.11**).

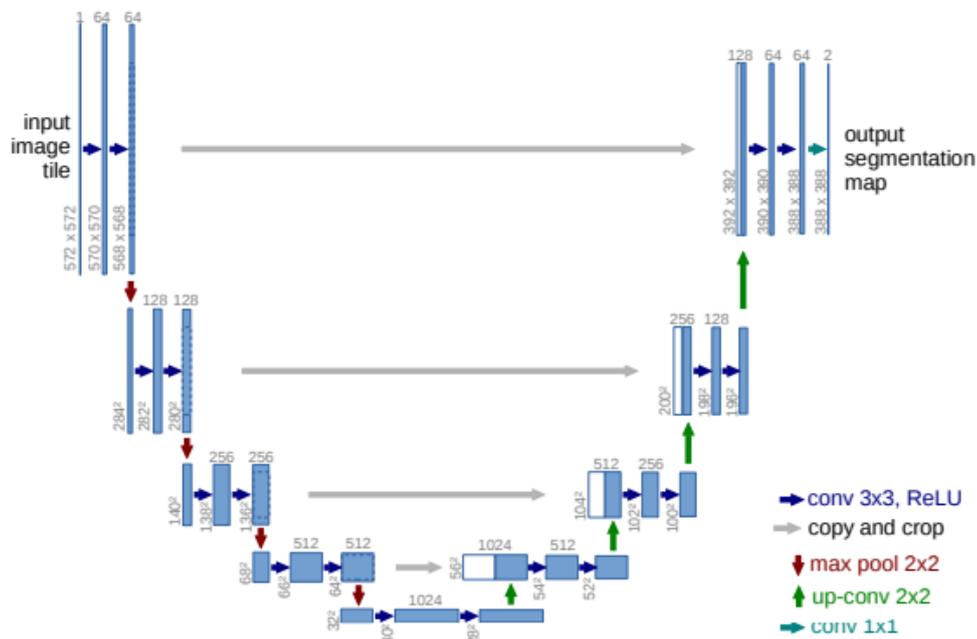


Figure 3.11 U-Net architecture. [65]

The contracting path enables feature extraction and follows the typical architecture of a CNN which consists of successive application of convolutions, activation functions and pooling. In this case specifically, two 3x3 convolutions (unpadded), ReLu activation function and a 2x2 max-pooling operation (stride 2) are used. After each downsampling, the number of feature maps is doubled while the map size halves, resulting in a reduction of spatial resolution. The expanding path increases resolution and enables precise localization through successive steps of: up-convolution (2x2 kernel), which double the dimension and halves the number of channels of the feature map tensor; the concatenation with the corresponding feature map from the contracting path (skip connection); and 2 convolutions (3x3 kernel) each

followed by a ReLu. For the skip connection, the feature map from the encoding portion needs to be cropped to the dimension of the corresponding map in the decoding portion due to the loss of border pixels after every convolution step. Finally, the network ends with a 1x1 convolution used to map the 64-channel feature map tensor obtained to the desired number of classes. In total the network has 23 convolutional layers. It is important to note that the improved performance obtained with the U-Net in comparison to other FCNN is the result of having each step of the expanding path concatenate activation maps from the previous lower layer and corresponding encoder layer. This allows each decoding step to obtain semantic information from each encoding step and thus avoid the loss of fine details which are especially important in medical imaging. For the present thesis, the U-Net network was applied in Python using the TensorFlow software library for machine learning and artificial intelligence.

3.5 Post-processing

The final step present in the process of LV wall segmentation is the post-processing of the mask obtained as output from the U-Net. In this project, this implicates both a step for improving the segmentation accuracy as well as a step for preparation of the mask for the succeeding feature extraction phase.

3.5.1 Morphological operations

If the network has a high performance, only some masks will present minimal imperfections that can easily be corrected with simple post-processing morphological operations which process images based on shapes. Such operations consist of the application of a structuring element (kernel), of desired shape and dimension, to an input mask to produce a same sized output mask where the value of each pixel is based on a comparison between the corresponding pixel in the input mask and its neighbors [67]. The most basic morphological operations are erosion and dilation. The first shrinks the object of interest by removing pixels from its boundary, and the second expands the

object of interest by adding pixels to its boundary. Specifically, when used with binary images, each operation obeys the following rules:

Erosion:

The value of the output pixel is 0 if any of the neighboring pixels (defined by the kernel) have a value 0. This eliminates background isolated pixels and thin lines and makes the object of interest smaller and thinner.

Dilation:

The value of the output pixel is 1 if any of the neighboring pixels (defined by the kernel) have a value 1. This fills in holes present inside the object of interest and makes it appear larger and thicker.

Usually, for image processing both operations are used in combination. Depending on the order in which they are employed, morphological opening or closing is obtained.

Opening:

Consists of the application of erosion followed by dilation, with the same kernel for both operations. Opening allows to eliminate pixels attributed to noise while preserving the size of the larger objects of interest in the mask.

Closing:

Contrary to opening, closing consists of the application of dilation followed by erosion, again with the same kernel for both operations. This is useful for filling small holes in the mask while preserving the size of larger holes and objects of interest.

Examples of the results obtained after applying each one of the four previous operations can be observed in **Figure 3.12**. For this thesis project, the morphological operation of closing was chosen as a post-processing strategy since some of the segmentations generated by the network presented small holes inside the wall of the LV.

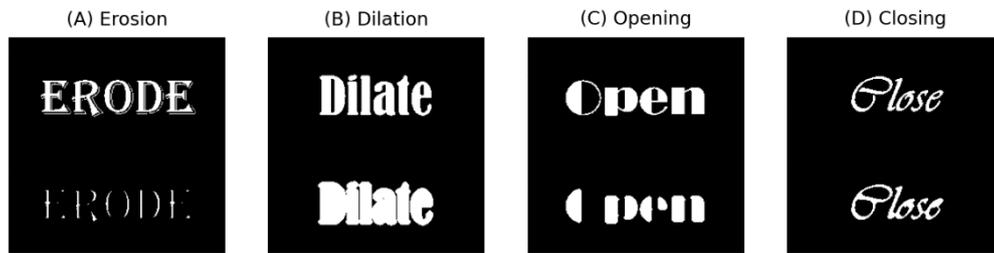


Figure 3.12 Overview of morphological transformations. A) Erosion B) Dilation C) Opening D) Closing. The original image is shown at the top, while the processed part is at the bottom in each case. [68]

3.5.2 Segment division

Since global and segmental features are extracted for classification, it is first necessary to divide the mask after post-processing in its representative segments. Given that the mask corresponds to the whole LV wall which includes epicardium and endocardium, it is first reduced to obtain only the middle lining which corresponds to the myocardium. This is obtained through a process of erosion which results in the morphological thinning of the object until it becomes a 1 pixel wide representation of the original [69]. Specifically, the skeletonization algorithm proposed by Lee *et al.* [70] was used and obtained from the implementation by Scikit-image image processing python library. The algorithm iteratively sweeps over the image evaluating the 3x3x3 (or 3x3 in 2D images) neighborhood of each pixel examined and removing pixels at each iteration until the image stops changing. Pixels chosen for removal are first checked in order to preserve connectivity of the image.

After obtaining the skeleton of the LV wall segmentation, it is divided into segments. Different standardized models for the segmentation of the LV exist but the American Heart Association recommends the use of the 17-segment model [18] which is shown in **Figure 3.13**. This model results in the division of the LV in the A4C view into 7 segments. In this specific view, only 6 segments exhibit a uniform motion, the top most segment corresponding to the apical cap does not present inward motion [71]. For this reason, the dataset HMC-QU offers the classification labels only of the 6 moving segments,

therefore, in this thesis project the segmentation follows the 17-segment model but the apical cap is not included in succeeding calculations.

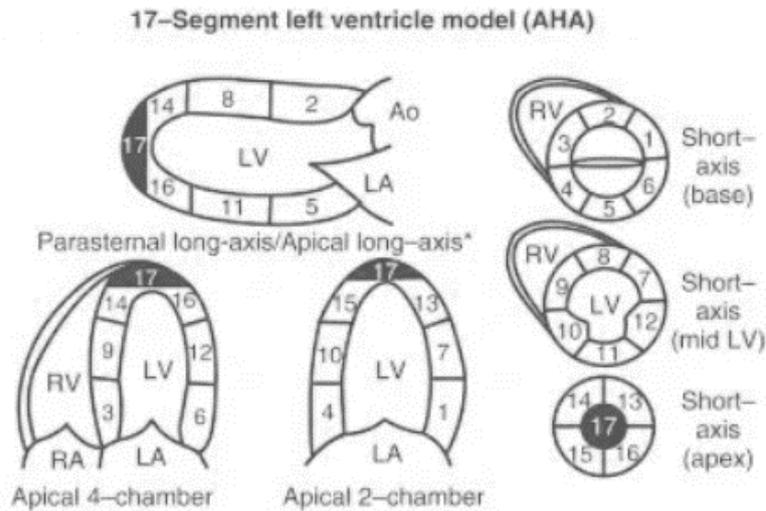


Figure 3.13 17-segment model of the left ventricle. [71]

Following the guidelines offered by the article presenting the HMC-QU dataset [23], the LV segmentation is divided in its left and right portion using the apical cap as mid-point. Afterwards, each of the 6 segments of interest is assigned $2/7$ of the length of its corresponding portion while the apical cap takes the remaining $1/7$ at each side (as is shown in **Figure 3.2**). This results in a left portion composed by segments 3,9,14 and half of segment 17 and right portion composed by segments 16,12,6 and half of segment 17 (from the 17-segment model in **Figure 3.13**). From now on, these segments will be referred to as segments 1,2,3,4,5,6,7 respectively, where segment 4 corresponds to the apical cap.

3.6 Classification

After obtaining the myocardial segment division from the final LV wall mask, this project proceeds to its final step, MI classification. As stated earlier, a machine learning approach was selected and 3 different methods were tested. Before explaining each method in greater detail, the process of feature

extraction and its specific application to this project will be described, given that this step produces the inputs used for training and testing of the ML models.

3.6.1 Feature extraction

For training and testing of the ML models, global and regional features were extracted from the skeleton and segment division obtained from the myocardial wall tracing. This thesis project focuses on the use of features that correspond to real life clinical parameters currently used in echocardiography evaluation in order to supply to the cardiologist values that he/she knows how to interpret and to train the classifiers on physiological information. With this in mind, the chosen features were LVEF, GLS and for each segment, LS, segment displacement and segment wall thickness. It is important to note that all features were obtained from the end systolic and end diastolic frames, defined respectively as the frames with minimum and maximum LV cavity area. Finally, all features were normalized with respect to end diastolic measurements in order to be able to make a direct comparison between different patients and eliminate the effects of variability between echo exams in regard to dimension. Next will be a description of each feature, its relevance and method of calculation.

3.6.1.1 Left ventricle ejection fraction

Ejection fraction is the most used clinical parameter in 2D echocardiography for assessing LV function [72]. LVEF corresponds to a volumetric measure calculated as $LVEF = [(EDV-ESV)/EDV] \times 100$. In order to approximate this value from 2D videos, which only offer information regarding area, different methods exist. The American Society of Echocardiography recommends the use of the biplane method of disks (Biplane Simpson's method) (**Figure 3.14**). This technique consists of obtaining the volume of the LV cavity as the sum of the volume of three sections of equal height: a cylinder (shape assumed for the base of the heart to the mitral valve), a truncated cone (from the mitral valve to

the level of the papillary muscles) and a full cone reaching until de apical cap [28]. For calculating these volumes, the Simpson's method requires tracing the endocardial border in end-systole and end-diastole of both the A4C and A2C views [29].

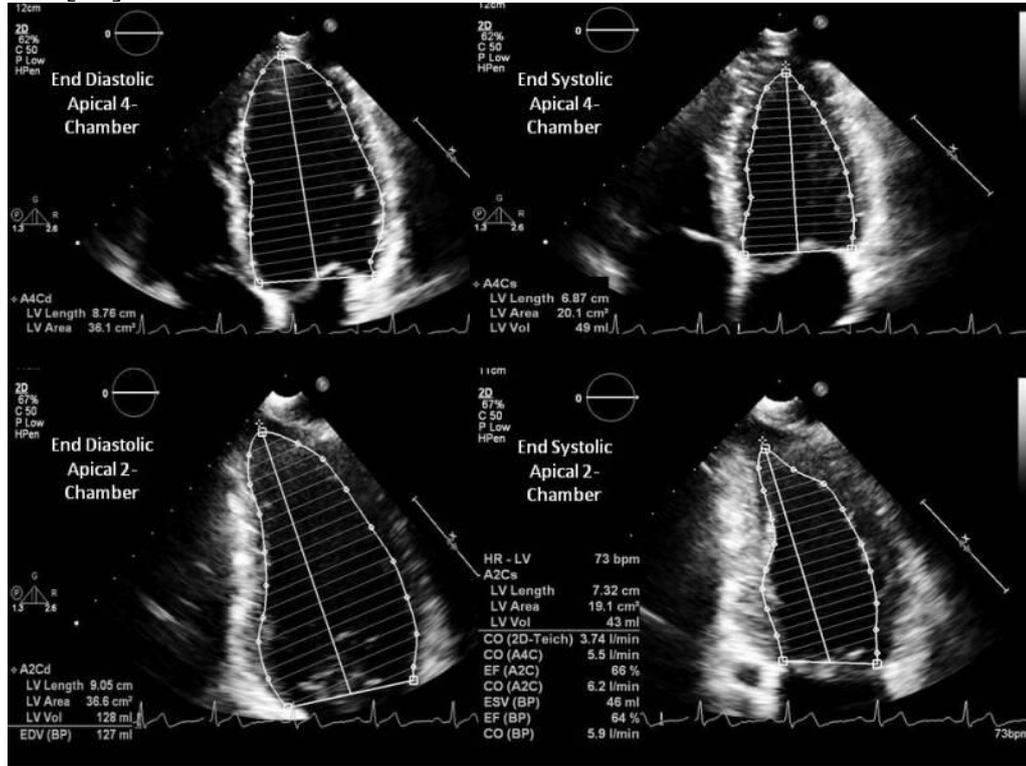


Figure 3.14 Biplane Simpson Method using the end diastolic and end systolic apical 4- and 2- chamber views for estimation of LV volume and calculation of the ejection fraction. [28]

Since only A4C echoes are used in this thesis project, the approximation of LVEF requires the use of other methods. In agreement with Kiranyaz *et al.* [23], this thesis also approximates the volumetric fraction with the area fraction as presented in (3.3).

$$LVEF \approx \frac{A_{ED} - A_{ES}}{A_{ED}} = 1 - \frac{A_{ES}}{A_{ED}} \quad (3.3)$$

In this thesis, in addition, other geometric considerations are also tested by using different exponential factors added to the area approximation equation. Specifically, an exponential factor of 2 was tested which assumes an ellipse shape approximation of the LV and, since this usually leads to an overestimation, an exponential factor of 3/2 was also tested. Exploiting the availability of real LVEF values for the CAMUS patients, the method that offered the smallest error in respect to the expected ejection fraction was chosen for the extraction of this feature for the HMC-QU patients (which do not include real EF values).

$$DISP = \frac{Avg(dispatch)}{Min(dist)}$$

3.6.1.2 Longitudinal strain

In addition to LVEF, global longitudinal strain is another commonly used value for the evaluation of the general systolic function. GLS measures the fractional length change between systole and diastole calculated as (3.4):

$$GLS = \frac{L_{ED} - L_{ES}}{L_{ED}} = 1 - \frac{L_{ES}}{L_{ED}} \quad (3.4)$$

This provides added information to EF by measuring myocardial contractility. In fact strain parameters have been shown to be more sensitive than LVEF for the assessment of early myocardial dysfunction and have been demonstrated to detect abnormal contraction patterns in the setting of apparently normal LVEF [73] [23]. What is more, this feature can be extrapolated for the study of regional function by evaluating the strain of each individual segment. In the case of this parameter, no approximation needs to be made as it can be directly calculated from the 2D echocardiography myocardial trace.

