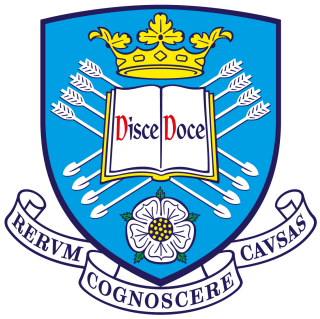

The University Of Sheffield.

University of Sheffield

Department of Computer Science

PhD Thesis

SOCRATIS: Segmentation Of Cardiac Regions And Total Infarct Scar

Author:
Michail Mamalakis

Supervisor:
Prof. Richard H. Clayton

January 25, 2022

SOCRATIS: Segmentation Of Cardiac Regions And Total Infarct Scar

Thesis presented
by
Michail Mamalakis
to
The Department of Computer Science

in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy
in the subject of
Cardiac Image Analysis

University of Sheffield
Sheffield, United Kingdom

January 2022

This PhD thesis is dedicated to my supervisor Prof. Richard H. Clayton; for his help and his support to me. To my family. To my brother Dr Antonios Mamalakis. He always was my motivation and inspiration to be a better researcher and person. To my mother Annita Pitsidinaki, who always let me sing to her and play guitar from virtual call. To my father Theodosios Mamalakis, the SINGER. The man who start the legacy of singers in Mamalakis family tree. To all my friends were help me with their company, the night walks and the barbeques.

In the end of this journey, I just want to remind to myself:

'Whatever you feel, it can be true or lie, correct or wrong. We are facing highly uncertain events in life. In the same time the statistical and individual view of things are existed. Live your life from your perspective and learn how to avoid drama and non-productive situations by using empathy. The universe is another view of your understanding. Learn good, fight hard, decide effectively.'

SOCRATIS: Segmentation Of Cardiac Regions And Total Infarct Scar

Abstract

Patients with myocardial infarction are at elevated risk of sudden cardiac death, and scar tissue arising from infarction is known to play a role. The accurate identification of scars therefore is crucial for risk assessment, quantification and guiding interventions. Typically, core scars and grey peripheral zones are identified by radiologists and clinicians based on cardiac late gadolinium enhancement magnetic resonance images (LGE-MRI). Manual segmentation of scar from LGE-MRI varies in size, shape, heterogeneity, artifacts, and image resolution. The process is time consuming, and influenced by the observer's experience (bias effect). Here we propose a framework that delivers an automatic segmentation pipeline to develop 3D anatomical models of the left ventricle with border zone and core scar areas that are free from bias effect. We deliver the SOCRATIS framework for automatic segmentation of cardiac left ventricle and total infarct scars. SOCRATIS involves two main pipelines; i) deep learning segmentation of cardiac regions and total infarct scar (DL-SOCRATIS) and ii) multi-atlas segmentation of cardiac regions and total infarct scar (MA-SOCRATIS). In this PhD thesis we discovered that MA-SOCRATIS was a superior technique providing high accuracy of the segmentation. However, DL-SOCRATIS was faster (less than 10 min per patient) and it does not require expert supervision, contrary to the MA-SOCRATIS. Since both pipelines are included in the SOCRATIS framework, it is possible to configure the segmentation so that duration of the analysis is traded off against high accuracy of the segmentation. SOCRATIS utilizes LGE-MR images, without the need of further multi-modal supervision (bSSFP, T2, LGE images) or training/tuning in other cohorts, contrary to the majority of the state of the art techniques. About the scar segmentation we utilized a border zone and core scar segmentation of left ventricle and total infarct scar (BZ-SOCRATIS). We tested two different segmentation AI approaches; i) an automatic and ii) semi-automatic, to segment the core scar and border zone of unhealthy tissue from LGE-MRI of the left ventricle. Our framework delivers a fully automatic segmentation pipeline to extract 3D specific patient anatomical models of left ventricle, border zone and core scar region, which can be used by clinicians to determine treatment and evaluation of the catheter radiofrequency ablation procedure.

Contents

Listing of figures

Table 1: TABLE OF ACRONYMs

AAM	Active Appearance Model
AV	Atrio-Ventricular
AI	Artificial Intelligence
b-SSFP	balanced Steady State-Free Precession
BZ	Border Zone
BZ-SOCRATIS	Border-Zone Segmentation Of Cardiac Regions And Total Infarct Scar
BZ-TH	Border Zone Threshold
CGAN-SOCRATIS	CycleGAN Segmentation Of Cardiac Regions And Total Infarct Scar
cine-MRI	Cine-Magnetic Resonance Imaging
cine-SOCRATIS	Cine Segmentation Of Cardiac Regions And Total Infarct Scar
CMR	Cardiac Magnetic Resonance
CNN	Convolution Neural Network
CNR	Carrier-to-Noise Ratio
CRF	Cinditional Random Field
CS	Core Scar
CT	Computerised Tomography
DCE	Dice Cross Entropy
DL	Deep Learning
DL-SOCRATIS	Deep-Learning Segmentation Of Cardiac Regions And Total Infarct Scar
DLc-SOCRATIS	Deep-Learning-cine Segmentation Of Cardiac Regions And Total Infarct Scar
DLl-SOCRATIS	Deep-Learning-LGE Segmentation Of Cardiac Regions And Total Infarct Scar
DICOM	Digital Imaging and Communications in Medicine
DNA	DeoxyriboNucleic Acid
ECM	ExtraCellular Matrix
ED	End-Diastolic
EM	Expectation Maximization
EP	ElectroPhysiology
ES	End-Systolic
FWHM	Full Width Half Maximum
GAN	Generative Adversarial Network
Gd	Gadolinium
GMM	Gaussian Mixture Model
GP	Gaussian Process
GT	Ground Truth
IE	Implicit Exploration
KNN	K-Nearest Neighboor
LA	Left Atrium
LGE-MRI	Late Gadolinium Enhancement Magnetic Resonance Imaging
LV	Left Ventricle

Table 2: TABLE OF ACRONYMs

MAP	Maximum A Posteriori estimation
MA-SOCRATIS	Multi-Atlas Segmentation Of Cardiac Regions And Total Infarct Scar
ML	Machine Learning
MLP	Multi Layer Perception
MRF	Markov Random Field
MRI	Magnetic Resonance Imaging
NN	Neaural Network
ReLu	Rectified Linear Unit
RFCA	Radio-Frequency Catheter Ablation
RGMM	Rician Gaussian Mixture Model
ROI	Region Of Interest
SA	Sino-Atrial
SKIZ	Skeleton by Image Zones
SNR	Signal-to-Noise Ratio
SOM	Self Organization Maps
SS-DCE	Sensitivity Specificity Dice-weighted Cross Entropy
STACS	Stochastic Active Contour
STD	Standard Deviation
SVM	Support Vector Machine
TG	Threshold Gaussian
TR	Threshold Rician
VA	Ventricular Arrhythmia
VT	Ventricular Tachycardia

1

Introduction

1.1 Summary

Abnormalities of electrical activation in the heart result in arrhythmia. Sudden death associated with cardiac arrhythmias remains a significant global health problem ([208]). Patients with regions of ventricular fibrosis or scar, often as a result of prior myocardial infarction, are at elevated risk of ventricular arrhythmias ([121]). Accurate identification of scar and its border zone (gray zone) is

therefore important for risk stratification and for guiding interventions ([147]).

A range of modalities are used in the clinical setting for imaging anatomy, and they are used as a baseline to determine treatment decisions and diagnosis. Ascha et al. ([11]) compare and contrast imaging modalities for mapping cardiac anatomy and function, including X-rays, computed tomography (CT), ultrasound, and magnetic resonance imaging (MRI). Chest X-ray is the baseline modality for patients with pulmonary hypertension. CT is a method characterized by high temporal resolution, spatial resolution, and speed of image acquisition. Although CT images enable diagnosis, they are associated with an X-ray dose. In contrast, MRI is less harmful to the health of the patient, however it is expensive and needs an expert operator. The most common protocol for cardiac imaging is cine-MRI, in which a series of images are taken throughout the cardiac cycle. The scan orientation is usually the short axis of the heart, with 8-mm or 2-mm slice intervals.

Late gadolinium enhancement (LGE-MRI) is a technique in which using gadolinium (Gd) increases the contrast visibility of MRI, and diethylentriaminopentacetic acid defines the in vivo behaviour of the whole heart. As a result, the imaging of cardiac chambers relative to the blood pool and other structures is improved. Ordovas et al. ([141]) demonstrated the utility of LGE-MRI to increase the difference in contrast between healthy and diseased myocardium. As a result, LGE-MRI is widely used in clinical scanning to detect myocardial scars. Moreover, using LGE-MRI, a gray zone of partially viable tissue can also be detected ([160]). The gray zone is a complex of necrotic (white intensity) and healthy (black intensity) zones which in the last years has been used to aid MRI-guided treatment for ventricular arrhythmia ([125]).

LGE-MRI automatic segmentation: Cardiac MRI with late gadolinium enhancement (LGE) is routinely used for imaging regions of scar in patients. However, scar segmentation is a challenging topic as there is great variability in the size, heterogeneity, intensity of scar regions in LGE images, low resolution and image noise artifacts ([175]).

The current standard practice is manual segmentation, which relies on sketching contours slice

by slice using pointing devices such as a mouse or trackball. Three dimensional MRI can include as many as one hundred slices. Thus, manual segmentation is time consuming, tedious, and limited by inter-observer and intra-observation variability ([200,181]). Automated segmentation can overcome these limitations, and methods have been developed for automated scar segmentation in atrial ([137,46,90,87]) and ventricular ([4,97,50,116]) images.

These automated segmentation techniques are typically pixel-to-pixel techniques such as support vector machines, c-means, and other machine learning techniques. These methods performance is highly related with the signal to noise ratio (SNR) of the image ([184]). As a result, these algorithms tend to suffer from low accuracy. Other approaches are model-based methods like atlas, multi-atlas, and deformable models ([13]). However, these approaches usually involves a large dataset to capture the geometrical shape distribution of the heart with sufficient generalisation.

To sum up, an automatic segmentation pipeline is a way to identify both the myocardium and regions of scar in LGE images of the left ventricle. The main focus of this PhD project is the automatic segmentation of cardiac chambers and more specifically of the left ventricle (LV) with scars based on magnetic resonance images (CMRI). We are interested to verify if it is possible to segment the regions of healthy and unhealthy myocardium (scar) of the LV robust and quickly by a fully automated segmentation framework without expert radiology supervision.

In the next sections, we will present the motivation and aim of our study, the contribution of our study in the medical image analysis field, and the future work of this study. Finally, we will deliver an overview of the PhD thesis.

1.2 Motivation and Aim

A normal heart beat arises from spontaneous electrical activity in the heart's natural pacemaker located in the right atrium. Electrical activity spreads through the atria, and then through the cardiac

conduction system into the ventricles. This electrical activation acts to initiate mechanical contraction. Abnormalities of the excitation sequence are arrhythmias which can be deadly. Arrhythmias are defined based on their location, atrial fibrillation, ventricular fibrillation, supra-ventricular tachycardia. Ventricular arrhythmia (VA) is an important cause of mortality and sudden cardiac death (75-80% of cases)([140], [117], [208]).

Ventricular scar is a main substrate for reentry arrhythmias and can sustain VA ([22]). Scars are created in myocardial tissue, when there is inadequate supply of blood to cardiac muscle. The slow conduction of electrical activation inside and around the scars generates reentry. Re-entrant activation usually exits sites along the border zone of scar. Thus, the spatial distribution of myocardial scars for the accurate treatment of VA is a major importance ([22], [121]). Ventricular tachycardia (VT) can be treated with radiofrequency catheter ablation surgery (RFCA) ([2], [166], [26]). The main reason is the large size of possible re-entrant circuits plus the fact that scars in the endocardium can be deep and sometimes difficult to detect. There is evidence that the size, shape, and location of the scar can be useful to guide ventricular tachycardia ablation ([88], [54], [7]). Thus, valid and accurate mapping of the scars is crucial for correct guidance and treatment of VT in catheter ablation (CA). Moreover, the anatomical mesh based on realistic mapping of the scar is an import contribution about the study of cardiac arrhythmia mechanisms ([147]).

Segmentation of scars from images of the left and right ventricle is a challenging topic (Figure 1.1), as the large variability of internal (e.g., size heterogeneity of scars, spatial distribution, intensity distribution) and external (e.g. resolution, noise) situations, cannot be fitted by a simple model ([175]). Manual segmentation of the scar based on threshold definition is a typical technique. Manual segmentation of the scar relies on sketching contours slice by slice using pointing devices such as a mouse or trackball. Three dimensional MRI can include as many as one hundred slices. Thus, manual segmentation is time consuming, tedious, and limited by inter-observer and intra-observation variability. Moreover, many algorithms used in manual segmentation depend on initial-

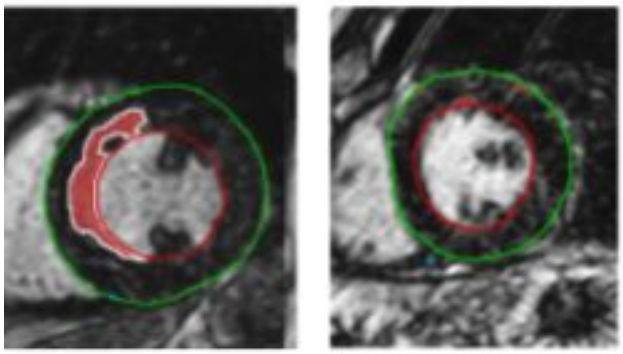

Figure 1.1: An example of MRI with the scar and myocardium segmented. Green color the epicardial contour, red the endocardial contour and red area the scar regions.

ization. When only one expert is involved, the segmentation can be biased. Moreover, manual segmentation affects the reproducibility of the results ([200], [181]). Automatic segmentation can overcome the limitations of manual methods. There are many studies based on automatic scar segmentation in atrial regions ([137], [46], [90]), [87]) and others in ventricular regions ([4], [97], [50], [116]).

Typical automated segmentation techniques are classification techniques such as support vector machines, c-mean, and other machine learning techniques. The main drawback of these methods is that the results are highly dependent on signal to noise ratio image ([184]). As a result, these algorithms suffer from low accuracy and robustness ([90]). Other methods are model-based methods like atlas models, which need a large dataset to capture the geometrical shape distribution of heart for sufficient generalization ([13]).

1.3 Contribution

During the initial part of this project it became apparent that there is a need to increase the certainty of scar detection in order to address the question: **"Is expert supervision crucial?"**. The need for expert supervision could be considered compulsory, as in the majority of AI methods the user in-

teraction is needed. The observation and segmentation of scar regions from LGE-MRI images can be easier than in cine-MRI because scar tissue is better highlighted in LGE-MRI images ([80]). However, it can be difficult even for experts to conclude if a region is scar or not, from LGE-MRI images alone. Thus, if the LGE-MRI modality is enough to justify the existence of a scar in LV, it will better for an automated method or even the experts if information about possible scar location is extracted from more than one observation, as an individual observation has a bias error. For this PhD project, we would like to test the following ideas:

a. Can an automatic segmentation pipeline deliver accurate segmentation of the left ventricle with scars from LGE-MRI images?

b. Can the implemented pipeline work well in different patient cohorts? The reason for this is that the diversity of MRI scanning machines, the variability of cardiac shape and size, and the external-internal conditions of MRI scanning procedure can affect the accuracy of AI.

c. Is expert supervision crucial for the accurate detection of scar and myocardium of left ventricle?

To this end, we implemented an automatic segmentation pipeline of the myocardium and scar regions of LV in LGE-MR images.

Myocardium segmentation involves two steps, initial and re-estimate segmentation, the initial step includes a deep learning structure ([13]) to detect the region of interest (ROI), and a multi-atlas to initial segment the ROI ([16], [102]). The re-estimate step combines k-means thresholding and level-set active contour segmentation method ([136], [105]). For the scar automatic segmentation, we use an algorithm that combines mixture model histogram fitting and constrained watershed segmentation by an automatic seed point scar region technique. Our pipeline refers to ([72]) with amendments in the watershed segmentation technique. A limitation was the use of Rayleign-Gaussian mixture models, healthy myocardium described as Rician distribution ([175], [164]). Our approach involves the

use of Rician-Gaussian mixture models.

The novel contribution of this thesis to cardiac automatic segmentation topic is:

- MA-SOCRATIS machine learning pipeline, which is:

  i. An unsupervised automatic segmentation pipeline, which does not need training or tuning (MA-SOCRATIS).

  ii. A comparison of the efficiency of deep learning approaches compared to probabilistic multi-atlases.

- DL-SOCRATIS deep learning pipeline, which is:

  i. A fully automatic segmentation pipeline using deep learning established techniques.

  ii. A sensitivity and specificity dice-weighted cross-entropy (SS-DCE) cost function to increase the dice accuracy of endocardial and epicardial automatic segmentation in a cross-vendor validation (85.8 to 89.8 % for the epicardial regions).

  iii. A less accurate but faster and no need of expert's supervision results compared with MA-SOCRATIS.

  iv. A flexible enough pipeline to be used as a quick initial evaluation of scar and myocardial regions in cross-modality analysis tasks; such as when expert radiologists have different or constrained modality scan images for a specific patient (cine-MR, LGE-MR, or a mixed of cine and LGE MR images).

- BZ-SOCRATIS scar segmentation pipeline, which is:

  i. A thorough study regarding automatic and semi-automatic segmentation of both border zone and core scar regions in LV with infarct LGE-MR images.

ii. An unsupervised automatic segmentation pipeline (combination of Rician-Gaussian mixture model and watershed techniques), which does not need training or tuning for fully automatic scar region segmentation.

iii. A highlight of the importance of correct detection of myocardium boundaries in LGE-MRI in patients with infarct tissue.

iv. A 3D geometric representation/mapping of the border zone, core scar and LV fully automatically.

- Robust and unbiased method for evaluation :

  i. By unbiased ground truth for the evaluation (intra-observation and inter-observer variability) of the automatic segmentation results.

  ii. By cross-institutional evaluation (our LGE dataset and MS-CMRSeg 2019 dataset) of the automatic segmentation results.

  iii. By gathering a challenging multidomain MRI dataset of 100 patients (LGE-MR, cine-MR) with manual segmentation of the left ventricle and scars.

1.4 Future work

As a future work, we believe that i) a more thoroughly study in transfer learning techniques, and ii) an additional training in cross-vendor and institute cohorts; they can increase the accuracy of DL-SOCRATIS pipeline. Finally, cross-institute validation of BZ-SOCRATIS pipeline to generalize the use and robustness of pipeline would be useful. In addition, to the imaging modality and increase robustness of the results, we suggest the implementation of an artificial intelligence - motion analysis of cardiac region and total infarct scar (AI-MACT) to detect the scar position and portion based on the deformation and strain of ES and ED phases of cine-MRI images.

1.5 Basic terminology

For this thesis, when we mention the word **unsupervised** methods, we reference at methods evaluated in an unseen cohort without training, tuning or using of transfer learning techniques.

1.6 Publications

During the three years of this PhD project, we published the following studies:

- MA-SOCRATIS: An automatic pipeline for robust segmentation of the left ventricle and scar (published Computerized Medical Imaging and Graphics journal).
- Traditional computer vision and deep learning approaches for optimal automatic segmentation of left ventricle with scar (submitted in journal).
- 3D anatomical models to assess cardiac arrhythmia risk: from manual segmentation to a new era of automation (submitted in journal).

1.7 Overview of thesis

The focus of this PhD project is the automated analysis of cardiac magnetic resonance images (CMR) and more specifically of the left ventricle (LV) with scars. Moreover, we would like to establish if it is possible to detect the regions of scar in the LV, using different techniques (segmentation, image analysis) and different MRI modalities (LGE-MRI, cine-MRI). Lastly, we want to test if a combination of these techniques could deliver full automated detection of the scar in a patient.

In the implementation part, we developed two software pipelines: multi-atlas segmentation of cardiac region and total infarct scar (MA-SOCRATIS) and deep learning segmentation of cardiac region and total infarct scar (DL-SOCRATIS). The MA-SOCRATIS is an unsupervised automatic

segmentation pipeline of LV with scar in LGE-MRI images. Moreover, we modified the pipeline with some deep learning parts (DL-SOCRATIS); and we compared the two different frameworks. As a final step, we used the scar segmentation pipeline which we implemented for two tasks. First, to the segmented core of the scar region and second extracted the gray zone and the core of scar tissue of LV (BZ-SOCRATIS).

This thesis is organized as follows. Chapter 2 describes the motivation and background of this study. Chapter 3 describes the related studies and literature review in medical imaging analysis and segmentation. Chapter 4 and 5 describes our research regarding automatic segmentation of the myocardium and scar of left ventricle (LV) based LGE-MRI images. Chapter 6 describes the evaluation of the results based on clinical data. Chapter 7 describes the contribution, limitations, and future work of this study. Chapter 8 presents the conclusions of this PhD thesis.

2

Background

2.1 Background overview

The aim of this PhD thesis is the automated segmentation of the LV with scars. We use techniques of machine learning and deep learning to detect the myocardium and scar regions of the left ventricle. A combination of these techniques could deliver automated detection of scar in a patient, without the need of an expert's interaction.

This chapter is organized as follows. Section 2.2 describes the cardiac anatomy background. Section 2.3 describes the kind of cardiac arrhythmias and specifically the ventricle (VA). Section 2.4 presents different imaging modalities techniques. Section 2.5 magnetic describes resonance about physics and scanning. Lastly, Section 2.6 summarizes the machine learning and deep learning techniques in medical applications.

2.2 Cardiac physiology

2.2.1 Cardiac anatomy

The cardiovascular system consists of the heart, circulation loops, blood vessels, coronary circulation, and blood. The role of the cardiovascular system is to transport nutrients, hormones, and oxygen. The circulation loops are classified as pulmonary (heart to lungs) and systemic (heart to all tissues of body). Blood vessels consist of arteries and arterioles, capillaries, veins, and venules. The heart is a muscular pumping organ which consists of the left-right atrium, ventricles, valves, aorta, superior cava, pulmonary artery, vein, septum, and inferior cava (figure 2.1), it is electrically excitable, and specialized conduction system. The main parts of the system are the sino-atrial (SA) node, Bachman's bundle, atrio-ventricular (AV) node, the His bundle, Tawara branches and Purkinje fibres. The SA is the main pacemaker of the heart which creates and propagates electrical waves in all conduction systems. The propagation of the electrical wave in the AV is slow, thus a delay of the excitation of the ventricle by the Purkinje fibres can be observed ([182]). This pattern is the excitation of a normal heart beat. Abnormal formation of the electrical impulse is an arrhythmia.

At the microstructure level, the heart is composed of a large number of cells with different functionalities and properties are existed. The most common cardiac cells are myocytes and fibroblasts. Myocytes are located mainly in heart muscle and generate mechanical tension. Fibroblasts from the largest volume fraction of the myocardium, and have passive mechanical properties, which play an

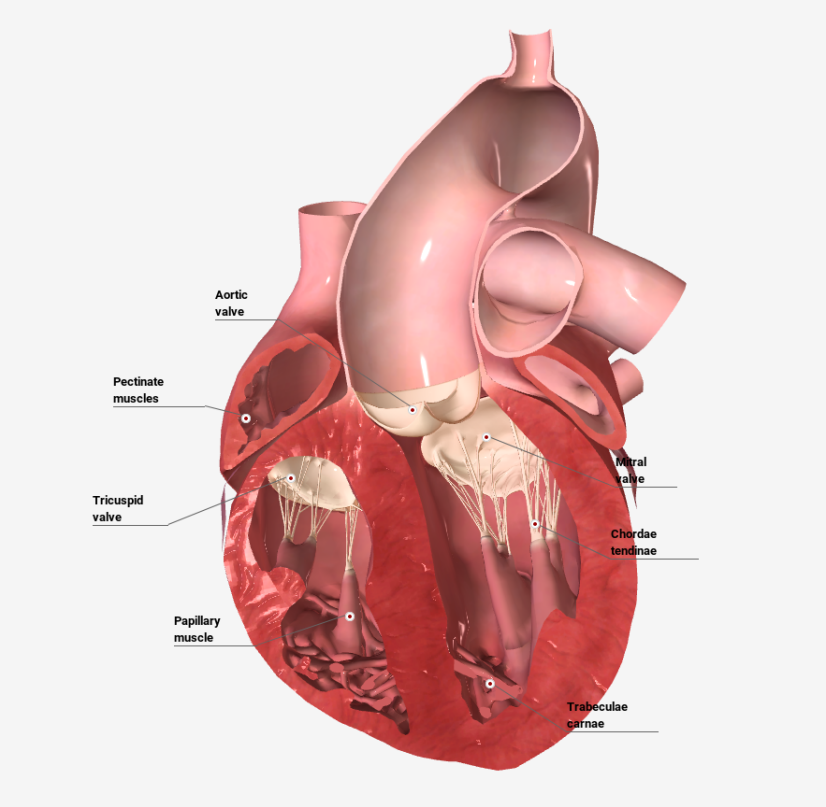

Figure 2.1: Anatomy of heart (reproduce with permission of ([23])

important role in the maintenance of the extracellular matrix (ECM). Other cell types are endothelial, neural cells and vascular smooth muscle cells ([39]).

There are many diseases about heart function. One of the major causes of sudden cardiac death is acute myocardial ischemia, in which narrowing of one or more coronary arteries results in reduced blood flow to the myocardium, leading to a decrease in oxygen and nutrients in regions of the myocardium. Moreover, myocardial infarction occurs when blood flow is obstructed, and this may result in myocardial scar.

2.2.2 Cardiac arrhythmias

Excitation continuously propagates into the recovering tissue. Re-entry is favoured by regions of slow conduction, slow conduction occurs in regions of the scar, and so reentry is often associated with scar ([8]). Re-entry is a mechanisms that sustains an arrhythmia, where excitation continuously propagates into recovering tissue. Regarding anatomical factors, obstacles like scars or infarct circus movement reentry around them. Mines et al. ([128]) established that the interruption of reentry at any point may terminate the process of circus movement and as an extension can stop an arrhythmia. This is the general idea of catheter ablation surgery.

Ventricular arrhythmias are one of the most important causes of mortality and sudden cardiac death (75-80% of cases) ([140], [117], [208]). More specifically, ventricular tachycardia (VT) is one of the main risks of sudden death in structural heart diseases ([98]). In 2003, Monserrat et al. ([132]) showed that VT is connected with sudden death risk in young patients with hypertrophic cardiomyopathy (HCM). In 2016, there were 50-100 sudden cardiac deaths (SCDs) per 100.000 people per year in Europe and USA due to VA ([70]). In 2017, the cardiovascular statistics based on Alkatib et al.([2]) estimate that there were many out-of-hospital cardiac arrest incidents (357.000 in adults over 18 years of age) due to VT. The survival rate was only 10 % (6% survival occurring in the home).

2.3 Scars and border zone regions

Sudden death associated with cardiac arrhythmias remains a significant global health problem. Patients with regions of ventricular fibrosis or scar are at elevated risk of ventricular arrhythmias ([180,121]), and accurate identification of the scar is important for both risk stratification and guiding interventions ([147]). Cardiac MRI with late gadolinium enhancement (LGE) is routinely used for imaging scars in patients ([102]). During the last decade, the radiology society explored the correlation of the border zone region (gray intensity pixels around the core scar) of the cardial myocardium with

cardiac arrhythmias, to predict and evaluate the existent strategies of treatment ([131,194]).

The border zone (BZ) tissue, surrounding a core infarct scar, determines a remodeling tissue with significant variation in the electrophysiological (EP) properties compared with the healthy myocardium. Two different remodeling phenomena occur, one associated with active remodeling (ion channels across the cell membrane) and another with passive remodeling resulting from structural changes (re-organization of gap junction distribution) ([41]). As a result the EP and structural features of the myocardial scar region (border zone and the core infarct) can contribute to the risk of VT in patients. However, this statement is still poorly understood and debated. The extraction of robust and accurate two-and three-dimensional computational models of different core scar regions and border zones can enhance the assist to develop improved methodologies. Thus, it will be possible the detection and differentiation of arrhythmogenic and nonarrhythmogenic scars. The initial step of these computational models is the segmentation of the border zone, core scar and myocardial regions.

2.4 The role of core scar and border zone in VT

By using LGE-MRI, expert radiologists can identify peri-infarct zones located between normal and infarcted tissue, surrounding the core scar. Border zones appear as regions of intermediate intensity on LGE CMR, thus they are referred to as gray zones. Histological studies have verified that intermediate intensity regions usually represent a mixture of viable myocardium and scar ([9]). It has been shown that the extent of the heterogeneous zone is associated with spontaneous and inducible VT, which can lead to mortality after myocardial infarction ([161,196]). The post-infarction VT conducting channels were associated with heterogeneous zones, which can be detected by LGE CMR images ([148]). Even if the experts can justify that re-entry is sustained by the region around the myocardial infarction, identification of the specific region (border zone or core scar) that initiates and

sustains re-entry in VT is a challenge ([68]).

To understand the pathophysiology of VT and the relation with border zone and core scar a different range of animal studies were presented in the literature. Some studies show that heterogeneous zones were located at successfully ablated VT sites, and that incomplete ablation of these zones was associated with VT recurrence ([12,55,188]). On the other hand, several studies point to the scar core on CMR as the substrate for VT ([49,167,150]). Gucuk Ipek et al. ([68]) mentioned that both sets of studies described above are likely correct, and portions of the heterogeneous border-zone are involved in VT circuits, most commonly as entrance or exit sites.

2.5 Imaging modalities

A range of modalities are used and are recommended as a baseline to determine treatment decisions and disease detection for the clinicians. Ascha et al. ([11]) are imaging modalities for mapping cardiac anatomy and functionality including X-ray, computed tomography (CT), doppler ultrasound scan, and magnetic resonance imaging (MRI).

The chest X-ray is a common scanning modality for initial detection and evaluation of pulmonary hypertension. Even the X-ray is a standard imaging modality, it is not preferable to cardiac scanning as it does not yield detailed information about the heart. CT is a scanning imaging technique involves high temporal, spatial resolution and speed. CT is a common technique using by clinicians in a huge range of cardiovascular and lung diseases. An important drawback is the exposure of the patient to CT is used to image the cardiac chambers for signs of right ventricle failure.

Doppler ultrasound is a noninvasive test that can be used to estimate the blood flow through your blood vessels by bouncing high-frequency sound waves (ultrasound) off circulating red blood cells. Regular ultrasound uses sound waves to produce images, however they can not show blood flow. Major principles of ultrasound imaging are the reflection of a sound wave. The ultrasound

probe detects these reflected waves to form the desired scan. Reflection of a sound wave observes when the wave travels between two tissues of different acoustic speeds and a fraction of the wave returns back. Doppler ultrasound may help diagnose many conditions, including blood clots, poorly functioning valves in leg veins, heart valve defects congenital heart disease, blocked arteries, and bulging arteries (aneurysms) ([69]).

Magnetic resonance imaging (MRI) scanners utilize strong magnetic fields, magnetic field gradients, and radio waves to generate images of the organs in the body. MRI is safe and it has a less harmful effect on the health of the patient composed of CT, however it is expensive and needs an expert operator. The most common protocol is with 8-mm slice intervals and 2-mm inter slice gap of a short-axis MRI (Figure 2.2). When a MRI scan utilizes to produce a proton density (PD) weighted image, the result is that the tissues with the higher concentration of protons (hydrogen atoms) produce the strongest signals and appear the brightest on the image. Proton density weighted sequence produces contrast mainly by minimizing the impact of T1 and T2 differences with long TR (2000-5000ms) and short TE (10-20). The easiest way to identify PD-weighted images is to compare the fluid with the fat signal. Fluids normally appear as grayish white, almost similar appearance as the fat in the body. If an accurate cardiac segmentation of cardiac chambers is needed, a late gadolinium enhancement LGE-MRI is used in the majority of cases. LGE-CMR scanning is a multi-parametric, multi-planar medical image technique with prognostic potential in patients with chronic myocardial infarction. Moreover, in LGE-MRI the scan is using gadolinium (Gd), which increases the contrast visibility, and diethylentriaminopentacetic acid which defines the in vivo behaviour of the whole heart. This procedure increase the detection accuracy of cardiac chambers.

Manually, scar recognition is most often based on the image characteristics of the scarred area, such as signal-to-noise ratio (SNR) and contrast-to-noise ratio (CNR) ([78]). Ordovas et al. ([141]) demonstrate the utility of LGE-MRI to increase the difference of contrast between the healthy and

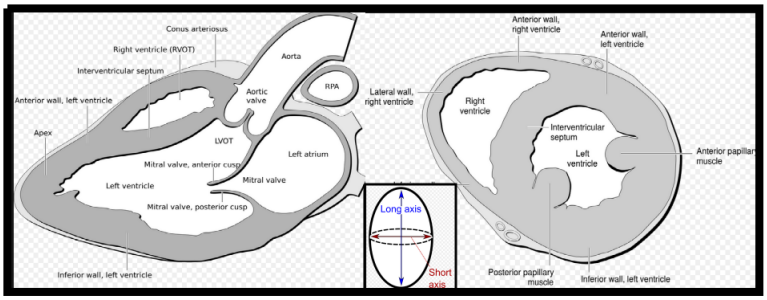

Figure 2.2: Long axis (left) and short axis (right) of cardiac anatomy

necrotic myocardium. Thus, by using LGE-MRI, the grey zone can efficiently be detected ([160]). The gray zone is a complex of necrotic (white intensity) and healthy (black intensity) zones which in the last years has been used to aid MRI-guided anatomic VT ablation ([125]). The myocardial scar usually is six standard deviations above the average myocardial intensity.

The established col of LGE-CMR is executed with breath holding in two-dimensional orientation and the final files are restored by inversion recovery gradient echo sequence methods ([141]). Except the two-dimensional LGE-CMR, radiology community is starting to use the three-dimensional LGE-MRI (3D) which can scan the whole heart with smaller gaps between the slices (higher resolution of z-axis) and increase the signal-to-noise ratio (SNR) ([93], [37]). There are studies verifying that the use of 3D LGE-MRI can reduce acquisition time compared the 2D ([64]) and present a higher mean scar volume after manual segmentation than the 2D ([56]). The main disadvantage of 3D LGE-MRI is the breath holding because of the large duration time is needed (more than twenty minutes). For this reason, there are studies in the literature that try to minimize the respiration effect of breath which creates miss alignment between the MRI slices in z-axis ([94], [149]). Bizino et al. ([24]) establish that the use of navigated free-breathing 3D LGE-MRI extracts is more reliable for scar mass compared to breath-hold techniques.