3.6.1.3 Segment displacement

Since RWMA are directly linked to the presence of MI due to the process of scarring and replacement of myocardium with fibrotic tissue, a displacement feature is included to measure wall motion. A normalized maximum displacement value is calculated for each of the 6 segments of interest in the same way as it is calculated by Kiranyaz *et al.* [23], this in order to facilitate posterior model and results comparison. The computation of the maximum displacement of each segment proceeds as follows:

1. 5 points equally spaced out are selected for each segment for end-systole and end-diastole frames.
2. The displacement of each point between the two frames is calculated using the L2 norm.
3. The maximum of the 5 measured displacements for each segment is selected, which corresponds to that segment's maximum displacement during the whole cardiac cycle.
4. Maximum displacement is normalized using the minimum distance between opposing segments (distance in end-systole frame) which is calculated using the L1 norm.

For maximum displacement calculation, the L2 norm is used since it assesses motion in both x and y direction in the way in which human perception naturally evaluates distance. On the other hand, the minimum interval between opposing segments uses the L1 norm in order to fit the ratios of segments in [0, 1] more precisely. Both measurements are shown in **Figure 3.15**.

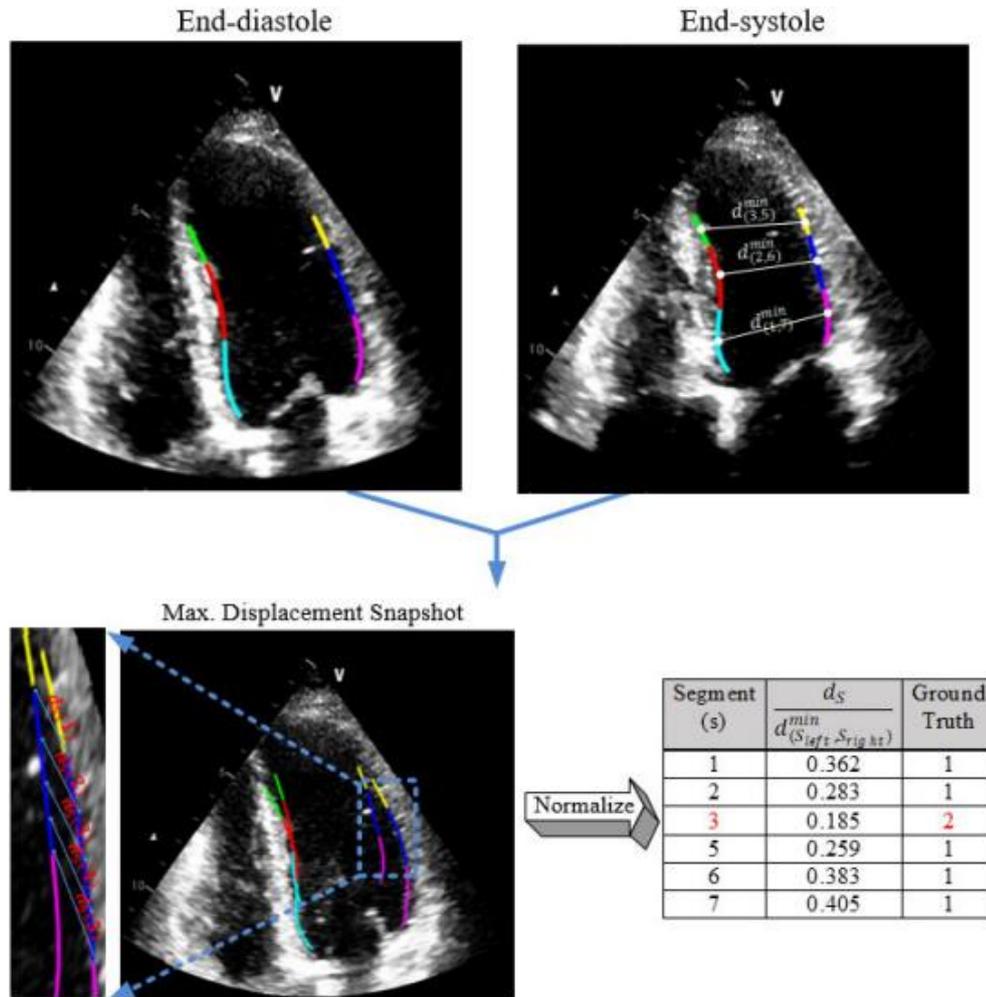


Figure 3.15 Computation of the normalized maximum displacement of the 6 segments of the A4C view echo with the ground-truth labels (normal = 1, infarcted = 2). [23]

3.6.1.4 Wall thickness

Myocardial wall thinning is found in MI patients as a result of damage and loss of tissue due to the presence of ischemia. For this thesis project, wall thickness is evaluated for each of the 6 segments of interest and is calculated as the fractional thickness difference between end-systolic and end-diastolic frames. Thickness at each frame for each segment is obtained as the average wall thickness of 3 equally spaced out points in each segment. For this feature, the

3 points are taken from the skeleton with the divided segments but the value is obtained using the complete mask which contains the whole LV wall.

Although wall thinning can signify the presence of infarction, this can also be the result of normal anatomical and congenital LV wall thinning or other non-coronary heart diseases [74]. The feature here measure does not correspond to an absolute distance measure but rather a percentage change between systole and diastole. This measure of dynamic change in wall thickness has been found to average between 70% and 106% in healthy patients where the wall thickens during systole in comparison to diastole. In most cases of CAD patients, the percentage is reduced or even no change in wall thickness is observed. However, for some patients expansion is obtained during diastole and as a result a negative percentage change is obtained. What is more, the behavior is not constant among the different infarcted segments meaning that the percentage change can vary significantly from one to the other [52].

3.6.2 Machine learning models

With the previously extracted features, three different machine learning algorithms were tested for the classification of MI. The different classification methods have been chosen among those most commonly used in literature: Support Vector Machine (SVM), K-Nearest Neighbors (KNN) and Random Forest (RF). Each of the 3 algorithms will be described in more detail in the next paragraphs. All were implemented in python through the use of the Scikit-learn machine learning library.

For each of these algorithms, three different models were generated. The first model consisted of the horizontal concatenation of features for each patient. In this way the model is trained to use all features from all segments as input and offer as output the general classification of the patient. This results in the patient classified as presenting MI or not, without specifying which segments are involved. The second model uses the vertical concatenation of features, meaning that the model is trained to classify each segment individually. In this way, for a single patient six classifications are obtained, one for each segment,

if at least one segment is classified as presenting MI then the patient will be classified with MI. The third model also uses individual segments as input but in this case a different model is trained for each segment, meaning that in total 6 models are generated, each one specific for one segment. Again, if at least one segment is classified as MI, the patient will be classified as MI.

3.6.2.1 K-Nearest Neighbors

K-Nearest Neighbors (KNN) is one of the simplest supervised ML algorithms [75]. It generates classifications based on proximity between data points. Distance between the sample being tested and each sample of the training data is calculated, commonly with the Euclidean distance between feature vectors. The classification of this new point will be defined as the most repeated class among its k nearest neighbors, with k being a positive integer (**Figure 3.16**)

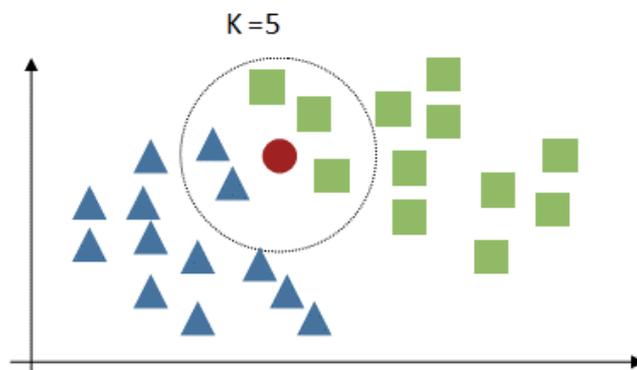


Figure 3.16 KNN. Selected neighbors to a new point when $k = 5$. [76]

The choice of this parameter will depend on the data used for training, a larger value will help reduce the effect of noise but the distinction between classes will be harder. If $k=1$ then the sample will be assigned the class of the single nearest neighbor which increases overfitting. In general, k is chosen to be small and, in the case of binary classification, an odd number to avoid tied votes between the selected nearest neighbors. The main drawback of this method is that it is highly affected by imbalanced data since it leads to predominance of classes with more objects. Additionally, since K-NN is based on distance

calculation, features used need to be normalized to prevent the scale of different physical quantities to impact on the classification of the tested samples.

3.6.2.2 Support Vector Machine

The support vector machine (SVM) algorithm aims to efficiently identify a function that allows the separation of the data in different classes [77] [78]. This function, or kernel, will represent the mathematical modification of the data into a higher dimension feature space where a Support Vector Classifier, or hyperplane, can obtain an accurate separation of the data. This allows for learning of non-linear models. The points closest to the hyperplane are called the *support vector points* and the distance of the vectors from the hyperplane are called the *margins* (**Figure 3.17**). SVM seeks to obtain the maximum number of correctly classified samples but at the same time increase the margin as much as possible. This can result in some misclassification but eventually allows the creation of a more generalized model. To balance these two requirements, SVM defines a penalty parameter for misclassifications. This parameter, typically called C , if chosen of a large value will highly penalize misclassification and thus result in a tight margin. On the contrary, if C is small, misclassification is allowed and a larger margin is obtained.

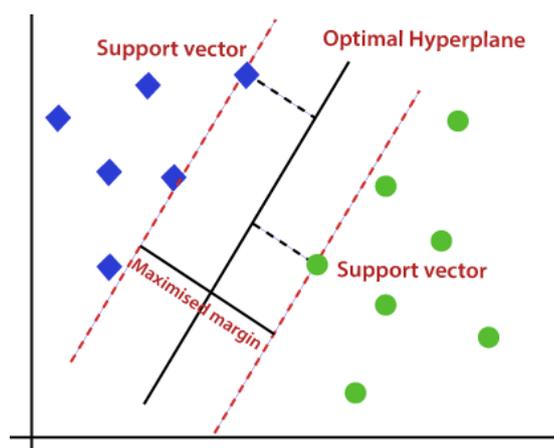


Figure 3.17 Division between classes using SVM. [79]

Most commonly, polynomial or radial basis function (RBF) kernels are used. The polynomial kernel has the form $(a * b + r)^d$ where a and b are two different observations in the data set, r determines the coefficient of the polynomial and d sets the degree of the polynomial. The RBF is expressed as $e^{-\gamma(a-b)^2}$ and behaves like a Weighted nearest neighbor model, meaning that the closest observations have a large influence on how new observations are classified while further away observations have little influence on the classification. Once again, a and b refer to two samples and the level of influence among them is calculated using the square distance which is scaled by the factor *gamma* (γ). With the use of the exponential function, points that are relatively far away from each other will result in a value very close to zero signifying a weak high dimensional relationship. Although SVM can handle high-dimensional data and obtain high classification accuracy, it does not scale well to large numbers of training samples or large numbers of features.

3.6.2.3 Random Forest

The last method tested for classification in this thesis project is Random Forest (RF) [80]. This approach is a generalization of Decision Trees (DT), where the output of multiple trees is combined to obtain a single final output (**Figure 3.18**). Decision trees are highly sensitive to training data, which results in high variance. Instead, RF while still maintaining the simplicity of DT, grants flexibility when classifying new samples which increases robustness and model accuracy.

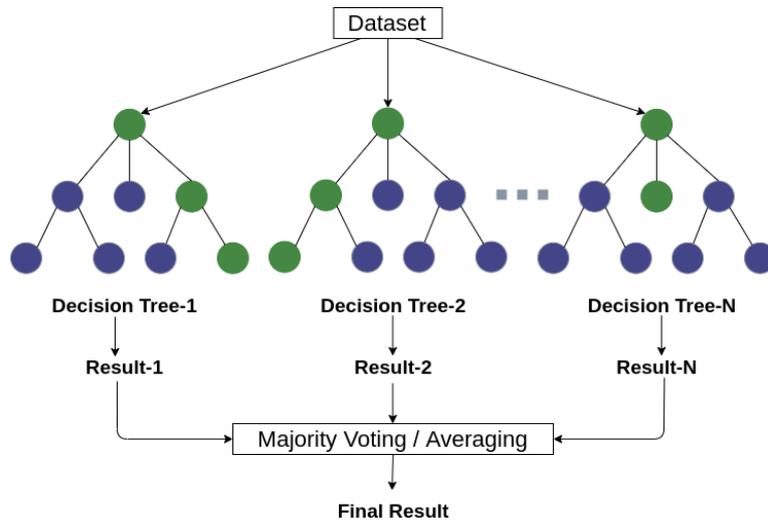


Figure 3.18 General architecture of a Random Forest model. [81]

The construction of the Random Forest model consists in generating multiple non-correlated Decision Trees by training each tree with a different subset of the original data. These “new” datasets (known as bootstrap datasets) contain the same number of samples as the original datasets and are built by randomly selecting samples from it (allowing repetition). Additionally, each tree is built using a random subset of features at each step (feature bagging), usually equal to the square root of the total number of features. In the end, bootstrapping and feature bagging result in a wide variety of trees. For new samples, classification is obtained by running the data down all of the trees in the random forest and choosing the label with more votes.

3.6.3 Data augmentation

Data augmentation is a technique used to artificially increase the training set in order to prevent overfitting, improve model accuracy, tackle the problem of data availability and reduce costs of obtaining, labeling and cleaning raw data. Data can easily be augmented by making minor changes to the available dataset in order to increase its size and diversity. Empirically, this can result in substantial increase in a model’s performance by, for instance, increasing the

relevance of certain informative features that are hard to learn and thus making the model more likely to capture them in the learning process [82].

In this thesis project, specifically, data augmentation is applied for the final classification process. It proves to be beneficial in this step since only the HMC-QU dataset contains labels for infarct detection, meaning that less data is available for classification in comparison to the step of segmentation. Since classification in this thesis uses as input the features described in the previous sections, data augmentation consists in generating new artificial features and their corresponding labels. This process is only applied to the tested algorithm (KNN, SVM or RF) that obtains the best results in order to evaluate the effect of data augmentation on model performance.

Given that the main goal is correctly identifying MI, the added data corresponded to the MI class. In total 15 patients (accounting to 90 segments) were designed. For this, all real 100% MI patients (patients with all 6 segments labeled as MI) were evaluated and the average, minimum, maximum and standard deviation values for each feature were calculated. Based on these values and values found in literature for abnormal ranges of each feature, different combinations were made to generate the artificial patients.

3.6.4 Validation Metrics

Validation metrics are used to quantitatively evaluate the performance of any ML model (deep learning models included). Among the many existing metrics some of the most common are the confusion matrix, F1 score and Area under the ROC curve (AUC).

The confusion matrix is an $N \times N$ table where N corresponds to the number of classes being predicted. For the case of 2 classes, the matrix follows the design shown in **Figure 3.19** which also presents the different parameters that can be calculated from it [83]:

Confusion Matrix		Target			
		Positive	Negative		
Model	Positive	a	b	<i>Positive Predictive Value</i>	$a/(a+b)$
	Negative	c	d	<i>Negative Predictive Value</i>	$d/(c+d)$
		<i>Sensitivity</i>	<i>Specificity</i>	Accuracy = $(a+d)/(a+b+c+d)$	
		$a/(a+c)$	$d/(b+d)$		

Figure 3.19 Confusion matrix with corresponding metrics.

- **Accuracy** : the fraction of the total number of correct predictions
- **Positive Predictive Value or Precision**: the fraction of positive cases that were correctly classified.
- **Negative Predictive Value**: the fraction of negative cases that were correctly classified.
- **Sensitivity or Recall**: the fraction of actual positive cases that were correctly classified.
- **Specificity**: the fraction of actual negative cases that were correctly classified.

The F1 score, also known as Dice metric, is calculated as the Harmonic Mean between precision and recall. It offers information on the precision and robustness of the model by combining the measure of the ability to classify positive instances as well as not miss a significant number of instances. These previous methods were used for the validation of the deep learning model used for LV segmentation.

Finally, the Area under the Receiver Operating Characteristic (ROC) curve indicates how well the model is able to identify the different classes and corresponds to the probability that the model ranks a random positive instance more highly than a random negative instance. The ROC curve (**Figure 3.20**) plots sensitivity (also known as true positive rate) and (1- specificity), which is also known as false positive rate [84].

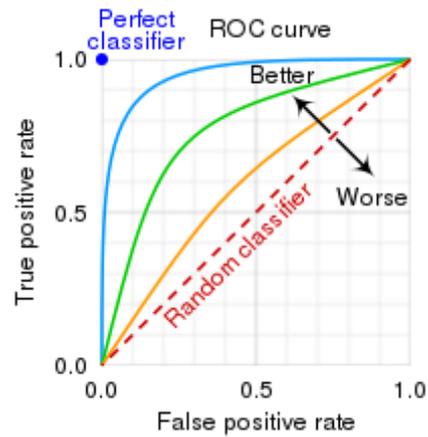


Figure 3.20 ROC curve with model comparison.

All the previous parameters have a range of $[0, 1]$ where 1 indicates an ideal performance of the model. The parameters chosen for evaluating a certain algorithm and the importance given to each depends on the specific case at hand. In the medical context, usually greater importance is given to the identification of positive cases (for example signifying the presence of a disease) for which precision, sensitivity, dice and AUC become more significant.

CHAPTER 4

4 Results

The present chapter illustrates the results obtained for each of the four main phases of this thesis's architecture after applying the methods discussed in Chapter 3. For the pre and post processing steps the analysis will focus on qualitative visual results, while for the segmentation and classification steps, the specific parameters chosen during training of the models will be detailed and performance of the algorithms will be quantified by calculating the evaluation metrics presented in **paragraph 3.6.4**. Regarding the classification of MI, the results are used to compare the three different models tested in order to select the one with the highest performance for the following step of data augmentation. Furthermore, the use of data augmentation is evaluated in order to assess its effect on model performance and a cascade approach is proposed in order to classify first the patient as a whole and, if MI is detected, then proceed to the identification of the specific segments involved. Finally, inference time for the whole architecture is presented both with and without the use of a GPU.

4.1 Pre-processing

As indicated before, HMC-QU images are initially cropped and converted to gray scale to imitate more closely the images of the other dataset. On the other hand, the masks from CAMUS are modified to present only the segmentation corresponding to the LV wall. The images obtained afterwards, alongside their manual segmentations, are shown in **Figure 4.1**.

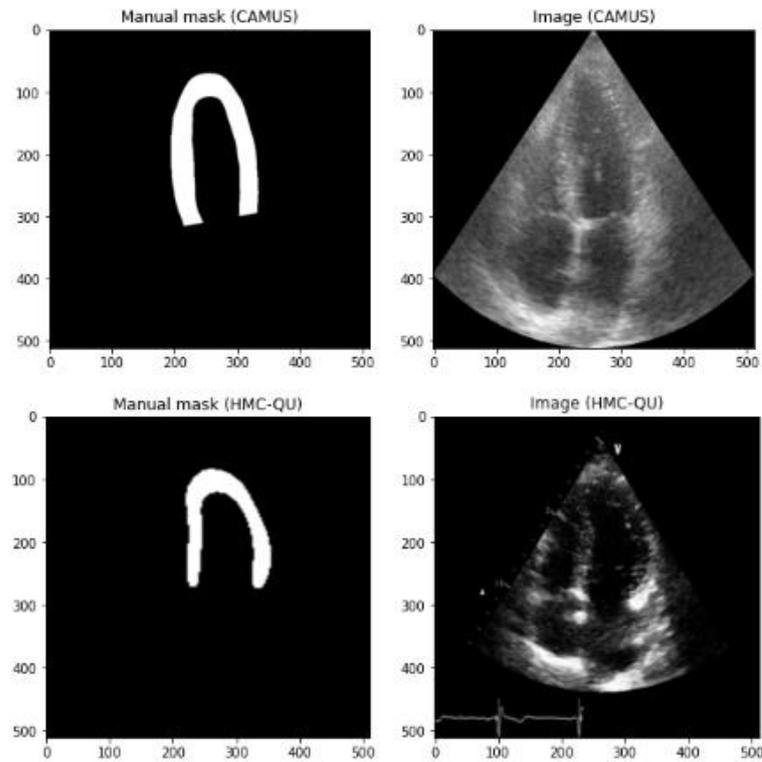


Figure 4.1 Sample CAMUS and HMC-QU images and corresponding masks after initial formatting.

From these images the histograms were extracted and, as can be seen in the example images in **Figure 4.2** and **Figure 4.3**, the datasets present very different pixel intensity distributions. In order to homogenize the distributions and highlight the area of interest, equalization was applied to all images which results in the modifications also shown in **Figure 4.2** and **Figure 4.3**. The equalized images and their masks, all resized to the dimension 512x512, were used as input for the deep learning segmentation model.

Results

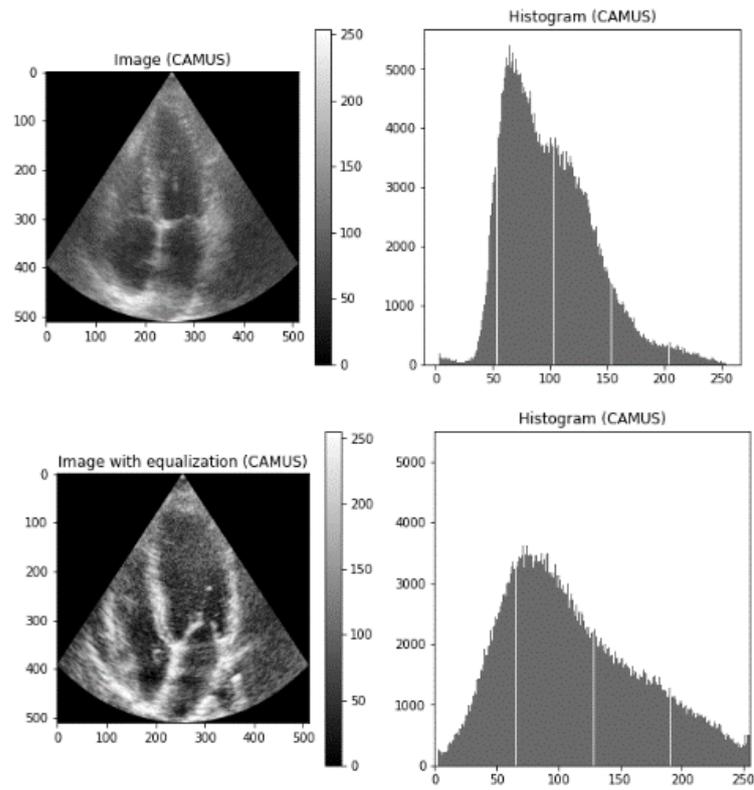


Figure 4.2 CAMUS sample image and histogram before and after equalization.

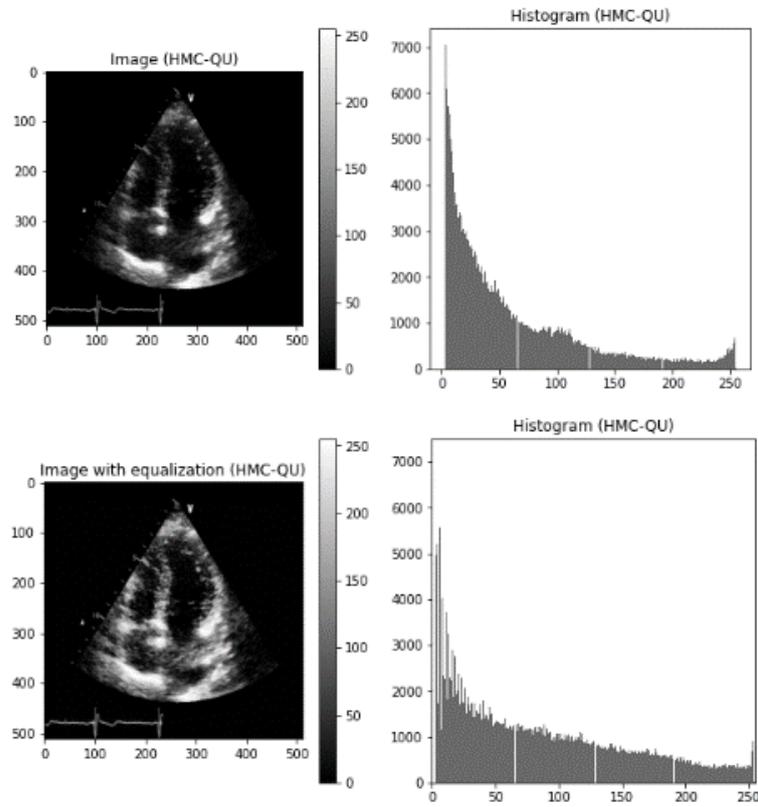


Figure 4.3 HMC-QU sample image and histogram before and after equalization.

4.2 Segmentation

In a U-Net architecture, the feature detection and decodification depends on the kernel weights, upsampling and downsampling blocks, concatenations and skip connections and how these layers are arranged. The backbone is the architectural element that defines this arrangement. Different backbones are available for the implementation of deep learning in Python, in this project the EfficientNet-b4 backbone [85] was chosen after testing different architectures since it obtained the best performance. The whole network (both encoding and decoding paths) was trained from scratch so no transfer learning was required. Additionally, data augmentation offered by the Tensorflow library was used in which input images were augmented by applying rotation in a range from 0-10° and zoom in a range from 0-10%. Training parameters are summarized in

Table 4.1. With a Quadro RTX 6000 GPU, training of this model takes close to 4 minutes and reaches early stopping after completing 6 epochs.

Table 4.1 U-Net model training hyper-parameters

Training parameter	Value
Loss function	Jaccard Loss
Optimizer and learning Rate	Adamax [10^{-4}]
Evaluation metric and min. delta	IoU score [10^{-4}]
Batch size	5
Max. number of epochs	20
Shuffling	Yes

Performance over the complete training, validation and test sets are presented in **Table 4.2**. The following 2 tables present the results of this same network when generating automatic masks for the HMC-QU frames (**Table 4.3**) and for the CAMUS (**Table 4.4**) images separately.

Table 4.2 Network performance over complete dataset.

(all data)	Dice	Sensitivity	Specificity
Training	0.9628	0.9685	0.9977
Validation	0.9435	0.9502	0.9964
Testing	0.9598	0.9645	0.9976

Table 4.3 Network performance over images belonging to the HMC-QU dataset.

(HMC-QU)	Dice	Sensitivity	Specificity
Training	0.9741	0.9827	0.9984
Validation	0.9553	0.9680	0.9970
Testing	0.9747	0.9839	0.9986

Table 4.4 Network performance over images belonging to the CAMUS dataset.

CAMUS	Dice	Sensitivity	Specificity
Training	0.9027	0.8898	0.9941
Validation	0.8807	0.8552	0.9931
Testing	0.9029	0.8935	0.9938

In all cases, parameter values are balanced between the three subsets, which indicates that overfitting is not present. In regards to evaluation using only CAMUS images, it is expected that lower values be obtained since this dataset represents only 15% of the total images used for segmentation.

4.3 Post-processing

4.3.1 Morphological closing

After obtaining the automatic segmentation mask as output of the FCNN, the morphological closing operation is applied in order to correct small defects

found in some of the masks (see **Figure 4.4** for an example). In particular, a 10x10 rectangular kernel was used and closing was done using the OpenCV python library. By recalculating evaluation parameters after closing, only minimal changes occur (**Table 4.5**) which indicates that not many masks presented defects.

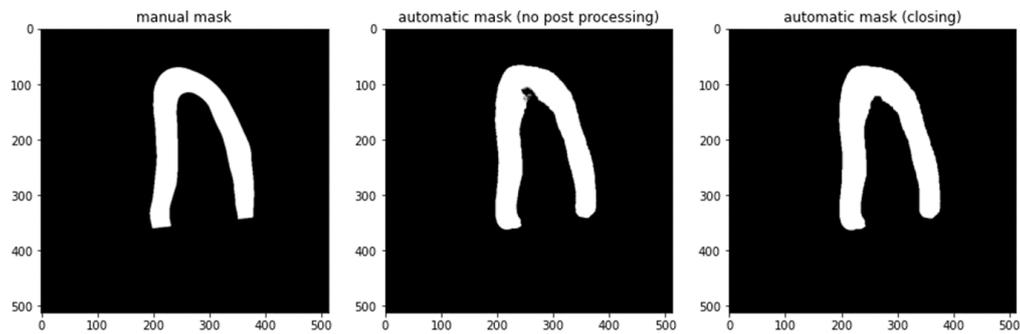


Figure 4.4 Comparison of manual annotation and automatic mask before and after post-processing for an example patient.

Table 4.5 Network performance over complete dataset after post-processing.

(all data)	Dice	Sensitivity	Specificity
Training	0.9628	0.9666	0.9978
Validation	0.9431	0.9457	0.9968
Testing	0.9429	0.9436	0.9968

4.3.2 Segment division

The process of segment division occurs in four main steps as seen in **Figure 4.5**. First, the trace of the myocardial layer is obtained from the LV wall segmentation through the skeletonization algorithm. The extreme points are

obtained for the skeleton: lowest right and left points to mark the start and end points of the myocardial tracing and the topmost point which marks the apex. These points are then used to divide the tracing in its left and right portions which are then partitioned into the individual segments as was described in **paragraph 3.5.2**.