2.6 Magnetic resonance, physics and scanning

2.6.1 MR physics

A hydrogen atom (part of water molecule) is composed of two sub-particles the nucleus (proton) and an electron. MR uses the gyromagnetic properties of hydrogen atoms to measure the ratio of magnetic dipole moment to angular momentum ([187]). If a strong magnetic field B_0 exists the random orientation of dipoles (spinning protons) in space tend to align to the magnetic field direction, and a net magnetisation is produced. This net magnetisation can be used to generate images ([187]). Moreover, protons spin around their own axes. As spin is combined with electrical charge, the hydrogen atoms generates magnetic dipole moment, like a small bar magnet ([21]). As each microscopic volume in the human body containing crowd behaviour of protons can be modelled by the Larmor frequency f ([187,158]) which is proportional to the gyromagnetic ratio γ and the magnetic field strength Bo, by: $f = \gamma * B_0$. As a result of the spin behaviour within microscopic volumes, the net magnetisation vector has an angular precession frequency, and a phase, which describes the angular position relative to the MR magnets ([158]). In preparation for MR scanning, the system measures the Larmor frequency of the water in the region of interest and then tunes the spins of net magnetisation to avoid a out-of-phase shift or dephasing of the two spins ([187]). By rotating the net magnetisation into the x–y plane, the constituent spins will also acquire energy from the radiofrequency (RF) pulse. After that, the net magnetisation will return to the lower energy levels (initial equilibrium point) spin-lattice relaxation T1 ([187]).

2.6.2 From MR signals to image

The MR signals are used to localize and encode to build up a cross-sectional 2-dimensional image (or image slice). In this sub-section we will describe the most commonly methodology to build up a

cross-sectional 2D image using a combination of RF pulses and gradient magnetic fields.

The first step is the selection of an image slice ([21]). By applying a gradient of magnetic field at the same time as the RF excitation pulse transmission, the resonance of protons is limited to a slice of ROI tissue. This can be done as the frequency of the RF pulse corresponds to the Larmor frequency of the slice volume and the direction corresponds to the applied gradient. The scan resonance only detects protons in the plane that are in the correct angles to the gradient direction of the MR magnetic field. After the slice selection, a phase encoding gradient (GP) is applied for a specified period ([21]). The protons change the frequency of precession based on the gradient behaviour (if gradient increases the magnetic field then the protons develop a higher frequency). The protons change their relative phase by an amount depending on their position along the gradient. Following the phase encoding, the frequency encoding gradient (GF) is applied in a direction at right angles to it. As in phase encoding, the protons are tuned to rotate at different frequencies with respect of gradient relative direction. The final step is the image reconstruction. The frequency encoded signal is analysed using a Fourier transform (FT) ([21]) and based on an inverse FT an image is reconstructed.

2.6.3 Late gadolinium enhancement imaging

Late gadolinium enhancement (LGE) scanning is a highly spin-lattice relaxation sensitive gradient-echo imaging ([158]). This modality scanning take on during the administration of a gadolinium-based contrast agent (Gd). Gd helps to distributes in greater volumes in fibrotic myocardium, where it demonstrates slower washout times compared to normal myocardium. Using a T1-sensitive gradient echo sequence, some minutes (around 10-15) after the administration of Gd, it is possible to demonstrate the abnormal deposition of the contrast agent. Because of the Gd injection the focal regions of fibrosis become enhanced, indicating replacement fibrosis, whereas normal myocardium remains black. Expert radiologists are needed to adjust the inversion–recovery null time so that normal myocardium is null and appears dark ([158]). LGE has been an established scanning technique

in CMR, as the presence of enhancement indicating various kinds of focal fibrosis has been validated histologically ([158]). The CMR focal fibrosis has been shown to be associated with higher risk of worse outcome in diseases such as dilated cardiomyopathy, hypertrophic cardiomyopathy and arrhythmias conditions.

2.7 Machine learning and deep learning in medical applications

The use of machine learning (ML) and deep learning (DL) in medical applications is difficult, as the majority of medical data sets are characterized by incompleteness, incorrectness (systematic noise), inexactness (no complete parameters for a certain task) and sparseness. However, there are many studies in the literature that involve the solution of a medical problem using machine learning or deep learning techniques ([81], [171]). Jiang et al. ([81]) deduce that AI can assist clinicians to identify and treat medical diseases more effectively because of the advantage of AI to learning and self-correcting from data and to extract information from a patient population. Below we will introduce some fundamental ML and DL terminologies and ideas.

2.7.1 Machine learning

Overview

Artificial Intelligence (AI) is an agent that reacts to a specific environment. It can identify the state of the environment through sensors and affect the state of its output. AI is a superset of machine learning (ML) as it does not use explicit programming. ML is a computer program in which the performance of some tasks evaluated by performance measurements improves by experience learning ([47]).

Machine learning applications

Machine learning techniques have been applied in many different fields, and these include automation, industry, automotive, and aerospace. The contribution of ML to solve and help in the diagnosis and prognosis of medical problems is significant due to the unbiased results, the saving time and the automatic procedures which ML can provide. Machine learning techniques, and Bayesian statistics ([169]) have the advantage to handle efficiently missing information due to imputation (e.g., multilayer perceptron (MLP), self-organization maps (SOM) and k-nearest neighbour (KNN)) and fewer assumptions of the models. Bayesian statistics is a combination of data analysis technique and model fitting, where the available knowledge about parameters in a statistical model is updated with the information in the observed data based on Bayes' theorem. ML techniques used in topics such as treatment strategies and detection of cardiac disease. ML is utilized by clinicians and radiologists to classify cardiac diseases and guide them in pre-surgery procedures ([170], [83]). This arise the question "Does the ML works in clinical applications? And how robust the results can be?"

AI and ML techniques applied to cardiovascular diseases during the years 2013-2016 were one thousand papers and was the third most popular field after neoplasms and nervous (2800,1800 respectively) ([81]). Jiang et al. ([81]) in their survey discuss the prospect of ML systems in health care. Moreover, they mention that ML literature focuses mainly on cancer, nervous system disease, and cardiovascular diseases. Shameer et al. ([170]) state the limitations and challenges of applying machine learning techniques in cardiovascular medicine. They highlight that after 2000 studies of machine learning and cardiology (data have PubMed) from zero to five thousand per year. Lastly, Johnshon et al. ([83]) review common ML algorithms with supervised and unsupervised learning and cardiology and ML techniques, so more accurate personalized medicine can be delivered by clinicians. They mention that unsupervised techniques will be used for better classification of patient cases and reinforcement learning algorithms can be an advisor for clinicians in difficult cases, regarding

treatment methodologies and prognosis of cardiac disease. Reddy et al. ([155]) concludes that the use of a preacquired MRI 3D mapping system is accurate to guide an left ventricle catheter ablation procedure. Jiang et al. ([81]) estabilised that the most common AI algorithms are the Support Vector Machine (SVM) and Neural Networks (NN). Moreover, they demonstrate that deep learning (Convolutional Neural Network) is used mainly in diagnostic imaging. Feng et al. ([57]) presents a learning from a demonstration framework which takes into account previous cardiac mapping procedures and a Gaussian process (GP) model-based learning for the detection of voltage mapping in right ventricle operations. The framework was validated and the regression error was reduced compared to an existing geometry-based method. The learning problem is solved to maximize the insight of an optimization cost function by minimizing the constrains related to this. Feng et al. ([57]) used a kernel training method for the 24/25 different folds datasets and the remaining one was used to test the mapping strategy by an "Implicit exploration (IE) approach". One of the most important challenges in machine learning techniques is avoiding the overfitting or underfitting of the data sets (we will discuss more in section 2.6.3). Owning to the restricted engagement of the predictive models with the input training data, the use of orthogonal and general different techniques of validation with different formality and independent data is needed ([170]).

2.7.2 Deep learning

Overview

Deep learning (DL) is a subset of the ML family techniques based on artificial neural networks with representation learning and multiply layers (deep). Das et al. ([47]) classified the DL algorithms as supervised learning, unsupervised learning, and reinforcement learning. In this subsection, we will present different deep learning strategies, some regularization techniques, and some network architectures. In the last subsection we will discuss about supervised and unsupervised methods and

different types of applications.

Deep learning networks training

For the training of a deep learning network the determination of the optimizer algorithm, the cost function, the form of output units, and the type of the activation function is needed. Regarding the optimization techniques, the compute of the gradients of the cost function is necessary, and usually the backward-propagation algorithm is used. To train effectively a neural networks, feedforward and backward propagation depend on each other. Below we will discuss about feedforward and backward propagation, regularization, and network architectures.

Feedforward propagation

In deep learning, the major goal is to fit a data set to a function, so can approximate the functions f^* that describe the general population of the data set. The feedforward networks parameterize $y=f(x;\theta)$ and it tries to learn θ so f can approach f^*. In the feedforward network, there are no feedback connections. The feedback connections presented only in recurrent neural networks. f^* is approached most of the time by a chain of functions $f_1,f_2,f_3, \ldots$ with $f(x)=f_3(f_2(f_1(x)))$. Each f function expresses a layer of the network. The length of the layers gives the **depth** and the final layer called **output layer**. The layer between the input (data) and output layer are called **hidden layers** as the user have no clue of the outputs of them. The **width** is the dimensionality of the hidden layers. Each layer is consisted by units.

In the scenario that f^* is a nonlinear function, then in the feedforward network it can be approached by a combination of linear chain functions with nonlinear transformed input $\varphi(x)$. The

function φ in deep learning is defined through the learning process to define the θ of each layer:

$$y = f(x; \theta, w) = f(x; \theta)^T w) \tag{2.1}$$

Backward propagation

Backward propagation computes the gradient of neural network parameters. This method traverses the network in reverse order with respect of feedforward chain, from the output to the input layer, according to the chain rule from calculus. The algorithm stores any intermediate variables (partial derivatives) required while calculating the gradient with respect to some parameters. If we assume functions Y=f(X) and Z=g(Y), in which the input and the output X,Y, and Z are tensors of arbitrary shapes, by using the chain rule, we can compute the derivative of Z with respect of X. By this way the update of network parameters values during the epoch is computed by:

$$\frac{dZ}{dX} = prod(\frac{dZ}{dY}, \frac{dY}{dX}) \tag{2.2}$$

where prod operator is utilized to multiply its arguments after the necessary operations, such as transposition and swapping input positions, have been carried out.

Regularization

An important limitation of ML and deep network techniques is that the model can overfit or underfit the data of a cohort. Overfitting is when the model fits the data so closely that it also follows the noise in the dataset. Underfitting occurs when the model is too simple to capture the trend of the data. Both underfitting and overfitting will not lead to models that generalise well.

There exist a lot of different strategies to avoid overfitting like drop out, data augmentation, and synthetic data ([171]). Dropout is a regularization technique for reducing overfitting in neural net-

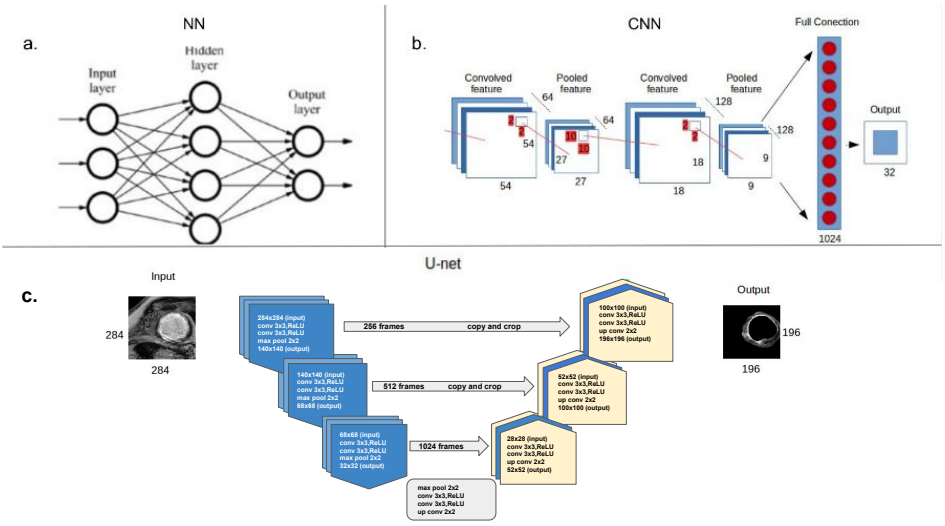

Figure 2.3: a) Neural network basic architecture (NN) b) Convoluted deep learning neural network (CNN) c) u-net modified structure

works by preventing complex co-adaptations on training data. Dropout relates a systematic dropping out of both hidden and visible units in a neural network during the training process. It is a way of performing model averaging on deep learning neural networks. Data augmentation and synthetic data techniques that focus on increasing the variation of shape, size, artifacts, and intensity histogram of the a specific dataset.

Networks architecture

There are plenty of different networks in deep learning and machine learning methods. Figure 5.1 shows typical neural networks including: b) a convolution neural network and c) an u-net structure. Figure 5.1 a) shows a convolution neural networks (CNNs) layer which is the most common layer used in deep high-parameter networks in the computer vision and medical image analysis fields. U-net is an established deep learning network for supervised medical image segmentation of different human's organs (lung, liver, heart, etc.).

Supervised and Unsupervised learning applications

Supervised learning algorithms are trained based on computing and adjusting the estimated error with respect to computing output and expected output. Unsupervised learning uses a clustering prepossessing step where it learns and adopts a pattern of input data (deep learning algorithm). Lastly, the reinforcement learning is based on an adjustment of the output of the algorithm with a reward for correct and a penalty for wrong outputs.

Supervised learning has been applied in fields including e-mail data, handwriting-speech-face recognition, information retrieval, operating system, computer vision, intrusion detection, signals detection, text categorization, text processing, hierarchical categorization of web pages, data optimization, dimensionality reduction, classification, computational finance, semantic scene classification and sound signal processing ([47]). Some unsupervised learning applications are DNA classification, social network analysis, organizing large clusters, astronomical data and analysis, anomaly detection of astronomical data, medical records, computational biology, medical problems, and speech activity processes ([47]). Finally, reinforcement learning is applied in computer games, machinery applications, traffic forecasting services, stock market analysis, and semantic annotation of environments. Das et al. ([47]) justifies that the ML will be always in the state of the art processes, as it makes machines autonomous, creates autonomous computing, and reduce the vigilance of users.

2.7.3 Training, Testing and validation

In machine learning and deep learning techniques is very important to split the dataset in training, and testing sub-samples ([186]). The standard approach in DL networks is to use a portion of 80-20 % or 70-30 % of the dataset samples in training, and testing respectively ([10]). In ML methods cross-validation (CV) strategies are used, where the dataset is split into training and testing subsets ([186]). Splitting data into training and testing subsets can be done using various methods, such as leave-

one-out, leave-p-out, k-fold, and Monte-Carlo sampling ([10]). The network or model is trained on the training set, while its evaluation performance is measured on the testing set. The aim of this process is to assess the ability of the machine learning algorithm to generalize to new data. For CV to be valid, the training and the testing sets need to be independent ([186]).

During the training process of ML and DL techniques, we try to avoid overfitting or underfitting of the dataset ([186]). Overfitting is when the model fits the data so accurate that it also follows the noise in the dataset. Noise is occurred as a result of the random fluctuations or the offsets generate from the true values in the independent, and dependent variables of the cohort. This can be observed when the performance of the method is better in the training samples than in the testing samples. Underfitting occurs when the model cannot detect a pattern in the cohort. This means that the model is not fitting the training data very accurate. When a model overfits the training data, it is said to have high variance, and when underfits high bias. To avoid the overfitting of the dataset, regularization techniques are used (discussed above in the subsection 2.6.2 REGULARIZATION). On the other hand, when underfitting is observed the model may not be complex enough, in terms of the features, the type of features, the depth, and other parameters of it.

2.8 The manual segmentation of myocardium and other organs regions.

Manual segmentation is the gold standard for segmentation of human organs from images obtained using different modalities. The segmented regions can be used to evaluate and detect diseases, or schedule and determine strategies of treatment in each specific patient. Manual segmentation of organs can be done based on manual annotation of points/contours (full manual approach) or by threshold techniques like full width half maximum (FWHM) or standard deviation (SD) methods (semi-automatic segmentation). In the full manual approach, radiologists sketch contours slice by slice using pointing devices such as a mouse or trackball. In 3D MRI or CT scanning image stacks

for manual segmentation can include as many as one hundred slices. In these cases, the manual segmentation is time consuming, thus semi-automatic approaches can be used. Semi-automatic segmentation of organs can be done based on threshold definition, where full width half maximum (FWHM) is a typical technique ([88]). FWHM is a technique where the expert has to define the darkest intensity (minimum) and the lightest (maximum) intensity of the region of interest, and the algorithm automatically thresholds all the regions that are higher or equal of the half of the two intensities ([89]).

Manual segmentation has some important limitations like the time consuming process and the bias effect from the expert's knowledge and biased observations. Moreover, in some cases manual segmentation or semi-automatic segmentation of organs can be very challenging even for experts. The main reason is the variability of internal (e.g., heterogeneity, spatial distribution, intensity distribution) and external (e.g. resolution, noise) factors ([175]). Lastly, because some algorithms for manual segmentation depend on initialization; when only one expert is involved the segmentation can be biased ([200]).

3

Related work and our pipelines

3.1 Overview

In this chapter, we will review the literature on automatic segmentation, image analysis, and processing of cardiac magnetic resonance images.

This chapter is organized as follows. Section 3.1 reviews the automatic segmentation of heart and scar regions. Section 3.2 surveys the automatic image analysis and detection techniques of heart

regions. Section 3.3 describes the different ways of clinical evaluation of automatic pipelines. Lastly in 3.4 we present our framework and the different pipelines we developed in the Phd thesis.

3.2 Review of automatic segmentation techniques of LV and scar

3.2.1 Review of myocardium segmentation methods

An important step so that automatic segmentation of scars from cardiac MRI can be attainable is the efficient and accurate segmentation of the myocardium, which can involve different AI techniques, such as active appearance models, edge and contour detection, region-based, graph-cuts, model fitting and deep learning.

The first technique that is used in the literature for myocardium and whole cardiac segmentation is the three dimensional **active appearance models** (AAM) ([44]). This technique is defined to locate rigid objects. Moreover it is focused on the learning patterns of the training set, so the specificity score of the model is decreased. This is a very good method to detect a general shape of a region of interest (ROI) but less accurate in the detection of the details of the specific ROI ([44], [185]).

Mitchell et al. ([129]) highlighted a comprehensive design of a 3D AAM ([44]). Andreopoulos et al. ([6]) presented a 3D AAM with short axis of cardiac MRI (CMR). They used a Gauss-Newton optimization technique and a time active shape model (ASM) for constraint 3D AAM segmentation. They indicated that the utilized of hierarchical 2D and time ASMs method improved significantly the results of 3D AAM method for the automatic segmentation of heart.

The main limitation of 3D AAM is misalignment. Because of errors arising from movement during breath-hold, the misalignment is widespread in the 3D AAM technique. Another important drawback is the dependence on the training set. The model learns only the variation of the training MRI cohort. Therefore, if there are no abnormal samples in the dataset, it is not possible to detect an abnormal case correctly (increase of cohort complexity). Lastly, the computational memory is a

concern due to the three dimensions of model.

Another promising technique of edge and contour detection is the **active contour method** ([92]). Kass et al. ([92]) introduced the idea of 'snake' contour models. Kass et al. ([92]) described a technique related to energy minimizing to detect contours and edges in images. It combines the energy function relating the spatial topology of the image. Geodesic contour is a method based on active contours and minimal distance curves ([28]). These two techniques create a combination of classical energy flow and the theory of curve evolution (a curve in a Riemannian space derived from the image content). As a result, the evolving curves are more adaptable regarding the topology, and different objects can be detected.

An active contour technique is used by Rananath et al. ([154]), as an extraction contour algorithm applied to CMR images. Pluempitiwiriyawej et al. ([151]) present a novel stochastic active contour scheme (STACS). They use the level set method to solve the minimization problem of the contour evolution with Euler-Lagrange equation. The energy function is combined stochastic region-based and edge-based information. With this scheme, they achieved a robust method regarding the initialization, contour, and noise segmentation problems. Ranganath et al. ([154]) uses the idea of Kass et al. ([92]) to detect the cardiac contours of the left ventricle. Pluempitiwiriyawej et al. ([151]) demonstrate an active contour structure that uses Euler-Lagrange equation to describe the contours. Again, the application was in cardiac MRI (CMR). Lee et al. ([105]) utilize a combination of iterative thresholding and active contour model. The main drawback of this method is the needed manual intervention when the intensity between the epicardium and the surrounding tissues (lungs and liver) is zero. The reason is the sensitivity of the computed gradient value to noise (the gray area of a MRI image magnifies noise). Paragios et al. ([145]) used level set framework to track the left ventricle. The level set framework of Paragios et al. ([145]) was a combination of geodesic active contour ([28]) using Heaviside as distribution ([204]). Geodesics contours can detect objects of complex topology with

high noise, artifacts, and shadow effects ([28]).

An advantage of classical active is that there is not a need to specify a different object to detect them (contrary to the geodesic contour). The main drawbacks of active contour algorithms are the sensitivity of the computed gradient value to noise and the need to initialize the energy functions near at the contour of interest.

Region based techniques is related to the active contour region of interest by global homogeneity region properties. Some known region-based techniques are k-means, fuzzy c-means, and the expectation-maximization (EM) algorithm using the Gaussian mixture model.

Chen et al. ([31]) describe a combination of adaptive k-mean clustering with spatial based on Gibbs random fields. The use of Gibbs random fields helps decrease the effect of noise and artifacts within the image in the k-means segmentation procedure. Ng et al. ([135]) use the k-means method to improve the performance of the watershed segmentation technique. The k-means are used to reduce the over-segmented phenomenon of watershed methods. The main drawback of k-means is the need of priory knowledge. By priory information, we mean the number and initial mean intensity values of segmented regions. Rezaee et al. ([157]) utilize a pyramidal image segmentation using fuzzy c-means clustering algorithm of the left ventricle (LV) in 140 CMR. They achieved a 90 per cent accuracy LV segmentation by using pyramidal segmentation instead of 86 per cent that achieved c-means alone. Pyramidal segmentation is a hierarchical way to sort the length of each cluster region in an image from maximum to minimum. Zhang et al. ([203]) describe a method based on hidden Markov random fields. Markov random field is a method related to region-based algorithms. Hidden Markov random field combines stochastic function and Markov. Zhang et al. ([203]) used to create synthetic images using stochastic sampling methods (Gibs sampler ([63])) to determine the sample distribution. The main drawback of this method is that the gradient can be trapped in the local minimum and does not convert the optimization technique.

The main limitations of region-based methods are: the sensitivity of the initialized region, and

the low performance because of the noise and the inhomogeneity of image intensities.

Another common automatic segmentation technique is **model fitting algorithms**. A predefined statistical shape model is used to fit a specific patient image. The most common algorithms using this technique are deformable models, atlas model, and statistical shape model ([85]).

Deformable models follow the principles of deformable theory, elasticity theory, and soft tissue mechanics ([146]). In deformable models are used as a combination of other techniques ([165]). Ecabert et al. ([51]) apply deformable model technique for automatic segmentation of the whole heart in CT images. They use a hybrid model of ASM and shape priors with a deformable model. As shape priors, they utilize a principal component analysis (PCA) method. In their study Ecabert et al. ([52]) present an automatic segmentation framework for heart detection. They use a linear transformation to initialize a deformable model ([139]). The main disadvantages of this method are: the needed object, the learning of training shape variation, and the high memory requirements.

Probabilistic atlases are shape models based on a training dataset, a generalized statistical model is created with specific standard deviation and mean values of edges shape ([118]). In some cases multi-atlases are used for better generalization. Multi-atlases are a complex of atlases from different vendors and patient shape variation.

Lorenzo et al. ([114]) use an atlas of 14 normal subjects and deformable registration techniques to fit the prior shape of atlas. In 2004, Lorenzo et al. ([114]) utilize an automatic atlas 4D segmentation framework. They use as prior shape, an atlas to register the edge shape of ROI in each patient. Finally, using 4D Markov random fields, they classify accurate regions of heart. Rikxoort et al. ([159]) use a multi-atlas scheme to segment the heart. They compare three different cases of multi-atlases. Firstly, they registered all multi-atlases prior shapes. Secondly, they register the most similar atlases to the MRI images. Moreover, they use the atlas with the best registration results (smaller error) as a reference. Based on the accuracy of other atlases in the local edge of interest. Bai et al. ([15]) use

multi-atlas to automatically segment the left ventricle of heart. They utilize an efficient refine registration to increase the accuracy of atlas fitting. Moreover, they use a patch-based label fusion model. As a result, the registration error is decreased. The main drawback of atlas segmentation is the lack of reproducibility of the results because of the assumption optimization problem. Moreover, shape variation of atlas in the training cohort ([17]). Lastly, some other limitations are: the high computational memory and time cost and the needed landmark determination by clinicians.

The advantage of model fitting algorithms is the ability to deformate edges with poor visualization of boundaries (weak edges). Because of this, these models are more flexibly than the rigid representation of AAM. However, this method has high computational cost regarding memory time, needing landmarks from expert or other complex prior shape algorithms such as AAP or ASP. Lastly, the adaptability of the method is restricted based on vendor and patient variation (etc., cardiac shape, age, sex). Thus, generalization is a tricky task ([71]).

Greig et al. ([67]) introduce the idea of **graph-cut technique**. Greig use maximum a posteriori estimation (MAP) to determine the maximum flow of a binary image. In the graph, the cut technique determines the seed points of the background and the region of interest. Based on these solvers, MAP determines the best optimal solution. Graph cut is a network graph of nodes (pixels) and edges (regions of pixels) that are connected to each other. A common technique linked with graph cut is *watershed* ([127], [115]). In the literature, there are studies that combine a watershed with other segmentation techniques, so more accurate results can be delivered ([72], [45], [15]).

Couprie et al. ([45]) introduce an extended version of the traditional watershed segmentation. This uses graph cuts, random walker, and shortest-path segmentation ([172]). Kang et al.([86]) use watershed algorithm of Couprie et al. ([45]) combined with the active contour model. Moreover, Bai et al. ([15]) utilize a combination of active contour and watershed algorithms to segment the whole heart. They use the active contour method to detect heart contours. Active contour is initialized by the results

of watershed segmentation. Thus, the drawback of noise and over-segmentation of the watershed method is reduced.

Generally, graph-cut techniques are accurate. However, the main limitations of watershed techniques are the need to manually seed points and the trend of the method to include more regions in the segmentation results than the ground truth (over-segmentation effect) ([85]).

In the literature, it is common to used more than one of the above techniques as an image processing pipeline to deliver more accurate results; as these frameworks can use combination of methods to strengthen and reduce the advantages and disadvantages of them. These pipelines are known as **image driven** segmentation frameworks.

Cocosco et al. ([40]) use a combination of ROI detection, Otsu's binary thresholding, and filters to increase SNR and region growing techniques. Huang et al. ([77]) described a pipeline to segment the left ventricle and papillary muscles. Again, they use a network to locate the ROI and a combination of image filtering, edge detection, thresholding, and image processing techniques. Moreover, Markov random field (MRF) and Conditional random fields (CRF) are used as supervised pipelines in automatic medical image segmentation ([91], [84]). MRF tries to estimate the likelihood distribution. It follows the theorem of independent observation regarding the labels. Probabilities uses uni-modal Gaussian distribution. CRF follows graphical notation. It is a discriminative model with relaxed independent assumptions of observation and labels. Moreover, in the discriminative model, there is no need of computing likelihood. This is an advantage as in some cases the real likelihood is very complex to be estimated.

There is an increase in the use of **deep learning algorithms** for cardiac automatic segmentation ([131,205,32]). Deep learning is a multilevel (deep) neural network (convolution neural network for imaging analysis) with a high number of fitting parameters which can learn a large amount of data (Big Data). Deep learning for cardiac automatic segmentation of LV is a widespread method.

The results of supervised methods are better than those obtained using machine learning and computer vision algorithms. In the cardiac segmentation challenge of MICCAI 2017, nine out of ten segmentation methods were based on deep learning ([205]). In addition, during the cardiac segmentation challenge of MICCAI 2019, the methods were used for image to image translation and segmentation of LGE-MRI, cine-MRI, and T1-MRI were deep learning networks. The automated myocardium segmentation methods are described by ([131,192,111]), all use deep learning networks applied to LGE-MRI images. The approach is described by Chen et al. ([32]) use a image to image translation network from balanced steady-state free-precession (bSSFP) images to LGE-MR images. They achieve an unsupervised segmentation of the LGE-MRI images.

One of the main drawbacks of deep learning (DL) techniques is the possibility of overfitting during the training process. As a result, the model fits the noise of the training set and loses the generalisation of the model that can be robust to variations in vendor and cardiac shape structure. To achieve a valid generalisation, a large labelled training set is required from different vendors and with variation of patients (such as abnormal or healthy cases, sex and age). Obtaining this kind of dataset is difficult. Moreover, because there is a need of manual segmentation or classification from expert radiologists.

Table 3.1 summarizes the different frameworks and pipeline techniques for automatic and semi-automatic segmentation of the myocardium of the LV with scar.

3.2.2 Review scar segmentation

Segmentation of a cardiac scar, and the surrounding region, often called as *grey zone* of a specific patient is challenging, because of the uncertain variation of size, shape, and location. The scars could be located in the atrial or ventricular wall of heart. Below we will review some machine learning and deep learning techniques.

Table 3.1: Comparison of automatic segmentation state of the art and established techniques of myocardial region based on average and standard deviation (SD) dice values (NM:not mentioned).

Region	**dice +/- SD (%)**	**Technique**
Myocardium	86.0 +/- NM	DL [131]
Myocardium	83.0 +/- NM	Unsupervised Multi-modal [32]
Myocardium	82.7 +/- 6	SK-Unet [192]
Myocardium	81.2 +/- 6.1	Combining Multi-Sequence and Synthetic-LGE [27]
Myocardium	80.1 +/- NM	Adversarial Domain Adaptation Network (GFRM) [191]
Myocardium	78.0 +/- NM	Multi-Atlas-CNN noise images [163]
Myocardium	74.9 +/- 10.0	Supervised-Domain-Adaptation [189]
Myocardium	71.4 +/- 10.0	Deep-Learning-Framework [206]
Myocardium	68.6 +/- 7.8	Shape-Transfer GAN network [177]
Myocardium	65.4 +/- 7.0	Multi-Atlas [16]
Myocardium	61.7 +/- 8.6	Pseudo-3D network [111]
Myocardium	61.7 +/- NM	CNN weak Domain Transfer [32]
Myocardium	61.0 +/- NM	Adversarial-CNN-Domain-Adaptation [33]
Myocardium	57.3 +/- NM	2D/3D-UNET Framework [195]
Myocardium	47.0 +/ 11.7	UNET++ Framework [156]

Machine learning techniques

Machine learning techniques are used for automated segmentation of atrial wall ([137],[46],[90]) . Karim et al. ([90]) surveyed different algorithms about automatic segmentation of the left atrium with scar. The data of this study were 60 images from pre-post ablation subjects. Segmentation techniques from seven different institutions were examined. These were hysteresis thresholding (IC), Mevis Region growing with mixture model fitting (MV), graph cuts with fuzzy c-means clustering (SY), mix of active contours and mixture model fitting (HB), combination of thresholding selection with manual wall delineation (YL) and unsupervised learning using k-means clustering (UTB). The best scores of the evaluation metrics were YL = 48%, UTA = 42%, UTB = 45% for pre-ablation and MV = 85%, YL = 84%, KCL = 78%, UTA = 78% for post-ablation cases. Karim et al. ([90]) suggested that a fixed model (SD and FWHM) is not a good candidate for segmentation of the atrium and ventricle. The reason is the different variability of internal characteristics (shape-size, histogram distribution, heterogeneity) and external conditions (resolution,image noise, and inversion time) of the scar region.

The left ventricle plays a critical role in generating cardiac output, but can be affected by infarct scar. Thus, the majority of segmentation studies focus on the left ventricle (LV).

Amado et al. ([4]) evaluated a region-growing algorithm based on two different kind of *semi-automatic segmentation* criteria of scar. The criteria were a N-standard deviation remote myocardium region (N=1,2,3,4,5,6) and a semi-automatic full-width at half-maximum (FWHM). FWHM method was a more accurate segmentation technique compare N-standard deviation techniques. Kolipaka et al. ([97]) evaluated a thresholding semi-automatic technique to segment the scar regions of LV. They used two to three standard deviations above the average intensity value.

Detsky et al. ([50]) used fuzzy clustering techniques in their study to identify infarct zones in multicontrast delayed enhancement images of the left ventricle. Lu et al. ([116]) utilized the graph cut

segmentation technique to separate the scar region from the healthy myocardial region of the left ventricle. The advantages of this automatic method were: high speed, comparing the manual segmentation, no false detection of scar region (common in FWHM manual methods), and automatic detection of scar regions. The limitation of this study was the needed manual correction of 66/136 images, due to misalignment deformation between the DE-MRI. Tao et al. ([175]) evaluated a segmentation method which combined intensity and spatial information to segment LV region. Their approach was based on theoretical density model theory. They assumed that the tissue follows Rician distribution ([78]) and a Gaussian distribution. The authors recommended this method for accurate myocardium scar delineation. Moreover, they suggested this as a useful tool for quantitative assessment in late gadolinium enhancement (LGE) MRI.

Karim et al. ([88]) evaluated six fixed-threshold methods and five algorithms about scar segmentation of LV. The dataset was from human's and pocine's hearts. Karim et al. ([88]) used an overlap metric and volumetric-based metric (dice metric) to evaluate the algorithms with respect to an expert's manual segmentation. The best segmentation results of human's hearts were those of MCG = 85%, KCL = 74%, and FWHM = 78%, 6-SD = 64%. On the other hand, porcine subject best results were observed in AIT = %86, KCL = 80% and 6-SD = 76%, FWHM = 69%. Karim et al. ([88]) elucidated that the best segmentation algorithm was the conditional Markov random field (MCG)([84], [99]). Mixture model with watershed transformation algorithm ([72]) had low accuracy. Some limitations of this study were the small number of subjects (30 patients) and the bias effect of ground truth (one expert). Moreover, there was high intensity variation across the images due to the coil shading effect. These segmentation methods were the portion of scar tissue. However, in cases of cardiac disease treatment, the specific location of the scar of LV is crucial.

Solis-Lemus et al. ([174]) developed a systematic pipeline to quantify the size and location of scar regions in the atrium based on LGE-MR images to evaluate the ablation procedures. The framework was semi-automatic and it was focusing in three phases: calculate the area of the scar, identify

the percentage of the pulmonary vein encircled by the scar, and compare the preprocessing and post-processing results of ablation. The main limitation is the manual segmentation of the atrium and scar regions.

Kurzendorfer et al. ([101]) utilized a fractal analysis and a random forest classifier to detect scar segmentation in LGE-MR images. They used fractal analysis to binary decompose the images and export information related to fractal dimension, the average gray value, and the size of binary objects. Lastly, they utilized a random forest to classify the features. The main limitation was the low accuracy (64 %) and the high standard deviation (17 %) that they can not verify the robustness and accuracy in other datasets, vendors, or institutes.

Deep learning techniques

In the last decade, deep learning networks have been used for automatic supervised scar segmentation ([131,176,201]). Deep learning is a promising method that is capable of accurate results. However, there is a need for preprocessing steps with training-testing datasets. Moreover, there can be a bias in the training dataset if the labelled images for supervised approaches rely on only one expert.

Abramson et al. ([1]) used a three level layer deep learning model for the detection of region of interest segmentation and correct anatomy of the regions in both slices and volumes. The main limitation of this study was the supervised training and multi-model scanning inputs (cine-MRI and LGE-MRI). Yang et al. ([197]) implemented an automatic segmentation pipeline of the left atrium (LA) with scars in LGE-MR images. They used a multi-atlas technique to deliver the LA and a deep learning network to deliver the scar regions.Main limitation of the pipeline, the supervision DL network that needs retraining or tuning in new institute or vendor datasets. Moccia et al. ([130]) implemented a supervised DL network and they studied two different protocols of fully automatic left ventricle scar segmentation. The limitations were: the lack of a true GT for algorithm training and testing (performed by one clinician) and the supervised nature of the deep learning technique.

Table 3.2: Comparison of established machine learning techniques for automatic and semi-automatic scar segmentation.

Region	Median (%)	Technique (cohort)
Scar	85.0	MCG[88]
Scar	74.0	KCL[88]
Scar	73.0	AIT[88]
Scar	44.0	MV[88]

Campelo et al. ([27]) delivered a deep learning automatic pipeline for segmenting scar regions using registered multisequence CMR image sequences (LGE, bSSFP, and T2). The limitations were the need of multi-modality scanning information and supervised training or tuning in new cohorts.

Table 3.2 summarizes the different frameworks and pipeline techniques for automatic and semi-automatic segmentation of the scar reg of the LV with scar. To conclude, in all deep learning approaches, the main limitations are the need of multi-modality scan images, the supervised network that needs retraining or tuning in new institute or vendor datasets.

3.2.3 Automated border zone and scar segmentation

Ukwatta et al. ([184]) developed an analytical framework of the automatic segmentation of LV with scar and border zones to use in an electrophysiology study of treatment strategies involving VT. They deliver very good results, however, they had to extract all LV myocardium manually annotated and their method evaluated only in 51 short-axis LGE-MR images. To author knowledge, MyoPS 2020 (https://zmiclab.github.io/projects/myops20/)is the first challenge where they study the automatic segmentation of the total scar region including the border zone (edema) and the core scar. Ukwatta et al. ([184]) achieved one of the best results of the challenge with a dice value of 60.4, 68.7, and 48.8 for the core scar, total scar, and border zone regions. The drawback of pipeline is the use of multi-modal scans (T2, LGE-MRI, cine-MRI) and they did not validate their code in cross-institute cohort.

Automated tumor segmentation

Another similar topic such as the automatic scar segmentation of the cardiac myocardium is the automatic tumor segmentation in organs like brain or lung. In the literature, there are plenty of studies related to the automatic segmentation of total tumor edema, non-enhancing tumor core, and enhancing tumor regions ([124,190,19,38,202,199]). The main limitations arising in this medical application of automatic segmentation regions have a serious overlap with the CMR scar segmentation (total scar, border zone, and core zone scar). Once again, the issue of robust and accurate results in the core and gray zones of the regions of interest, either tumor or scar, is crucial. Finally, the need of unsupervised techniques without any further need for retraining or even tuning in a new cohort is another important topic arising.

3.2.4 Image-to-image translation review

The unsupervised image-to-image translation (UNIT) problem is a hot topic in the deep learning community, especially after the introduction of generalized adversarial networks (GAN) based on the study of ([65]). Moreover, GAN can be used for the creation of synthetic images. These techniques increase the generalization of the networks and decrease the overfitting effect when the number of images are limited ([30]). This is happening as the synthetic images increase the variation of different shapes, sizes, and artifacts in the training process of model. In this section, we will review some of the state of the art networks for modality scan translation and some studies related to translation from cine-MR to LGE-MR images.

Hoffman et al. ([74]) offers a combination of mapping images across domains (pixel-level adaptation), mapping images from source to target and back to the source (cycle-consistency), mapping feature level method which discriminates the features or semantics of two images based on a specific task (feature level adaptation) and a semantic segmentation task. By this way, they deliver an accu-

rate translation from image to image space and semantic segmentation. Kim et al. ([95]) implement an unsupervised image-to-image translation. The novelty of this work was the use of an attention module to learn semantically important areas, like geometrical shape transformation. Moreover, they utilize a learnable normalization function to determine the scale level of geometrical shape variation in the image to image translation. Zhu et al. ([207]) implement an unpaired image-to-image translation using cycle-consistent adversarial networks (CycleGAN), an approach for learning to translate an image from a source to a target domain in the absence of paired trained data. Their network translation task which involved color and texture changes, was successful. However, in extreme transformations, such as geometric changes, the network failed. This can be done as in some cases the distribution characteristics of the training datasets do not include all geometrical shape variations of the translation task. Ming-Yu Liu et al. ([110]) designed an conditional GAN to deliver the task of translation. They use an u-net as a generator and a PatchGAN network as a discriminator to penalize structures at the scale of image patches.

Translation from cine-MR to LGE-MR image scan modalities is essential as an alternative registration of the LGE-MR images to cine-MRI and vice versa. Tao et al. ([175]) implement a shape-transfer GAN for cardiac segmentation of LV, RV, and myocardium from LGE-MRI images. Their model could generate realistic LGE images with anatomical shape information. Their method was validated on a dataset of 40 LGE-MRI patients and achieved a 68.6% myocardium segmentation dice metric accuracy. Gao et al. ([61]) propose a universal MR image standardization method which used an optimized CycleGAN to standardize images. Based on a weak pair approach, an accurate intensity transformation between two randomly paired MR images. There are some studies related image to image translation from other scan modalities, LGE-MRI or cine-MRI. Chartsias et al. ([30]) demonstrate the use of CycleGAN to translate CT images to realistic MR images. Welander et al. ([193]) evaluate two unsupervised GAN models (CycleGAN and UNIT) for translation of T1 to T2-weighted MRI. Both networks show similar results. They conclude that translation images to be

visually realistic can not be done with just the use of quantitative measurements.

CycleGAN is a fast trained and accurate enough network to deliver an unsupervised unpaired image-to-image translation from cine-MRI to LGE-MRI, which is the task of this study.

3.3 Review of clinical evaluation of automatic pipelines

3.3.1 Observer variability - evaluation metrics

Observer variability is related to Measurement Systems Analysis. A thorough study about the collection of data and evaluation of the automatic algorithm is needed as part of quality control. Parameterization of the uncertainty sources and the bias effect of observers ([152]). Popovic et al. ([152]) highlighted the needed observer variability in the cardiovascular imaging field. McAlindon et al. ([122]) defined intra-observation and inter-observer variability. Furthermore, they determined scan variability (interscan) in MRI imaging, regarding different manual segmentation techniques. FWHM is an established semi-automatic segmentation technique to be used in the majority of cases. Regarding different manual techniques, inter-scan variability has the highest variation followed by intra-observation and inter-observer variability ([122]).

Inter/intra-observer variability can determine the variation of the results of different observers (inter), and regarding the time of extraction of the same observer (intra). The need for a generalized unbiased ground truth to evaluate an automatic segmentation technique is crucial. Li et al. ([109]) compared different methods of ground truth extraction with respect to inter-observer variability (with eight experts), and the best voting threshold method was found to be the most accurate ([109]).

3.4 Our pipelines

In this thesis we developed the SOCRATIS framework, which is a fully automatic segmentation methodology for LV with scar in LGE-MR images. SOCRATIS has two implemented pipelines;

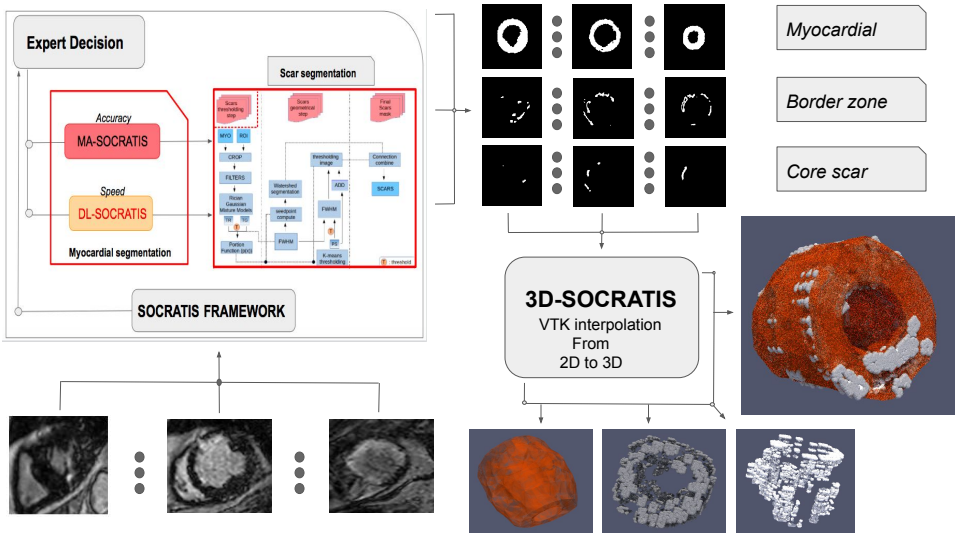

Figure 3.1: Our automatically three dimensional model extraction method. The method use as input LGE-MR images of specific patient. (Left top corner) SOCRATIS framework for automatic myocardial and scar region segmentation (BZ-SOCRATIS). The expert can decide about robust and accurate results using the MA-SOCRATIS pipeline or fast results using DL-SOCRATIS pipeline. (Left bottom corner) three different LGE-MRI slices of different depth and different patient. (Right top corner) automatic results of two dimensional slices of myocardial, border zone and core scar regions. (Right bottom corner) 3D-SOCRATIS pipeline to extract three dimension models of myocardial, border zone and core scar region of left ventricle.

a multi atlas - segmentation of cardiac region and total infarct scar (MA-SOCRATIS) and a deep learning - segmentation of cardiac region and total infarct scar (DL-SOCRATIS). This section describes the mathematical background of the MA-SOCRATIS and DL-SOCRATIS automatic segmentation pipelines and the overview of SOCRATIS framework. A comprehensive presentation of all the myocardium segmentation techniques is in Appendix A.

3.4.1 SOCRATIS framework

The outcome of the SOCRATIS framework aims to optimise the automatic segmentation of the LV with scars based on LGE-MRI images. SOCRATIS framework involves two alternative pipelines DL-SOCRATIS or MA-SOCRATIS, for quick or robust automatic segmentation results,

respectively. Fig. 3.1 illustrates the SOCRATIS framework and the expert user interaction. BZ-SOCRATIS is an extension of the MA-SOCRATIS pipeline which identifies both the border zone (BZ) and core scar regions of the LV cardiac chamber in LGE-MR images. The 3D-SOCRATIS pipeline was used to extract from 2D to 3D the border zone, core, scar, and myocardial of LV geometry.

MA-SOCRATIS

Based on the review of the methods summarised above, we have constructed an automatic pipeline for LV myocardium and scar segmentation as described below. Myocardium segmentation involves two steps; initial and reestimate segmentation. The initial step includes a multi-atlas initial segmentation technique ([16,102]). The re-estimate step uses a combination of k-means thresholding and a level-set active contour segmentation method ([136,105]). Combining these methods enables their limitations to be overcome. The parameters of the active contour are determined from the initial step, which overcomes the limitation of the active contour scheme, and the use of active contours overcomes the effects of noise and image artifacts on a k-means pixel-wise classifier.

DL-SOCRATIS

Based on the review of the methods summarised above and the need to use only LGE-MRI images as the based input pipeline, we have modified the initial myocardium segmentation of MA-SOCRATIS, by two different aspects. We utilized a combination of image-to-image translation network (LGE-MRI to cine-MRI) cycle-GAN ([207]) and an u-net ([162]) network trained on cine-MR images. In the second aspect, we used an u-net network trained on LGE-MR images. In both assumptions, we used a deep learning structure ([13]) to detect the region of interest (ROI).

3.4.2 BZ-SOCRATIS

For automated segmentation of scar, we use an algorithm that combines mixture model, histogram fitting and constrained watershed segmentation using an automatic seed point for the scar region. This approach is based on Hennemuth et al. ([72]) with some amendments, so the performance of watershed segmentation is improved. A limitation of the method described by Hennemuth et al. ([72]) was the use of Rayleign-Gaussian mixture models. Other studies have assumed that the healthy myocardium is described by a Rician distribution ([175,164]). Thus, our approach involves the use of Rician-Gaussian mixture models. In addition, the re-estimation of the myocardium can deliver automatic seed points for the scar region with good spatial accuracy. Regarding the automatic border zone and core scar segmentation, we add a border zone (BZ-TH) thresholding segmentation technique. This technique is a BZ-TH threshold determination between the rician threshold (TR) and the Gaussian threshold (TG); based on different portions. The region between the TR and BZ-TH thresholds is the border zone region and between the BZ-TH and TG are the core scar areas.

In the next chapters, we will describe the methods that are involved in MA-SOCRATIS, BZ-SOCRATIS, and DL-SOCRATIS pipelines.

4

Methodology of MA-SOCRATIS and BZ-SOCRATIS pipelines

4.1 Overview

In this section, we describe in detail the MA-SOCRATIS pipeline, and the extension of MA-SOCRATIS's scar segmentation pipeline (BZ-SOCRATIS) to automatically and semi-automatically the border

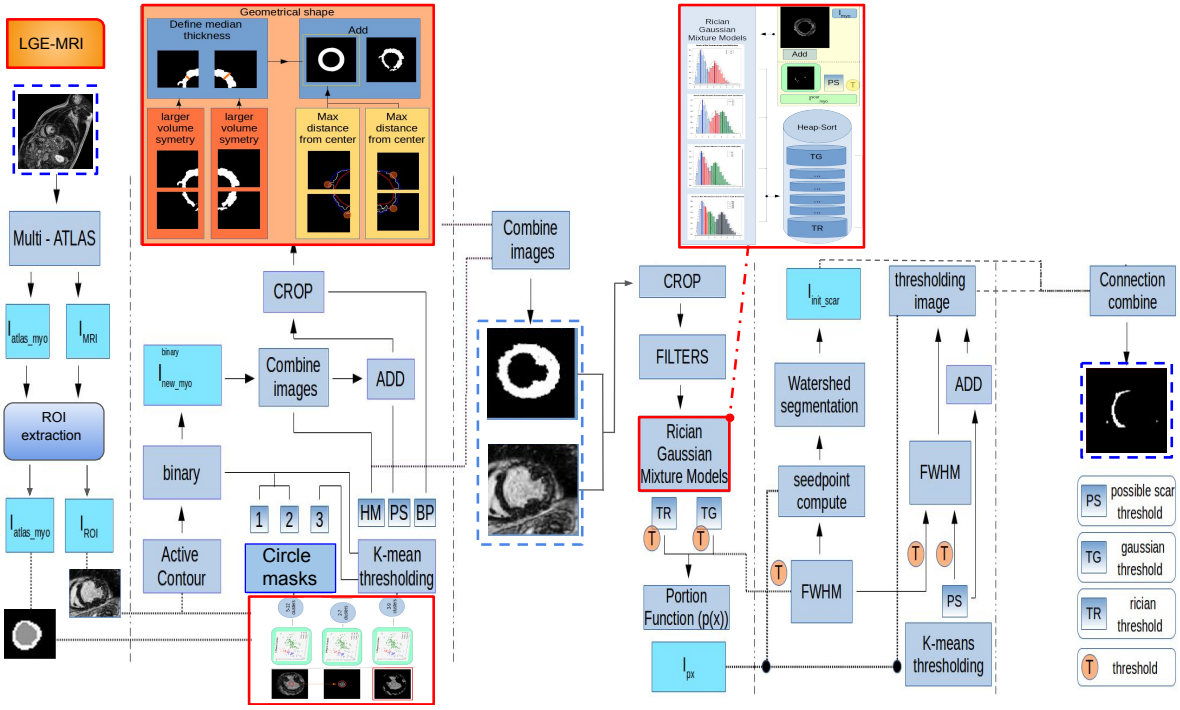

Figure 4.1: MA-SOCRATIS pipeline for automatic segmentation of myocardium (left) and scar (right) regions of the LV.

zone can be extracted.

4.2 Methods of MA-SOCRATIS

This study aimed to create a computational pipeline able to process a stack of LGE MR images with 8-bit resolution and size 256×256 pixels, to obtain contours representing the LV endocardium, LV epicardium, and scar edge. We separated the problem into two stages; myocardial segmentation and scar segmentation. These pipelines are illustrated in Figure 4.1, and we describe in turn.

4.2.1 Overview

One of the aims of this thesis is to test if a combination of AI techniques can deliver full automated segmentation of the myocardium and scar regions in a patient, without the need for expert's super-

vision.

In the implementation part of the first automatic pipeline, we implemented two different pipelines for myocardial and scar segmentation (Figure 4.1). Myocardial segmentation was split in two steps, initial segmentation and re-estimation.

As an initial segmentation, we used the pipeline of multi-atlas segmentation with registration refinement ([16]) to estimate the myocardium in LGE-MRI images. The idea of using an atlas pipeline as part of initial myocardium segmentation was from Kurzedorfer et al. ([102]) study. They used a cine-MRI atlas with deformable registration to estimate the endocardium and epicardium of LV with scars in 3D LGE-MRI images. However, the result was less than 80% accurate when measured using the dice metric. A multi-atlas segmentation technique was used by Yang et al. ([198]) for left atrium segmentation with higher results than Kurzedorfer et al. ([102]). Even using the multi-atlas approach ([16]), the result was not high accurate, as the dice average and standard deviation values were 65.4 ± 7.0. A possible explanation can be that multi-atlas method was created based on the balanced steady state free precession (b-SSFP) sequence of a patient without LV scar regions. Additionally, the cohort was composed of LGE-MRI images with LV scar. Because of the low accuracy of the multi-atlas segmentation, we included a re-estimation step to increase the accuracy of the automatic myocardium segmentation pipeline.

The re-estimation step (Figure 4.1) was a combination of k-means clustering and a geometric median shape variation. k-means was used as a thresholding process to define potential scar tissue based on histogram intensity variation technique (Figure 4.1). On the other hand, geometric median shape variation was utilized to estimate the median shape of the myocardium based on space variation technique (Figure 4.1). Finally, we used an active contour technique to determine the unhealthy and healthy myocardial wall.

Scar segmentation was produced by a combination of a Rician-Gaussian mixture model (thresholding technique) to determine the minimum and maximum intensity of the FWHM criterion

(Figure 4.1). Potential scar intensity threshold of the myocardium re-estimate step (k-mean thresholding process), and the intensity threshold of FWHM criterion were compared (Figure 4.1). The minimum value was then used as the final intensity threshold of the scar region. Lastly, a watershed with automatic seed point framework was used to segment the final scar region (Figure 4.1).

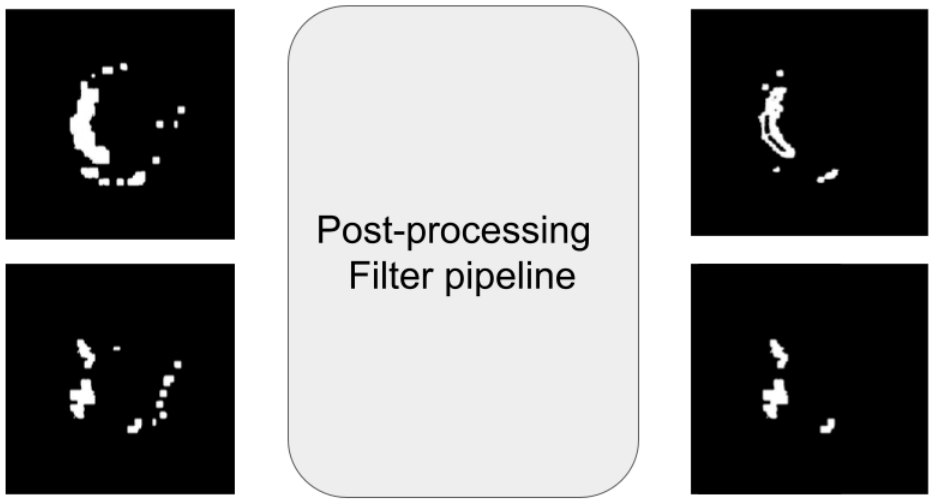

Figure 4.2: Post-processing pipeline results of the images before (left) and after (right) the processing.

The initial parameters of the MA-SOCRATIS pipeline do not need to redefine in a new cohort of LGE-MRI for segmentation of LV with scars. By this way MA-SOCRATIS is an unsupervised pipeline which needs no training or tuning in new datasets.

4.2.2 Pre-processing and post-processing frameworks

The pre-processing framework is used to reduce noise and artifacts in the images. We used a framework with a series of binomial deconvolution filters, Landweber deconvolution, and curvature anisotropic diffusion image filters ([120]). The post-processing framework is a combination of ob-

ject removal filter ([106]), median smooth filter, and binary filter (Figure 4.2). The median filter is a nonlinear filter. It is used to smooth an image without being affected by random noise or shot noise, and it is implemented in the ITK library. We used post-processing techniques to make the segmentation results smooth, eliminate artifacts of small regions (object removal filters), and reduce the SNR ratio. Image pre-processing is a field of computer vision theory related to decades of research and study. As a possible future work, we will use optimization techniques to deliver smoother results from a clinical point of view, which could possibly increase the performance of the SOCRATIS framework.

4.2.3 Myocardium segmentation

Myocardium segmentation (Figure 4.3) involved two steps; an initial segmentation step and a re-estimate segmentation step. Below we will thoroughly present the structure of these two steps.

Myocardium segmentation: Initial step

The pipeline for the initial segmentation step is shown in the left hand column of Figure 4.3. The comprehensive review of Chen et al. ([32]) describes a combination of machine learning techniques and shape models as a way to increase the robustness of a segmentation pipeline. This is a very common idea in the literature. The idea of using an atlas pipeline for initial myocardium segmentation was described by Kurzendorfer et al. ([102]). A cine-MRI atlas with deformable registration was used to estimate the LV endocardium and epicardium with scars from 3D LGE-MRI images. A multi-atlas segmentation technique was used by Yang et al. ([198]) for left atrium segmentation. Since this approach yielded better performance than a single atlas, we used multi-atlas segmentation to initially estimate the myocardium in LGE-MRI images.

The multi-atlas segmentation was based on a method with registration refinement described by

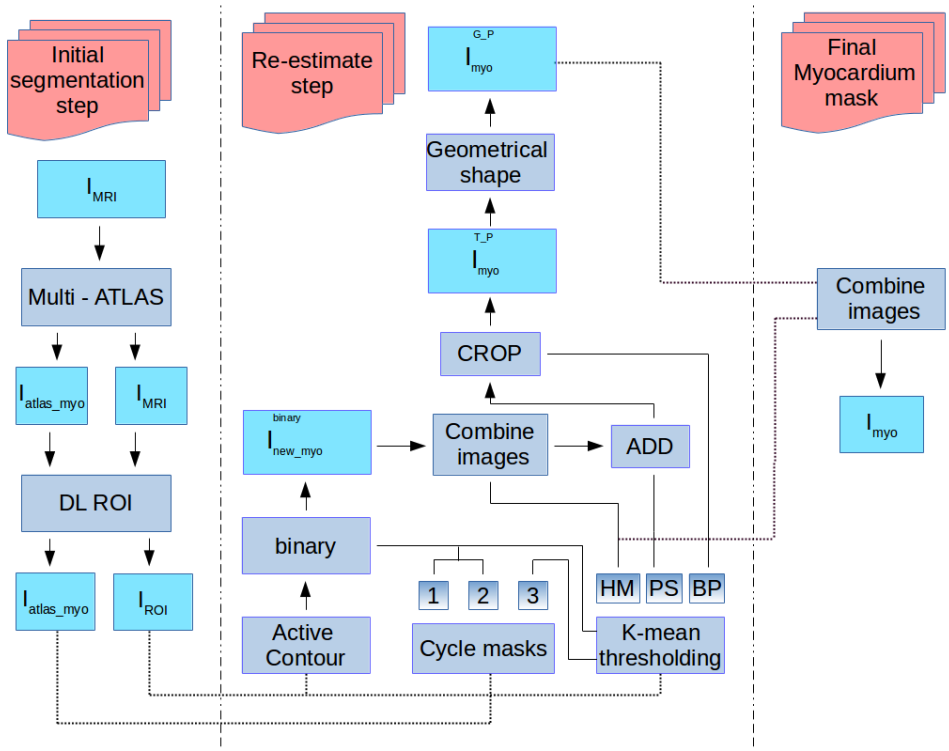

Figure 4.3: The myocardium segmentation phase. The myocardium segmentation is separated in three steps; initial segmentation step, re-estimation step, and final myocardium mask extraction step. The initial segmentation step involves the semi-automatic segmentation of multi-atlases approach and the detection of region of interest (ROI) using a convolution neural network. The re-estimation step involves a combination of active contour method and a k-mean thresholding technique to estimate the mean intensity of healthy, unhealthy and blood pool regions. The circle mask algorithm uses the information of the multi-atlas segmentation results to apply epicardium and endocardium boundaries in the outputs of k-mean thresholding technique. After this phase, the last images combined with the results of active contour. A geometrical shape approach is applied to verify the symmetrical shape of LV and fill the holes of the myocardium region extracted from the combined stage. The final mask is extracted by a combination of the healthy myocardium image (HM) and the image from the geometrical shape output. BP: blood pool, PS: possible scars, HM: healthy myocardium, 1,2,3: the three binary circle masks with different radius.

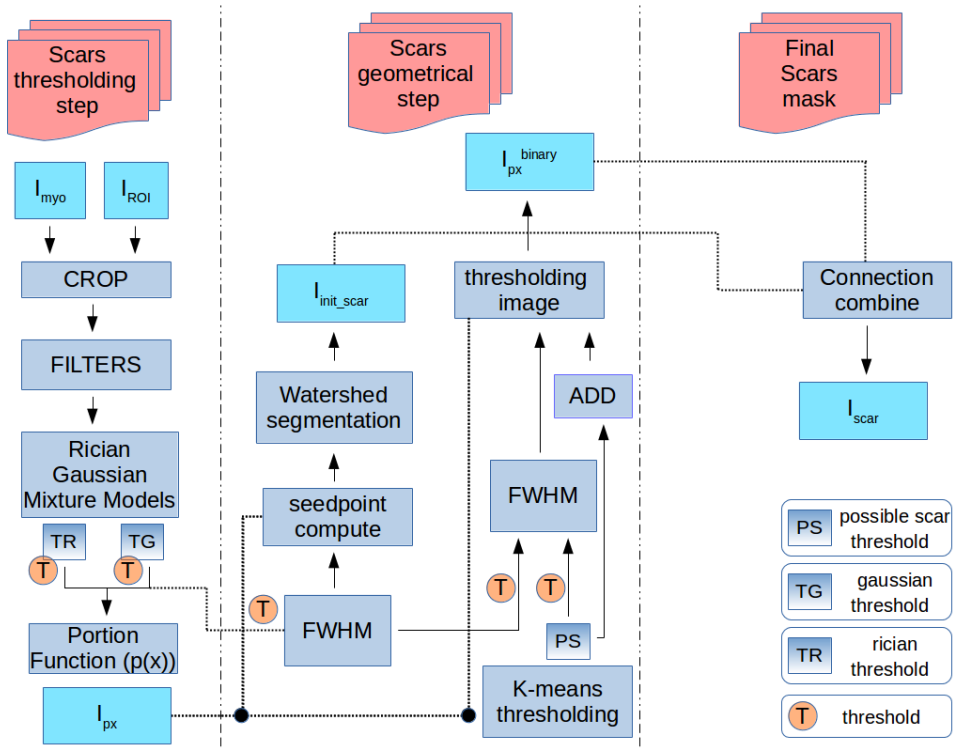

Figure 4.4: The scars segmentation phase. The scar segmentation pipeline for automatic has three steps; scars thresholding step, scars geometrical step, and final scars mask step. For the implementation of this pipeline we used as first step, four different combination of Rician-Gaussian (RG) mixture model (RG, RRG, RGG, RRGG) and a heap sort algorithm to sort the mixture model results from minimum to maximum values. As second step, we extract a partial volume model with the Rician threshold TR (minimum value of heap-sort block) and the Gaussian threshold TG (maximum value of heap-sort block). As third step, we use a watersheed technique with an automatic seed point generation algorithm to determine the core scar. Finally, we utilize the partial volume model to create a binary model of the border zone region. As last step, we connect the border zone and the core zone to extract the total scar region. PS: possible scars, TR: threshold rician, TG: threshold gaussian

Bai et al. ([16]). Landmarks in the LGE-MRI images related with the epicardium and endocardium boundaries were manually annotated by experts. The annotation landmark points ([16]) were six; two points in the first and last short-axis slices of LGE-MRI, and four in the middle slice, to determine the epicardium, and endocardium boundaries of left and right ventricle. The experts need about 30 seconds per patient for the manual annotation of the landmarks. Moreover, an automated process for the generation of landmarks has been described by Oktay et al. ([138]). The multi-atlas segmentation yielded a 256 $\times$ 256 myocardial segmentation $I_{atlas-myo}$. A 100 x 100 pixels ROI image I_{ROI} was then obtained from the center of $I_{atlas-myo}$ image by cropping the raw image I_{MRI}.

Myocardium segmentation: Re-estimation step

Figure 4.5 shows three example LGE images in the first column, and the results of LV myocardium segmentation based on the initial step alone in the second column. The third column is the manual segmentation of the myocardium based on expert radiologists and the forth MA-SOCRATIS results after the re-estimation step. Table 4.1 shows the dice accuracy of MA-SOCRATIS and the initial myocardium step (multi-atlas), demonstrating the improved accuracy of our pipeline (MA-SOCRATIS) compared with the multi-atlas. This improvement of the resolution and performance of myocardium segmentation was based on the re-estimationn step shown in the central column of Figure 4.3. There were four main phases to this re-estimation step; active contour, circle estimation, k-means thresholding, and a geometric phase.

Active Contour: The region-based techniques involved the selection of a global model so the estimation of the myocardium could be accomplished. Some hybrid techniques of boundary-based and region-based algorithms were generated. Kang et al.[85] classified this technique into three subcategories, parametric representation algorithms, level-set, and clustering. In the parametric representation, the field was dominated by the Mumford-Shah functional and the Bayesian, maximum α posteriori (MAP) estimators. The Mumford-Shah function ([48]) utilizes a piecewise smooth model

Table 4.1: Improved accuracy of myocardial segmentation by MA-SOCRATIS with re-estimation compared to the Multi-Atlas segmentation alone, using the MS-CMRSeg 2019 dataset.