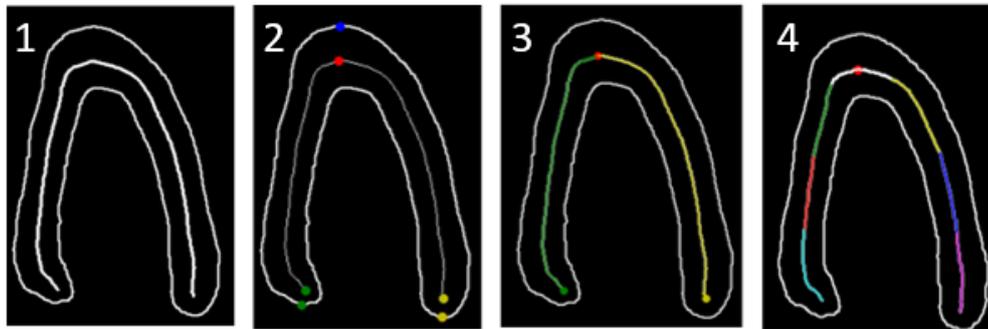


Figure 4.5 Segment division. 1) Skeleton of segmentation 2) Location of extreme points 3) Division of left and right portions 4) Division of each portion

4.4 Classification

4.4.1 Feature extraction

In order to identify end systolic and end diastolic frames from which the features are extracted, the area of the LV cavity for all frames was first calculated as the number of pixels contained inside the LV wall segmentation. End-systole and end-diastole are then chosen as the frames with the smallest and largest number of pixels, respectively.

4.4.1.1 LVEF

Three different methods were used for evaluating LVEF, all three volumetric approximations using the fractional area change between systole and diastole but with varying exponents applied to the fraction. Since the CAMUS dataset includes the actual LVEF of each patient, these values were used to compare the errors obtained by the three approximations. Such errors are presented in

Table 4.6. At first only method 1 and 2 were tested, since the first resulted in an underestimation and the second in an overestimation of the real values, method 3 was added as a midpoint. In fact, method 3 obtained the lowest relative and absolute errors and was thus selected as the approximation to be used for the LVEF feature. With this approximation, the following box plot distributions were obtained for the CAMUS data (**Figure 4.6 and Table 4.7**) and the HMC-QU data (**Figure 4.7 and Table 4.8**).

Table 4.6 LVEF approximation methods with corresponding relative and absolute errors.

Method	Mean relative error	Mean absolute error
1. $LVEF = 1 - (A^{ES} / A^{ED})$	32.78%	16.72%
2. $LVEF = 1 - (A^{ES} / A^{ED})^2$	27.42%	14.60%
3. $LVEF = 1 - (A^{ES} / A^{ED})^{3/2}$	17.30%	8.54%

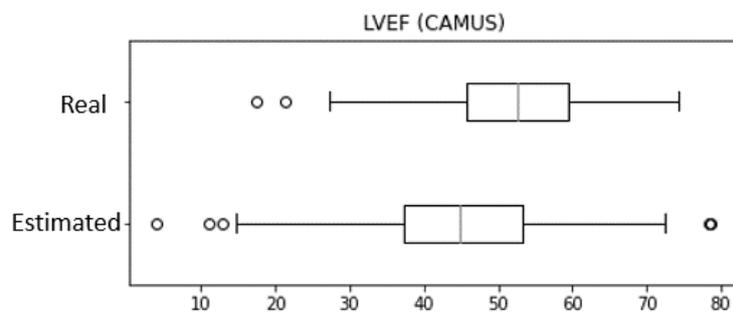


Figure 4.6 Box plots of real and estimated LVEF for CAMUS images.

Table 4.7 Statistical values for real and estimated LVEF for CAMUS images.

CAMUS	Average	Maximum	Minimum
Real LVEF	51.89%	74.4%	17.4%
Estimated LVEF	44.95%	78.75%	4.03%

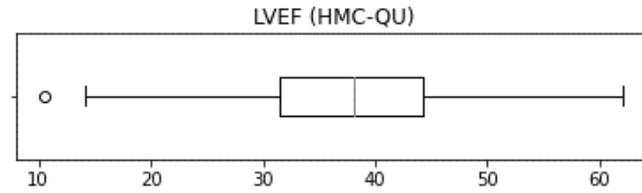


Figure 4.7 Box plot of estimated LVEF for HMC-QU images.

Table 4.8 Statistical values for estimated LVEF for HMC-QU images.

HMC-QU	Average	Maximum	Minimum
Estimated LVEF	37.89%	62.13%	10.54%

4.4.1.2 Longitudinal strain

Longitudinal strain is calculated globally and regionally as the fractional length change between end-systole and end-diastole. GLS is calculated using the LV axis length measured as the distance (L2 norm) from the top apical point (topmost point of the LV wall mask) to the middle bottom point (midpoint between lowest left and right points of the LV wall mask) (**Figure 4.8**). Results obtained for GLS are presented in **Figure 4.9** and **Table 4.9**; **Figure 4.10** and **Table 4.10**. LS for each individual segment uses as length measurements the number of pixels that compose each segment. Since the number of pixels are equally distributed between segments (each corresponds to $2/7$ of the length of the right or left portion of the LV wall), LS is almost identical for each of the three segments of each half of the LV (**Figure 4.11** and **Table 4.11**; **Figure 4.12** and **Table 4.12**).



Figure 4.8 LV axis length measurement.

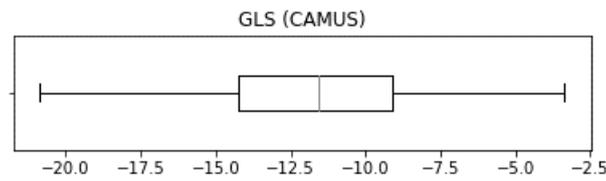


Figure 4.9 Box plot of estimated GLS for CAMUS images.

Table 4.9 Statistical values for estimated GLS for CAMUS images.

CAMUS	Average	Maximum	Minimum
Estimated GLS	-11.71%	-20.87%	-3.36%

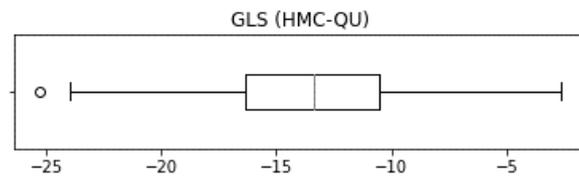


Figure 4.10 Box plot of estimated GLS for HMC-QU images.

Table 4.10 Statistical values for estimated GLS for HMC-QU images.

HMC-QU	Average	Maximum	Minimum
Estimated GLS	-11.71%	-20.87%	-3.36%

Estimated GLS	-13.43%	-25.28%	-2.61%
------------------	---------	---------	--------

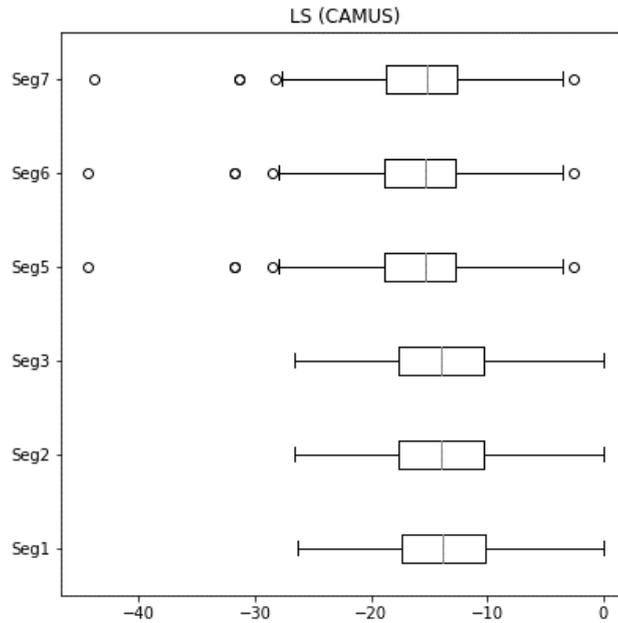


Figure 4.11 Box plots of estimated LS for all segments of CAMUS images.

Table 4.11 Statistical values of estimated LS for all segments of CAMUS images.

CAMUS	Average	Maximum	Minimum
Seg1	-13.60%	-26.37%	0
Seg2	-13.77%	-26.67%	0
Seg3	-13.77%	-26.67%	0
Seg5	-15.79%	-44.44%	-2.56%
Seg6	-15.79%	-44.44%	-2.56%
Seg7	-15.60%	-43.84%	-2.53%

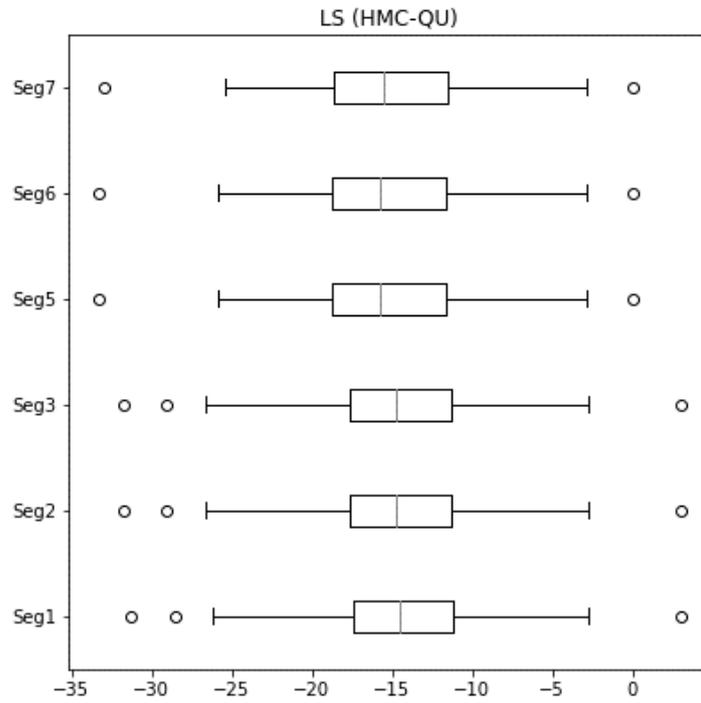


Figure 4.12 Box plots of estimated LS for all segments of HMC-QU images.

Table 4.12 Statistical values of estimated LS for all segments of HMC-QU images.

HMC-QU	Average	Maximum	Minimum
Seg1	-14.25%	-31.33%	2.99%
Seg2	-14.46%	-31.71%	3.03%
Seg3	-14.46%	-31.71%	3.03%
Seg5	-15.30%	-33.33%	0
Seg6	-15.30%	-33.33%	0
Seg7	-15.09%	-32.91%	0

4.4.1.3 Segment displacement

Segment displacement was calculated and normalized following the same guidelines as Kiranyaz *et al.* [23] as was described in **paragraph 3.6.1.3**. The boxplot distribution of the values obtained for each segment for the images corresponding to the CAMUS dataset are present in **Figure 4.13** and **Table 4.13** and for those from HMC-QU in **Figure 4.14** and **Table 4.14**. For the most part, the values obtained for both types of images are very similar, with an absolute variation of 2% or 3%.

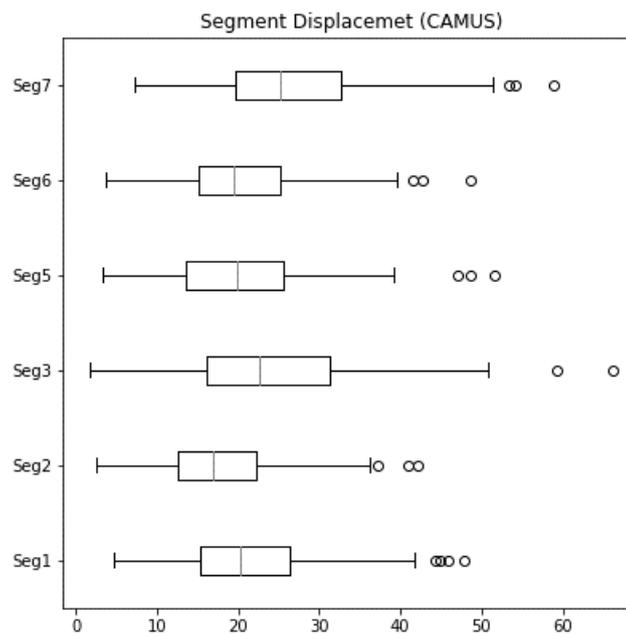


Figure 4.13 Box plots of estimated maximum displacement for all segments of CAMUS images.

Table 4.13 Statistical values of estimated maximum displacement for all segments of CAMUS images.

CAMUS	Average	Maximum	Minimum
Seg1	21.25%	44.77%	4.66%
Seg2	17.54%	42.24%	2.51%
Seg3	24.16%	66.18%	1.71%
Seg5	20.39%	51.62%	3.27%
Seg6	20.18%	48.69%	3.82%
Seg7	26.38%	58.88%	7.34%

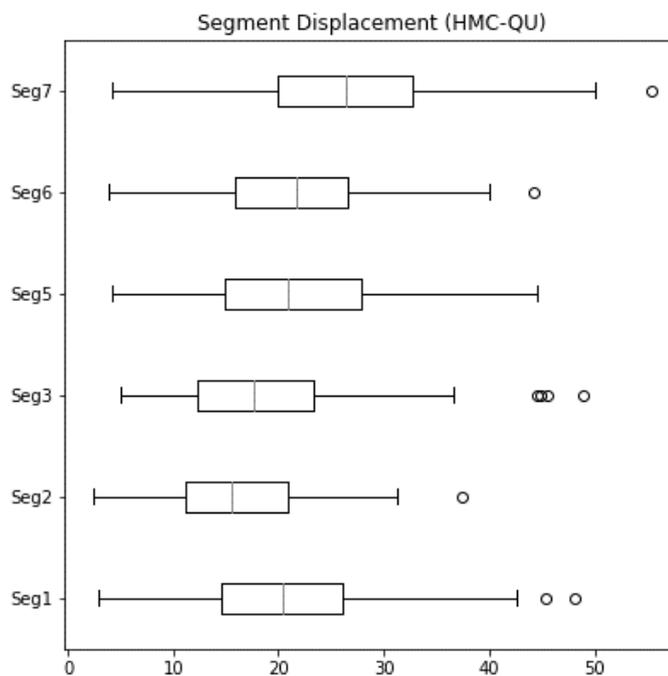


Figure 4.14 Box plots of estimated maximum displacement for all segments of HMC-QU images.

Table 4.14 Statistical values of estimated maximum displacement for all segments of HMC-QU images.

HMC-QU	Average	Maximum	Minimum
Seg1	20.45%	48.01%%	2.97%
Seg2	15.79%	37.50%	2.36%
Seg3	18.55%	49.01%	4.97%
Seg5	21.75%	44.50%	4.16%
Seg6	21.52%	44.29%	3.82%
Seg7	26.68%	55.42%	4.26%

4.4.1.4 Wall thickness

Lastly, wall thickness was calculated as the fractional average thickness change for each segment as was described in **paragraph 3.6.1.4**. The boxplot distribution of the values obtained for each segment for the images corresponding to the CAMUS dataset are present in **Figure 4.15** and **Table 4.15** and for those from HMC-QU in **Figure 4.16** and **Table 4.16**. For both datasets, wall thickness presents a wide range of values from negative to positive.

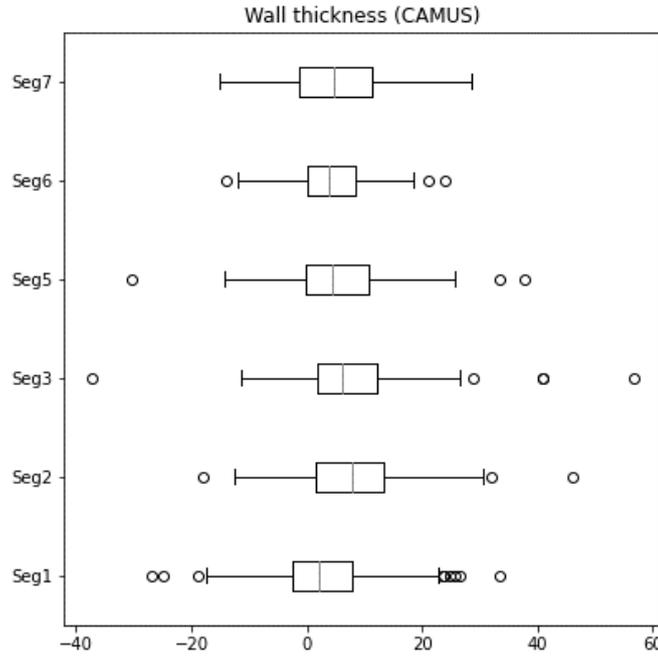


Figure 4.15 Box plots of estimated dynamic wall thickness for all segments of CAMUS images.

Table 4.15 Statistical values of estimated dynamic wall thickness for all segments of CAMUS images.

CAMUS	Average	Maximum	Minimum
Seg1	3.08%	33.32%	-26.80%
Seg2	8.21%	46.22%	-17.90%
Seg3	7.39%	56.71%	-37.34
Seg5	5.22%	37.83%	-30.41%
Seg6	3.98%	23.99%	-13.86%
Seg7	4.87%	28.69%	-15.05%

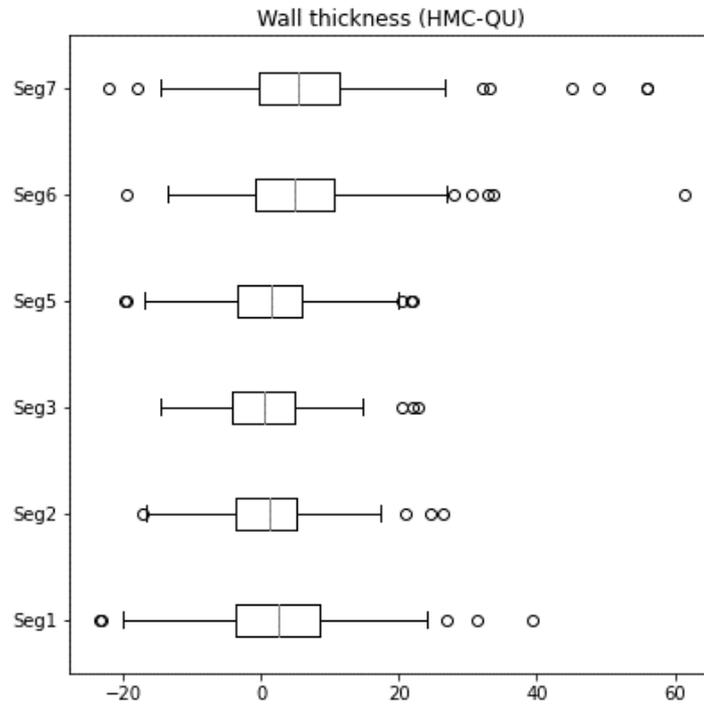


Figure 4.16 Box plots of estimated dynamic wall thickness for all segments of HMC-QU images.

Table 4.16 Statistical values of estimated dynamic wall thickness for all segments of HMC-QU images.

HMC-QU	Average	Maximum	Minimum
Seg1	2.72%	39.48%	-23.43%
Seg2	0.87%	26.45%	-17.22%
Seg3	0.53%	22.74%	-14.58%
Seg5	1.43%	21.95%	-19.64%
Seg6	5.97%	61.41%	-19.38%
Seg7	6.95%	55.93%	-22.07%

4.4.1.5 Cross-correlation

After obtaining all the features, cross correlation between each one of them and the label MI was calculated (**Table 4.17**). From this analysis it was clear that LVEF and segment displacement were the most relevant features while wall thickness did not offer high correlation for classification. It was therefore decided to remove the thickness feature and thus, for training of the classification models, only LVEF, GLS, LS and segment displacement were used.

Table 4.17 Cross-correlation between MI label and features extracted.

	MI Label	LVEF	GLS	LS	Thickness	Displacement
MI Label	1	-0.479	0.384	0.404	-0.128	-0.447

4.4.2 KNN

After testing multiple k values for the K-Nearest Neighbor algorithm, k=3 was chosen as the optimum value. Therefore, the results about to be presented for the 3 different K-NN models trained, all use k=3 as parameter. The detailed confusion matrices obtained for the 3 models are presented in **Appendix A.1**.

Regarding the model trained for general classification, the input sample corresponds to a patient and all the 14 features calculated: 2 global features (LVEF and GLS) and 2 features for each of the 6 segments (LS and maximum displacement). **Table 4.18** summarizes the results obtained.

Table 4.18 Results for KNN Model 1

Model 1	TP	FN	TN	FP	Accuracy	Sensitivity	Specificity
Training	65	6	38	9	87.3%	91.5%	80.9%
Test	17	1	10	2	90%	94.4%	83.3 %

TP: true positive ; FN: false negative ; TN: true negative; FP: false positive
 *positive corresponds to classification as MI, negative to classification as Non-MI

Segment specific classification, which is obtained with the second model, expects as input a single segment with 4 features: 2 global features (LVEF and GLS) and 2 features of the segment (LS and maximum displacement). By including global features for each segment, the model classifies a local phenomenon taking into account the global picture. The results obtained for training and test data are presented in **Table 4.19**.

Table 4.19 Results for KNN Model 2

Model 2	TP	FN	TN	FP	Accuracy	Sensitivity	Specificity
Training	143	61	471	35	86.5%	70.1%	93.1 %
Test	39	12	99	28	77.5 %	76.5%	77.95 %

Model 3 which differentiates between segments, requires training of 6 different classifiers (each one specific for a single segment). The performance obtained over the test data for each of the 6 classifiers are summarized in **Table 4.20**.

Table 4.20 Results for KNN Model 3

Model 3	TP	FN	TN	FP	Accuracy	Sensitivity	Specificity
Seg1	5	0	20	2	92.6%	100%	90.9%
Seg2	8	2	23	2	88.6%	80%	92%
Seg3	12	6	15	1	79.4%	66.7%	93.8%
Seg5	8	4	12	2	76.9%	66.7%	75%
Seg6	3	1	14	1	89.5%	75%	93.3%
Seg7	2	0	28	7	81.1%	100%	80%

4.4.3 SVM

The SVM technique uses more model parameters in comparison to K-NN. After testing different combinations, it was found that the best results were obtained when using a RBF kernel with a gamma factor equal to 0.1 and a penalty parameter C of 60. These values were used for the 3 models tested

(patient classification, segment classification, segment specific classification) and resulted in the confusion matrices presented in **Appendix A.2**. In **Table 4.21****Table 4.22****Table 4.23** are summarized the results obtained. It can be seen, in fact, that SVM obtained lower performance for all 3 models in comparison to the K-NN approach.

Table 4.21 Results for SVM Model 1

Model 1	TP	FN	TN	FP	Accuracy	Sensitivity	Specificity
Training	66	5	33	14	83.9%	92.9%	70.2%
Test	16	2	8	4	80%	88.8%	66.7%

Table 4.22 Results for SVM Model 2

Model 2	TP	FN	TN	FP	Accuracy	Sensitivity	Specificity
Training	111	93	457	49	80%	54.4%	90.3%
Test	27	24	114	13	79.2%	52.9%	89.8%

Table 4.23 Results for SVM Model 3

Model 3	TP	FN	TN	FP	Accuracy	Sensitivity	Specificity
Seg1	1	3	20	6	70%	25%	76.9
Seg2	4	7	14	5	60%	36.4%	73.8%
Seg3	13	4	12	1	83.3%	76.5%	92.3%
Seg5	16	4	4	6	66.7%	80%	40%
Seg6	10	4	14	2	60%	28.6%	87.5%
Seg7	3	6	20	1	76.7%	33.3%	95.2%

4.4.4 RF

Finally, the Random Forest model tested was designed to have 50 different trees trained with bootstrap datasets and feature bagging considering a number of features equal to the square root of the total number of features. RF obtained the best results as a whole in comparison with SVM and K-NN, as can be seen explicitly from the confusion matrices in **Appendix A.3** and summarized in **Table 4.24****Table 4.25****Table 4.26**. This was therefore the model chosen to test the next step of data augmentation.

Table 4.24 Results for RF Model 1

Model 1	TP	FN	TN	FP	Accuracy	Sensitivity	Specificity
Training	70	1	0	47	99.2%	98.6%	100%
Test	18	0	10	2	93.3%	100%	83.3%

Table 4.25 Results for RF Model 2

Model 2	TP	FN	TN	FP	Accuracy	Sensitivity	Specificity
Training	204	0	506	0	100%	100%	100%
Test	35	16	111	16	82%	68.6%	87.4%

Table 4.26 Results for RF Model 3

Model 3	TP	FN	TN	FP	Accuracy	Sensitivity	Specificity
Seg1	5	1	22	2	90%	83.3%	91.7%
Seg2	10	1	14	5	80%	90.9%	73.7%
Seg3	14	0	13	3	90%	100%	81.3%
Seg5	10	1	16	3	86.7%	90.9%	84.2%
Seg6	5	1	22	2	90%	83.3%	91.7%
Seg7	3	1	26	0	96.7	75%	100%

4.4.5 Data augmentation

After adding the 15 artificial patients to the training set and moving 15 real patients to the test set, the 3 RF models were retrained (RF + DA). The greatest advantage of this method is allowing to increment the test set, always with real patients, without reducing the training set. This procedure is especially useful in cases, like the one in this thesis, where few data is available and in addition it is not well balanced among classes. The performances obtained are presented in **Table 4.27**, **Table 4.28** **Table 4.29** (confusion matrices available in **Appendix A.4**). These results demonstrate that the procedure of data augmentation leads to improved and more truthful classifications since more samples are being tested.

Table 4.27 Results for RF+DA Model 1

Model 1	TP	FN	TN	FP	Accuracy	Sensitivity	Specificity
Training	77	0	40	1	99.2%	100%	97.6%
Test	26	1	15	3	91.1%	96.3%	83.3%

Table 4.28 Results for RF+DA Model 2

Model 2	TP	FN	TN	FP	Accuracy	Sensitivity	Specificity
Training	267	1	441	1	99.7%	99.6%	99.8%
Test	60	17	167	24	84.7%	77.9%	87.4%

Table 4.29 Results for RF+DA Model 3

Model 3	TP	FN	TN	FP	Accuracy	Sensitivity	Specificity
Seg1	8	1	32	4	88.9%	88.9%	88.9%
Seg2	15	1	22	7	82.2%	93.8%	75.9%
Seg3	22	0	20	3	93.3%	100%	87%
Seg5	16	0	25	4	91.1%	100%	86.2%
Seg6	8	1	33	3	91.1%	88.9%	91.7%
Seg7	5	0	38	2	95.6%	100%	95%

Additionally, the ROC curves with the test sets for each classifier are also presented and the AUC is calculated (**Figure 4.17** **Figure 4.24**). As confrontation, the ROC curves relative to a simple threshold method are also presented. Following the approach of Kiranyas *et al.* [23], which obtained sensitivity and specificity values of 80.16% and 91.24%, respectively, segments are classified applying a threshold to the maximum displacement value. In the article, segment maximum displacement is calculated for patients with LVEF between 15% and 55%. Then, if the displacement is less than 19% the segment is classified as infarcted and patients with at least 1 infarcted segment are classified as presenting MI. Even though Chapter 2 **paragraph 2.3**, discusses the reasons why classification should not be based on one single feature and much less by using a threshold value, this analysis was conducted in this thesis for confrontation.

In fact, from the ROC curves of the threshold method, different ideal threshold values are obtained for each model (ranging from 15%-20%), which further supports the statement that a fixed threshold cannot be defined for classification. From the ROC curves it can also be seen that the RF obtained greater AUC for all the models meaning it is a better strategy for classification. Actually, the average AUC of the 6 models for segment specific classification is 0.79 for the threshold method, while with the RF+DA method an average AUC of 0.92 is obtained.