Intra-Inter observer results				
Region	dice (%)	Sensitivity (%)	Specificity (%)	Cohort
MA-SOCRATIS	70.0 ± 5.5	90.5 ± 7.0	93.0 ± 1.7	MS-CMRSeg 2019 dataset
Multi-Atlas	65.4 ± 7.0	57.8 ± 4.6	97.8 ± 1.4	MS-CMRSeg 2019 dataset

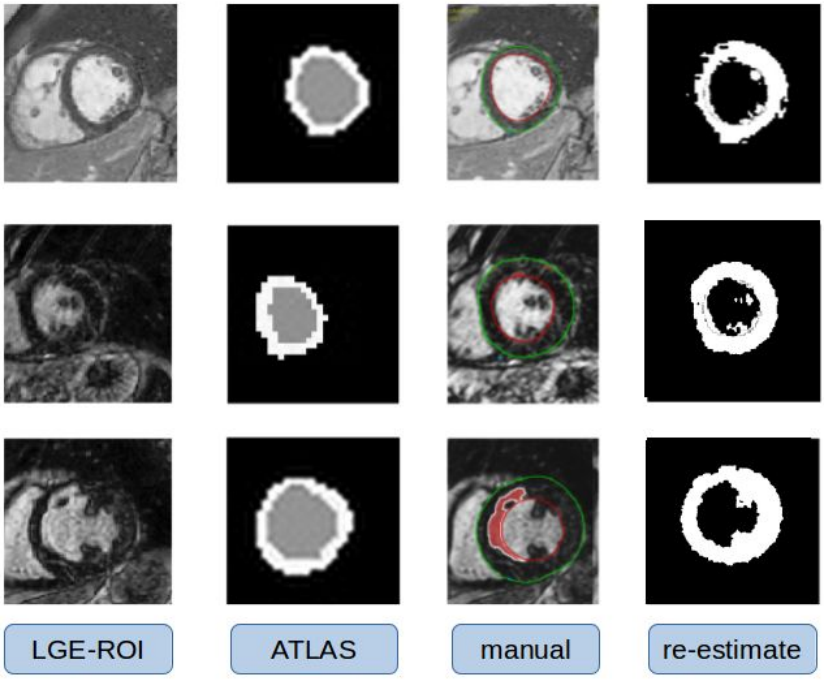

Figure 4.5: Comparison of initial and re-estimation steps in three example images. LGE-ROI shows the region of interest, with the LV towards the centre of the image. ATLAS label shows the result of initial segmentation with the multi-atlas method. Manual label shows manual segmentation of myocardium (green) and scar (red). Re-estimate label shows the result of the re-estimation step.

given by:

$$\mathrm{E(f,C)} = \mu\mathrm{Length}(C) + \lambda \int \int_R (f(x,y) - I(x,y))^2 \, dx\, dy + \int \int_{R-C} ||\nabla(f(x,y)||^2 \, dx\, dy + \mu|C| \tag{4.1}$$

where λ and μ were positive parameters, C was the boundary length, R was a domain, I the image, and f the piecewise smooth function which estimated the image of interest. Chan et al. ([29]) modeled the active contour based on techniques of Mumford-Shah segmentation and level set method. The novelties of this model were that it did not utilize an edge function to converge the detection of the desired boundary, there was no need to smooth the initial image even with a lowsignal to noise ratio, the model could detect objects without the use of gradient definition and it could be determined the interior curve automatically, starting from only one initial curve. A level set formulation of a curve C based on a $\varphi(x,y)$ Lipschitz function was defined by[29], by:

$$C = d\omega = (x,y) \in \Omega : \varphi(x,y) = 0 \tag{4.2}$$

$$\mathrm{inside}(C) = \omega = (x,y) \in \Omega : \varphi(x,y) > 0 \tag{4.3}$$

$$\mathrm{outside}(C) = \Omega/\omega = (x,y) \in \Omega : \varphi(x,y) < 0 \tag{4.4}$$

where Ω was the pixels of the image and x,y the index of a pixel. Chan et al. ([29]) utilized the energy equation to approach the optimization problem:

$$F(c_1, c_2, C) = \mu\mathrm{Length}(C) + \nu\mathrm{Area(inside(C))} + \lambda_1 \int_\Omega (u_o(x,y) - c_1)^2 \, dx\, dy + \lambda_2 \int_\Omega (u_o(x,y) - c_2)^2 \, dx\, dy \tag{4.5}$$

And based on the above equation, the Mumford-Shah function ([48]) and the Heaviside function H,

they established the following energy equation:

$$F(c_1, c_2, \varphi) = \mu \int_{\Omega} \delta_{\varepsilon}(\varphi(x,y))|\nabla\varphi(x,y)|\,dx\,dy + \nu \int_{\Omega} H_{\varepsilon}(\varphi(x,y))\,dx\,dy + \lambda_1 \int_{\Omega} (u_o(x,y) - c_1)^2 H_{\varepsilon}(\varphi(x,y))\,dx\,dy$$

$+ \text{”}\lambda_2 \int_{\Omega}(u_o(x,y) - c_2)^2(1 - H_{\varepsilon}(\varphi(x,y)))\,dx\,dy$(4.6)where H_{ε} was a regularization form of H given by:

$$H_{\varepsilon}(z) = \begin{cases} 1, & \text{if } z > \varepsilon \\ 0, & \text{if } z < -\varepsilon \\ \frac{1}{2}[1 + \frac{z}{\varepsilon} + \frac{1}{\pi}\sin(\frac{\pi z}{\varepsilon})], & \text{if } |z| \leq \varepsilon \end{cases} \tag{4.7}$$

and $\delta_{\varepsilon} = \frac{dH_{epsilon}(z)}{dz}$. Lastly, they used Euler-Lagrance equation to update the $\varphi(t,x,y)$:

$$\begin{cases} \frac{d\varphi}{dt} = \delta_{epsilon}[\mu div(\frac{\nabla\varphi}{|\nabla\varphi|}) - \nu - \lambda_1(u_o - c_1)^2 \\ +\lambda_2(u_o - c_2)^2] = 0 \in (0,\infty) \times \Omega, \\ \varphi(0,x,y) = \varphi_o(,x,y) \in \Omega \quad \text{(initial contour)} \\ \frac{\delta_{\varepsilon}(\varphi)}{|\nabla\varphi|}\frac{d\varphi}{d\vec{n}} = 0 \in d\Omega \end{cases} \tag{4.8}$$

where $\vec{n}$ is the external normal to the boundary $d\Omega$.

The I_{ROI} initialized the active contour without the edge detection method as discussed above ([29]) . Following the active contour method, we used the first and second cycle binary masks to isolate the endocardium and epicardium area.

Circle estimation: The initial myocardium segmentation was used to estimate three circular binary masks enclosing different parts of the chamber. The centre of each circle was the origin of

the endocardial contour in the initial multi-atlas segmentation. We defined two additional radii, a maximum and a minimum. The maximum was calculated as the half of the Euclidean distance between the minimum and maximum points of the initial epicardium plus an offset. The minimum radius was calculated from the endocardial contour in a similar way (equation below). The endocardial and epicardial circles (first and second masks) were used as binary masks to constrain the active contour segmentation method described below. The third circle was used to define the maximum intensity of the blood pool region as part of the k-meanss thresholding. Fig4.6 shows how the minimum radius was computed.

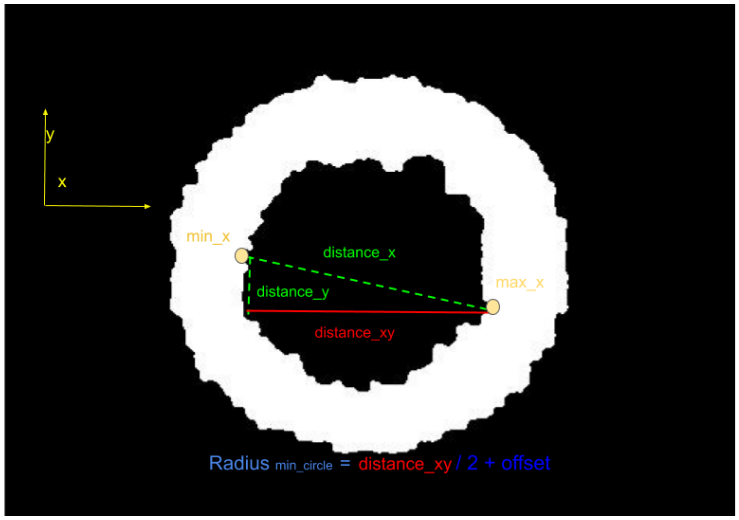

Figure 4.6: The computation process of minimum radius for the circle estimation step.

$$radius_{circle}^{max} = |max_{point}(C_{epi}^{atlas}) - min_{point}(C_{epi}^{atlas})| + offset_{epi} \tag{4.9}$$

$$radius_{circle}^{min} = |max_{point}(C_{endo}^{atlas}) - min_{point}(C_{endo}^{atlas})| + offset_{endo} \tag{4.10}$$

The third radius circle is given by the difference of the maximum and minimum radius:

$$radius_{circle}^{BP} = radius_{circle}^{max} - radius_{circle}^{min} \tag{4.11}$$

$$offset_{epi} = rate * |(max_{point}(C_{epi}^{atlas}) - min_{point}(C_{epi}^{atlas}))| \tag{4.12}$$

$$offset_{endo} = rate * |(max_{point}(C_{endo}^{atlas}) - min_{point}(C_{endo}^{atlas}))| \tag{4.13}$$

The endocardial and epicardial circles (first and second masks) were used as binary masks to restrict the active contour segmentation method described below. The third circle was used to define the maximum intensity of the blood pool region as part of the k-mean thresholding.

k-means thresholding: Figure 4.7 shows the k-means structure we used in our pipeline. The framework combines three k-means clusters with a variation in the initial number of regions. The first k-means cluster has 5 to 12 regions, the second 2 to 7, and the third 3 to 9. The variation in the number of regions is determined by the histogram intensity in each image. The first k-means cluster was used to create an homogeneous image, so an FWHM threshold could be computed (using $radius_{circle}^{max}$ mask). The second k-means cluster was used to define the maximum intensity of the blood pool area (using $radius_{circle}^{BP}$ mask). Given the homogeneous image from the first k-means cluster, a third k-means cluster is computed as the FWHM threshold for a possible scar. Based on this FWHM threshold, the blood pool and regions of possible scar were defined. If the threshold of the blood pool in the second k-means cluster was greater than the highest mean value of the first cluster, then it was assumed that there was no scar present. In this case, the FWHM threshold of the first k-means cluster defines only healthy myocardium and blood pool regions (lower and higher respectively) and the third k-mean was not computed.

The outputs of the k-means framework were therefore estimates of the pixel intensity thresholds for healthy myocardium, blood pool, and possible scar regions.

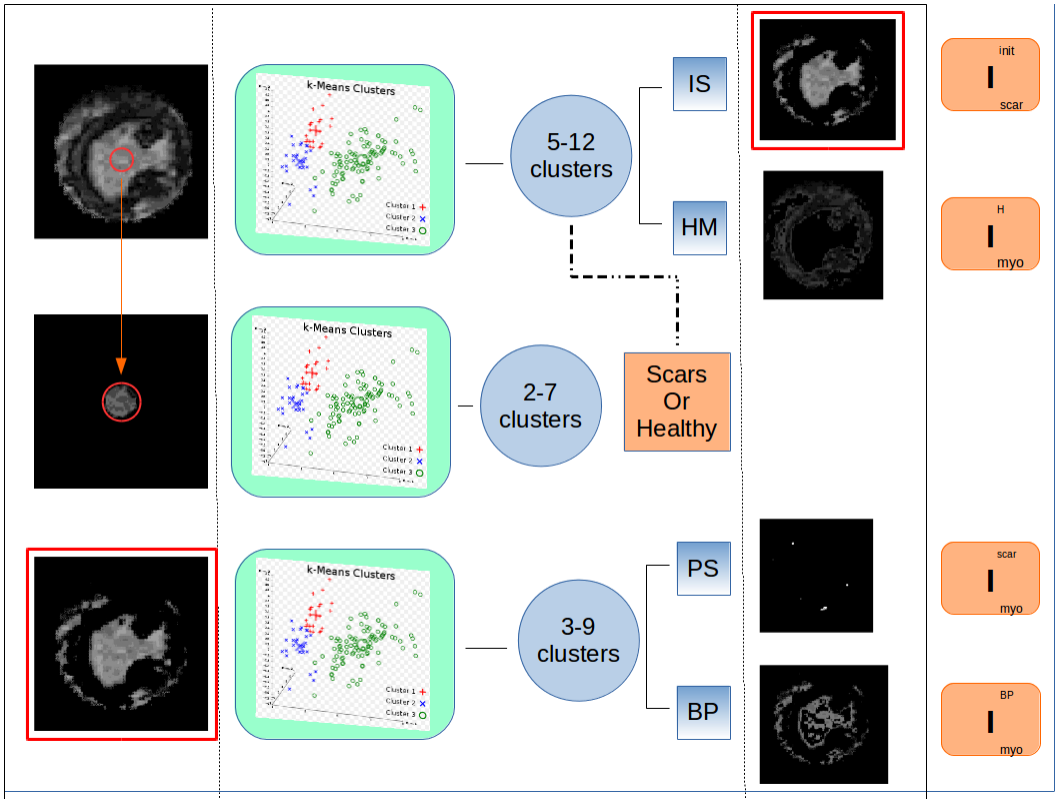

Figure 4.7: This is the re-estimate myocardium phase based on k-means. There are three different k-mean clusters of a variation of different (5-12, 2-7 and 3-9) areas of determination. Each of these clusters are focusing in a different region of interest (ROI). Respectively from up to down, the first is focusing in the whole LV region histogram, the second in blood pool histogram and the third in initial scar region. Based on these 3 k-mean clusters the BP, PS, HM and IS are defined in the end of this step, where: BP: blood pool, PS: possible scars, HM: healthy myocardium, IS: initial scars

Geometric phase: The geometric phase is shown in Figure 4.8. The thresholds estimated from the k-means framework were used to compute a healthy myocardium image I^{H}_{myo}, a possible scar surface image I^{scar}_{myo} and a blood pool image I_{BP}. I_{ROI} was used to initialise an active contour without edge detection method ([29]). This was combined with the first and second binary circle masks to restrict endocardium and epicardium areas. A framework to process redefined binary myocardium mask I^{binary}_{newmyo} based on an object removal filter ([106]), median smooth filter and binary filter was used. Lastly, pre-processing filters were used on I^{binary}_{newmyo} to reduce noise and artifacts in the image.

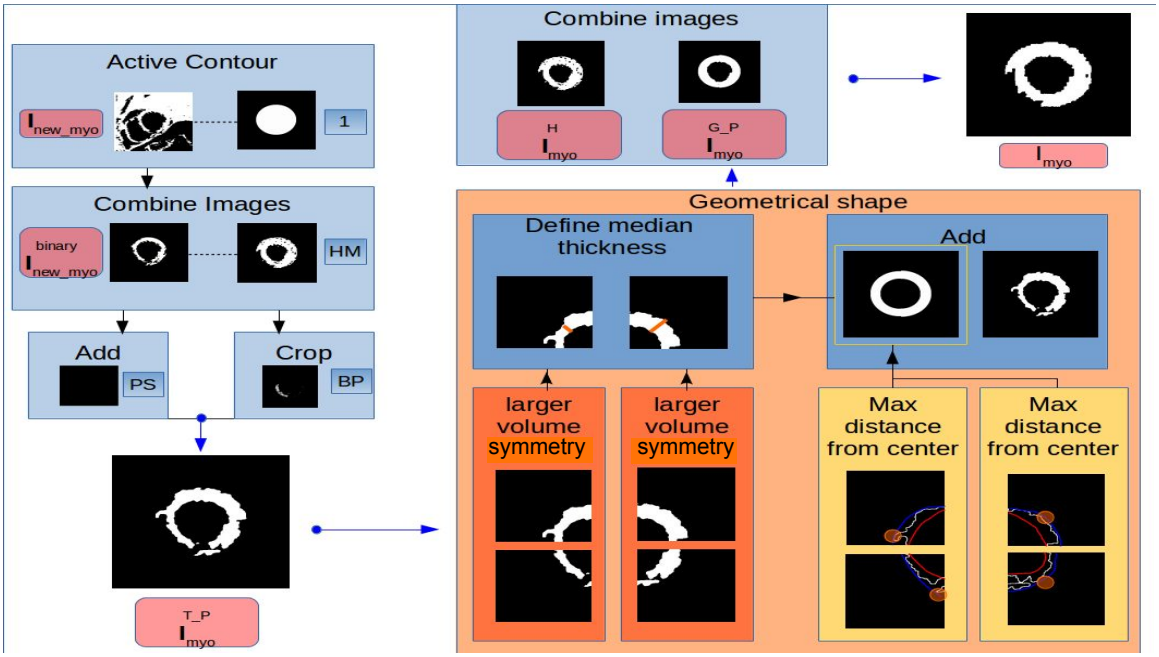

Figure 4.8: Detailed workflow for final step of re-estimation of myocardium segmentation. The outer circle mask (1), healthy myocardium (HM), possible scar (PS), and blood pool (BP) were combined, a further geometrical mask was applied (see text for details), and a final myocardial binary mask I_{myo} was computed.

After *k-means thresholding* we estimated median myocardium thickness (Figure 4.8), the last component of the re-estimation step. We separated $I_{myo}^{T_P}$ and its associated epicardial contour into four quadrants. We extracted two maximum volumes of $I_{myo}^{T_P}$ between the mirror sub-parts. A median ventricular wall thickness was then computed. Four arc masks were created based on the two median thicknesses. These arc masks were located at the maximum distance of the correspond epicardium sub-parts. Each arc mask was added to $I_{myo}^{T_P}$ so the $I_{myo}^{G_P}$ myocardium mask of geometric phase was extracted. Finally I_{myo}^{H} and $I_{myo}^{G_P}$ were combined, so the myocardial binary mask I_{myo} could be computed.

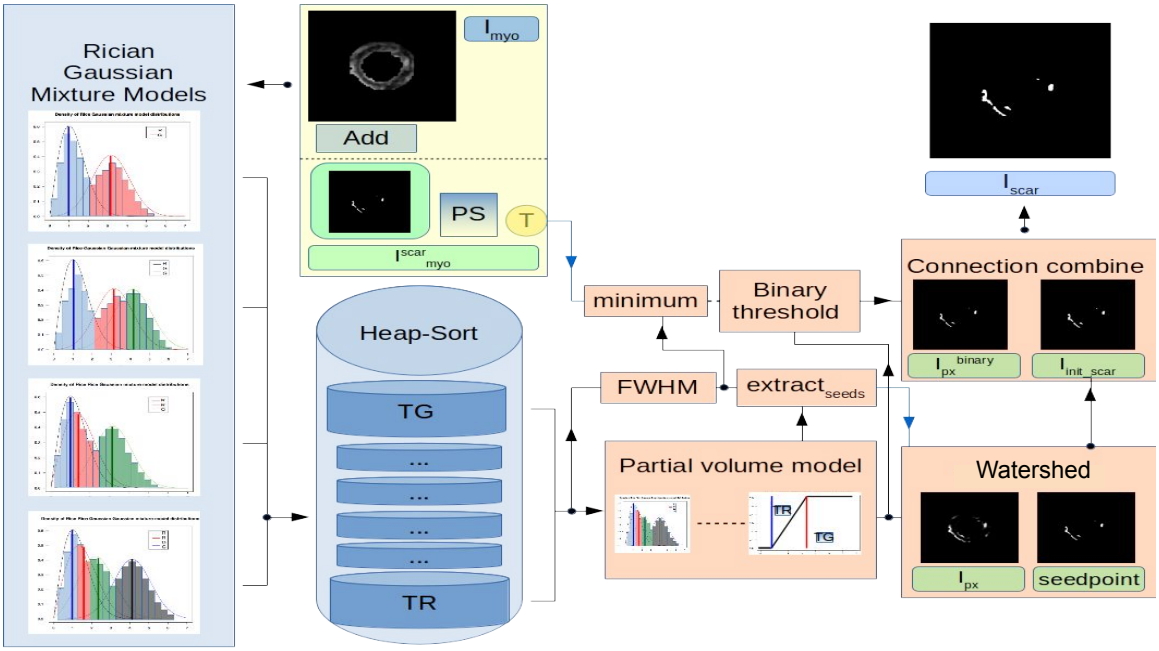

Figure 4.9: The BZ-SOCRATIS pipeline for automatic segmentation of total scar, or border zone and core scar regions of the left ventricle. For the implementation of this pipeline we used as first step, four different combination of Rician-Gaussian (RG) mixture model (RG, RRG, RGG, RRGG) and a heap sort algorithm to sort the mixture model results from minimum to maximum values. As second step, we extract a partial volume model with the Rician threshold TR (minimum value of heap-sort block) and the Gaussian threshold TG (maximum value of heap-sort block). Moreover, we define a border zone threshold (BZ-TH) to split the intensity range of TR and TG in intensity areas of border zone area and core scar. As third step, we use a watersheed technique with an automatic seed point generation algorithm to determine the core scar. Finally, we utilize the partial volume model to create a binary model of the border zone region. As last step, we connect the border zone and the core zone to extract the total scar region.

4.2.4 Scar segmentation

The pipeline for scar segmentation is shown in Figure 4.4. The pipeline has two detection components, threshold and geometric, and is designed to take into account the variation in pixel intensity histograms and the geometrical shape size of the LV.

Threshold detection component

We used a Rician-Gaussian mixture model ([175,164]) to estimate the intensity threshold of normal myocardium and scar. First, we defined a threshold for scar in each myocardium region. We used four different mixture models of Rician-Gaussian (RG) distributions (RG, RGG, RRG and RRGG), and these are shown in Figure 4.9. For each LGE-MRI image, we tested each combination of RG mixture model. Then we sorted the mean values of each RG model from maximum to minimum. The maximum value was defined as Gaussian threshold (TG) and the minimum as Rician threshold (TR). We assumed two different cases of mixture model initialisation; fixed mean values (pixel intensity of: 0, 60, 100, 200) and initial mean values based on Hennemuth et al. ([72]). The above strategy is defined after manual sensitivity evaluation of different initial conditions (initial values of the RG mixture model), and different number of mixture models. The presented strategy outperformed all the others. We defined the threshold for the scar region as the FWHM of maximum (Gaussian mean value) and minimum (Rician mean value) sorted mean values. Lastly, these thresholds of Rician (TR) and Gaussian (TG) distributions were used to define a linear partial volume model ([72]).

Rician Gaussian mixture models technique: Mixture models are a statistical tool related to fitting distribution in the density histogram of an image (graph-cut technique). The most common is a Gaussian mixture models (GMM). Lets assume a GMM of M components and a probability a vector $(\pi_i..., \pi_M)$ (where the summation of the vector is equal to one) to be the weight to each component. The main assumption regarding the components is that each component k is randomly and independently sampled, and a observation $p_k(x)$ is generated. Let assumed X to be a random sample with observed values x and Z a random sample of the probability vectors with observed values π_j. Then $P(Z = j) = \pi_j$ and $P(X = x|Z = j) = p_j(x)$. In GMM the $p_j(x)$ followed a Gaussian distribution, thus through the joint distribution of Z and X and the marginalization of Z the

$(X = x) = \sum_{j=1}^{M} \pi_j p_j(x) = \sum_{j=1}^{M} \pi_j N(x; \mu_j, \Sigma_j$ where μ_j and Σ_j denoted the mean and covariance matrix of the j^{th} component ([183]). Using the Bayes theorem, the posterior probability is then given by:

$$P(j|x) = \frac{p(x|Z = j)p(Z = j)}{p(x)} = \frac{\pi_j N(x; \mu_j, \Sigma_j)}{\sum_{l=1}^{M} \pi_l N(x; \mu_l, \Sigma_l)} \tag{4.14}$$

To fit the GMM, an expectation maximization (EM) optimization algorithm was used. The assumptions were that the mixture component number was specified and the mixture parameters of the model were $\theta = \pi_j, \mu_j, \Sigma_j | j = 1....M$. Then the parameters of the model were defined by solving the maximization of the log-likelihood ([183]):

$$\theta_M L = arg_\theta max log(p(X; \theta)), \tag{4.15}$$

assuming independent and identically distributed (i.i.d.) of the input sample dataset, so the likelihood could be written as:

$$p(X; \theta) = \Pi_{n=1}^{N} \sum_{j=1}^{M} \pi_j N(x_n; \mu_j, \Sigma_j). \tag{4.16}$$

The EM has two steps, the estimation (E) and the maximization (M) step. In E step, the hidden values z_{jn} were computed based on:

$$\langle z_{jn}^t \rangle = \frac{\pi_j^t N(x_n; \mu_j^t, \Sigma_j^t}{\sum_{j=1}^{M} \pi_j^t N(x_n; \mu_j^t, \Sigma_j^t} \tag{4.17}$$

In the M step, the expected values of the above equation are defined as the likelihood equation which tries to estimated the $\theta^(t)$ parameters by the:

$$Q(\theta, \theta^t) = \sum_{n=1}^{N} \sum_{j=1}^{M} \langle z_{jn}^t \rangle log \pi_j + \sum_{n=1}^{N} \sum_{j=1}^{M} \langle z_{jn}^t \rangle log N(x_n; \mu_j, \Sigma_j) \tag{4.18}$$

The updates of the parameters were given from:

$$\begin{cases} \pi_j^{(t+1)} = \frac{1}{N}\sum_{n=1}^{N} < z_{jn}^t >, \\ \mu_j^{(t+1)} = \frac{\sum_{n=1}^{N} < z_{jn}^t > x_n}{\sum_{n=1}^{N} < z_{jn}^t >}, \\ \Sigma_j^{(t+1)} = \frac{\sum_{n=1}^{N} < z_{jn}^t > (x_n - \mu_j^t)((x_n - \mu_j^t)^T}{\sum_{n=1}^{N} < z_{jn}^t >}, \end{cases} \tag{4.19}$$

Hennemuth et al. ([72]) apply a Raleigh-Gaussian mixture model in cardiac scar segmentation of LV. The mixture model intensity equation was given by:

$$h(x) = \alpha_R \frac{x}{\sigma_R} e^{-\frac{x^2}{2\sigma_R^2}} + \alpha_G \frac{1}{\sqrt{2\pi}\sigma_G} e^{-\frac{1}{2}(\frac{x-\mu}{sigma_G})^2} \tag{4.20}$$

Finally, Hennemuth et al. ([72]) use a Rician-Gaussian mixture model (RGM) with density function:

$$h(x) = \alpha_R \frac{x}{\sigma_R^2} e^{-\frac{x^2+m^2}{2\sigma_R^2}} I_o(\frac{xm}{\sigma_R^2}) + \alpha_G \frac{1}{\sqrt{2\pi}\sigma_G} e^{-\frac{1}{2}(\frac{x-\mu}{sigma_G})^2} \tag{4.21}$$

where Io is the zero modified Bessel function, α_R and α_G the mixture model portions of Rician and Gaussian distributions.

Geometric segmentation component

For the geometric component, we used a graph cut technique to identify the geometrical shape of scars. We used a morphological watershed segmentation technique, with initial seed points computed automatically with respect to the linear partial volume model equation and Euclidian distances $dist_{variation}$ between the endocardium and epicardium contours ([18]).

The seed points and the image from the p(x) equation were the input of a watershed segmentation algorithm ([126]), which provided an initial estimate of the scar region $I_{init_{scar}}$. The minimum of the

RGMM threshold and the k-mean threshold obtained in the re-estimation phase of myocardial segmentation was used as a threshold to convert the image from the p(x) equation to a binary image as shown in Figure 4.9. The final scar binary image I_{scars} was then extracted by combining the binary p(x) image and $I_{init_{scar}}$.

Watershed technique: Watershed is an efficient graph cut segmentation algorithm based on mathematical morphology theory that used for segmentation and region detection. The main drawback is the over-segmentation of the image. However there exist a lot of techniques in literature to reduced this effect ([20], [126]). In morphology theory, morphological gradient is a gray-tone function related to the highest and lowest pixel intensity of a given neighborhood of an index pixel x. The morphological definition of gradient was the difference between the dilated function ($f \oplus B$) and eroded function ($f \ominus B$) of a morphological function f where B was a disk with specific radius. Let y be the intensity value of a pixel in an image. The distance function was defined by the distances of each pixel point (y) from its neighbour pixels set Y^c with ($d(y) = dist(y, Y^c)$). There are many different distance function types such as Euclidean or geodesic (in our case we used geodesic). Maps of these distances could be computed such as $X_{i(d)} = y : d(y) \geq i = Y \ominus B$ where i was the level of threshold. The geodesic distance was defined as the shortest path between two points α, β of a X set of pixels in an image (dis(a,b)). The geodesic zone ($z(Y_i)$) was determined as the set of points of a region X which had finite distance from the Y_i subregion of X and are closer than the other Y_j sub-regions. Some other morphological important futures in watershed transformation, are the minimum and maximum regional functions. These were functions that compute the minimum (or maximum) of a region and a map of these values based on a pixels intensity threshold level i was created. Regarding the watershed line segmentation, if intensity function f (morphological gradient of image), geodesic zone ($z(Z_i(f))$) of the sub-regions ($Z_i(f)$) of image ($X_i(f)$) and minimal of the function f ($m_i(f)$) in the i thresholding levels were defined, then the computation of the watershed

subsection was implemented by:

$$\begin{cases} W_{i+1}(f) = [\cup_{i+1} z_{Z_{i+1}(f)}(X(f)_{i+1})] \cup m_{i+1}(f) \\ m_{i+1}(f) = Z_{i+1}(f) / R_{Z_{i+1}}(Z_i(f)) \\ DL(f) = W_N^{\tau}(f)(\text{with} max(f) = N) \end{cases} \tag{4.22}$$

where the iterative algorithm was initiated with ($W_{-1}(f) = 0$), $R_{Z_{i+1}}(Z_i(f)$ the geodesic distances of $Z_i(f)$ and DL(f) were the final watershed line. The total sum of ($m_i(f)$) equation gave the geodesic skeleton by zones (SKIZ).

For more robust results, we used seed points in the area of possible scars. The automatic seed points were computed based on the p(x) image and the Euclidian distances($dist_{variation}$) between the endocardium and epicardium contours by:

$$\begin{cases} \frac{distance(x, C_{endo}^{final})}{distance(C_{epi}^{final}, C_{endo}^{final})} \leq dist_{variation}, & \text{if } p(x) = 1 \end{cases} \tag{4.23}$$

4.3 Implementation of pipeline

4.3.1 Cardiac MR images and image analysis

In this study, we used cardiac LGE MR images for training different components of our processing pipeline, and for comparing the performance of the pipeline against a ground truth (GT) determined from manual segmentation.

4.3.2 Image analysis of the datasets

The effect of image noise was reduced with filters comprising binomial deconvolution, Landweber deconvolution, and curvature anisotropic diffusion image filters ([153]).

The manually segmented GT binary images of myocardium and scar and the automatic segmentation image results were modified by a post-processing framework to reduce the artifacts and noise of the images. This framework is an iterative process of a removal filter ([106]), a median smooth filter, and a binary filter. The removal filter erases any binary region smaller than 500 and 1000 full connected pixels regarding scar and myocardium binary images, respectively.

Finally, all automatic segmentation binary images had to be aligned with each corresponding manual inter-intra-ground truth image.

4.3.3 Training and testing data

We assembled a new dataset of short-axis LGE-MR images from 20 patients (LGE-20). All images were acquired, anonymised, and stored in DICOM format. All patients gave informed consent. The region of interest (ROI) was defined on the center of the endocardium region of each image. For myocardium regions, we evaluated our system on the data of 20 patients we collected and on the data of 40 patients of MS-CMRSeg 2019 challenge (STACOM at MICCAI 2019). The LGE-MRI dataset was obtained from a collection of three different MRI machine vendors with a variation in the initial settings used for a Philips MRI scanner. The images from each patient comprised between 8 and 12 short axis slices. Three experts segmented these images by identifying the LV endocardial and epicardial surfaces, and the borders of scar. Segmentation was done using MASS (research version 2017; Leiden University Medical Center, Leiden, the Netherlands).

4.3.4 Ground truth estimation

To assess *inter-* observer variability, each of the three experts segmented the myocardium and scar in each slice from each of the 20 patients. The experts were independent and they did not know the structure of our automatic segmentation pipeline. For the manual segmentation, we used a

voxel-by-voxel marking for the myocardium GT, and semi-automatic approach using thresholding filters (FWHM) for the scar GT. To assess *intra*-observation variability, one expert repeated the segmentation three times. Ground truths (GTs) were then determined based on the best voting threshold assessment described by Li et al. ([109]). GT masks were given by $k = (J+1)/2$ where k was the threshold value of agreement among observations, and J the total manual observer observations.

To evaluate the performance of the processing pipeline, for each scar and myocardium region, we computed the dice, sensitivity, and specificity metrics as follows:

$$\text{dice}(B_{model}, B_{ground}) = \frac{2(|B^1_{model}| \wedge |B^1_{ground}|)}{|B^1_{model}| + |B^1_{ground}|}, \tag{4.24}$$

$$\text{Sensitivity}(B_{model}, B_{ground}) = \frac{|B^1_{model} \wedge B^1_{ground}|}{|B^1_{ground}|}, \tag{4.25}$$

$$\text{Specificity}(B_{model}, B_{ground}) = \frac{|B^0_{model} \wedge B^0_{ground}|}{|B^0_{ground}|}, \tag{4.26}$$

where B_{model} represents the model prediction for each image, B_{ground} GT label, and the superscripts denote the set of pixels in the image that are classified as either scar or myocardium (1) or outside these regions (0).

Pipeline parameters definition

The parameters of the MA-SOCRATIS pipeline defined as the circle rate of equation (4.13) was 0.2 and for equation (4.12) was 0.75. Regarding the endocardium and epicardium boundary constrains, the maximum radius of the epicardium circle region was 160 pixels and the minimum radius of the endocardium was 25 pixels. The minimum thickness of the myocardium was 50 pixels. These parameters were extracted from the statistical mean values of normal epicardium and endocardium

radius. These parameters do not need to be redefined in a new cohort of LGE-MRI for segmentation of LV with scars. In this way MA-SOCRATIS is an unsupervised pipeline which needs no training or tuning in new datasets.

Regarding the FWHM ratio, instead of using the half distance (0.5) between the maximum and minimum intensity pixels, we used 0.45. This amendment was justified as the scar segmentation pipeline performed better with the amendment compared to 0.5. Finally, there was a mean intensity pixel threshold of 80. This threshold was determined based on image quality analysis of the available cohorts. After a statistical analysis of the histogram datasets we observed that the 80 mean intensity threshold is the lowest intensity level in which experts can perform good quality of manual segmentation in an image. For this study, three different experts evaluated and verified this observation (80 mean intensity pixels) about the good quality of a LGE-MR image. If the FWHM was less than this threshold, we assumed that there was no scar.

These parameters do not need to redefine in a new cohort of LGE-MRI for segmentation of LV with scars.

4.3.5 Software and hardware details

The code implementation is available on a public repository with url: https://github.com/INSIGNEO/MA-SOCRATIS. The code is implemented in C++ and Python. For automatic segmentation of myocardium scar regions of our dataset, we used an Intel(R) Core(TM) i5-4570 CPU @ 3.20GHz with 8GB RAM. The automatic segmentation pipeline needed 8.55 minutes for each patient of the LGE-MRI data set (8-12 slices of LGE-MRI images, around 40 s per slice).