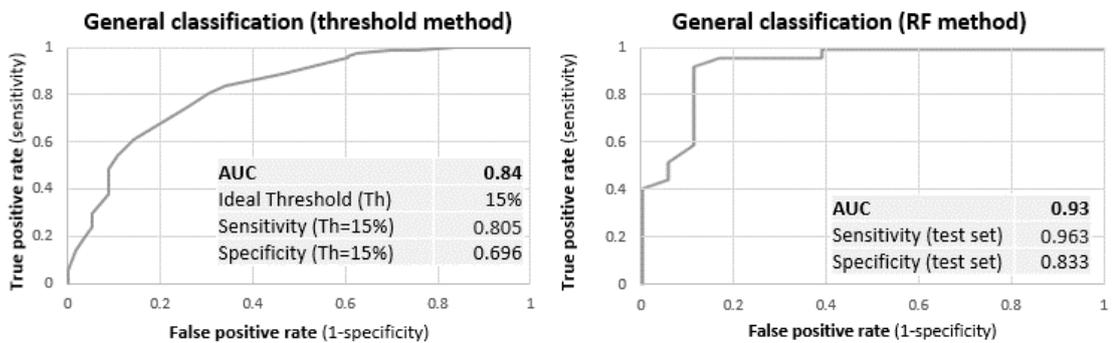


Figure 4.17 ROC curves general classification: Threshold method vs RF+DA

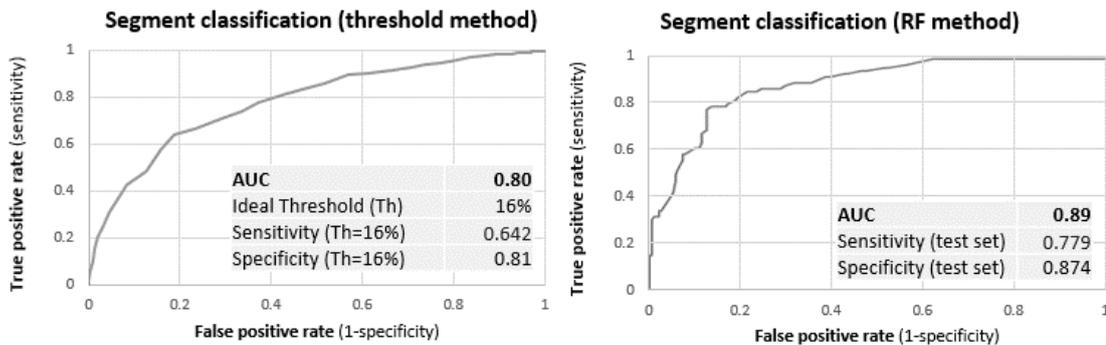


Figure 4.18 ROC curves segment classification: Threshold method vs RF+DA

Results

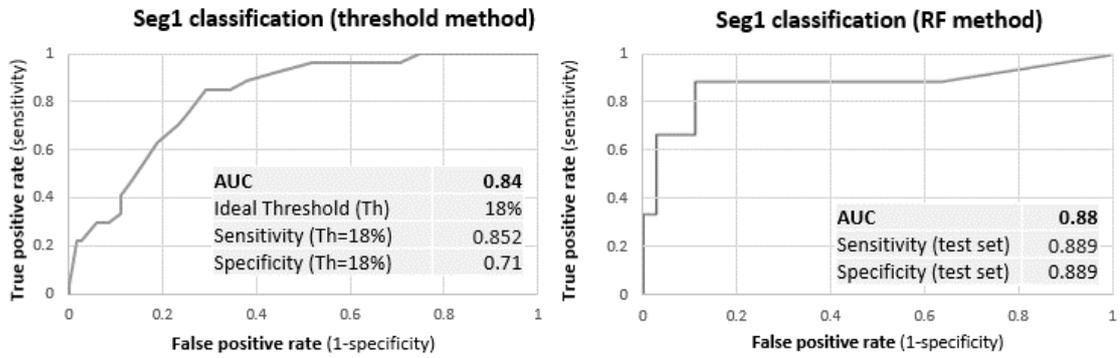


Figure 4.19 ROC curves Seg1 classification: Threshold method vs RF+DA

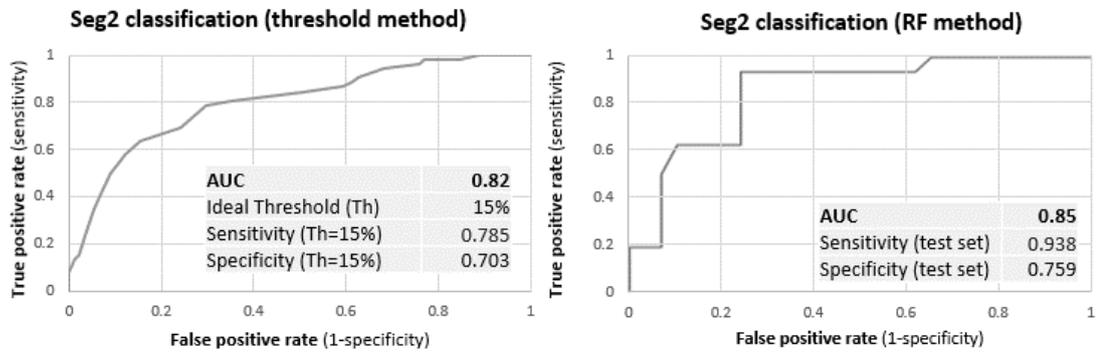


Figure 4.20 ROC curves Seg2 classification: Threshold method vs RF+DA

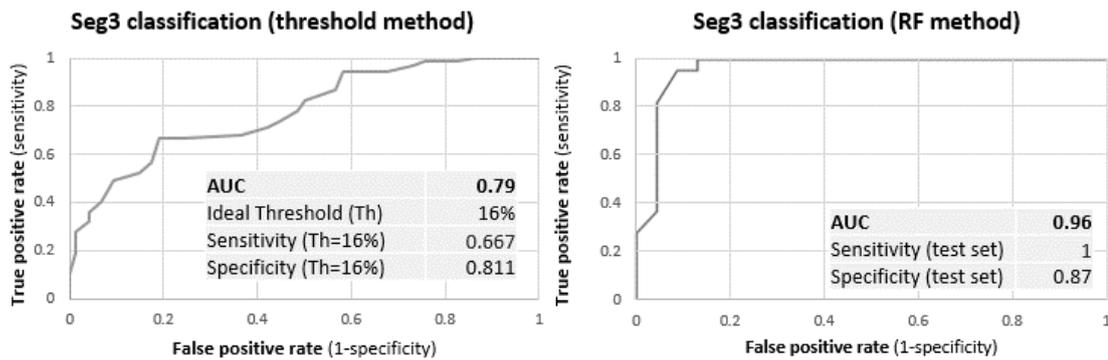


Figure 4.21 ROC curves Seg3 classification: Threshold method vs RF+DA

Results

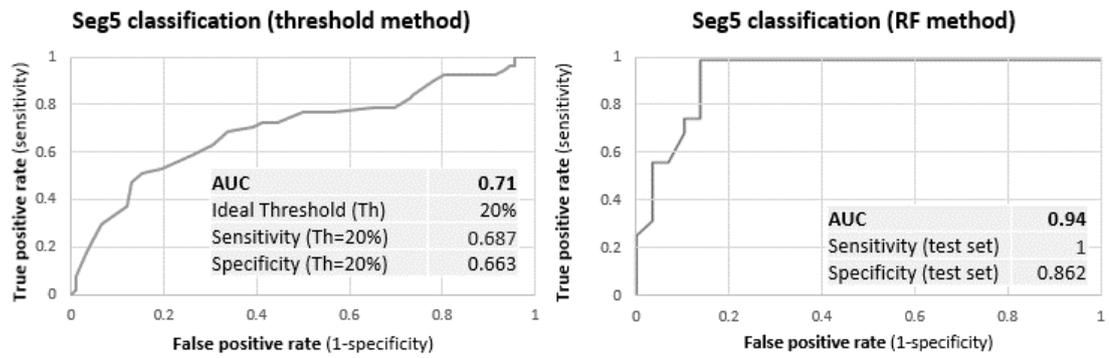


Figure 4.22 ROC curves Seg5 classification: Threshold method vs RF+DA

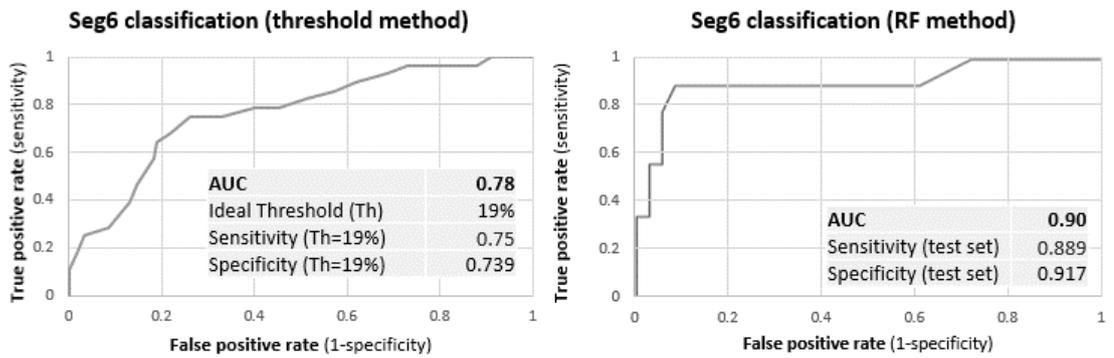


Figure 4.23 ROC curves Seg6 classification: Threshold method vs RF+DA

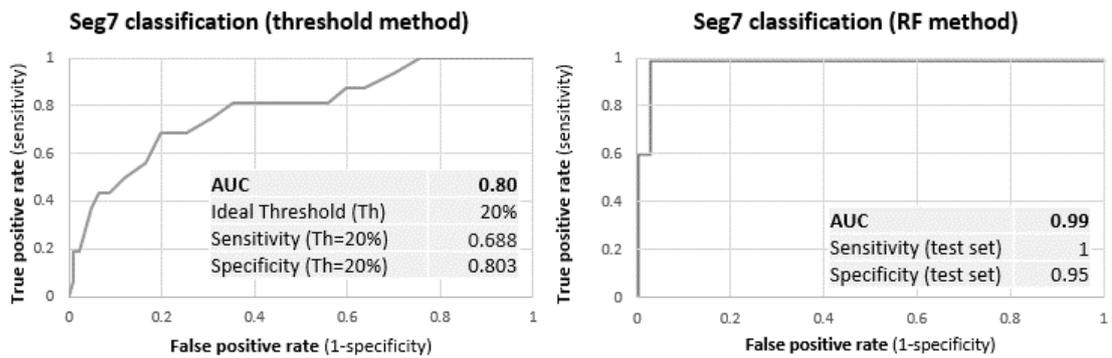


Figure 4.24 ROC curves Seg7 classification: Threshold method vs RF+DA

4.5 Cascade approach

Based on the results obtained with the ML classifiers, the highest sensitivity and specificity are obtained by classifying the patient as a whole. However, specifying which segment, or segments, present the infarction is of great importance when diagnosing MI since this can aid in the identification of the specific coronary artery affected. Therefore, the models trained to classify the segments should also be included in the proposed solution. Taking both of these things into account, this thesis project finally proposes the implementation of a cascade approach in which the general classification of the patient is first made with the first RF model and, in the case the patient is labeled as MI, the classification of the individual segments follows. For this second classification, the RF model that comprehends one classifier for each segment is used, since it demonstrates higher performance compared to the single classifier trained with all segments. The complete confusion matrices obtained for each segment can be reviewed in **Appendix A.5**. Following, will be presented the summary of the results.

When applying the segments classifiers to the 29 patients classified as MI the results presented in **Table 4.30** are obtained.

Table 4.30 General results for segment classification with cascade approach

Model 3	TP	FN	TN	FP	Accuracy	Sensitivity	Specificity
Seg1	10	0	18	1	96.6%	100%	94.7%
Seg2	15	0	13	1	96.6%	100%	92.9%
Seg3	19	0	9	1	96.6%	100%	90%
Seg5	17	0	9	3	89.7%	100%	75%
Seg6	10	0	18	1	96.6%	100%	94.7%
Seg7	5	0	22	2	93.1%	100%	91.7%

Since this evaluation includes segments used during training of each of the 6 classifiers, **Table 4.31** presents the results if only segments belonging to the test set of each model are evaluated. Even when evaluating only test set segments, sensitivity in all cases reaches 100%, meaning that all MI cases were

correctly identified. It is important to note that not a lot of data was available for classification. This highly affects parameter values because for example if only 4 samples are available, by misclassifying only 1 sample the parameter already reduces 25%. The same occurs with correct classification where a parameter can reach 100% but only 2 samples were tested.

Table 4.31 Results for segment classification with cascade approach evaluating test segments

Model 3	TP	FN	TN	FP	Accuracy	Sensitivity	Specificity
Seg1	4	0	7	1	91.7%	100%	87.5%
Seg2	8	0	3	0	100%	100%	100%
Seg3	5	0	2	0	100%	100%	100%
Seg5	5	0	3	1	88.9%	100%	75%
Seg6	3	0	6	0	100%	100%	100%
Seg7	2	0	8	0	100%	100%	100%

4.6 Processing time

Processing time depends on the availability or not of a GPU. In both cases, the majority of the time is engaged in the segmentation process (obtaining the automatic mask) and the feature extraction process. For the latter, the step that takes the longest is the calculation of the LV cavity area for all frames in order to identify the two corresponding to end-systole and end-diastole. From then on, all processes are done exclusively with these two frames and are therefore done very fast. As reference, the following table (**Table 4.32**) presents the time required to complete each step of the whole thesis solution architecture for a sample HMC-QU echo video lasting 1.64 seconds (at 25fps this accounts to 41 frames total). The use of a GPU greatly reduces the time necessary for preprocessing and for obtaining the segmentation mask. However, for feature extraction the time reduction is not as significant because the calculations are done with the NumPy Python library which does not make use of the GPU even when available.

Table 4.32 Processing time with and without GPU for a sample echo video

Procedure	Step	Without GPU	Quadro RTX 6000 GPU
Preprocessing	Frames extraction	4s ($\approx 97\text{ms/frame}$)	8ms
	Equalization	10s ($\approx 0.24\text{s/frame}$)	1.23s
Segmentation (U-Net)	Automatic mask generation	45s ($\approx 1\text{s/frame}$)	2.06s ($\approx 50\text{ms/frame}$)
Post processing	Closing	58ms ($\approx 1.4\text{ms/frame}$)	26ms
Feature extraction	Area and LVEF	37s	24.6s
	GLS	161ms	100ms
	Segment division and LS	427ms	260ms
	Displacement	19ms	5ms
Classification (RF)	Patient (general) classification	19ms	9ms
	Segment classification	56ms	32ms
TOTAL		96.74s	28.4s

CHAPTER 5

5 Conclusions

The purpose of this thesis project was the realization of an automatic algorithm for the early diagnosis of myocardial infarction from A4C echocardiography videos with the aid of artificial intelligence models. This involved two main steps, segmentation of the LV wall for myocardial layer identification and classification based on features extracted from the segmentation. Regarding segmentation, despite the abundance of articles with a deep learning approach, this thesis offers, to the knowledge of the author, the first solution combining two publicly available datasets, CAMUS and HMC-QU. This is significant since it allows the generation of a network capable of greater generalization when exposed to 2D echocardiography recordings obtained with different equipment. Additionally, the performances obtained were at the level, if not higher, of the articles reviewed. The obtained segmentation is not only useful for the following automatic feature extraction but can, by itself, be used as an enhanced visualization platform over the raw echo to assist cardiologists and allow them to validate the diagnosis offered by the algorithm. Regarding feature extraction, this thesis project uses for the first time real clinical parameters of cardiac function for classification using the HMC-QU dataset. Both global and segmental physiological parameters are obtained which once again, by themselves, serve the cardiologist for validating the diagnosis. Finally, a random forest classifier was built to identify the presence of myocardial infarction, first diagnosing the patient as a whole and then, if MI is detected, identifying the specific segments involved. This offers a double verification when the algorithm detects MI and, by classifying the individual segments, a more informative diagnosis is obtained.

5.1 Open points

Although this thesis project obtains high performances both in segmentation of the LV and classification of MI, some improvements and future developments can still be considered. Regarding segmentation, the proposed network only works with A4C view echocardiography. The main drawback from this is that features can only be obtained from this view which, for example, results in the impossibility of calculating LVEF with the modified Simpson's rule recommended by The American Society of Echocardiography. This leads then to the use of an approximated LVEF value for classification. Concerning feature extraction and specifically the time it takes to be completed, in future works the NumPy library should be replaced with one that takes advantage of GPU. This would substantially decrease the time in the feature extraction process and thus of the solution pipeline as a whole. In regards to the classification process, the main difficulty present is the low availability of data since only the HMC-QU dataset can be used for this step. The obtained model performance was in general very high but it is greatly affected by a bias introduced from the lack of sufficient samples. Lastly, in order to verify that the solution is effective and robust, it should be validated with a dataset different from the ones used in this thesis.

A Appendix

A.1 KNN classification results

Table A.1 KNN model 1: Training set confusion matrix.

Confusion matrix Training set		Target (labels)	
		MI	Non-MI
Model 1 (KNN)	MI	65	9
	Non-MI	6	38
	Accuracy	Sensitivity	Specificity
	87.3%	91.5%	80.9 %

Table A.2 KNN model 1: Test set confusion matrix.