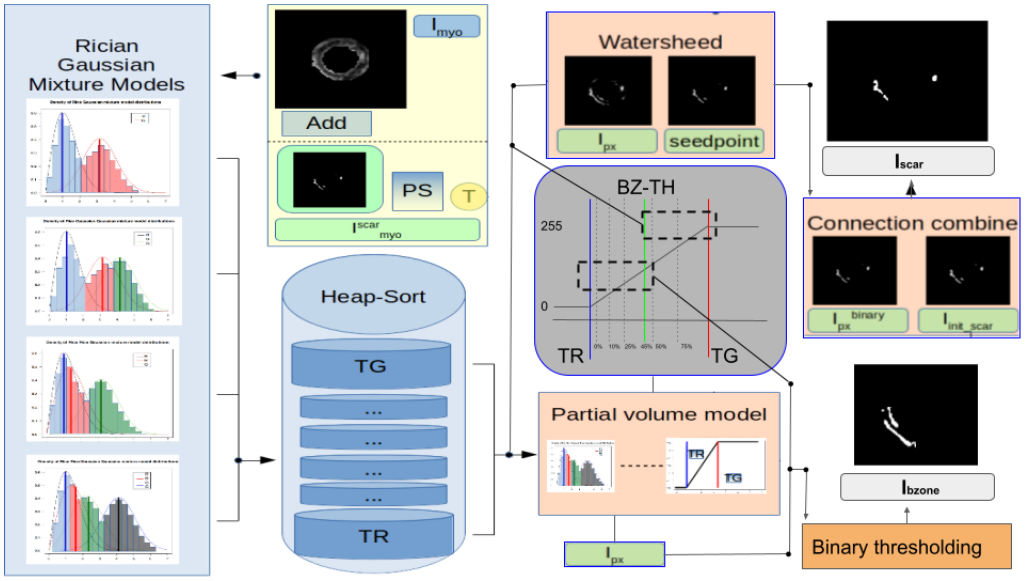

Figure 4.10: The BZ-SOCRATIS pipeline for automatic segmentation of total scar, or border zone and core scar regions of the left ventricle. For the implementation of this pipeline we used as first step, four different combination of Rician-Gaussian (RG) mixture model (RG, RRG, RGG, RRGG) and a heap sort algorithm to sort the mixture model results from minimum to maximum values. As second step, we extract a partial volume model with the Rician threshold TR (minimum value of heap-sort block) and the Gaussian threshold TG (maximum value of heap-sort block). Moreover, we define a border zone threshold (BZ-TH) to split the intensity range of TR and TG in intensity areas of border zone area and core scar. As third step, we use a watersheed technique with an automatic seed point generation algorithm to determine the core scar. Finally, we utilize the partial volume model to create a binary model of the border zone region. As last step, we connect the border zone and the core zone to extract the total scar region.

4.4 Methods of BZ-SOCRATIS

4.4.1 Overview

In this section, we implement an extension of MA-SOCRATIS's scar segmentation pipeline so the border zone extraction could be possible. Initially, the scar segmentation pipeline uses a combination of a Rician-Gaussian mixture model and a portion thresholding technique to determine the intensity pixels in scar regions. Following this step, a watershed method with an automatic seed point framework segments the final scar region. BZ-SOCRATIS uses an extra border zone threshold between the Rician and Gaussian thresholds to segment the border zone and core scar regions.

4.4.2 Total scar segmentation

The pipeline for total scar segmentation is shown in Figure 4.10. The pipeline has two detection components, threshold and geometric, and is designed to take into account the variation in pixel intensity histograms and the geometrical shape size of the LV.

Threshold detection component:

For the threshold detection component we followed the same process as in MA-SOCRATIS pipelines, be using a Rician-Gaussian mixture model ([175,164]) to estimate the intensity threshold of normal myocardium and scar.

4.4.3 Geometric segmentation component:

For the geometric component, we used a graph cut technique to identify the geometrical shape of scars. We used a morphological watershed segmentation technique, with initial seed points computed automatically with respect to the linear partial volume model equation and Euclidian distances $dist_{variation}$ between the endocardium and epicardium contours ([18]).
The seed points and the image from the p(x) equation were the input of a watershed segmentation algorithm ([126]), which provided an initial estimate of the scar region $I_{init_{scar}}$. The minimum of the RGMM threshold and the k-mean threshold obtained in the re-estimation phase of myocardial segmentation was used as a threshold to convert the image from the p(x) equation to a binary image as shown in Figure 4.10. The final scar binary image I_{scars} was then extracted by combining the binary p(x) image and $I_{init_{scar}}$.

4.4.4 BZ-TH and different portions segmentation

In the border zone segmentation, the geometric phase was the same. Regarding the thresholding segmentation, we used again the Rician-Gaussian distributions (Figure 4.10). The only difference of the two tasks is the BZ-TH box of Figure 4.10. In the border zone and core scar segmentation tasks, we had to compute an extra threshold of the border zone (BZ-TH) so we could split the two regions of interest. We used a variation of different portions regarding this threshold between the TR and TG, such as 10, 25, 45, 50, 75 %. The border zone area was determined as the region between the TR and BZ-TH and the core scar region between the BZ-TH and TG thresholds.

4.5 Implementation of pipeline

Training and testing data

We utilized a new dataset of short-axis LGE-MR images from 20 patients (LGE-20). All images were acquired, anonymised, and stored in DICOM format. All patients gave informed consent. The region of interest (ROI) was defined on the center of the endocardium region of each image. We followed the same pre-and post-processing of the dataset in MA-SOCRATIS training and testing application (Section:4.2.5, Training and testing data).

To evaluate the performance of the processing pipeline, for each scar and myocardium region, we computed the dice, sensitivity and specificity metrics.

4.6 Next chapter

In the next chapter, we will present the DL-SOCRATIS myocardial segmentation architecture.

5

Methodology of DL-SOCRATIS pipeline

Another aim of this study was to compare the performance of DL-SOCRATIS and MA-SOCRATIS computational pipelines when used for processing stacks of LGE MR images with 8-bit resolution and size 256 $\times$ 256 pixels, to obtain contours representing the LV endocardium, LV epicardium, and scar edge. We separated the problem again into two stages; myocardial segmentation and scar segmentation (explained in the first section of MA-SOCRATIS pipeline). These pipelines are illustrated in Figure 5.1, and we describe each in turn. Based on these results, in this chapter we describe

our integrated SOCRATIS framework; a framework to identify both the myocardium and regions of scar in LGE images of the left ventricle that can be tuned to optimise either speed or robustness.

5.1 Overview

Deep learning networks are increasingly used for automatic cardiac segmentation ([205]). The myocardium segmentation pipeline of DL-SOCRATIS (Figure 5.1) has the same architecture as MA-SOCRATIS except for the initial segmentation step. In DL-SOCRATIS, the multi-atlas segmentation part is replaced by two different deep learning networks DLc-SOCRATIS and DLl-SOCRATIS. DLc-SOCRATIS was a network trained using cine-MRI data, and DLl-SOCRATIS was a network trained using LGE-MRI data. In DL-SOCRATIS, the multi-atlas segmentation part is replaced by two different deep learning networks DLc-SOCRATIS and DLl-SOCRATIS. DLc-SOCRATIS was a network trained using cine-MRI data (cine-SOCRATIS and CGAN-SOCRATIS), and DLl-SOCRATIS was a network trained using LGE-MRI data (DL-SOCRATIS).

This pipeline was evaluated using LGE-MRI (more details in implementation subsection 5.3.8), as the scar regions are highlighted best in this modality. We chose to train the u-net structure using both cine-MR and LGE-MR images (more details in implementation subsection 5.3.8) . The reason for this choice was that in some cases the extraction of the myocardium region from LGE-MRI images was tricky and biased, even for experts. The main reason is there is not a clear distance between the intensity values of the scar, the blood pool, and the healthy myocardium regions, and the organs (lung, liver) around the heart. Moreover, the collection of an extensive LGE-MRI dataset to train a deep learning network is difficult as in some cases patients have to hold their breath for an extended time during a LGE-MRI scan.

To conclude, we would like to study the different accuracy and speed of a deep learning network in a translation image to image framework and in a network trained in the domain space of interest

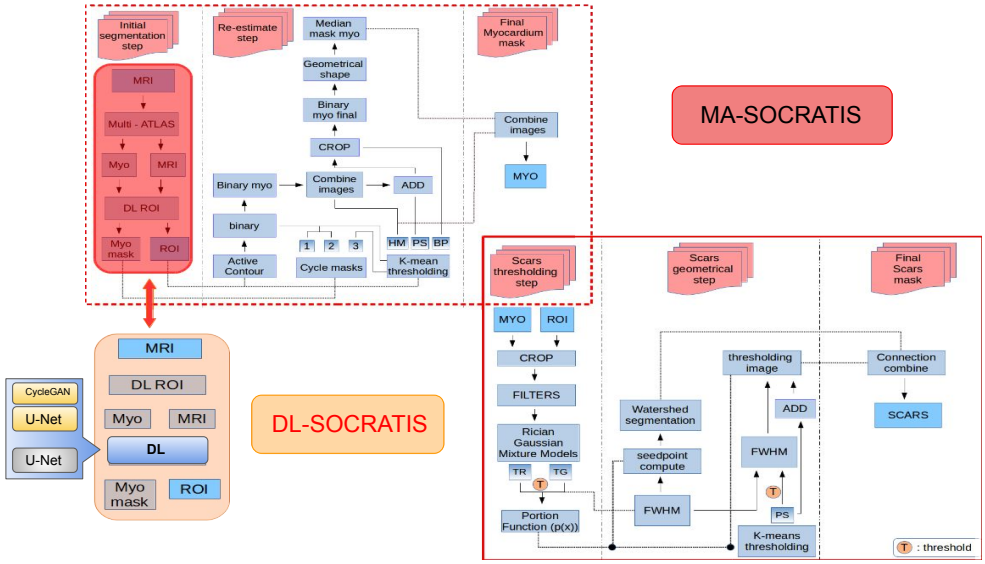

Figure 5.1: DL-SOCRATIS and MA-SOCRATIS ([119]) pipelines for automatic unsupervised segmentation pipeline. Red colour the myocardium segmentation and scar segmentation frameworks of MA-SOCRATIS. Orange colour the modified initial segmentation step of DL-SOCRATIS framework. In top Figure 1,2,3, are the three binary circle masks with different radius, BP is blood pool, PS is possible scar, and HM is healthy myocardium. Bottom right Figure highlights the scar segmentation framework, which is common in both pipelines. R is threshold Rician, TG is threshold Gaussian, and PS are possible scars. Bottom left Figure is the deep learning network structure to detect and segment the left ventricle. We used two different networks a combination of cycle-GAN/u-net network, and a u-net network.

(LGE-MRI). The aim of the next subsection is to develop a deep network that can train on cine-MRI and be deployed on LGE images because of the limited availability of published LGE datasets.

5.2 Methods of DLc-SOCRATIS

For the DLc-SOCRATIS method, we utilized a combination of a deep learning structure ([13]) to specify the region of interest (ROI), a cycle-GAN network ([207]) to translate the images from LGE-MR to the cine-MR space and an u-net ([162]) network trained on cine-MR images to segment the left ventricle (Figure 5.2) (CGAN-SOCRATIS and cine-SOCRATIS respectively).

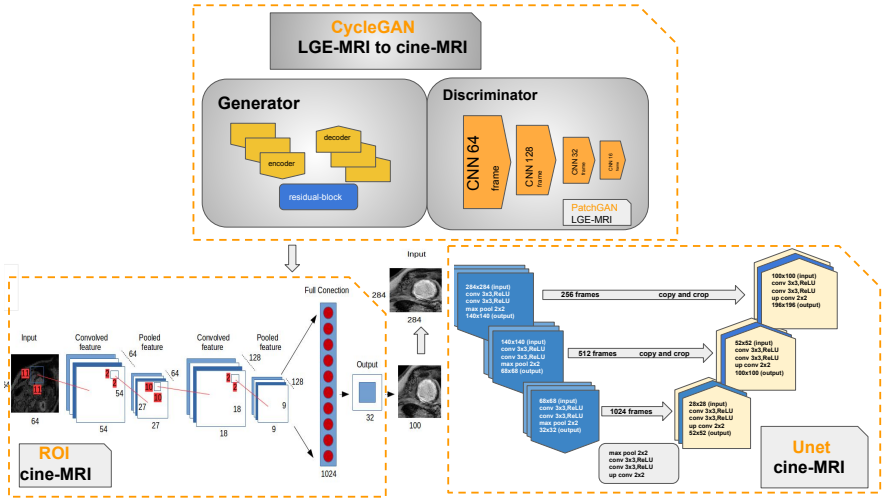

Figure 5.2: CGAN-SOCRATIS architecture. Top: CycleGAN structure to translate LGE-CMR to cine-CMR images. Discriminator is a patchGAN architecture and Generator is a encoder, decoder with residual block transformer network. Bottom left: CNN to detect the left ventricle. from cine-MR images Bottom right: Unet network to segment the myocardium from cine-MR images.

5.2.1 CYCLE-GAN

Architecture: For our first deep learning model, we needed an image to image translation method ([73]) as the DL network trained on cine-MRI data and the input scanning data were a LGE-MR images. In the original translation methods, there was a consistence on the domain and initial space datasets to be paired. In our case, the cine-MR and LGE-MR images were not paired. Thus, there was a needed unpaired translation method. Finally, we want to take into account both the forward and backward cycle consistency (eq.5.2) so the final loss of the network can be computed more efficiently accurately. To deliver the above needs, we used the Cylcle-GAN network ([207]). Figure 5.2 shows the network and the idea behind the CGAN-SOCRATIS pipeline.

The generalized adversarial network (GAN) theory ([79]) references a generator (G) and a discriminator (D). The generator is a network which tries to mimic the distribution of the domain dataset. The domain dataset initializes a discriminator network. In the majority of cases, a Gaussian noise

distribution initializes the generator. Based on the loss function, the GAN tries to train the weights of G to mimic the behaviour of D. As generator we used a ResNet deep learning network ([82]). This network contains three convolutions, several residual blocks, and two fractionally strided convolutions with stride 1. We use 6 blocks for 128 x 128 images and 9 blocks for 256 x 256 and higher-resolution training images. We use instance normalization. For the discriminator networks we used an 70 x 70 PatchGANs ([79], [108], [104]).

Training: The cycleGAN networks were trained using two challenging cohorts of LGE-MR and cine-MRI images (more details of the datasets are given in subsection 4.3.8). The LGE-MR and cine-MR images in end-diastolic phase of the cardiac cycle.

For the training and testing simulation, we set $\lambda = 10$ in equation 5.3. The loss function was optimized using the Adam method ([96]) a batch size of 4. All networks were trained from scratch with a learning rate of 0.0001. We keep the same learning rate for the first 100 epochs and linearly decay the rate to zero over the next 100 epochs (following the process of ([207])).

Optimisation function: Equation 5.1 is a pixel level adaptation cost function, so the network learns to map samples across domains. In the training process, we need to include the information of pixel level adaptation and both forward and backward cycle consistency (eq.5.2) so we used the loss function of equation 5.3.

$$\begin{cases} L_{GAN}(G, D_Y, X, Y) = E_{y \approx p_{data}}(y)[logD_Y(y)] \\ +E_{x \approx p_data}(x)[log(1 - D_Y(G(x))], \end{cases} \tag{5.1}$$

$$\begin{cases} L_{cycle}(G, F) = E_{x \approx p_{data}}(x)[F(G(x)) - x_1] \\ +E_{y \approx p_{data}}(y)[G(F(y)) - y_1]. \end{cases} \tag{5.2}$$

$$\begin{cases} L_{loss}(G, F, D_X, D_Y) = L_{GAN}(G, D_Y, X, Y) \\ +L_{GAN}(F, D_X, Y, X) + \lambda * L_{cycle}(G, F) \end{cases} \tag{5.3}$$

where G generator generates an image G(x), images from domain Y, while D_Y aims to distinguish between the translated sample G(x) and real samples y. controls the relative importance of the two objectives. The mapping functions G link the space of X to Y and F Y to X, and they associated adversarial discriminators D_Y and D_X.

5.2.2 Deep learning

Architecture: A deep learning structure ([13], [162]) was used to identify ROI in LGE-MRI images. The initial image size is 256×256. We downsampled it to 64×64 by interpolation/shrink technique (ITK software) to reduce the complexity of the network. The down-sampling was by a factor of 4, which helps to define the region of interest more accurately. After the ROI detection, we used up-sampling techniques to return to the original size of LGE-MR images. Figure 5.3(a) shows the structure of network ([13]).

For myocardium segmentation, we used a u-net structure with reduced depth compared to the original ([162]) network (cine-SOCRATIS). The modified structure of the u-net is shown in Figure 5.3(b). These modifications were based on the fact that the dataset was in 256×256 thus we need to remove the initial level of the original u-net network. There are contracting, expanding, and base layers on the left, right, and bottom of the Figure, respectively. Every step of the contraction part was based on a 3×3 convolution, batch normalisation, rectified linear unit (ReLu) and a 3×3 convolution layer. The expanding part consisted of a 3×3 convolution, a ReLU, a 3×3 convolution, and two consecutive blocks. The upsampling of the images was a sequence of 2×2 up-convolution with stride 2 blocks. The input image had a 284×284 size and the segmented myocardium region was 196×196.

Training: The deep learning networks were trained using the MICCAI 2017 whole cardiac cine-MRI segmentation challenge dataset (more details of the datasets in subsection 4.3.8). As the LGE-MRI images are in end-diastolic phase of the cardiac cycle, we trained using only the end-diastolic time frame of this training dataset.

The ROI network was trained based on the mean squared error cost function. The optimisation technique we used was the Adam method ([96]) with a fixed learning rate 0.001. We used data augmentation techniques details mentioned in subsection 5.3.2) and we trained the model for 100 iterations. The accuracy of model was 91% on the validation dataset and 92% on the training dataset. For the u-net network, the loss function was optimized using the Adam method ([96]) with an exponential decrease in the learning rate. The initial value of learning rate was 0.1, with 200 epochs. Moreover, classical methods of data augmentation are applied.

Optimisation function: We used the mean squared error for training the ROI detection model. Regarding the u-net train, we used two different cost functions. First, a dice-weighted cross-entropy loss combination (DCE):

$$\begin{cases} L_{final} = \lambda_{CE} * L_{CE} + \lambda_{dice} * L_{dice} \\ L_{dice} = \sum_i w_i * \log(2 - \frac{\sum_x t_i(x) p(x,i) + \varepsilon}{\sum_x t_i(x) + p(x,i) + \varepsilon} \\ L_{CE} = \sum_i w_i * (- \sum_x w_{t_i(x)} * \log(p(x, t_i(x))) \end{cases} \tag{5.4}$$

And second, a sensitivity, specificity dice-weighted cross entropy (SS-DCE) cost function:

$$\begin{cases} L_{DCE} = \lambda_{CE} * L_{CE} + \lambda_{dice} * L_{dice} \\ L_{sensitivity} = \frac{tp}{(tp+fn)} \\ L_{specificity} = \frac{tn}{(tn+fp)} \\ L_{SS} = 2 - L_{sensitivity} - L_{specificity} \\ L_{final} = \lambda_a * L_{DCE} + \lambda_b * L_{SS} \end{cases} \tag{5.5}$$

where t_i is the binary map of the i class (in our case the classes are two: endocardium and epicardium) The λ are weights. λ_{CE} is equal with 0.4, λ_{dice} with 0.6, λ_a with 0.2 and λ_b with 0.8. The tn, tp, fn, and fp are true negative, true positive, false negative, and false positive pixels, respectively.

Validation: For the DCE cost function training, we used a validation bootstrapping scheme of five cycles. We split the dataset to validate and train the data with a 0.3 and 0.7 rate, respectively. In each cycle, we replaced the selected data (images) and we shuffled them by a random seed point initialization. All validation accuracy results were higher than 90 % when measured with the dice metric, and higher than 82 % when measured with the jaccard metric. The accuracy of the model is shown in Table 5.1.

Regarding the SS-DCE cost function training, we used a validation bootstrapping scheme of ten cycles, with the same rate as before (0.3 and 0.7). All validation accuracy results were higher than 93 % when measured with the dice metric, and higher than 87 % when measured with the jaccard metric. The accuracy of the model is shown in Table 5.2. As we would like to know which of the two training u-net has the best accuracy, we tested in a cross-vendor cohort of 40 patients with cine-MR images (cine-40 mentioned in Implementation section). Table 5.4 shows the performance of the two different cost functions in a cross-vendor clinical cine-MRI dataset. In the first pipeline, the SS-DCE cost function delivered better performance than DCE. DCE achieved almost 86 percent dice accuracy and with SS-DCE 90% in the forty patient (436 images) cross-vendor clinical dataset.

Table 5.1: u-net first segmentation results of endocardium (EN) and epicardium (EP) regions in cine-100 dataset using the dice cross entropy (DCE) cost function.

DCE cost function results of cine-MRI					
Metric	1st cycle	2nd cycle	3rd cycle	4th cycle	5th cycle
jaccard (EP)(%)	83.3	81.9	82.2	84.4	83.8
jaccard (EN)(%)	92.5	92.8	92.6	93.0	93.2
dice (EP)(%)	90.9	90.0	90.2	91.5	91.2
dice (EN)(%)	96.2	96.3	96.2	96.5	96.6

Table 5.2: u-net segmentation results of endocardium (EN) and epicardium (EP) regions in cine-100 dataset using the sensitivity, specificity, dice cross entropy (SS-DCE) cost function.

SS-DCE cost function results of cine-MRI					
Metric	1st cycle	2nd cycle	3rd cycle	4th cycle	5th cycle
jaccard (EP)(%)	87.1	87.0	87.5	87.2	86.8
jaccard (EN)(%)	92.5	93.0	92.6	93.1	93.2
dice (EP)(%)	93.1	93.0	93.4	93.2	92.8
dice (EN)(%)	96.1	96.4	96.1	96.5	96.6
SS-DCE cost function results of cine-MRI					
Metric	6st cycle	7nd cycle	8rd cycle	9th cycle	10th cycle
jaccard (EP)(%)	87.1	86.7	86.4	87.2	86.8
jaccard (EN)(%)	93.6	93.4	93.7	93.5	93.2
dice (EP)(%)	93.1	92.8	92.7	93.3	93.0
dice (EN)(%)	96.8	96.6	96.9	96.7	96.4

Thus, in the cine-SOCRATIS, we will use the u-net model trained by SS-DCE cost function.

5.3 Methods of DLl-SOCRATIS

DLl-SOCRATIS is a deep learning pipeline with a combination of two deep learning structures, a CNN ([13]) to detect the region of interest (ROI) and an u-net network trained on LGE-MR images to segment the myocardium (Figure 5.3). Below we explain the analytical architecture and the training evaluation strategies.

Table 5.3: Different cost function segmentation results; In cross-institute and cross-vendor, end-diastolic cine-40 dataset for endocardium (EN) and epicardium (EP) regions. The sensitivity, specificity, and dice cross-entropy (SS-DCE) highly improve the performance compared with the dice cross-entropy (DCE) function.

Results of cine-MRI				
Metric	jaccard (EP) (%)	jaccard (EN) (%)	dice (EP) (%)	dice (EN) (%)
u-net DCE	75.5	92.2	85.8	96.0
u-net SS-DCE	81.7	93.5	89.8	96.7

5.3.1 Deep learning

Architecture: A deep learning structure ([13], [162]) was used to identify ROI in LGE-MRI images. The initial image size is 256×256. We downsampled it to 64×64 by a cubic interpolation/shrink technique (ITK software) to reduce the complexity of network. Figure 5.3(a) shows the network based on the assumption of Avendi et al. ([13]).

For myocardium segmentation, we used a u-net structure with reduced depth compared to the original ([162]) network. The modified structure of the u-net is shown in Figure 5.3(b), and it is identical to the u-net of DLc-SOCRATIS pipeline.

Training: The deep learning networks were trained using a challenging cohort of LGE-MRI images (more details of the datasets in subsection 4.3.8). The LGE-MRI images are in end-diastolic phase of the cardiac cycle.

The ROI network was trained based on the mean squared error cost function. The optimisation technique we used was the Adam method ([96]) with a fixed learning rate 0.001. We used data augmentation techniques (details mentioned in subsection 5.3.2) and we trained the model for 100 iterations. The accuracy of model was 91.6% on the validation dataset and 92.5% on the training dataset. For the u-net network, the loss function was optimized using the Adam method ([96]) with an exponential decrease in the learning rate. The initial value of learning rate was 0.1, with 100 epochs. Moreover, classical methods of data augmentation are applied.

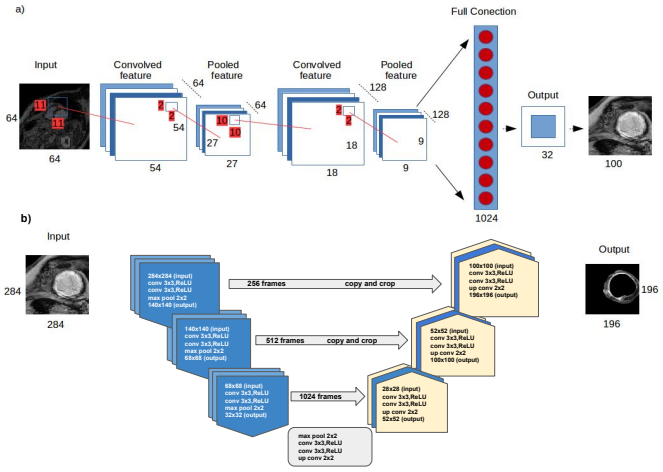

Figure 5.3: a) Convoluted deep learning newral network (CNN) for ROI detection b) u-net modified structure for myocardium segmentation

.

Optimisation function: We used the mean squared error for training the ROI detection model. Regarding the u-net train, we used the sensitivity, specificity dice-weighted cross entropy (SS-DCE) cost function as we defined in (5.5).

where t_i is the binary map of the i class (in our case the classes are two: endocardium and epicardium) The λ are weights. λ_{CE} is equal with 0.4, λ_{dice} with 0.6, λ_a with 0.2 and λ_b with 0.8. The tn, tp, fn, and fp are true negative, true positive, false negative, and false positive pixels, respectively.

Validation: Regarding the SS-DCE cost function training, we used a validation bootstrapping scheme of 5 cycles. Again, the portion of validation and training data was 0.3 and 0.7, respectively. All validation accuracy results were higher than 94% when measured with the dice metric, and higher than 88% when measured with the jaccard metric. Table 5.4 shows the performance of the model.

Table 5.4: u-net segmentation results of endocardium (EN) and epicardium (EP) regions in LGE-100 dataset, using the SS-DCE cost function .

SS-DCE cost function results of LGE-MRI					
Metric	1st cycle	2nd cycle	3rd cycle	4th cycle	5th cycle
jaccard (EP)(%))	88.60	88.70	88.86	88.18	89.95
jaccard (EN)(%)	90.41	90.84	92.56	90.07	94.86
dice (EP)(%)	94.03	94.06	94.12	93.63	94.75
dice (EN)(%)	95.01	95.23	96.11	94.85	97.51

5.3.2 Deep learning regularization techniques

Regarding the DL networks, we need to use regularization techniques to avoid any overfitting of the training datasets. We have used data augmentation techniques including rotation (rotation around the center of the image by a random angle in the range of 0°–15°), width shift range (width shift of the image by up to 20 pixels), height shift range (height shift of the image by up to 20 pixels), ZCA whitening (add noise in each image, [100]) and horizontal - vertical flips of images.

Finally, we constrained the weights incident to each hidden unit to have a norm less than or equal to 3 for each convolution layer of the modified u-net network.

5.3.3 Scar segmentation pipeline architecture

Figure 5.1 outlines the structure of the scar automatic segmentation pipeline. This pipeline is the same as MA-SOCRATIS. The main parts are the RGMM framework and the watershed seed point algorithm.

5.3.4 Ground truth estimation

The evaluation of automatic segmentation should be robust. Thus, a generalized and unbiased ground is crucial. In this study, three experts did manual segmentation of both myocardium and scar regions in each LGE-MR image with inter-observer variability, and one expert repeated the

manual segmentation three times, so that intra-observation variability could be determined. The intra-observation and inter-observer variability were then computed using the best voting threshold ([109]).

The results of automatic segmentation are compared against these GTs by computing the dice, sensitivity and specificity metrics.

5.4 Implementation of pipeline

5.4.1 Cardiac MR images and image analysis

In this study, we used cardiac cine-MR and LGE-MR images for training different components of our processing pipeline, and for comparing the performance of the pipeline against a ground truth determined from manual segmentation.

5.4.2 Training and test data

The ROI deep learning network of the initial segmentation step for myocardium segmentation was trained using the MICCAI 2009 LV segmentation challenge dataset. To train the u-net network for cine-MRI segmentation, we used the MICCAI 2017 whole heart segmentation challenge dataset.

In the present study, we assembled three new datasets (cine-40, cine-100, and LGE-100) comprising 900 short-axis LGE-MR images (LGE-100) from 100 patients and 900 short-axis cine-MR images from 100 patients (cine-100). All images were acquired, anonymized, and stored in DICOM format. The LGE/cine-MRI dataset was obtained from a collection of three different MRI machine vendors with a variation in the initial settings used for a Philips MRI scanner. All patients gave informed consent analysis of all LGE-MRI images as well as the manual segmentation of the endocardium and epicardium regions and scars of the left ventricle was performed offline using MASS software (Version 2016EXP, Leiden University Medical Center, Leiden, The Netherlands). The

region of interest (ROI) was defined on the center of the endocardium region of each image. Image analysis was performed for all slices to reduce the effect of noise and increase the signal to noise ratio (SNR). A single expert did manual segmentation of both epicardial and endocardial LV contours. Lastly, we assembled the cine-40 dataset comprising 436 short-axis cine-MR images from 40 patients randomly selected from the 100 patients of the cine-100 cohort above.

We used a short-axis LGE-MR image of 20 patients (LGE-20) to evaluate the DL-SOCRATIS. All images were acquired, anonymized, and stored in DICOM format. The LGE-MRI dataset was obtained from a collection of three different MRI machine vendors with a variation in the initial settings used for a Philips MRI scanner. All patients gave informed consent. The region of interest (ROI) was defined on the center of the endocardium region of each image. For myocardium regions, we evaluated our system on the data from 20 patients, described in Tao et al. ([176]). The images from each patient comprised between 8 and 12 short axis slices. Three experts segmented these images by identifying the LV endocardial and epicardial surfaces, and the borders of scar. Segmentation was done using MASS (research version 2017; Leiden University Medical Center, Leiden, the Netherlands).

MA-SOCRATIS and DL-SOCRATIS results of automatic segmentation of myocardial LV in LGE-MRI with scar was evaluated based on a cross-institution scheme; the LGE-20 and 40 patients of MS-CMRSeg 2019 (STACOM at MICCAI 2019) challenge datasets.

The results of both deep learning segmentation pipelines were compared against the manual ground truth (GT) by computing the dice and jaccard accuracy metrics.

5.4.3 Image analysis

We use the same pre-processing framework as in MA-SOCRATIS pipeline (Section 4.2.1). Again, the manual GT binary images of the myocardium - scar region and the automatic segmentation image results were modified using a post-processing framework as in MA-SOCRATIS. Finally, all

automatic segmentation binary images had to be aligned with each corresponding manual inter-intra-ground truth image. The code implementation is available on a public repository with url: https://github.com/INSIGNEO/DL-SOCRATIS.

5.5 Next chapter

In the next chapter, we will present the results of the MA-SOCRATIS and DL-SOCRATIS myocardial segmentation and BZ-SOCRATIS and MA-SOCRATIS total scar, border zone and crore zone scar of LV cardiac chamber in LGE-MR images.

6

Comparison and evaluation of pipelines

6.1 Overview

In this chapter, we present an evaluation of our unsupervised pipelines (MA-SOCRATIS, DL-SOCRATIS, and BZ-SOCRATIS) using cross-institute LGE-MRI cohorts (LGE-100, LGE-20, MS-CMRSeg 2019 in subsection 5.3.8) based on unbiased ground truths (intra-observation, inter-observer variability and each observation individual). This chapter includes Section 6.2, which presents the results of MA-SOCRATIS in LGE-20 and MS-CMRSeg 2019 challenge cohorts. Section 6.3 shows the results of DL-SOCRATIS two challenges and a comparison of DL-SOCRATIS, MA-SOCRATIS state of the art techniques. Section 6.4 describes the results of automatic extraction of the scar border zone using the BZ-SOCRATIS LGE-20 cohort images.

6.2 MA-SOCRATIS automatic segmentation of LV with scars

6.2.1 Automated segmentation of myocardium and scar

The automated segmentation of the myocardium and scar was evaluated against both intra-observation and inter-observer GTs, and these comparisons are summarised in Table 6.1. The overall accuracy is given as a mean dice, sensitivity, and specificity score with standard deviations, which were evaluated over all 167 MR images included in the test set. The dice score measures the overall similarity of the pipeline and ground truth classifications, whereas the sensitivity and specificity measure how well the pipeline identifies regions of myocardium and scar relative to the ground truth. To further evaluate the accuracy and robustness of our myocardium automatic pipeline in different cohorts, we tested it in the 40 patients cohort of the MS-CMRSeg 2019 challenge dataset (STACOM at MICCAI 2019). Table 6.1 shows the results.

Overall, the dice scores for myocardium segmentation were higher than for scar segmentation, with only small differences between comparisons of intra-observation and inter-observer ground

Table 6.1: Summary of automatic segmentation of myocardium and scar compared to intra-observation and inter-observer ground truth. Scar segmentation performance was measured with either manual (MM) or automatic myocardium segmentation (AM)

Intra-Inter observer results				
Region	dice (%)	Sensitivity (%)	Specificity (%)	Compared with
Myocardium	81.9 ± 9.4	91.9 ± 11.0	95.9 ± 1.2	Intra Ground Truth
Myocardium	78.1 ± 9.3	94.2 ± 11.8	95.1 ± 1.2	Inter Ground Truth
Myocardium	70.0 ± 5.5	90.5 ± 7.0	93.0 ± 1.7	MS-CMRSeg 2019 challenge
Scar	70.5 ± 11.7	75.0 ± 12.5	99.6 ± 0.3	Intra Ground Truth (MM)
Scar	29.5 ± 8.0	31.6 ± 9.0	98.1 ± 0.7	Intra Ground Truth (AM)
Scar	70.5 ± 17.2	85.4 ± 7.5	99.6 ± 0.2	Inter Ground Truth (MM)
Scar	24.4 ± 9.0	36.5 ± 11.2	98.0 ± 0.7	Inter Ground Truth (AM)

truth. In MS-CMRSeg 2019 dataset, the accuracy reduced to 70 %. MA-SOCRATIS achieved these results without the need of training, tuning, or transfer learning in both cohorts.

The accuracy of automated scar segmentation was the same for both intra-observation and inter-observer ground truth. The accuracy of scar segmentation depends on the myocardium segmentation, and so we used the intra-observation and inter-observer ground truths for myocardium segmentation before applying the scar segmentation pipeline. The sensitivity of scar segmentation was lower than for myocardium segmentation, whereas the specificity was very high. This can be explained because the binary scar region has a smaller number of pixels compared to the number of pixels classified as myocardium. The sensitivity will therefore have a higher variance as the sensitivity is calculated from the number of white pixels, and the shape, size, and location of the scars are less smooth than the myocardium. Figure 6.1 illustrates the performance of automated myocardium segmentation with respect to inter-observer variability in three cases where the accuracy (mean dice) was lower than average (<70%), close to the average (70-80%), and higher than average (>80%). Although the top two rows show good accuracy with respect to the manual observations, our automated method has identified the papillary muscle in the myocardium region for this case

Figure 6.1: Myocardium segmentation and inter-observer ground truth. Comparison of automatic segmentation, manual segmentation by 1st and 3rd observer, and overall inter-observer ground truth for images where the automated myocardial segmentation showed high accuracy (top row), mid-range accuracy (middle row), and low accuracy (bottom row) compared to ground truth. Right hand column shows the image ROI in each case.

(Figure 6.1), whereas this has been ignored by the manual observers. MA-SOCRATIS pipeline can include or exclude the papillary muscle in the automatic myocardial segmented area. For this study we extract the metric results (dice, sensitivity, and specificity) by excluding the pupillary muscle, as the experts did not include it in the manual segmentation GTs. This is an advantage of our pipeline, because in some cases ([53], [36]) it is crucial, the detection of papillary muscle as part of ventricle myocardium. However, the papillary muscle should not be included in calculations such as ejection fraction or myocardial mass. In those cases, our pipeline can ignore the papillary. The papillary muscles are detected in the active contour technique of the MA-SOCRATIS re-estimation step, and it can be removed by utilizing merge filters (ITK libraries). In the bottom row, the myocardium is very clear on the image, but the low intensity of the surrounding tissue has resulted in patchy automated segmentation. These observations are consistent with the examples of myocardial segmentation compared to intra-observation variability shown in Figure 6.2, where the detection of papillary mus-

Figure 6.2: Myocardium segmentation and intra-observation ground truth. Comparison of automatic segmentation, manual segmentation with 1st and 2nd observations, and overall intra-observation ground truth for images where the automated myocardial segmentation showed high accuracy (top row), mid-range accuracy (middle row), and low accuracy (bottom row) compared to ground truth. Right hand column shows the image ROI in each case.

cle by automated segmentation is even more pronounced.

Scar segmentation is illustrated in Figures 6.3 and 6.4, for inter-observer and intra-observation variability, respectively. For large scars (top row in both figures), the accuracy of the automated pipeline is high, and the consistency between and within manual observers is also high. For smaller scars, it becomes harder for both automated and manual segmentation to distinguish between scars and noise in the image.

6.2.2 Accurate myocardium segmentation is important for subsequent scar segmentation

In our pipeline, the segmentation of an LGE image proceeds in two stages: segmentation of the myocardium and then segmentation of scar regions. In the results reported above, we have treated these two parts of the pipeline separately. When we combined automated myocardium segmentation

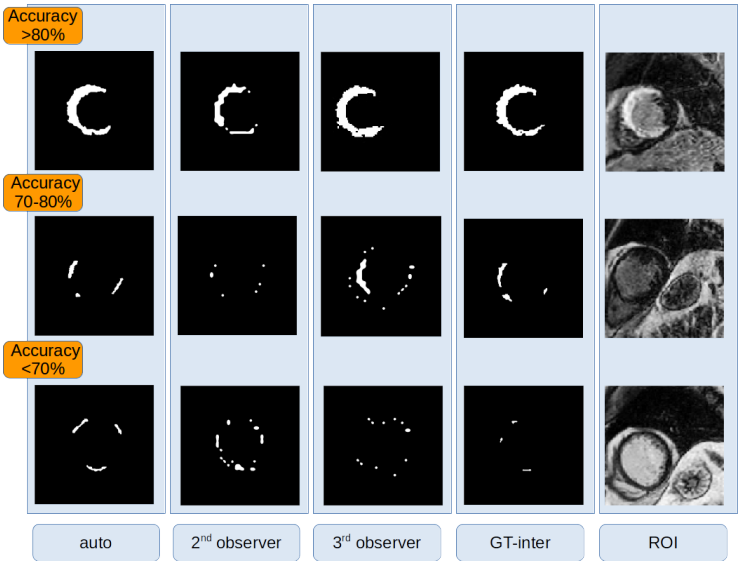

Figure 6.3: Scar segmentation and inter-observer ground truth. Comparison of automatic segmentation of scar, manual segmentation of scar with 1st and 3rd observer, and overall inter-observer ground truth for images where the automated scar segmentation showed high accuracy (top row), mid-range accuracy (middle row), and low accuracy (bottom row) compared to ground truth. Right hand column shows the image ROI in each case.

Figure 6.4: Scar segmentation and intra-observation ground truth. Comparison of automatic segmentation, manual segmentation with 1st and 2nd observations, and overall intra-observation ground truth for images where the automated scar segmentation showed high accuracy (top row), mid-range accuracy (middle row), and low accuracy (bottom row) compared to ground truth. Right hand column shows the image ROI in each case.

with automated scar segmentation, then we found that errors in automated myocardium segmentation were associated with a loss of accuracy for scar segmentation, and these findings are summarised in Table 6.1. The accuracy of our scar segmentation pipeline based on manual segmentation was 70.5% with respect to inter-observer or intra-observation ground truth (Table 6.1), but the accuracy fell to less than 30% when our automatic myocardial segmentation pipeline was used. The reasons for this finding are illustrated in Figure 6.5, which shows a set of images for a single patient where the myocardium was segmented automatically with high accuracy relative to manual myocardium segmentation. The four columns on the left of the figure highlighted in blue show that the myocardium is distinct in the raw LGE-MR image, and there is generally a good correspondence between the automatic segmentation and the inter-observer and intra-observation ground truths. The four columns highlighted in orange show the inter-observer and intra-observation ground truths for the scar, together with the automatic scar segmentation based on the manual myocardial (MM) segmentation shown in columns 2 and 3. The final column shows automatic scar segmentation based on the automatic myocardium (AM) segmentation shown in column 1.

With automatic myocardium (AM) segmentation, more scar pixels are identified close to the endocardial surface compared to both manual scar segmentation and automatic scar segmentation based on manual myocardium (MM) segmentation. In this example, the endocardial pixels account for the differences in scar segmentation with the AM and MM approaches, and this effect can be traced back to the myocardial segmentation in column 1 of Figure 6.5.

6.2.3 MA-SOCRATIS extra results

Extra results of automatic segmentation cohort

The automated segmentation of the myocardium and scar was evaluated against both inter-observer and intra-observation ground truth, and these comparisons were summarised in Table 6.1. Fig-

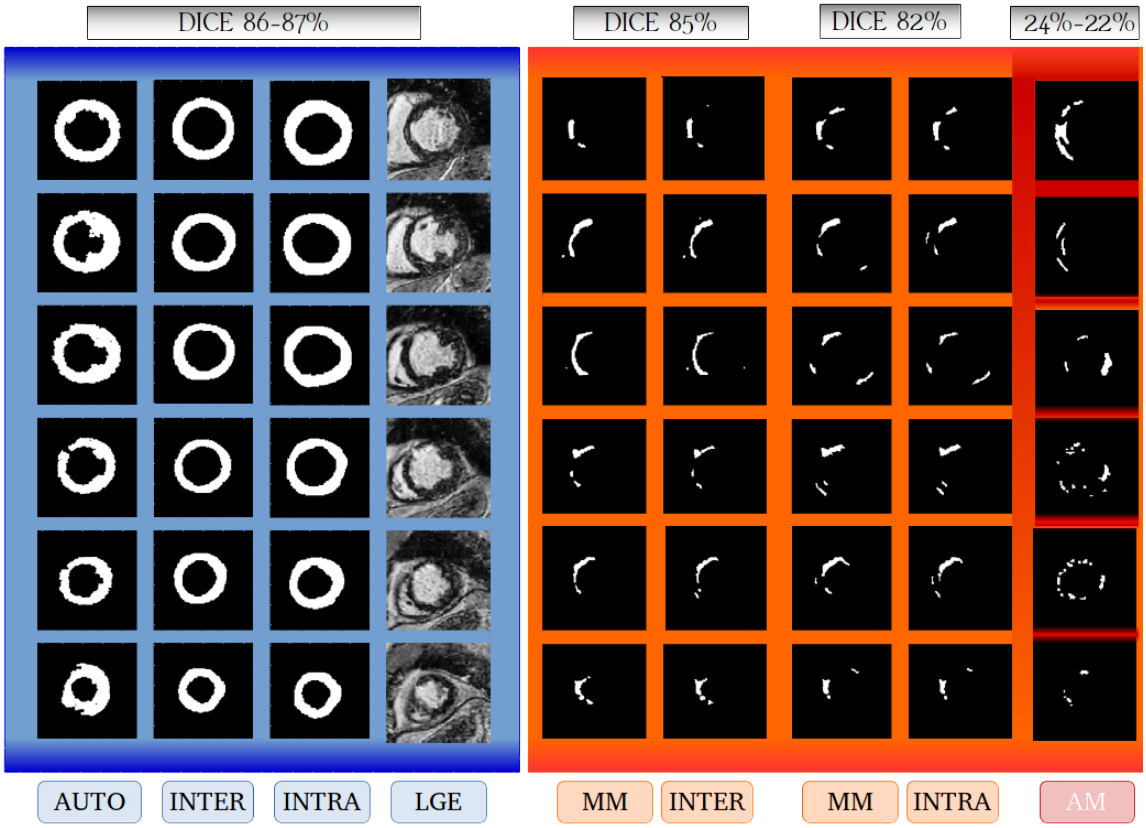

Figure 6.5: Images from a single patient showing comparison of: i) automated myocardium segmentation with manual inter-observer and intra-observation ground truth (blue), ii) automated scar segmentation with manual myocardium segmentation (MM) compared to inter-observer and intra-observation ground truth (orange), and iii) scar segmentation with automated myocardium segmentation (AM, red).

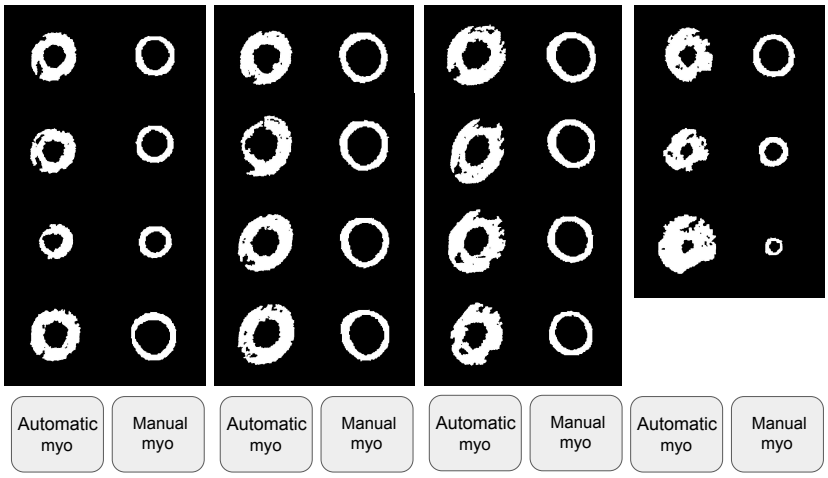

Figure 6.6: Myocardium and scar results of inter-observer manual segmentation and automatic segmentation pipeline.

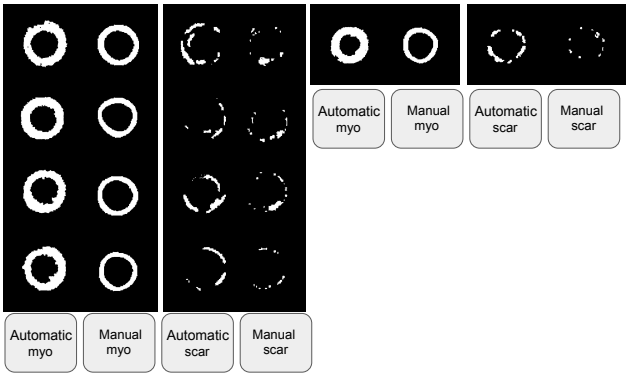

Figure 6.7: Myocardium and scar results of inter-observer manual segmentation and automatic segmentation pipeline.

ures 6.6-6.14 illustrate the performance of automated scar/myocardium segmentation with respect to inter-observer variability in the LGE challenge data of the 20 patients. Here we illustrate 7 patients. Based on these figures, it is clear that dice metrics are not the best metric to evaluate the scar regions even if the scar regions are detected. This can be explained as noise and artifacts between the images affect the accuracy of the dice metric. The over and under segment of LV could be the shape of multi-atlas process. As the MRI slices mis-align with multi-atlas slices (registration error) and the circle estimation process is under determined.

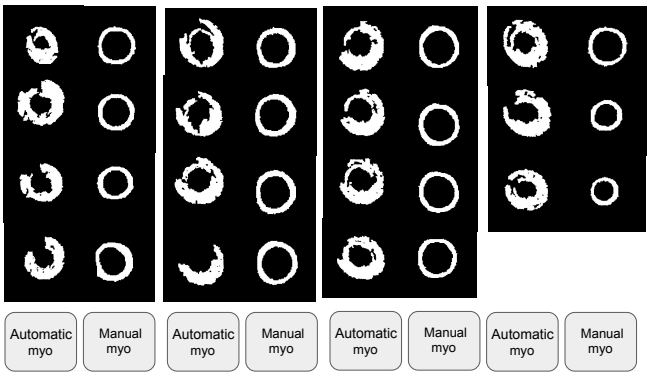

Figure 6.8: Myocardium and scar results of inter-observer manual segmentation and automatic segmentation pipeline.

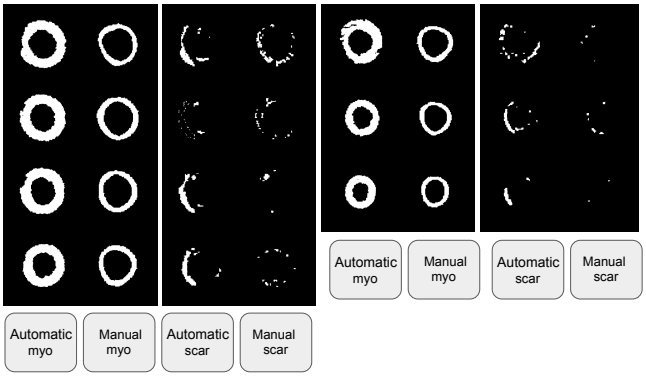

Figure 6.9: Myocardium and scar results of inter-observer manual segmentation and automatic segmentation pipeline.

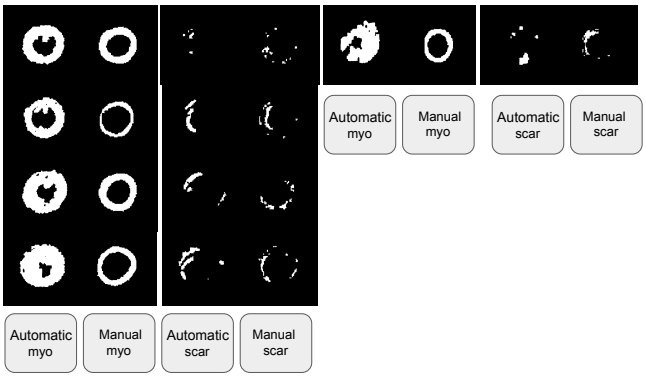

Figure 6.10: Myocardium and scar results of inter-observer manual segmentation and automatic segmentation pipeline.

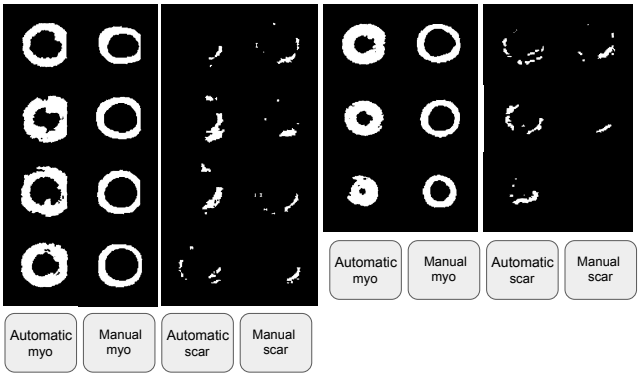

Figure 6.11: Myocardium and scar results of inter-observer manual segmentation and automatic segmentation pipeline.

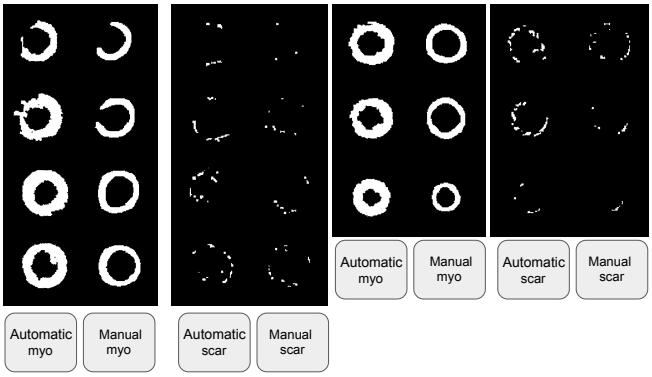

Figure 6.12: Myocardium and scar results of inter-observer manual segmentation and automatic segmentation pipeline.

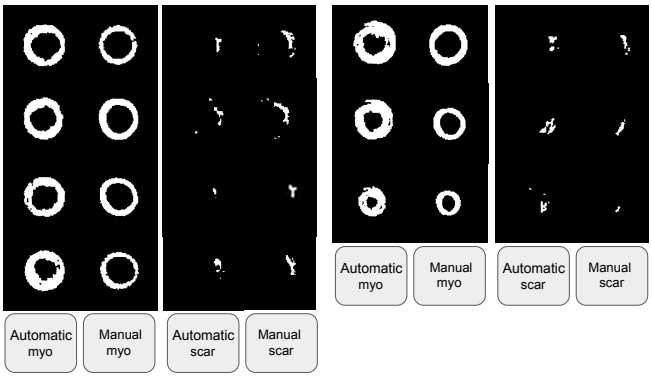

Figure 6.13: Myocardium and scar results of inter-observer manual segmentation and automatic segmentation pipeline.

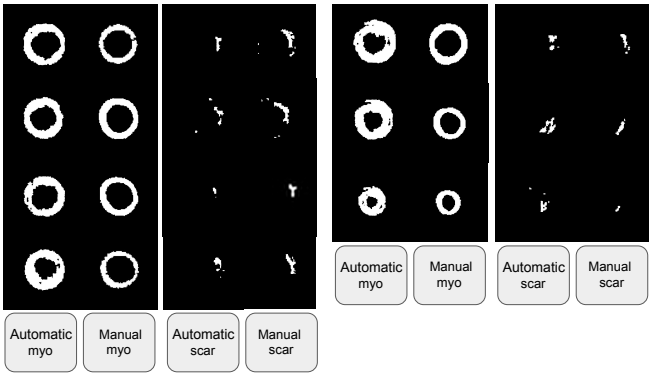

Figure 6.14: Myocardium and scar results of inter-observer manual segmentation and automatic segmentation pipeline.

6.3 DL-SOCRATIS automatic segmentation of myocardium of LV.

In DL-SOCRATIS, the multi-atlas segmentation part is replaced by two different deep learning networks DLc-SOCRATIS and DLl-SOCRATIS. DLc-SOCRATIS was a network trained using cine-MRI data (cine-SOCRATIS and CGAN-SOCRATIS), and DLl-SOCRATIS was a network trained using LGE-MRI data (DL-SOCRATIS).

6.3.1 Comparison with MA-SOCRATIS

We trained the u-net network using cine-MRI and evaluated in a LGE-MRI dataset. We estimate that this cross-modality application will probably reduce the accuracy of the segmentation task. The Fig. 6.15, 6.17 confirmed this estimation. The LGE-MRI captures the infarct region better than the cine-MRI scan modality. The scar region is not detectable in the cine-MR images. As a result, the cine-SOCRATIS network did not detect efficiently the infarct scar region for a cross-modality task. Thus, the dice accuracy of myocardium segmentation in the cine-SOCRATIS is worst than this of DL-SOCRATIS The circle binary mask of epicardium and endocardium is given by the u-net. Fig. 6.16 shows the results of DL-SOCRATIS and MA-SOCRATIS, respectively, from the same patient. We compare the automatic myocardium segmentation (AM) scar segmentation scheme, as

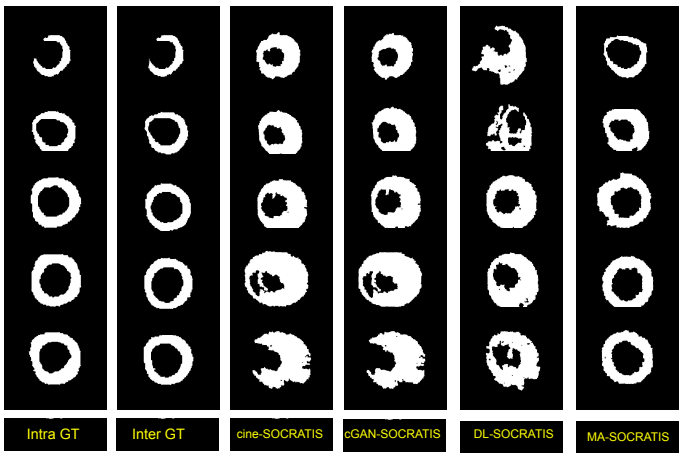

Figure 6.15: Results of MA-SOCRATIS, DL-SOCRATIS, cine-SOCRATIS and CGAN-SOCRATIS pipelines in LGE-20 dataset about myocardium lesions, compared with intra GTs.

it references above. The myocardium segmentation accuracy of this patient for intra-inter cases is 44% and 43% for the MA-SOCRATIS and 56 % and 71 % for the DL-SOCRATIS. This is the worst accuracy of MA-SOCRATIS. The main reason for these results is fault align of LGE-MRI images because of the multi-atlas techniques. The scar AM segmentation accuracy of this patient for intra-inter cases is 14% and 30% for the MA-SOCRATIS and 12 % and 17 % for the DL-SOCRATIS. In this specific case, the accuracy of DL-SOCRATIS is better than that of MA-SOCRATIS for the myocardium region and worse for the scar region. A possible explanation can be that the initial myocardium segmentation was more accurate in DL-SOCRATIS than MA-SOCRATIS. Figures 6.15, 6.17 highlight the results of myocardium segmentation for the cine-SOCRATIS, CGAN-SOCRATIS, DL-SOCRATIS, and MA-SOCRATIS for two random patients. Figure 6.15 shows that MA-SOCRATIS has the most accurate results based on intra-GT and inter-GT compared with others pipelines. Regarding the deep learning pipeline, CGAN-SOCRATIS is slightly more accurate than cine-SOCRATIS. DL-SOCRATIS is more accurate about the three last images and CGAN-SOCRATIS is more accurate about the first two. Figure 6.17 shows that again MA-

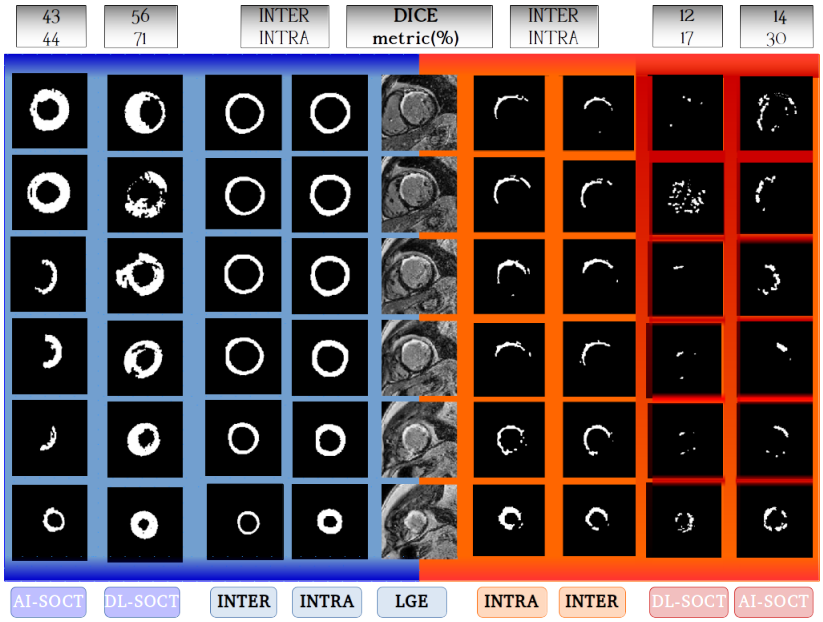

Figure 6.16: Worst patient results of MA-SOCRATIS pipeline; A comparison of MA-SOCRATIS and DL-SOCRATIS results of myocardial and scar lesions. Images from one patient showing comparison of: i) automated myocardium segmentation with manual inter-observer and intra-observation ground truth (blue), ii) automated scar segmentation with manual myocardium segmentation (MM) compared to inter-observer and intra-observation ground truth (orange), and iii) scar segmentation with automated myocardium segmentation (AM, red).

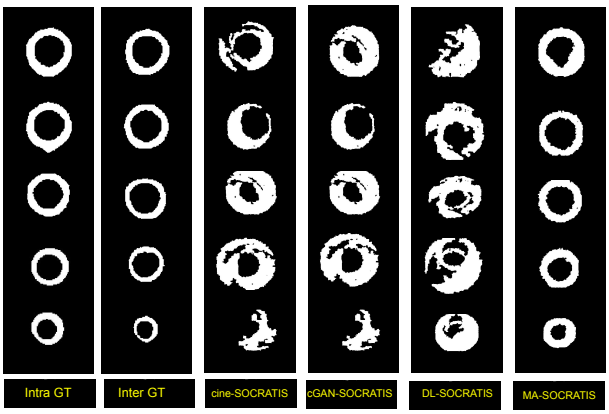

Figure 6.17: Results of MA-SOCRATIS, DL-SOCRATIS, cine-SOCRATIS and CGAN-SOCRATIS pipelines in LGE-20 dataset about myocardium lesions compared with inter GTs.

SOCRATIS has the most accurate results based on intra-GT and inter-GT compared with the three DL pipelines. Regarding the deep learning pipeline, CGAN-SOCRATIS is again slightly more accurate than cine-SOCRATIS. DL-SOCRATIS is more accurate about the second and last images. On the other hand, CGAN-SOCRATIS is more accurate about the first, third, and forth images. These behaviours between the three DL pipelines can be explained as when the myocardium has not a large portion of the scar region, it can be detected more accurately in the cine-MRI images. Thus, the CGAN-SOCRATIS and cine-SOCRATIS had better results comparing the DL-SOCRATIS model. The table 6.2 shows the total mean values and standard deviations of all patients for the two pipelines. The myocardium segmentation accuracy of MA-SOCRATIS is 81.9 and 78.1 % for intra-observation and inter-observer GTs, respectively, contrary to DL-SOCRATIS, which achieves 60.4 and 55.0 %. Regarding, the scar segmentation accuracy, MA-SOCRATIS achieves 24.4 and 29.5 % for intra and inter observer GTs respective contrary to DL-SOCRATIS, which achieves 8.5 and 12.5 %. Regarding the robustness of the pipelines, MA-SOSCRATIS has the smallest standard deviation (SD). DL-SOCRATIS has the lower SD compared to the CGAN-SOCRATIS and cine-SOCRATIS pipelines. This can be explained as the DL-SOCRATIS is one modality analysis

Table 6.2: Summary of myocardial scar segmentation results comparing MA-SOCRATIS ([119]), DL-SOCRATIS, cine-SOCRATIS and CGAN-SOCRATIS with intra/inter-observer ground truths. The results are the average values of dice, sensitivity and specificity metrics +/- the standard deviation (SD)

Intra observer variation results				
Region	dice +/- SD (%)	Sensitivity +/- SD (%)	Specificity +/- SD (%)	Technique
Myocardium	**81.9** +/- 9.4	91.9 +/- 11.0	95.9 +/- 1.2	MA-SOCRATIS
Myocardium	60.38 +/- 10.1	70.87 +/- 15.0	92.38 +/- 2.5	DL-SOCRATIS
Myocardium	57.8 +/- 10.8	61.4 +/- 15.9	93.5 +/- 2.8	CGAN-SOCRATIS
Myocardium	56.5 +/- 12.2	60.7 +/- 16.1	93.5 +/- 2.8	cine-SOCRATIS
Scar	24.4 +/- 9.0	36.5 +/- 11.2	98.0 +/- 0.7	MA-SOCRATIS
Scar	12.15 +/- 6.4	16.1 +/- 8.3	97.0 +/- 1.1	DL-SOCRATIS
Scar	10.4 +/- 5.5	13.6 +/- 7.9	97.0 +/- 1.6	CGAN-SOCRATIS
Scar	10.1 +/- 5.4	12.6 +/- 8.2	97.1 +/- 1.6	cine-SOCRATIS
Inter observer variation results				
Region	dice +/- SD (%)	Sensitivity +/- SD (%)	Specificity +/- SD (%)	Technique
Myocardium	**78.1** +/- 9.3	94.2 +/- 11.8	95.1 +/- 1.2	MA-SOCRATIS
Myocardium	55.0 +/- 11.1	70.0 +/- 18.8	91.6 +/- 3.0	DL-SOCRATIS
Myocardium	51.7 +/- 11.5	61.3 +/- 18.2	92.2 +/- 3.7	CGAN-SOCRATIS
Myocardium	50.0 +/- 12.6	60.0 +/- 20.0	91.3 +/- 5.5	cine-SOCRATIS
Scar	29.5 +/- 8.0	31.6 +/- 9.0	98.1 +/- 0.7	MA-SOCRATIS
Scar	8.5 +/- 5.2	17.7 +/- 10.3	96.9 +/- 10.7	DL-SOCRATIS
Scar	8.3 +/- 4.9	17.7 +/- 15.1	97.0 +/- 1.6	CGAN-SOCRATIS
Scar	8.0 +/- 5.7	17.3 +/- 15.5	96.9 +/- 1.7	cine-SOCRATIS

pipeline contrary to the CGAN-SOCRATIS and cine-SOCRATIS which are cross-modality networks (trained in cine-MRI applied in LGE-MRI).

To conclude, in this cohort the multi-atlas technique delivers more accurate results than the DL. However, the DL pipelines are less time consuming than multi-atlases (however, DL needs more time in training) and they do not need any landmarks. In addition, the deep learning pipeline can be used as a cine-MRI modality analysis network; in this case, the expert radiologists have only cine-MRI modalities to scan images for a specific patient (cine-SOCRATIS pipeline).

6.4 BZ-SOCRATIS automatic segmentation of border zone and core scar of LV

In the second task, we studied the semi-automatic scar segmentation accuracy of our pipeline in the border zone (BZ) and core scar (CS) regions. BZ and CS segmentation for intra-observation and inter-observer GTs are illustrated in Figures 6.18, 6.20 and Figures 6.19, 6.21 respectively. Figures 6.18 and 6.19 show that in the majority of the cases, where a large size of BZ and CS regions are detected, accurately segmented, and the consistency between and within the intra/inter-manual GTs are high. For small BZ and CS areas, it becomes harder in both GTs to distinguish between the region of interest (ROI) and noise in the image. Figures 6.20 and 6.21 show that in some cases the accuracy of our pipeline (Automatic) is higher than the GTs based on the LGE-MRI observation. This was happening as the manual scar segmentation was based on FWHM techniques and the experts determine 'by hand' only the epicardial and endocardial contours. As the manual contour methodology for border zone or core scar determination is very time consuming and highly dependent on the image resolution, the clinician and expert radiologists prefer to use thresholding techniques. As a result, the manual segmentation of the scar can not be considered a purely manual process. However, even if the human interaction and the human error can not be determined in this

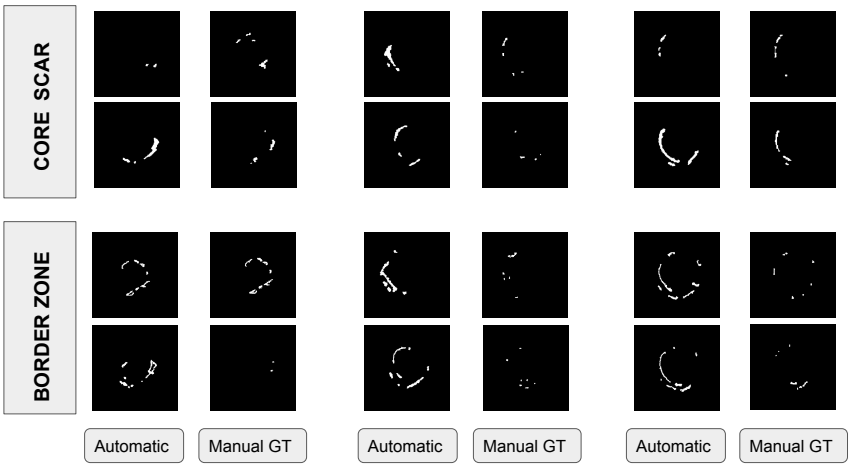

Figure 6.18: Results of border zone and core scar segmentation. 1^{st} row: the intra observation ground truth (Manual GT), 2^{nd} row: the automatic results of our method (Automatic) and 3^{rd} row: the left and right ventricle regions of LGE-MR input slice. The top two columns are the core scar area segmentation and the bottom two columns the border zone area segmentation.

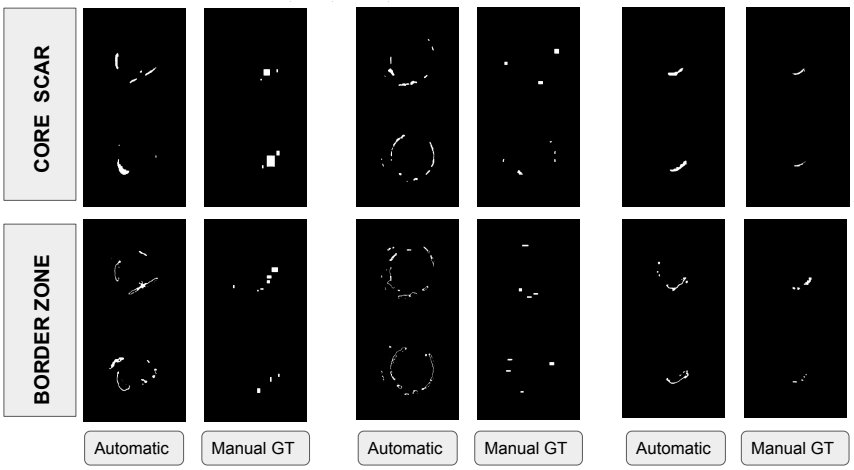

Figure 6.19: Results of border zone and core scar segmentation. 1^{st} row: the inter observation ground truth (Manual GT), 2^{nd} row: the automatic results of our method (Automatic) and $3^{rd} row$: the left and right ventricle regions of LGE-MR input slice. The top two columns are the core scar area segmentation and the bottom two columns the border zone area segmentation.