Confusion matrix Test set		Target (labels)	
		MI	Non-MI
Model 1 (KNN)	MI	17	2
	Non-MI	1	10
	Accuracy	Sensitivity	Specificity
	90%	94.4%	83.3 %

Table A.3 KNN model 3: Seg1 Test set confusion matrix

Confusion matrix	Target (labels)
------------------	-----------------

Test set (Seg1)		MI	Non-MI
Model 3 (KNN)	MI	5	2
	Non-MI	0	20
	Accuracy	Sensitivity	Specificity
	92.6%	100%	90.9%

Table A.4 KNN model 3: Seg2 Test set confusion matrix

Confusion matrix Test set (Seg2)		Target (labels)	
		MI	Non-MI
Model 3 (KNN)	MI	8	2
	Non-MI	2	23
	Accuracy	Sensitivity	Specificity
	88.6%	80%	92%

Table A.5 KNN model 3: Seg3 Test set confusion matrix

Confusion matrix Test set (Seg3)		Target (labels)	
		MI	Non-MI
Model 3 (KNN)	MI	12	1
	Non-MI	6	15
	Accuracy	Sensitivity	Specificity
	79.4%	66.7%	93.8%

Table A.6 KNN model 3: Seg5 Test set confusion matrix

Confusion matrix Test set (Seg5)		Target (labels)	
		MI	Non-MI
Model 3 (KNN)	MI	8	2
	Non-MI	4	12
Accuracy		Sensitivity	Specificity
76.9%		66.7%	75%

Table A.7 KNN model 3: Seg6 Test set confusion matrix

Confusion matrix Test set (Seg6)		Target (labels)	
		MI	Non-MI
Model 3 (KNN)	MI	3	1
	Non-MI	1	14
Accuracy		Sensitivity	Specificity
89.5%		75%	93.3%

Table A.8 KNN model 3: Seg7 Test set confusion matrix

Confusion matrix Test set (Seg7)		Target (labels)	
		MI	Non-MI
Model 3 (KNN)	MI	2	7
	Non-MI	0	28
Accuracy		Sensitivity	Specificity

	81.1%	100%	80%
--	-------	------	-----

Table A.9 KNN model 2: Training set confusion matrix.

Confusion matrix Training set		Target (labels)	
		MI	Non-MI
Model 2 (KNN)	MI	143	35
	Non-MI	61	471
		Accuracy	Sensitivity
		86.5%	70.1%
		Specificity	93.1 %

Table A.10 KNN model 2: Test set confusion matrix.

Confusion matrix Test set		Target (labels)	
		MI	Non-MI
Model 2 (KNN)	MI	39	28
	Non-MI	12	99
		Accuracy	Sensitivity
		77.5 %	76.5%
		Specificity	77.95 %

A.2 SVM

Table A.11 SVM model 1: Training set confusion matrix.

Confusion matrix Training set		Target (labels)	
		MI	Non-MI
Model 1 (SVM)	MI	66	14
	Non-MI	5	33
		Accuracy	Sensitivity
		83.9%	92.9%
			Specificity
			70.2%

Table A.12 SVM model 1: Test set confusion matrix.

Confusion matrix Test set		Target (labels)	
		MI	Non-MI
Model 1 (SVM)	MI	16	4
	Non-MI	2	8
		Accuracy	Sensitivity
		80%	88.8%
			Specificity
			66.7%

Table A.13 SVM model 2: Training set confusion matrix.

Confusion matrix Training set		Target (labels)	
		MI	Non-MI
Model 2 (SVM)	MI	111	49
	Non-MI	93	457

	Accuracy	Sensitivity	Specificity
	80%	54.4%	90.3%

Table A.14 SVM model 2: Test set confusion matrix.

Confusion matrix Test set		Target (labels)		
		MI	Non-MI	
Model 2 (SVM)	MI	27	13	
	Non-MI	24	114	
		Accuracy	Sensitivity	Specificity
		79.2%	52.9%	89.8%

Table A.15 SVM model 3: Seg1 Test set confusion matrix

Confusion matrix Test set (Seg1)		Target (labels)		
		MI	Non-MI	
Model 3 (SVM)	MI	1	6	
	Non-MI	3	20	
		Accuracy	Sensitivity	Specificity
		70%	25%	76.9

Table A.16 SVM model 3: Seg2 Test set confusion matrix

Confusion matrix Test set (Seg2)		Target (labels)	
		MI	Non-MI

Model 3 (SVM)	MI	4	5
	Non-MI	7	14
	Accuracy	Sensitivity	Specificity
	60%	36.4%	73.8%

Table A.17 SVM model 3: Seg3 Test set confusion matrix

Confusion matrix Test set (Seg3)		Target (labels)	
		MI	Non-MI
Model 3 (SVM)	MI	13	1
	Non-MI	4	12
	Accuracy	Sensitivity	Specificity
	83.3%	76.5%	92.3%

Table A.18 SVM model 3: Seg5 Test set confusion matrix

Confusion matrix Test set (Seg5)		Target (labels)	
		MI	Non-MI
Model 3 (SVM)	MI	16	6
	Non-MI	4	4
	Accuracy	Sensitivity	Specificity
	66.7%	80%	40%

Table A.19 SVM model 3: Seg6 Test set confusion matrix

Confusion matrix Test set (Seg6)		Target (labels)	
		MI	Non-MI
Model 3 (SVM)	MI	10	2
	Non-MI	4	14
		Accuracy	Sensitivity
		60%	87.5%

Table A.20 SVM model 3: Seg7 Test set confusion matrix

Confusion matrix Test set (Seg7)		Target (labels)	
		MI	Non-MI
Model 3 (SVM)	MI	3	1
	Non-MI	6	20
		Accuracy	Sensitivity
		76.7%	95.2%

A.3 RF

Table A.21 RF model 1: Training set confusion matrix.

Confusion matrix Training set		Target (labels)	
		MI	Non-MI
Model 1	MI	70	0

(RF)	Non-MI	1	47
	Accuracy	Sensitivity	Specificity
	99.2%	98.6%	100%

Table A.22 RF model 1: Test set confusion matrix.

Confusion matrix Test set		Target (labels)	
		MI	Non-MI
Model 1 (RF)	MI	18	2
	Non-MI	0	10
	Accuracy	Sensitivity	Specificity
	93.3%	100%	83.3%

Table A.23 RF model 2: Training set confusion matrix.

Confusion matrix Training set		Target (labels)	
		MI	Non-MI
Model 2 (RF)	MI	204	0
	Non-MI	0	506
	Accuracy	Sensitivity	Specificity
	100%	100%	100%

Table A.24 RF model 2: Test set confusion matrix.

Confusion matrix	Target (labels)
-------------------------	------------------------

Test set		MI	Non-MI
Model 2 (RF)	MI	35	16
	Non-MI	16	111
	Accuracy	Sensitivity	Specificity
	82%	68.6%	87.4%

Table A.25 RF model 3: Seg1 Test set confusion matrix

Confusion matrix Test set (Seg1)		Target (labels)	
		MI	Non-MI
Model 3 (RF)	MI	5	2
	Non-MI	1	22
	Accuracy	Sensitivity	Specificity
	90%	83.3%	91.7%

Table A.26 RF model 3: Seg2 Test set confusion matrix

Confusion matrix Test set (Seg2)		Target (labels)	
		MI	Non-MI
Model 3 (RF)	MI	10	5
	Non-MI	1	14
	Accuracy	Sensitivity	Specificity
	80%	90.9%	73.7%

Table A.27 RF model 3: Seg3 Test set confusion matrix

Confusion matrix Test set (Seg3)		Target (labels)	
		MI	Non-MI
Model 3 (RF)	MI	14	3
	Non-MI	0	13
		Accuracy	Sensitivity
		90%	100%
			Specificity
			81.3%

Table A.28 RF model 3: Seg5 Test set confusion matrix

Confusion matrix Test set (Seg5)		Target (labels)	
		MI	Non-MI
Model 3 (RF)	MI	10	3
	Non-MI	1	16
		Accuracy	Sensitivity
		86.7%	90.9%
			Specificity
			84.2%

Table A.29 RF model 3: Seg6 Test set confusion matrix

Confusion matrix Test set (Seg6)		Target (labels)	
		MI	Non-MI
Model 3 (RF)	MI	5	2
	Non-MI	1	22
		Accuracy	Sensitivity
			Specificity

	90%	83.3%	91.7%
--	-----	-------	-------

Table A.30 RF model 3: Seg7 Test set confusion matrix

Confusion matrix Test set (Seg7)		Target (labels)	
		MI	Non-MI
Model 3 (RF)	MI	3	0
	Non-MI	1	26
		Accuracy	Sensitivity
		96.7	75%
		Specificity	100%

A.4 RF + DA

Table A.31 RF+DA model 1: Training set confusion matrix.

Confusion matrix Training set		Target (labels)	
		MI	Non-MI
Model 1 (RF+ DA)	MI	77	0
	Non-MI	0	40
		Accuracy	Sensitivity
		99.2%	100%
		Specificity	97.6%

Table A.32 RF+DA model 1: Test set confusion matrix.

Confusion matrix	Target (labels)
------------------	-----------------

Test set		MI	Non-MI
Model 1 (RF +DA)	MI	26	3
	Non-MI	1	15
		Accuracy	Sensitivity
		91.1%	96.3%
			Specificity
			83.3%

Table A.33 RF+DA model 2: Training set confusion matrix.

Confusion matrix Training set		Target (labels)	
		MI	Non-MI
Model 2 (RF + DA)	MI	267	1
	Non-MI	1	441
		Accuracy	Sensitivity
		99.7%	99.6%
			Specificity
			99.8%

Table A.34 RF+DA model 2: Test set confusion matrix.

Confusion matrix Test set		Target (labels)	
		MI	Non-MI
Model 2 (RF +DA)	MI	60	24
	Non-MI	17	167
		Accuracy	Sensitivity
		84.7%	77.9%
			Specificity
			87.4%

Table A.35 RF+DA model 3: Seg1 Test set confusion matrix

Confusion matrix Test set (Seg1)		Target (labels)		
		MI	Non-MI	
Model 3 (RF + DA)	MI	8	4	
	Non-MI	1	32	
		Accuracy	Sensitivity	Specificity
		88.9%	88.9%	88.9%

Table A.36 RF+DA model 3: Seg2 Test set confusion matrix

Confusion matrix Test set (Seg2)		Target (labels)		
		MI	Non-MI	
Model 3 (RF +DA)	MI	15	7	
	Non-MI	1	22	
		Accuracy	Sensitivity	Specificity
		82.2%	93.8%	75.9%

Table A.37 RF+DA model 3: Seg3 Test set confusion matrix

Confusion matrix Test set (Seg3)		Target (labels)	
		MI	Non-MI
Model 3 (RF+DA)	MI	22	3
	Non-MI	0	20

	Accuracy	Sensitivity	Specificity
	93.3%	100%	87%

Table A.38 RF+DA model 3: Seg5 Test set confusion matrix

Confusion matrix Test set (Seg5)		Target (labels)	
		MI	Non-MI
Model 3 (RF+DA)	MI	16	4
	Non-MI	0	25
		Accuracy	Sensitivity
		91.1%	100%
		Specificity	86.2%

Table A.39 RF+DA model 3: Seg6 Test set confusion matrix

Confusion matrix Test set (Seg6)		Target (labels)	
		MI	Non-MI
Model 3 (RF+DA)	MI	8	3
	Non-MI	1	33
		Accuracy	Sensitivity
		91.1%	88.9%
		Specificity	91.7%

Table A.40 RF+DA model 3: Seg7 Test set confusion matrix

Confusion matrix Test set (Seg7)		Target (labels)	
		MI	Non-MI

Model 3 (RF + DA)	MI	5	2
	Non-MI	0	38
	Accuracy	Sensitivity	Specificity
	95.6%	100%	95%

A.5 Cascade approach

Table A.41 General results for Seg1 classification with cascade approach

Confusion matrix (Seg1)		Target (labels)	
		MI	Non-MI
Model 3 Cascade approach	MI	10	1
	Non-MI	0	18
	Accuracy	Sensitivity	Specificity
	96.6%	100%	94.7%

Table A.42 General results for Seg2 classification with cascade approach

Confusion matrix (Seg2)		Target (labels)	
		MI	Non-MI
Model 3 Cascade approach	MI	15	1
	Non-MI	0	13
	Accuracy	Sensitivity	Specificity

	96.6%	100%	92.9%
--	-------	------	-------

Table A.43 General results for Seg3 classification with cascade approach

Confusion matrix (Seg3)		Target (labels)	
		MI	Non-MI
Model 3 Cascade approach	MI	15	1
	Non-MI	0	9
Accuracy		Sensitivity	Specificity
96.6%		100%	90%

Table A.44 General results for Seg5 classification with cascade approach

Confusion matrix (Seg5)		Target (labels)	
		MI	Non-MI
Model 3 Cascade approach	MI	17	3
	Non-MI	0	9
Accuracy		Sensitivity	Specificity
89.7%		100%	75%

Table A.45 General results for Seg6 classification with cascade approach

Confusion matrix (Seg6)		Target (labels)	
		MI	Non-MI

Model 3 Cascade approach	MI	10	1
	Non-MI	0	18
	Accuracy	Sensitivity	Specificity
	96.6%	100%	94.7%

Table A.46 General results for Seg7 classification with cascade approach

Confusion matrix (Seg7)		Target (labels)	
		MI	Non-MI
Model 3 Cascade approach	MI	5	2
	Non-MI	0	22
	Accuracy	Sensitivity	Specificity
	93.1%	100%	91.7%

Table A.47 Results for Seg1 test segments classification with cascade approach

Confusion matrix Test segments (Seg1)		Target (labels)	
		MI	Non-MI
Model 3 Cascade approach	MI	4	1
	Non-MI	0	7
	Accuracy	Sensitivity	Specificity
	91.7%	100%	87.5%

Table A.48 Results for Seg2 test segments classification with cascade approach

Confusion matrix Test segments (Seg2)		Target (labels)		
		MI	Non-MI	
Model 3 Cascade approach	MI	8	0	
	Non-MI	0	3	
		Accuracy	Sensitivity	Specificity
		100%	100%	100%

Table A.49 Results for Seg3 test segments classification with cascade approach

Confusion matrix Test segments (Seg3)		Target (labels)		
		MI	Non-MI	
Model 3 Cascade approach	MI	5	0	
	Non-MI	0	2	
		Accuracy	Sensitivity	Specificity
		100%	100%	100%

Table A.50 Results for Seg5 test segments classification with cascade approach

Confusion matrix Test segments (Seg5)		Target (labels)	
		MI	Non-MI
Model 3 Cascade approach	MI	5	1
	Non-MI	0	3

	Accuracy	Sensitivity	Specificity
	88.9%	100%	75%

Table A.51 Results for Seg6 test segments classification with cascade approach

Confusion matrix Test segments (Seg6)		Target (labels)	
		MI	Non-MI
Model 3 Cascade approach	MI	6	0
	Non-MI	0	3
	Accuracy	Sensitivity	Specificity
	100%	100%	100%

Table A.52 Results for Seg7 test segments classification with cascade approach

Confusion matrix Test segments (Seg7)		Target (labels)	
		MI	Non-MI
Model 3 Cascade approach	MI	8	0
	Non-MI	0	2
	Accuracy	Sensitivity	Specificity
	100%	100%	100%

Bibliography

- [1] C. Clinic, "Heart," 26 August 2021. [Online]. Available: <https://my.clevelandclinic.org/health/body/21704-heart>. [Accessed 2023].
- [2] A. C. Guyton and J. E. Hall, Textbook of Medical Physiology, Philadelphia: W.B. Saunders, 2000.
- [3] H. Fukuta and W. C. Little, "The Cardiac Cycle and the Physiological Basis of Left Ventricular," *National Institute of Health*, vol. 4, no. 1, pp. 1-11, 2008.
- [4] D. B. Tran, C. Weber and R. A. Lopez, "Anatomy, Thorax, Heart Muscles," *National Library of Medicine*, 2022.
- [5] E. a. E. Learning, "The left ventricle in myocardial ischemia and infarction," p. Chapter 6, 2022.
- [6] Y.-T. Shen, L. Chen, J. M. Testani, C. P. Regan and R. P. Shannon, "Animal Models for Cardiovascular Research," in *Animal Models for the Study of Human Disease (Second Edition)*, 2017, pp. 147-174.
- [7] B. H. Foundation, "Global Heart & Circulatory Diseases Factsheet," 2023.
- [8] C. f. D. C. a. Prevention, "Coronary Artery Disease (CAD)," 2021.
- [9] A. Degerli, S. Kiranyaz, T. Hamid, R. Mazhar and M. Gabbouj, "Early Myocardial Infarction Detection over Multi-view Echocardiography," 2022.
- [10] K. Thygesen, J. S. Alpert and H. D. White, "Universal Definition of Myocardial Infarction," *Circulation*, vol. 116, no. 22, pp. 2634-2653, 2007.
- [11] M. Baggiolini, U. Bretz, B. Dewald and M. E. Feigenson, "The polymorphonuclear leukocyte," *Agents and Actions*, vol. 8, pp. 3-10, 1978.
- [12] A. E. Arai, "Healing After Myocardial Infarction," *JACC: Cardiovascular Imaging*, vol. 8, no. 6, pp. 680-683, 2015.