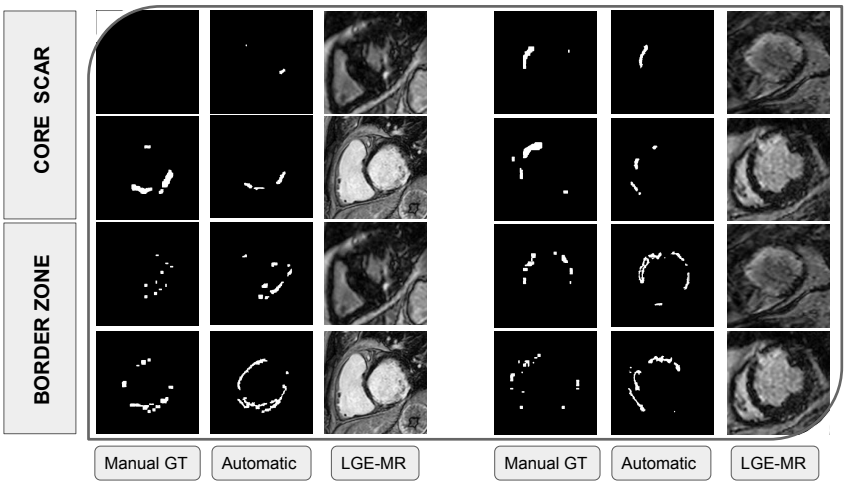

Figure 6.20: Results of border zone and core scar segmentation. 1^{st} row: the intra observation ground truth (Manual GT), 2^{nd} row: the automatic results of our method (Automatic) and 3^{rd} row: the left and right ventricle regions of LGE-MR input slice. The top two columns are the core scar area segmentation and the bottom two columns the border zone area segmentation.

process (as there is no contour drawing of scar regions); the error is still present as the FWHM needs a human selection of the highest intensity (scar) and the lowest intensity (healthy myocardium) pixels. In addition, the human error in the extraction of scar GTs is affected by the manual contour determination of endo and epi (as we will discuss in Section 6.1). Therefore, we can claim that in the process of GTs extraction, there exists a variation and uncertainty because of the inter-observer and intra-observation variability.

To conclude, we will end up comparing our automatic pipeline with the FWHM or other thresholding techniques. This is the reason that in some cases our pipeline performs more accurately than the GTs (Figures 6.20 and 6.21)

6.5 Next chapter

In the next chapter, we will discuss about the robustness of the MA-SOCRATIS pipeline regarding different GTs, and the statistical significant of the different deep learning methods we used in DL-

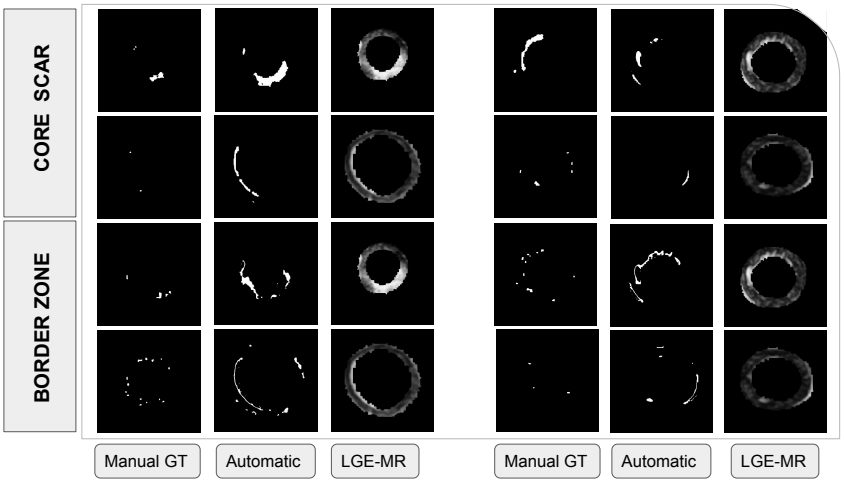

Figure 6.21: Results of border zone and core scar segmentation. 1^{st} row: the inter observation ground truth (Manual GT), 2^{nd} row: the automatic results of our method (Automatic) and 3^{rd}*row*: the left and right ventricle regions of LGE-MR input slice. The top two columns are the core scar area segmentation and the bottom two columns the border zone area segmentation.

SOCRATIS pipeline. Moreover, we will compare the results of MA-SOCRATIS, DL-SOCRATIS, and BZ-SOCRATIS pipelines in myocardial and scar automatic segmentation with established and state of the art techniques. We will conclude by summarizing the limitations, advantages, and future work of the SOCRATIS pipeline.

7

Discussion and future work

In this chapter, we present a statistical significant analysis of the different versions of the pipelines we used (DL-SOCRATIS, cine-SOCRATIS, CGAN-SOCRATIS, and MA-SOCRATIS). Furthermore, we compare the results of the performance of our two pipelines (DL-SOCRATIS and MA-SOCRATIS) with state of the art and established techniques, and we define the SOCRATIS framework. Moreover, we discuss the BZ-SOCRATIS pipeline and the significant analysis of the different thresholding portions in the BZ-TH technique. Lastly, we discuss the importance and

Table 7.1: Summary of intra-observation and inter-observer variability in manual segmentation of myocardium and scar compared to overall ground truth (GT)

Observation	Intra Myo dice (%)	Intra Scar dice (%)	Inter Myo dice (%)	Inter Scar dice (%)
First	90.9	50.5	90.5	44.4
Second	85.2	41.6	86.4	38.2
Third	88.7	44.8	89.6	43.6
GT	100	100	100	100

robustness of our automated pipeline and conclude by considering both the limitations of our approach and the next steps of development.

7.1 Discussion of MA-SOCRATIS pipeline

7.1.1 Robustness of the pipeline in different GT cases

To illustrate the differences in manual segmentation between and within individual experts, Figure 7.1 shows an example of manual myocardial and scar segmentation by three independent observers. In Figure 7.1(left), the inter-observer differences for myocardial segmentation can be clearly seen in the upper row, and the lower row illustrates how these differences were resolved using the best threshold approach to provide the ground truth. Figure 7.1(right) shows that the inter-observed differences for scar segmentation were greater than for myocardial segmentation. The intra-observation variability was generally less than inter-observer variability.

In addition Table 7.1 shows the effectiveness of GT determination. As it can be observed, there is a consistent variability within the three observers (inter-variability) and three observations (intra-variability). The manual myocardium segmentation was much more repeatable than scar segmentation as we expected (the artifact and noise of the scar region increase the complexity of a repeatable segmentation).

To check the robustness of our pipeline, we evaluated with respect to both single expert observa-

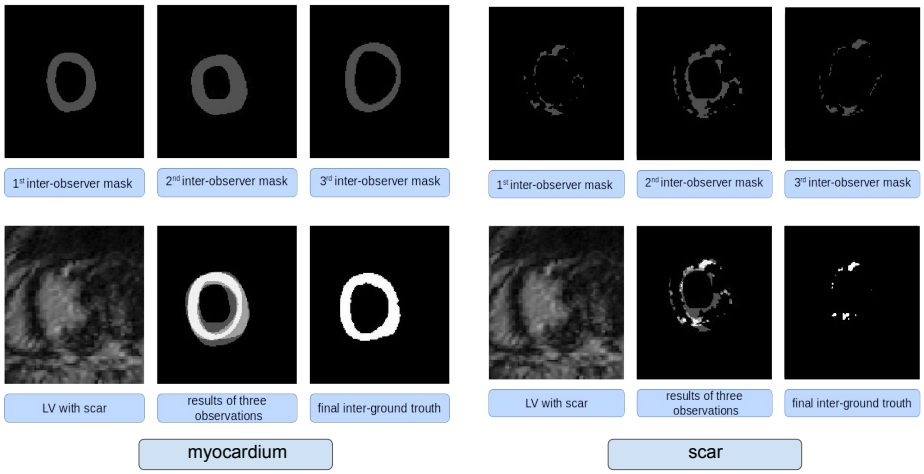

Figure 7.1: Extraction of unbiased inter-observer ground truth for myocardium (left) and scar (right).

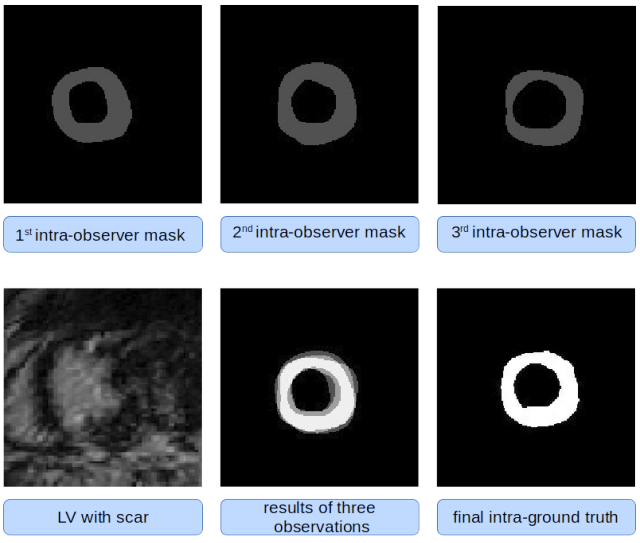

Figure 7.2: Extraction of unbiased myocardium intra-observation variation ground truth.

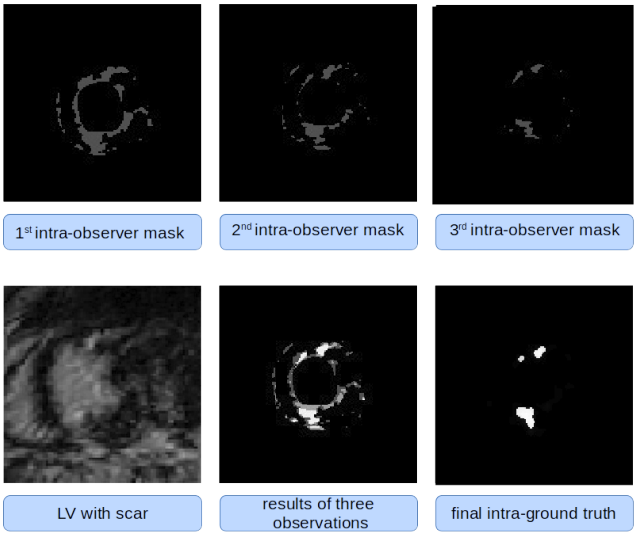

Figure 7.3: Extraction of unbiased scar intra-observation variation ground truth.

tions and each intra-observation of scar and myocardium lesions (single-case). By this method, we could capture the variation in sensitivity, accuracy, and specificity metrics of our two pipelines.

Figures 7.4 and 7.5 show the variation of sensitivity, specificity, and dice value results by changing the ground truth inter-observer and intra-observation variability. We used the terminology: Myo_Sens_GT for myocardial sensitivity total ground truth, Myo_Spec_O1 for myocardial specificity observer/ation one ground truth, and Myo_dice_O2 for myocardial dice observer/ation two ground truth.

Figure 7.4 and Figure 7.5 show that the sensitivity, specificity, and dice metrics are less robust in the cases of both one observation and one observer compared to the total GT. The standard deviation, especially in the sensitivity metric, was also higher in single observations and with a single observer compared to the total GT, and the mean values of the metrics are in some cases an overestimation or underestimation of the total GT.

Figures 7.1(left), and 7.2 show examples of myocardium segmentation with different observers and observations. By the extraction of an unbiased truth, we can evaluate our pipeline in a more

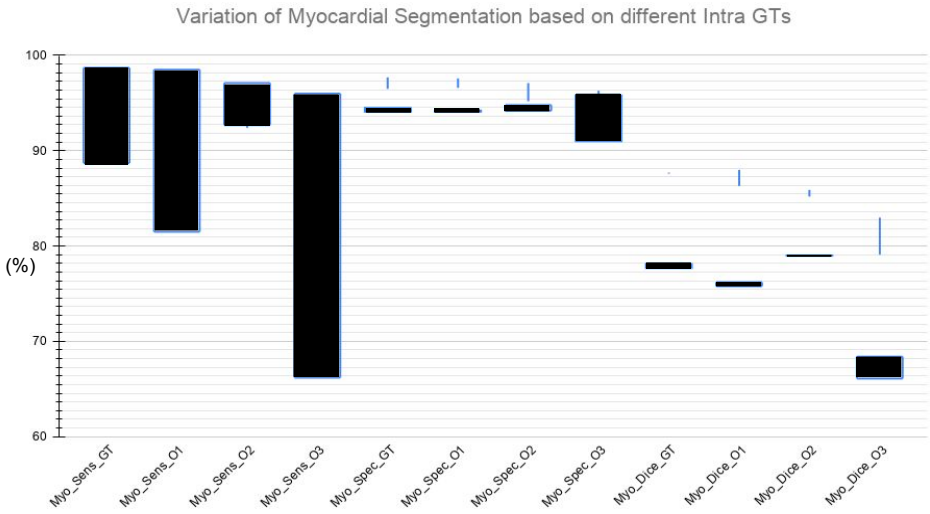

Figure 7.4: Variation in sensitivity, dice and specificity metrics of automatic myocardial segmentation by changing the intra observation ground truths.

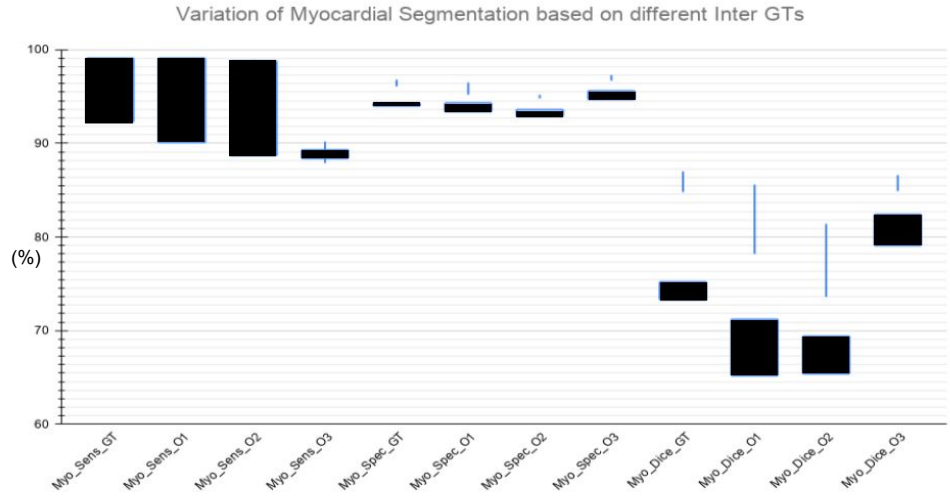

Figure 7.5: Variation in sensitivity, dice and specificity metrics of automatic myocardial segmentation by changing the inter observation ground truths

generalized way. Therefore, our pipeline accuracy is less sensitive to observation bias, something that can be a problem with other automated segmentation techniques such as deep learning networks. Moreover, by the extraction of the unbiased GT, the effect of noise is reduced between observations, especially in scar regions. In the examples shown in figures 7.1(right) and 7.3, the variability of the first, second, and third observation observers and the reduced variability of the GT can be seen in each case (intra-inter). The inter-observer and intra-observation variability that we observed highlights the difficulty of obtaining a GT, and the importance of at least two different experts and at least two observations by one expert for constructing a biased GT.

Table 7.2 summarises the robustness of our scar and myocardium segmentation pipelines. Myocardium is segmented with an accuracy of 0.819 assessed with the intra-observation GT and 0.781 with inter-observer GT. Moreover, there is a variation from 0.83 to 0.78 of accuracy in intra-observation case and 0.82 to 0.745 in inter-observer single cases. These results are consistent with our hypothesis that *the number of expert observations influences assessment of accuracy in an automatic pipeline*. The bias effect of an observation could overestimate or underestimate the accuracy of an automatic pipeline. Automatic techniques should therefore be evaluated with an unbiased ground GT if this is possible.

For automatic scar region segmentation, we utilised two manual myocardium GT binary segmentations to evaluate the scar accuracy in inter-observer and intra-observation cases (MM). Therefore, the scar segmentation pipeline delivers more robust results than the myocardium pipeline alone as in both GT the dice metric was 70.5 % and the accuracy variation among intra-observations and different experts were from 70.0 to 65.1 %.

All these findings demonstrate the importance of reducing bias in GTs used to evaluate automatic segmentation pipelines.

Table 7.2: Summary of myocardium scar segmentation results comparing different inter/intra- ground truths

Intra-observation variation results				
Region	dice (%)	Sensitivity (%)	Specificity (%)	Manually Case
Myocardium	81.9 ± 9.4	91.9 ± 11.0	95.9 ± 1.2	Ground Truth
Myocardium	80.0 ± 10.2	90.2 ± 12.5	96.4 ± 1.1	First observation
Myocardium	83.0 ± 8.8	92.4 ± 8.4	96.5 ± 1.0	Second observation
Myocardium	78.1 ± 11.7	86.3 ± 22.4	45.9 ± 1.9	Third observation
Scar	70.5 ± 11.7	75.0 ± 12.5	99.6 ± 0.3	Ground Truth
Scar	70.0 ± 8.2	76.2 ± 14.4	99.4 ± 0.3	First observation
Scar	65.8 ± 18.6	73.1 ± 21.1	99.3 ± 0.4	Second observation
Scar	68.2 ± 12.3	77.1 ± 10.5	99.4 ± 0.4	Third observation
Inter-observer variation results				
Region	dice (%)	Sensitivity (%)	Specificity (%)	Manually Case
Myocardium	78.1 ± 9.3	94.2 ± 11.8	95.1 ± 1.2	Ground Truth
Myocardium	79.3 ± 12.3	93.5 ± 12.7	94.8 ± 1.6	First observer
Myocardium	75.5 ± 10.8	93.0 ± 13.1	93.8 ± 1.1	Second observer
Myocardium	82.0 ± 10.0	90.3 ± 8.0	96.0 ± 2.0	Third observer
Scar	70.5 ± 17.2	85.4 ± 7.5	99.6 ± 0.2	Inter Ground Truth
Scar	68.9 ± 16.5	82.2 ± 10.6	99.6 ± 0.2	First observer
Scar	65.1 ± 21.6	75.6 ± 23.4	99.6 ± 0.2	Second observer
Scar	67.5 ± 18.3	87.1 ± 8.5	99.6 ± 0.1	Third observation

7.1.2 Comparison with established and state of the art techniques

There are few publicly available cardiac LGE-MRI datasets with labeled scar segmentation. Thus, comparison of MA-SOCRATIS with established and state of the art techniques using the same dataset was challenging. We therefore compared our pipeline with established automatic or semi-automatic segmentation methods of LV with scars; focusing on datasets with identically or almost the same experimental conditions as our cohort (20 patient testing cohort, LGE-MRI modalities). To the author's knowledge, the datasets used to evaluate the methods which we compared with MA-SOCRATIS are not published. An exception is the MS-CMRSeg challenge of 2019, which we included as a cross-institute validation dataset.

The established methods we used have been evaluated on datasets which followed a specific preprocessing and tuning methodology ([88], [131], [112]). So we had to utilize the same procedures or none in our datasets. We chose not to do any tuning in either our own LGE-MRI dataset of 20 patients, or in the MS-CMRSeg challenge of 2019. We did not use any supervision from experts, so the pipeline was totally automatic. Moreover, we used an initial noise reducing pre-processing step for both datasets, as described in the supplementary material (*Pre-processing and post processing frameworks* section).

Table 7.3 compares the performance of our MA-SOCRATIS pipeline with state of the art techniques for myocardium segmentation of LV with scar from LGE-MRI images based on deep learning techniques. All of these state of the art techniques are supervised to some extent, and the performance of our myocardial segmentation pipeline was comparable or better than the dice metrics quoted in these studies. The segmentation techniques described by ([131],[112]) all used deep learning myocardium segmentation networks applied to LGE-MRI images. They used supervised techniques and still they did not achieve more than 86.0 % dice value. In contrast, the method MA-SOCRATIS is an unsupervised method with no tuning or transfer learning techniques, yet achieved

Table 7.3: Comparison of our myocardium and scar segmentation pipeline and established machine/deep learning techniques

Region	**dice (%)**	**Technique (cohort)**
Myocardium	86.0	DL[131] dataset
Myocardium	**81.9**	MA-SOCRATIS (Intra GT dataset)
Myocardium	80.8	CNN[112] dataset
Myocardium	78.1	MA-SOCRATIS (Inter GT dataset 80.5 median)
Region	**Median (%)**	**Technique (cohort)**
Scar	85.0	MCG[88] dataset
Scar	**75.5**	MA-SOCRATIS (Inter GT dataset)
Scar	74.0	KCL[88] dataset
Scar	73.0	AIT[88] dataset
Scar	72.5	MA-SOCRATIS (Intra GT dataset)
Scar	44.0	MV[88] dataset

an 81.9 % dice.

In one of the studies listed in Table 7.3, Karim et al. ([88]) evaluated four different techniques with a test dataset of 20 patients. Of the methods evaluated, the conditional Markov random field (MCG) had the best results following by Graph-cuts with EM-algorithm technique (KCL) and support vector machines combined with a level set technique (AIT). All of these techniques were supervised and some were semi-automatic. This means that there is a need for training in a specific cohort. The cohort has to be split into testing (20 patient) and training (10 patient) patients. The only unsupervised technique was the EM-algorithm and watershed transformation (MV), and these achieved a modest median accuracy of only 44%. Our scar segmentation pipeline is unsupervised, and with our test dataset of 20 patients, we achieved a 72.5 and 75.5 % median dice accuracy for intra-observation and inter-observer ground truths respectively. The MV technique evaluated by Karim et al. ([88]) and our scar segmentation pipeline are both based on the study of Hennemuth et al. ([72]), although we have modified the approach of ([72]) as described in the methods section.

We also tested our pipeline using the open source MS-CMRSeg 2019 challenge dataset. Table

Table 7.4: Comparison of our myocardium segmentation pipeline and state of the art techniques in MS-CMRSeg 2019 published dataset based on average and standard deviation (SD) dice values (NM:not mentioned).

Region	dice +/- SD (%)	Technique (MS-CMRSeg 2019 dataset)
Myocardium	83.0 +/- NM	Unsupervised Multi-modal [32]
Myocardium	82.7 +/- 6	SK-Unet [192]
Myocardium	81.2 +/- 6.1	Combining Multi-Sequence and Synthetic-LGE [27]
Myocardium	80.8 +/- NM	Automatic Cardiac Segmentation Framework [111]
Myocardium	80.1 +/- NM	Adversarial Domain Adaptation Network (GFRM) [191]
Myocardium	78.0 +/- NM	Multi-Atlas-CNN noise images [163]
Myocardium	74.9 +/- 10.0	Supervised-Domain-Adaptation [189]
Myocardium	71.4 +/- 10.0	Deep-Learning-Framework [206]
Myocardium	**70.0 +/- 5.5**	MA-SOCRATIS
Myocardium	68.6 +/- 7.8	Shape-Transfer GAN network [177]
Myocardium	65.4 +/- 7.0	Multi-Atlas [16]
Myocardium	61.7 +/- 8.6	Pseudo-3D network [112]
Myocardium	61.0 +/- NM	Adversarial-CNN-Domain-Adaptation [33]
Myocardium	57.3 +/- NM	2D/3D-UNET Framework [195]
Myocardium	47.0 +/ 11.7	UNET++ Framework [156]

7.4 summarizes the results and compares the performance of MA-SOCRATIS with other published methods. The approach described by ([32,112,192]) used a multi-modal image translation network from balanced steady-state free-precession (bSSFP) images to LGE images. In all case these methods need tuning or supervised training with the training cohort of the MS-CMRSeg 2019 challenge dataset. Chen et al. ([32]) used a cascaded segmentation for the bSSFP images. Although they achieved an unsupervised segmentation of the LGE-MRI images, the cascade network required supervised training. Wang et al. ([192]) used manual ROI definition and they used all T2 cine-MRI and LGE in the training process. They used an image postprocessing process; such as hole filling technique, connected component analysis, and the largest connected component of all slices in each patient. Segmentation areas that exceed the largest connected range were removed. The approach described by ([32], [111], [192]) used a multi-modal image translation network from balanced steady-state

free-precession (bSSFP) images to LGE images. In all cases, they need tuning or supervised training in T2, bSSFP, and LGE-MR images in the training cohort of MS-CMRSeg 2019 challenge dataset.

In contrast to these studies, MA-SOCRATIS is free of any kind of multimodal training in bSSFP or LGE-MRI images or tuning, yet it still achieved 70.0 % accuracy against the MS-CMRSeg 2019 challenge dataset. It is important to mention that our pipeline achieved the most robust results as the standard deviation of the results was the smallest (5.5 %) of all methods.

7.2 Discussion of DL-SOCRATIS pipeline

7.2.1 Performance analysis of the models

The quantitative performance analysis of the intra/inter observation/er GTs for MA-SOCRATIS, cine-SOCRATIS, CGAN-SOCRATIS, and DL-SOCRATIS networks over the LGE-20 dataset has been presented as box-plots in Fig. 7.7 and 7.6. From the box-plots presented in Fig. 7.7, it is clearly visible that the MA-SOCRATIS network achieves higher performance for the myocardium segmentation, irrespective of the quantitative evaluation indices (sensitivity, specificity and dice) compared with the other networks. Moreover, it is clear that the MA-SOCRATIS delivered the most robust results as the variations of the three metrics are the smallest compared with the deep learning techniques in both intra-GT and inter-GT experiments. The most sensitive pipeline is the DL-SOCRATIS as it has the higher variation of all. CGAN-SOCRATIS even if it has the second average performance in the deep learning techniques; it highlights the lower variation with respect to the three quantitative evaluation indices among the DL networks.

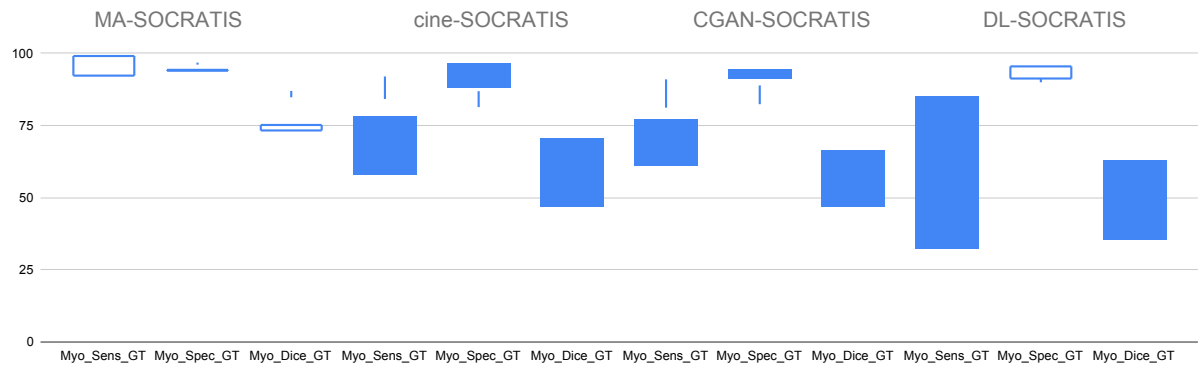

Figure 7.6: The sensitivity, specificity and dice metrics statistical analysis of the MA-SOCRATIS, cine-SOCRATIS, CGAN-SOCRATIS and DL-SOCRATIS for the inter-observer GTs dataset.

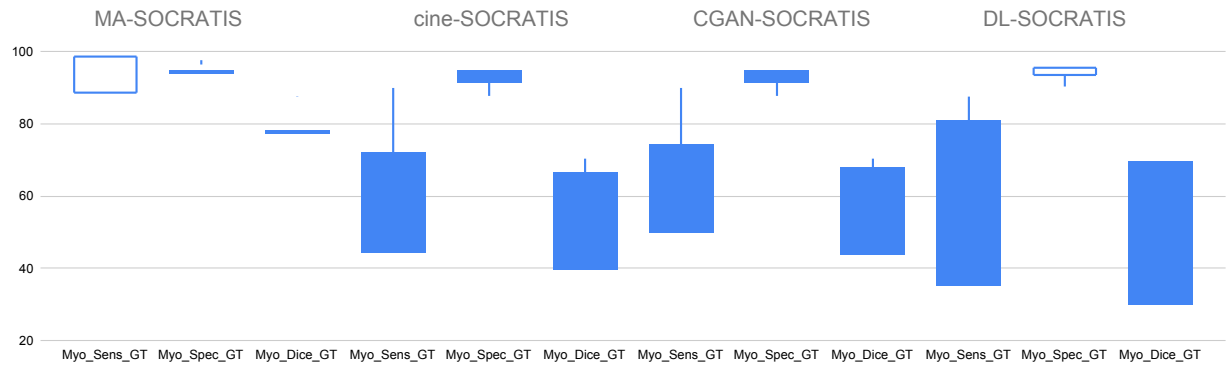

Figure 7.7: The sensitivity, specificity and dice metrics statistical analysis of the MA-SOCRATIS, cine-SOCRATIS, CGAN-SOCRATIS and DL-SOCRATIS for the intra-observation GTs dataset.

Table 7.5: Comparison of MA-SOCRATIS and DL-SOCRATIS myocardium segmentation pipelines and state of the art techniques in MS-CMRSeg 2019 published dataset based on average and standard deviation (SD) dice values (NM:not mentioned).

Region	dice (%)	Technique (MS-CMRSeg 2019 dataset)
Myocardium	83.0 +/- NM	Unsupervised Multi-modal [32]
Myocardium	82.7 +/- 6	SK-Unet [192]
Myocardium	80.8 +/- NM	Automatic Cardiac Segmentation Framework [112]
Myocardium	80.1 +/- NM	Adversarial Domain Adaptation Network (GFRM) [191]
Myocardium	74.9 +/- 10.0	Supervised-Domain-Adaptation [189]
Myocardium	**70.0 +/- 5.5**	MA-SOCRATIS
Myocardium	68.6 +/- 7.8	Shape-Transfer GAN network [177]
Myocardium	65.4 +/- 7.0	Multi-Atlas [16]
Myocardium	61.7 +/- 8.6	Pseudo-3D network [111]
Myocardium	57.3 +/- NM	2D/3D-UNET Framework [195]
Myocardium	**53.3 +/- 9.2**	DL-SOCRATIS-tuning
Myocardium	50.1 +/- 10.0	DL-SOCRATIS-no-tuning
Myocardium	47.0 +/- 11.7	UNET++ Framework [156]
Myocardium	44.1 +/- 10.8	CGAN-SOCT-no-tuning
Myocardium	43.1 +/- 13.2	cine-SOCT-no-tuning

7.2.2 Comparison of the DL-SOCRATIS and MA-SOCRATIS pipelines; with state of the art techniques

Table 7.5 compares the performance of the DL-SOCRATIS and MA-SOCRATIS pipelines with established and state of the art techniques for automatic myocardium segmentation. The approach described by ([32,111,192]) used a multi-modal image translation network from balanced steady-state free-precession (bSSFP) images to LGE images. In all these cases, tuning or supervised training in the training cohort of MS-CMRSeg 2019 challenge dataset was needed. Chen et al. ([32]) used a cascaded segmentation for the bSSFP images. Although they achieved an unsupervised segmentation of the LGE-MRI images, the cascade network required supervised training. Wang et al. ([192]) used manual ROI definition and they used all T2 cine-MRI and LGE in the training process. They used

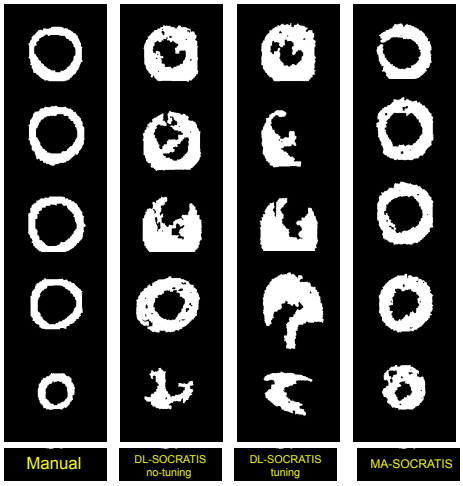

Figure 7.8: Results of MA-SOCRATIS, DL-SOCRATIS tuning and DL-SOCRATIS no tuning, pipelines in MS-CMRSeg 2019 dataset about myocardium lesions, compared with GTs.

an image postprocessing process; such as hole filling technique, connected component analysis, and the largest connected component of all slices in each patient. Segmentation areas that exceed the largest connected range were removed. The approach described by ([31,111,192,112]) used a multi-modal image translation network from balanced steady-state free-precession (bSSFP) images to LGE images. In all cases, they need tuning or supervised training in T2, bSSFP, and LGE-MR images in the training cohort of MS-CMRSeg 2019 challenge dataset.

Contrary to these studies, MA-SOCRATIS is free of any kind of supervised multi-modality training (bSSFP, cine-MRI, and LGE-MRI images) or even tuning and it still achieved 70.0 % accuracy in the MS-CMRSeg 2019 challenge dataset. DL-SOCRATIS without any need of tuning , multi-modality training (bSSFP, T2, LGE images) or manually assign of landmarks in the region of interest (as in MA-SOCRATIS); it succeed 50.1 % dice value. Additionally, it was important to study the variation of the result metrics if the u-net network was tuned in the new cohort (DL-SOCRATIS-tuning). For this task, we used five patients of the training cohort of MS-CMRSeg 2019 challenge dataset and evaluated the results again in the testing cohort of MS-CMRSeg 2019

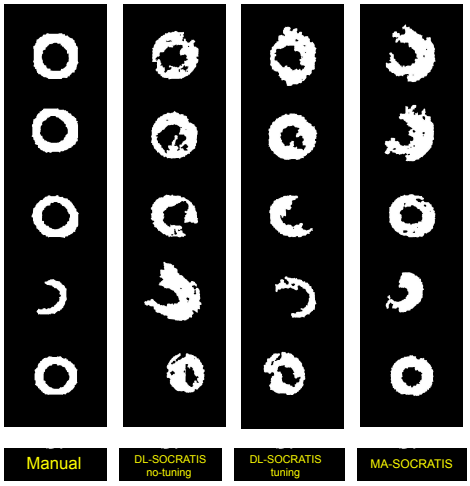

Figure 7.9: Results of MA-SOCRATIS, DL-SOCRATIS tuning and DL-SOCRATIS no tuning, pipelines in MS-CMRSeg 2019 dataset about myocardium lesions, compared with GTs.

dataset. The DL-SOCRATIS-tuning delivered an improved accuracy of 53.3 percent. This results is 6.3% higher than the UNET++ Framework ([156]) that succeed only 47.0% in an attempt to deliver results with almost the same conditions as our DL-SOCRATIS pipeline (no need of multi-modality training in bSSFP/T2/LGE images, or manually supervision in the region of interest from expert). Lastly, the DL-SOCRATIS is in both cases more robust than the UNET++ Framework ([156]) as the SD is 9.2 or 10.0 % (tunning, no tunning case respectively) against the 11.7% of the UNET++. Figures 7.8,7.9,7.10 and 7.11 show the results of MA-SOCRATIS and DL-SOCRATIS on MS-CMRSeg 2019 challenge dataset. The results of MA-SOCRATIS are the best estimation of the manual myocardial except the Figure 7.9 where the DL-SOCRATIS-tuning was more accurate. The DL-SOCRATIS-tuning is more accurate than the DL-SOCRATIS-no-tuning as we expected. An exception are the three cases of Fig. 7.8 image four and Fig. 7.9 images one and three.