Bibliography

- [13] K. Thygesen, J. S. Alpert, A. S. Jaffe, B. R. Chaitman, J. J. Bax, D. A. Morrow and H. D. White, "Fourth universal definition of myocardial infarction," *European Heart Journal*, vol. 40, no. 3, p. 237–269, 2018.
- [14] S. Vernon, ""T" is for T-wave Inversion," 2022.
- [15] U. F. & D. Administration, "Ultrasound Imaging," 2020.
- [16] J. H. Holmes, "Diagnostic ultrasound during the early years of A.I.U.M.," *Journal of Clinical Ultrasound*, pp. 299-308, 1980.
- [17] A. A. Mohamed, A. A. Arifi and A. Omran, "The basics of echocardiography," *Journal of the Saudi Heart Association*, vol. 22, no. 2, pp. 71-76, 2010.
- [18] M. D. Cerqueira, N. J. Weissman, V. Dilsizian, A. K. Jacobs, S. Kaul, W. K. Laskey, D. J. Pennell, J. A. Rumberger, T. Ryan and M. S. Verani, "Standardized Myocardial Segmentation and Nomenclature for Tomographic Imaging of the Heart," *Circulation*, vol. 105, no. 4, p. 539–542, 2002.
- [19] A. Shah, "Equipment and Knobology," 2022.
- [20] A. Koratala, "Focused Cardiac Ultrasound for the Nephrologist: The apical window," *Renal Fellow Network*, 2019.
- [21] J. L. Díaz and G. Via, "Focused cardiac and lung ultrasonography: implications and," *Romanian Journal of Anaesthesia and Intensive Care*, vol. 23, no. 1, pp. 41-54, 2016.
- [22] M. Zidan, "Apical four chamber view".
- [23] S. Kiranyaz, A. Degerli, T. Hamid and R. Mazhar, "Left Ventricular Wall Motion Estimation by Active Polynomials for Acute Myocardial Infarction Detection," *IEEE Access*, vol. 8, pp. 210301-210317, 2020.
- [24] X. Wu, X. Li, G. Mou, D. Wang, Y. He and Z. Li, "Identification of the left ventricle endocardial border on two-dimensional ultrasound images using deep layer aggregation for residual dense networks," *Applied Intelligence*, vol. 52, p. 16089–16103, 2022.

Bibliography

- [25] S. Deopujari, Y. Dubey and M. Mushrif, "Left Ventricular segmentation of 2-D echocardiographic image : A Survey," *International Conference on Communication Systems and Network Technologies (CSNT)*, 2011.
- [26] D. Carroll and J. Feger, "Wall motion score index (echocardiography)," *Radiopaedia.org*, 2021.
- [27] F. Johnson, "WHAT IS THE MEANING OF HYPOKINESIA, DYSKINESIA AND AKINESIA?," 2017.
- [28] A. Kosaraju, A. Goyal, Y. Grigorova and A. N. Makaryus, "Left Ventricular Ejection Fraction," *National Library of Medicine*, 2022.
- [29] R. M. Lang, L. P. Badano, V. Mor-Avi, A. Armstrong and L. Ernande, "Recommendations for Cardiac Chamber Quantification by Echocardiography in Adults: An Update from the American Society of Echocardiography and the European Association of Cardiovascular Imaging," *Journal of the American Society of Echocardiography*, vol. 28, no. 1, pp. 1-39, 2015.
- [30] M. Cameli, G. E. Mandoli, C. Sciacaluga and S. Mondillo, "More than 10 years of speckle tracking echocardiography: Still a novel technique or a definite tool for clinical practice?," *Echocardiography*, vol. 36, no. 5, pp. 958-970, 2019.
- [31] V. Morariu, D. Arnautu, S. Morariu, A. Popa, T. Parvanescu, M. Andor, S. Abhinav, V. David, A. Ionescu and M. Tomescu, "2D speckle tracking: a diagnostic and prognostic tool of paramount importance," *European Review for Medical and Pharmacological Sciences*, vol. 26, no. 11, pp. 3903-3910, 2022.
- [32] A. Degerli, M. Zabihi, S. Kiranyaz, T. Hamid, R. Mazhar, R. Hamila and M. Gabbouj, "Early Detection of Myocardial Infarction in Low-Quality Echocardiography," *IEEE Access*, vol. 9, pp. 34442-34453, 2021.
- [33] O. A. Smiseth, H. Torp, A. Opdahl, K. H. Haugaa and S. Urheim, "Myocardial strain imaging: how useful is it in clinical decision making?," *European Heart Journal*, vol. 37, no. 15, pp. 1196-1207, 2016.
- [34] Philips, "AutoStrain LV/RV/LA automated strain measurements," *Philips*.

Bibliography

- [35] S. Brown, "Machine learning, explained," *MIT*, 2021.
- [36] IBM, "Cos'è la Computer Vision?," 2022.
- [37] A. Kaplan and M. Haenlein, "Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence," *Business Horizons*, vol. 62, pp. 15-25, 2019.
- [38] Mathworks, "Deep Learning: 3 cose da sapere," 2022.
- [39] IBM, "What is supervised learning?," 2022.
- [40] M. KASS, A. WITKIN and D. TERZOPOULOS, "Snakes: Active Contour Models," *International Journal of Computer Vision*, pp. 321-331, 1988.
- [41] D.-J. Kroon, "Active Shape Model (ASM) and Active Appearance Model (AAM)," *MathWorks*, 2023.
- [42] Y. Ali, S. Beheshti and F. Janabi-Sharifi, "Echocardiogram segmentation using active shape model and mean squared eigenvalue error," *Biomedical Signal Processing and Control*, vol. 69, 2021.
- [43] H. Bi, Y. Jiang, H. Tang, G. Yang, H. Shu and J.-L. Dillenseger, "Fast and accurate segmentation method of active shape model with Rayleigh mixture model clustering for prostate ultrasound images," *Computer Methods and Programs in Biomedicine*, vol. 184, 2020.
- [44] S. Leclerc, E. Smistad, J. Pedrosa, A. Ostvik, F. Cervenansky, F. Espinosa, T. Espeland, E. J. Berg, P.-M. Jodoin, T. Grenier, C. Lartizien and J. Drhooge, "Deep Learning for Segmentation using an Open Large-Scale Dataset in 2D Echocardiography," *IEEE Transactions on Medical Imaging*, vol. 38, no. 9, pp. 2198-2210, 2019.
- [45] H. Mittal, A. C. Pandey, M. Saraswat, S. Kumar, R. Pal and G. Modwel, "A comprehensive survey of image segmentation: clustering methods, performance parameters, and benchmark datasets," *Multimedia Tools and Applications*, vol. 81, p. 35001–35026, 2022.

Bibliography

- [46] H. Seo, M. B. Khuzani, V. Vasudevan, C. Huang, H. Ren, R. Xiao, X. Jia and L. Xing, "Machine Learning Techniques for Biomedical Image Segmentation: An Overview of Technical Aspects and Introduction to State-of-Art Applications," *Med. Phys.*, vol. 47, no. 5, pp. 148-167, 2020.
- [47] P. Sabia, A. Afrookteh, D. A. Touchstone, M. W. Keller, L. Esquivel and S. Kaul, "Value of regional wall motion abnormality in the emergency room diagnosis of acute myocardial infarction. A prospective study using two-dimensional echocardiography," *Circulation*, vol. 84, no. 3, pp. 185-192, 1991.
- [48] I. Rácz, L. Fülöp, R. Kolozsvári, G. T. Szabó, A. Bódi, A. Péter, A. Kertész, I. Hegedüs, I. Édes, L. Balkay and Z. Köszezi, "Wall motion changes in myocardial infarction in relation to the time elapsed from symptoms until revascularization," *The Anatolian Journal of Cardiology*, vol. 15, no. 5, pp. 363-370, 2015.
- [49] L. Antonelli, M. Katz, F. Bacal, M. R. Makdisse, A. G. Correa, C. Pereira, M. Franken, A. N. Fava, C. V. Serrano Junior and A. E. Pesaro, "Heart Failure with Preserved Left Ventricular Ejection Fraction in Patients with Acute Myocardial Infarction," *Arquivos Brasileiros de Cardiologia*, vol. 105, no. 2, p. 145–150, 2015.
- [50] J. W. Cunningham, M. Vaduganathan, B. L. Claggett, J. E. John and A. S. e. a. Desai, "Myocardial Infarction in Heart Failure With Preserved Ejection Fraction: Pooled Analysis of 3 Clinical Trials," *JACC: Heart Failure*, vol. 8, no. 8, pp. 618-626, 2020.
- [51] S. Suma, "WHAT IS GLOBAL LONGITUDINAL STRAIN AND ITS CLINICAL IMPORTANCE?," 2021.
- [52] L. M. Eber, H. M. Greenberg and J. M. Cooke, "Dynamic Changes in Left Ventricular Free Wall Thickness in the Human Heart," *Circulation*, vol. 39, no. 4, pp. 455-464, 1969.
- [53] PulpLearning, "Train, Validation, Test: cosa sono e come si usano nel Machine Learning," 2020.

Bibliography

- [54] C. Chen, C. Qin, H. Qiu, G. Tarroni, J. Duan, W. Bai and D. Rueckert, "Deep learning for cardiac image segmentation: A review," *Frontiers in Cardiovascular Medicine*, vol. 7, no. 25, 2020.
- [55] A. Coste, *University of Utah: Image Processing Report*, 2022.
- [56] Scikit-Image, "Exposure module," 2022.
- [57] OpenCV, "Histogram Equalization," 2022.
- [58] T. Point, "Histogram Equalization," 2022.
- [59] S. Indolia, A. K. Goswami and P. Asopa, "Conceptual Understanding of Convolutional Neural Network- A Deep Learning Approach," *Procedia Computer Science*, vol. 132, pp. 679-688, 2018.
- [60] D. Breach, "Convolutional Neural Network - Deep Learning," 2022.
- [61] IBM, "Convolutional Neural Networks," 2022.
- [62] A. Amidi and S. Amidi, "Convolutional Neural Nets," *Stanford Vision Lab*, 2021.
- [63] C. Banerjee, T. Mukherjee and E. Pasilio, "An Empirical Study on Generalizations of the ReLU Activation Function," *Association for Computing Machinery*, pp. 164-167, 2019.
- [64] IndoML, *Indian Symposium on Machine Learning*, 2022.
- [65] O. Ronneberger, P. Fischer and T. Brox, "U-Net: Convolutional Networks for Biomedical," *Medical Image Computing and Computer-Assisted Intervention*, p. 234–241, 2015.
- [66] J. Long, E. Shelhamer and T. Darrell, "Fully Convolutional Networks for Semantic Segmentation," *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 3431-3440, 2015.
- [67] OpenCV, "Morphological Transformations," 2022.
- [68] P. Bankhead, "Morphological operations," *Bioimage Analysis Book*, 2022.

Bibliography

- [69] Scikit-Image, "Skeletonize," 2022.
- [70] T.-C. Lee, R. L. Kashyap and C.-N. Chu, "Building skeleton models via 3-D medial surface/axis thinning algorithms," *CVGIP: Graphical Models and Image Processing*, vol. 56, no. 6, p. 462–478, 1994.
- [71] T. Key, "Stress Echocardiography and Echo in Cardiopulmonary Testing," 2018.
- [72] L. H. Lund, B. Pitt and M. Metra, "Left ventricular ejection fraction as the primary heart failure phenotyping parameter," *European Journal of Heart Failure*, vol. 24, no. 7, pp. 1158-1161, 2022.
- [73] G. Joseph, T. Zaremba and M. B. e. a. Johansen, "Echocardiographic global longitudinal strain is associated with infarct size assessed by cardiac magnetic resonance in acute myocardial infarction," *Echo Research and Practice*, vol. 6, no. 4, pp. 81-89, 2019.
- [74] J. M. Cwajg, E. N. S. F. Cwajg, Z.-X. He, U. Qureshi and e. al, "End-diastolic wall thickness as a predictor of recovery of function in myocardial hibernation: Relation to rest-redistribution Tl-201 tomography and dobutamine stress echocardiography," *Journal of the American College of Cardiology*, vol. 35, no. 5, pp. 1152-1161, 2000.
- [75] IBM, "K-Nearest Neighbors Algorithm," 2022.
- [76] B. Mirbozorgi, "Theory of K-Nearest Neighbors (KNN)," *Analytics Vidhya*, 2020.
- [77] MathWorks, "Support Vector Machine (SVM)," 2022.
- [78] I. Steinwart and A. Christmann, *Support Vector Machines*, Springer, 2008.
- [79] J. T. Point, "Support Vector Machine Algorithm," 2022.
- [80] IBM, "What is random forest?," 2022.
- [81] H. Ampadu, "Random Forest Understanding," 2021.
- [82] R. Shen, S. Bubeck and S. Gunasekar, "Data Augmentation as Feature Manipulation," *International Conference of Machine Learning*, 2022.

Bibliography

- [83] T. Srivastava, "11 Important Model Evaluation Metrics for Machine Learning Everyone should know," *Analytics Vidhya*, 2023.
- [84] S. Narkhede, "Understanding AUC - ROC Curve," 2018.
- [85] M. Tan and Q. V. Le, "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks," *International Conference on Machine Learning*, 2019.