DL-SOCRATIS did not segment accurately the myocardium; as the u-net did not detect correctly the location and the shape of the left ventricle (LV). Thus, the initial segmentation underestimates or overestimates the circle radius of SOCRATIS reestimation phase and it detected a

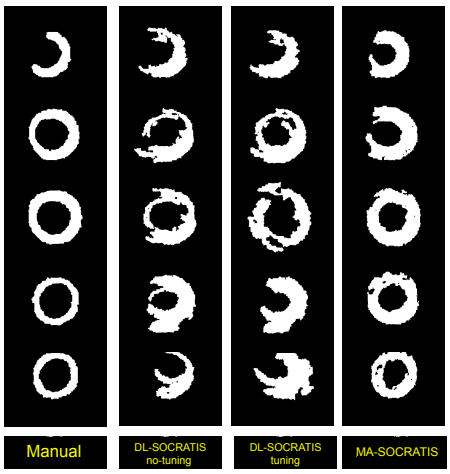

Figure 7.10: Results of MA-SOCRATIS, DL-SOCRATIS tuning and DL-SOCRATIS no tuning, pipelines in MS-CMRSeg 2019 dataset about myocardium lesions, compared with GTs.

wrong position of the LV center. As a result, in some images the pipeline detects some portions of the left and right ventricle myocardium like in Fig. 7.11 image four and five. In some other cases, it detects some peripheral organs/areas such as in Fig. 7.8 the images two, three, four, and five of DL-SOCRATIS-tuning and images three and four of DL-SOCRATIS-no-tuning pipelines.

7.2.3 Compare the tuning and no tuning deep learning network of initial myocardium segmentation of DL-SOCRATIS

In this subsection, we will discuss explanations and possible solutions to the problem of false detection of the myocardial region. Moreover, we will analyse the behaviour of tuning and no-tuning u-net network in the 40 patients of MS-CMRSeg 2019 dataset.

Figures 7.12,7.13 show some of the results for the tuning and non-tuning u-net network regarding the initial endocardial (ENDO) and epicardial (EPI) segmentation of the left ventricle. In most of the cases, the endocardial position was detected correctly in the tuning model, with no tuning (an exception is the image of Fig. 7.13).

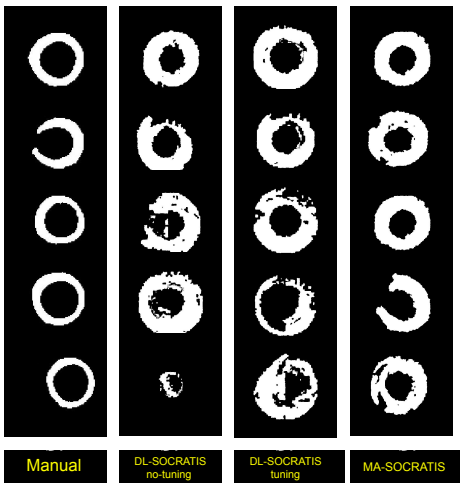

Figure 7.11: Results of MA-SOCRATIS, DL-SOCRATIS tuning and DL-SOCRATIS no tuning, pipelines in MS-CMRSeg 2019 dataset about myocardium lesions, compared with GTs.

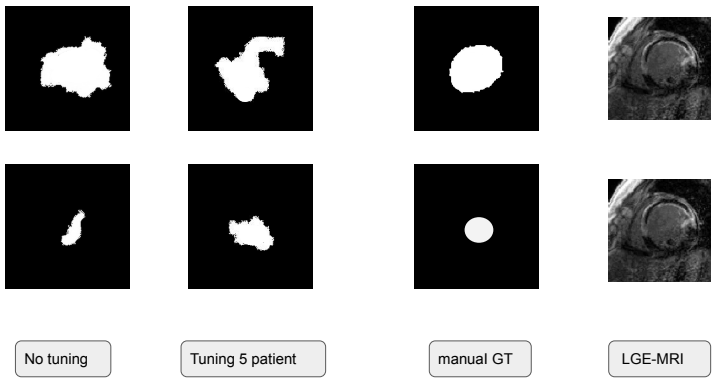

Figure 7.12: DL-SOCRATIS no-tuning and DL-SOCRATIS tuning results of patient about myocardial-scar lesions, with manual GTs.

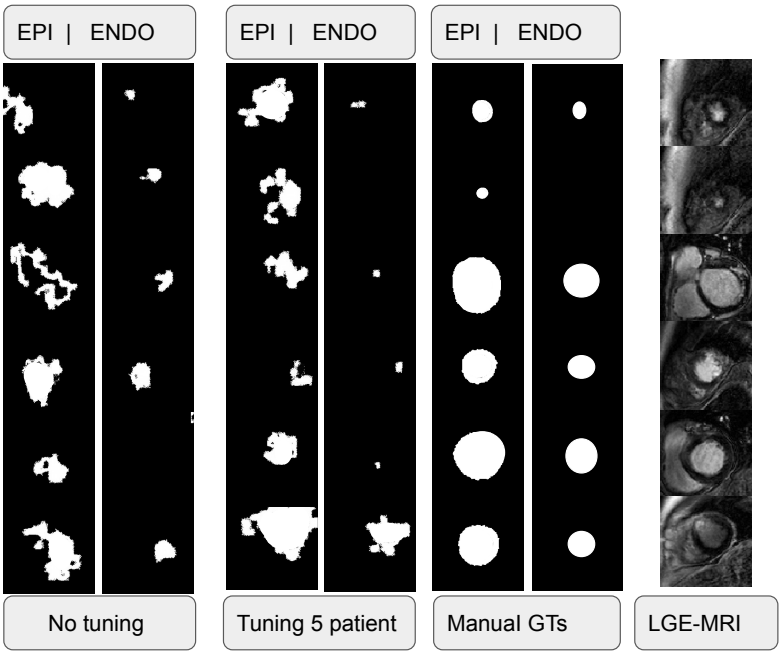

Figure 7.13: DL-SOCRATIS no-tuning and DL-SOCRATIS tuning results of patient about myocardial-scar lesions, with manual GTs.

Regarding the shape of ENDO, there are some cases where the size is more accurate in tuning model (Fig. 7.12) than the non-tuning and there are some cases where the size of ENDO is more accurate in non-tuning model to the tuning model (Fig. 7.13 second,forth and sixth images). However, in the case, the position of ENDO is false. In most cases, the epicardial position was detected correctly in the tuning model, contrary to the non-tuning (an exception is the image of Fig. 7.13). Regarding the shape of EPI, there are some cases where the size is more accurate tuning model (Fig. 7.12, Fig.7.13 second, third and fifth images) than the non-tuning and there are some cases where the size of EPI is more accurate in non-tuning model to the tuning model (Fig. 7.13 first and sixth images). However, in the last observation, the position of EPI is false. Figure 7.14 shows the results of the tuning model LGE images, the sixth patient of MS-CMRSeg 2019 challenge cohort. For patient number six, the epicardial and endocardial position was detected correctly in the tuning model.

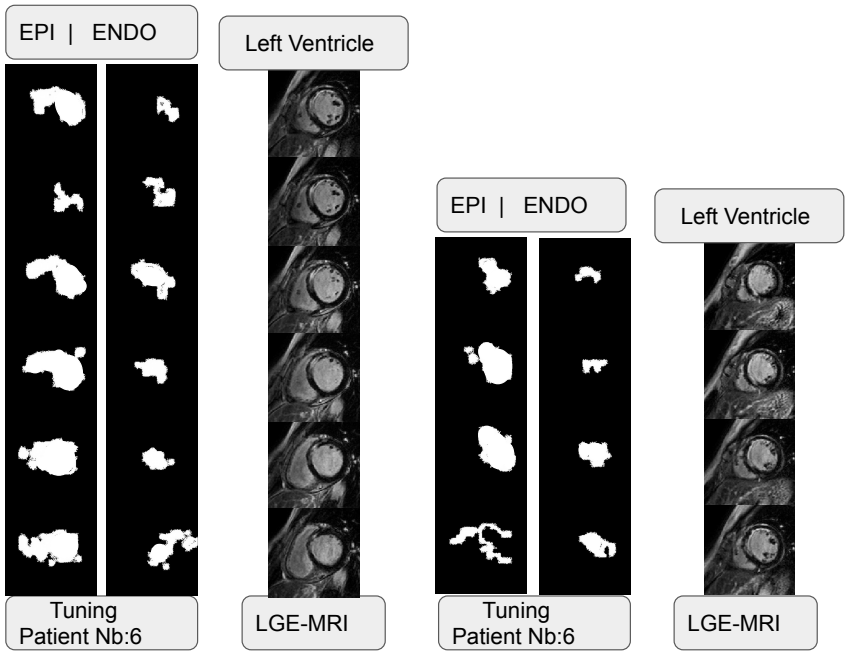

Figure 7.14: DL-SOCRATIS tuning automatic segmentation myocardial results of patient number 6 of MS-CMRSeg 2019 dataset.

Regarding the shape of ENDO, there are some cases where the size is accurate enough (second, forth, fifth, ninth, and tenth images of Fig. 7.14). EPI size is accurate in the first, seventh, eighth, and ninth images of Fig. 7.14. It can be observed that the tuning network was the segment LV and a portion of RV epicardium region. Thus, the epicardium size is not correct. For tuning the model, we used a transfer learning technique. We initialized the u-net weight-based training results of cohort LGE-100, so we can tune the model in the five training patient of MS-CMRSeg 2019 dataset. One limitation of this study is that the accuracy of DL-SOCRATIS can be the transfer learning technique and the size of the tuning network. In future work, we extend the study to other transfer learning techniques and different variations of the tuned number of patients in cross-institutional dataset.

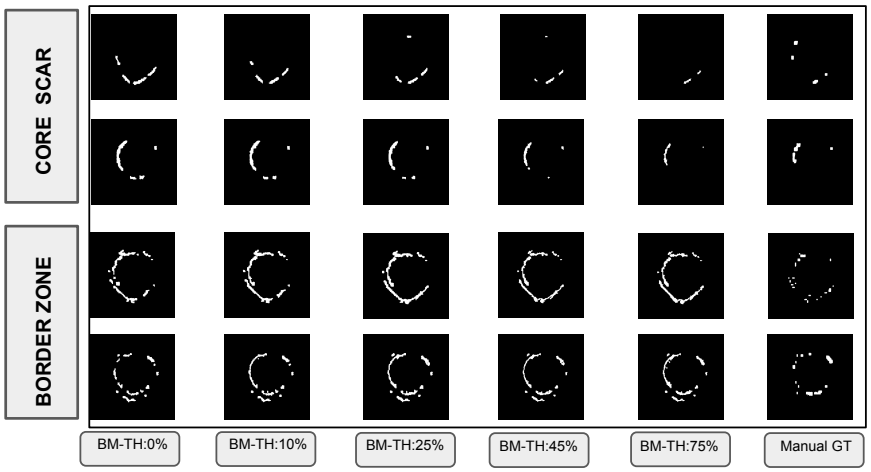

Figure 7.15: The border zone and core scar automatic results of ablation analysis for different values of border zone thresholding parameter. In this figure we present the automatic segmentation results of BZ-TH values in 0,10,25,45,75 % of the difference between minimum and maximum intensity (BZ-TH:0%, BZ-TH:10%, BZ-TH:25%, BZ-TH:45% and BZ-TH:75%) compare with the manual inter-observer ground truth (Manual GT).

7.3 Discussion of BZ-SOCRATIS pipeline

7.3.1 Performance analysis of the BZ-TH portion change in BZ-SOCRATIS

In this section, we discuss how significant are the analysis of different BZ-TH portions. Figure 7.15 illustrates the different results of border zone and core scar automatic segmentation for five different variations of BZ-TH portions (0, 10, 25, 45 and 75%). The most accurate result based on the manual GT is the BZ-TH:25%; however the results of the two regions in each of the different portions do not show significant difference. Table 7.6 summarizes the different dice, sensitivity and specificity scores of the border zone (BZ), and core scar (CS) areas for different portions of BZ-TH; as we determined in the methodology section. The overlap between intra-observation and inter-observer variability GTs as far as it concerns the highest results of the different portions of BZ-TH method for BZ and CS regions are in BZ-TH:10% and BZ-TH:25%. For intra-GTs, the mean dice value of BZ is 43.7 and 37.7% with a standard deviation (STD) 10.7 and 11.2% and of CS 58.7 and 57.7%

with a 17.5 and 17.0% STD in BZ-TH:10% and BZ-TH:25% respectively. For inter-GTs, the mean dice value of BZ is 31.7 and 32.4% with a standard deviation (STD) 13.4 and 13.2% and of CS 56.4 and 60.0% with a 24.2 and 21.8% STD in BZ-TH:10% and BZ-TH:25% respectively. Regarding the sensitivity metric for intra-GTs the BZ region was segmented with 55.1, 50.1% value (STD of 10.9, 7.2%) and the CS region achieved a 57.1 and 67.7% sensitivity value (STD of 22.2 and 17.2%) in BZ-TH:10% and BZ-TH:25% respectively. For inter-GTs, the sensitivity score of BZ is 57.1, 67.7% (STD of 13.4, 13.2%) and of CS 70.8 and 71.1% (STD of 17.6, 18.9%) in BZ-TH:10% and BZ-TH:25% respectively.

Figure 7.16 shows the boxplots for the quantitative performance analysis of the different portion of BZ-TH method in intra/inter observation/er GTs. The most robust and higher accuracy results are the one of BZ-TH:25% (red box); as they have the highest mean value and the same time the lowest variability of the three metrics (sensitivity, specificity, and dice) compared with the 50, 40 and 75 BZ-TH portions.

To conclude, the different portions of BZ-TH method do not change dramatically the accuracy and STD of the results, so we can claim that the best portion threshold for our pipeline is the BZ-TH:25% based on the above observations (Figures 7.15, 7.16 and Table 7.6).

7.3.2 Comparison with established and state of the art techniques of core scar segmentation.

There are few publicly available cardiac LGE-MRI datasets with scar segmentation. Thus, comparison of BZ-SOCRATIS with established and state of the art techniques using the same dataset was challenging. We therefore compared our pipeline with established automatic or semi-automatic segmentation methods of LV with scars; focusing on datasets with identically or almost the same experimental conditions as our cohort (20 patient testing cohort, LGE-MRI modalities). To the authors' knowledge, the datasets used to evaluate the methods which we compared with BZ-SOCRATIS are

Table 7.6: Summary of automatic segmentation of border zone (BZ) and core scar (CS) compared to intra-observation and inter-observer ground truth.

Intra observer results				
Region	dice (%)	Sensitivity (%)	Specificity (%)	parameter
Scar (BZ)	41.5 ± 10.7	53.3 ± 13.2	98.3 ± 0.5	BZ-TH:0%
Scar (BZ)	43.7 ± 10.9	55.1 ± 10.9	98.6 ± 0.5	BZ-TH:10%
Scar (BZ)	37.7 ± 11.2	50.1 ± 7.2	98.9 ± 0.4	BZ-TH:25%
Scar (BZ)	38.0 ± 11.0	50.7 ± 6.9	99.0 ± 0.4	BZ-TH:45%
Scar (BZ)	36.9 ± 11.5	49.7 ±7.8	98.8 ± 0.4	BZ-TH:50%
Scar (BZ)	38.8 ± 11.2	50.5 ± 6.9	99.0 ± 0.4	BZ-TH:75%
Scar (CS)	60 ± 18.2	72.8 ± 15.3	99.1 ± 0.5	BZ-TH:0%
Scar (CS)	58.7 ± 17.5	66.3 ± 15.8	99.2 ± 0.5	BZ-TH:10%
Scar (CS)	57.7 ± 17.0	70.3 ± 22.8	99.4 ± 0.5	BZ-TH:25%
Scar (CS)	55.7 ± 17.3	67.4 ± 25.6	99.4 ± 0.4	BZ-TH:45%
Scar (CS)	57.0 ± 17.8	68 ± 24.5	99.4 ± 0.5	BZ-TH:50%
Scar (CS)	56.2 ± 17.5	67.5 ± 25.6	99.3 ± 0.5	BZ-TH:75%
Inter observers results				
Region	dice (%)	Sensitivity (%)	Specificity (%)	parameter
Scar (BZ)	32.3 ± 14.2	50.7 ±24.2	97.5 ± 0.8	BZ-TH:0%
Scar (BZ)	31.7 ± 13.4	57.1 ± 22.2	97.0 ± 0.9	BZ-TH:10%
Scar (BZ)	32.4 ± 13.2	67.7±17.2	96.0 ± 1.6	BZ-TH:25%
Scar (BZ)	31.5 ± 12.1	64.8 ± 21.1	96.1 ± 1.2	BZ-TH:45%
Scar (BZ)	30.1 ± 12.6	65.1 ± 20.6	96.1 ± 1.2	BZ-TH:50%
Scar (BZ)	28.4 ± 13.3	64.0 ± 22.5	95.2 ± 2.3	BZ-TH:75%
Scar (CS)	55.3 ± 22.4	70.8 ± 18.5	98.7 ± 1.0	BZ-TH:0%
Scar (CS)	56.4 ± 24.2	70.8 ± 17.6	98.9 ± 0.8	BZ-TH:10%
Scar (CS)	60.0 ± 21.8	71.1 ± 18.9	98.8 ± 0.7	BZ-TH:25%
Scar (CS)	61.9 ± 21.1	66.0 ± 16.8	99.1 ± 0.7	BZ-TH:45%
Scar (CS)	60.9 ± 21.1	65.9 ± 16.9	99.1 ± 0.7	BZ-TH:50%
Scar (CS)	53.8 ± 15.6	46.2 ± 14.1	99.5 ± 0.2	BZ-TH:75%

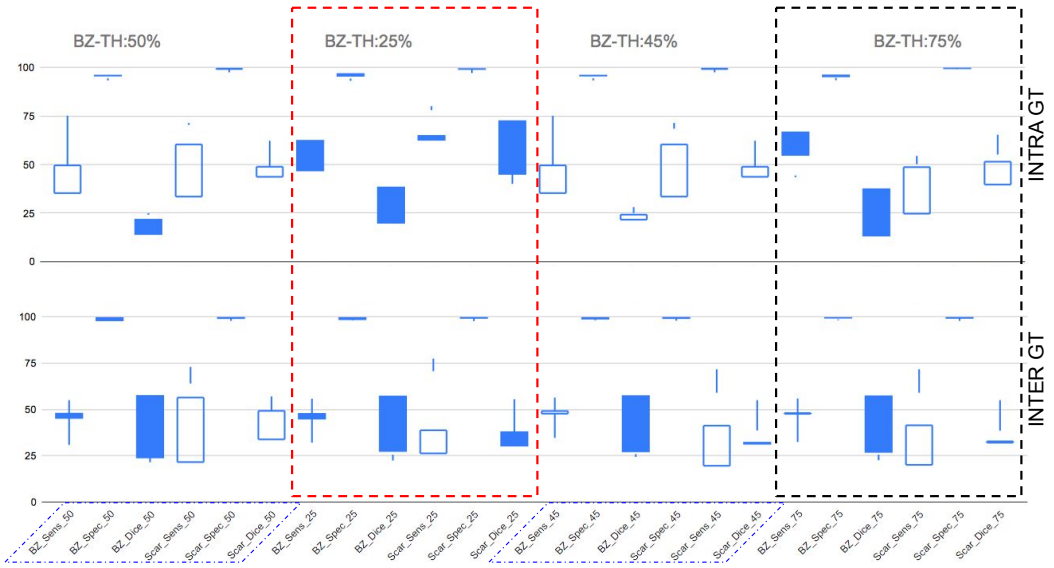

Figure 7.16: Box plots for the quantitative performance analysis of the different values of BZ-TH method in intra/inter observation/er GTs. In this figure we present the dice, sensitivity and specificity metrics results of BZ-TH values for 50,25,45,75 % of the difference between minimum and maximum intensity (BZ-TH:50%, BZ-TH:25%, BZ-TH:45% and BZ-TH:75% left to right).

Table 7.7: Comparison of our total scar segmentation pipeline and established machine/deep learning techniques

Region	Median (%)	Technique (cohort)
Scar	85.0	MCG[88] dataset
Scar	**75.5**	MA-SOCRATIS (Inter GT dataset)
Scar	74.0	KCL[88] dataset
Scar	73.0	AIT[88] dataset
Scar	72.5	MA-SOCRATIS (Intra GT dataset)
Scar	44.0	MV[88] dataset

not published.

In Table 7.7, Karim et al. ([88]) evaluated four different techniques with a test dataset of 20 patients. Of the methods evaluated, the conditional Markov random field (MCG) had the best results following by Graph-cuts with EM-algorithm technique (KCL) and support vector machines combined with a level set technique (AIT). All of these techniques were supervised and some were semi-

automatic. This means that there is a need of training in a specific cohort. The cohort has to be split in testing (20 patient) and training (10 patient) patients. The only unsupervised technique was the EM-algorithm and watershed transformation (MV), and these achieved a modest median accuracy of only 44.0%. Our scar segmentation pipeline is unsupervised, and with our test dataset of 20 patients, we achieved a 72.5 and 75.5 % median dice accuracy for intra-observation and inter-observer ground truths respectively.

7.3.3 Extraction of 3D models: a future perspective for automatic border zone, core scar and myocardial mapping.

After the 2D slice extraction of the BZ, CS, and LV regions; we utilized a 3D interpolation pipeline (3D-SOCRATIS) which automatically extracted the 3D representation of these regions. Based on these 3D geometries, the expert radiologist and clinicians have a three-dimensional anatomical model of a specific patient's LV with infarct. The 3D model is extracted fully automatically based on DICOM LGE-MR images of the patient. Figures 7.17, 7.18, 7.19 and 7.20 show the 3D anatomical models of the BZ, CS, and LV areas of different patients.The differences regarding the smoothness of the 3D anatomical models of 7.17 and 7.19 are a result of different number of iteration in the post-processing step (as mentioned in 4.2.2). These results were delivered after a manual sensitivity analysis for different post-processing iterations of the 3D volume extraction pipeline (3D-SOCRATIS). In the future more investigation regarding optimal number of iteration is needed to deliver smoother 3D anatomical models.

7.4 Abstraction of SOCRATIS framework

In this section of the thesis, we will discuss the idea behind the SOCRATIS framework. We implemented a framework for fully automatic, and semi-automatic segmentation of LV with scars

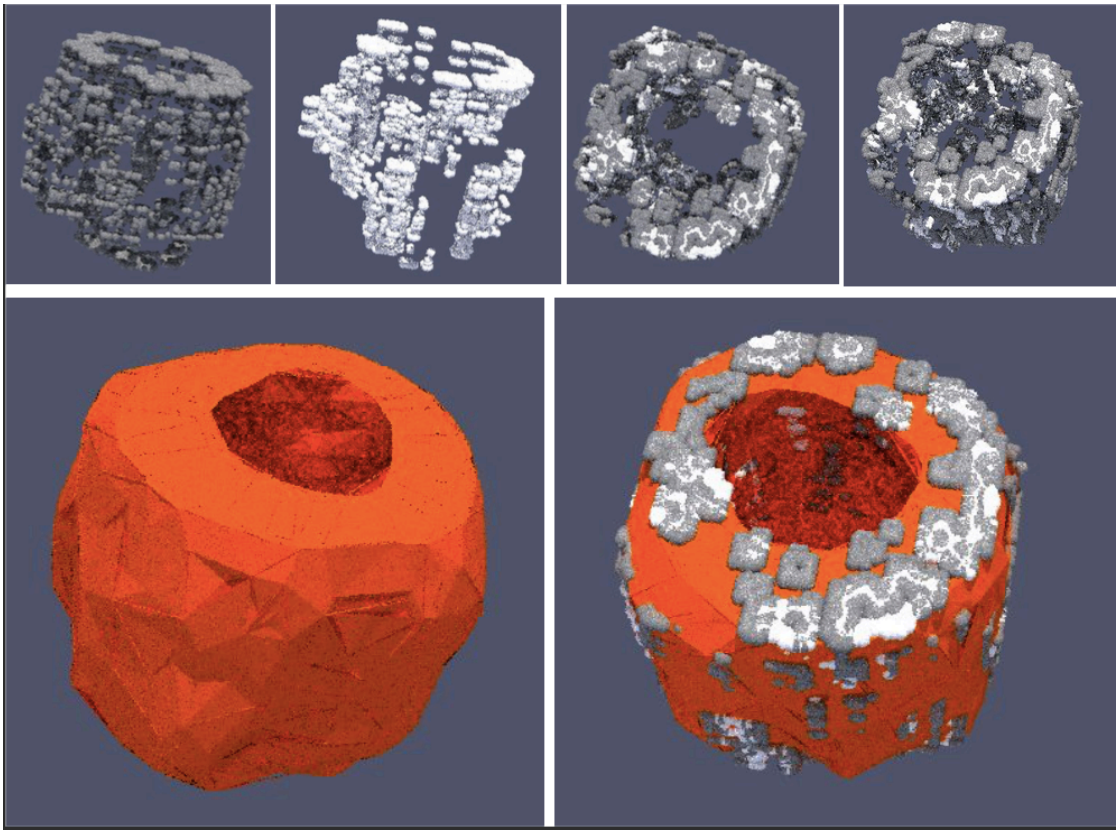

Figure 7.17: Three dimension representation models. Top left corner: left ventricle (LV), top middle: LV with border zone, top right corner: LV with core scar, bottom left corner: total scar, bottom middle and right corner: two different orientations of LV with total scar regions.

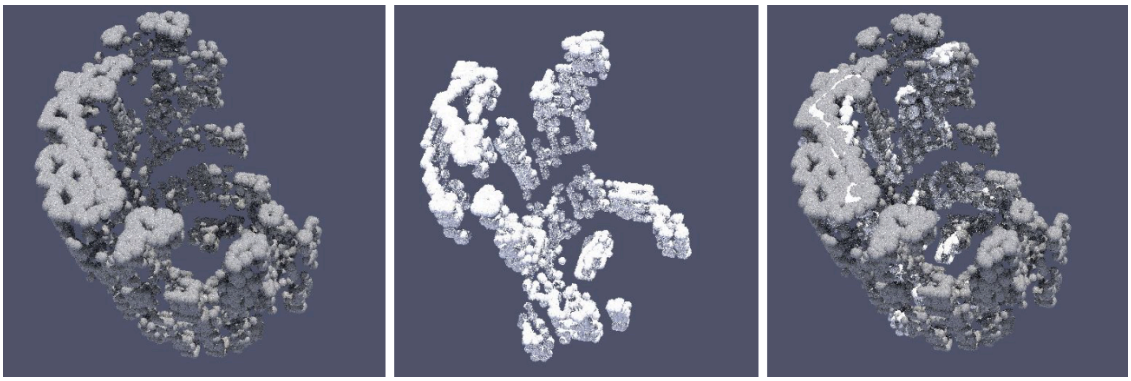

Figure 7.18: Three dimension representation of border zone, core scar and total scar (from left to right).

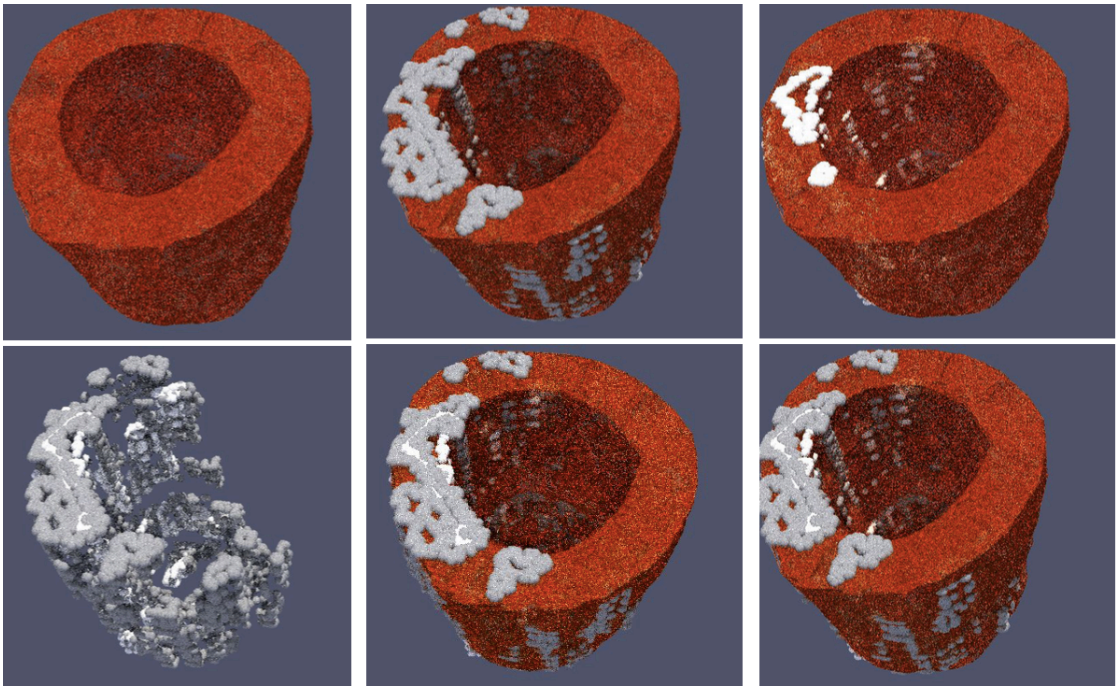

Figure 7.19: Three dimension representation of left ventricle (LV), LV with border zone, LV with core scar, total scar and two different orientations of LV and total scar regions (from up to down and left to right).

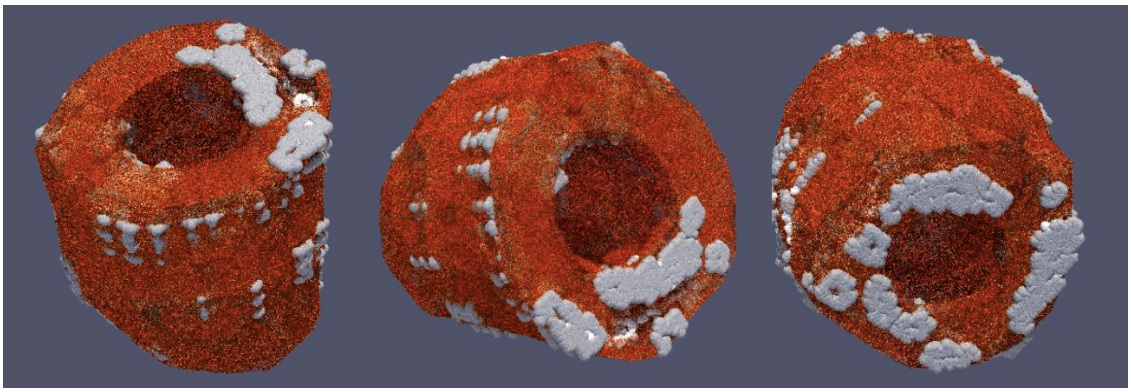

Figure 7.20: Three dimension models of two different patients. Patient 1: left and middle 3D geometry and Patient 2: right 3D geometry. The 3D models represent the left ventricle, border zone and core scare regions of the two specific patients.

optimized based on speed or either on accuracy of the results. We will mention the importance and robustness of MA-SOCRATIS automated method and in some cases the benefit of the high speed

that DL-SOCRATIS can offer. Lastly, we will present the strategy that SOCRATIS framework follows.

7.4.1 Robustness and accuracy of the two pipelines

MA-SOCRATIS without any need for tuning or multi-modality training, delivered dice values of 81.9 and 78.1 % accuracy in our challenging LGE cohort and a 70.0 % accuracy in the cross institute validation of the method (MS-CMRSeg 2019 challenge dataset). DL-SOCRATIS operates without any need of tuning, multi-modality training (bSSFP, T2, LGE images) or manually assign of landmarks in the region of interest (as in the initial segmentation step of multi-atlases myocardium of MA-SOCRATIS pipeline); it succeeded 60.4 and 55.0 % in our challenging LGE cohort and a 50.1% accuracy in the cross institute validation of the method (MS-CMRSeg 2019 challenge dataset). Regarding the robustness of the dice standard deviation (SD) values of the methods were: 9.4, 9.3 and 5.5 % for the MA-SOCRATIS and 10.0, 11.1 and 9.2 % for the DL-SOCRATIS with respect to the intra/inter 20 LGE-MRI patient cohort and the 40 patient of MS-CMRSeg 2019 challenge dataset. To conclude, it is obvious that the results of MA-SOCRATIS are more robust and accurate.

7.4.2 Speed of the two pipelines

MA-SOCRATIS needed almost 1 hour after the multi-atlas registration step and 8.55 minutes for the automatic segmentation of the myocardial and scar regions of LV for each patient (8-12 slices of LGE-MRI images, around 40 s per slice after the multi-atlas registration step) in a Intel(R) Core(TM) i5-4570 CPU @ 3.20GHz with 8GB RAM. Moreover, the time expert supervision for the initial manual landmark assignment does not included in the computation of speed. On the other hand, DL-SOCRATIS is fully automated, there is no need of expert supervision, and it

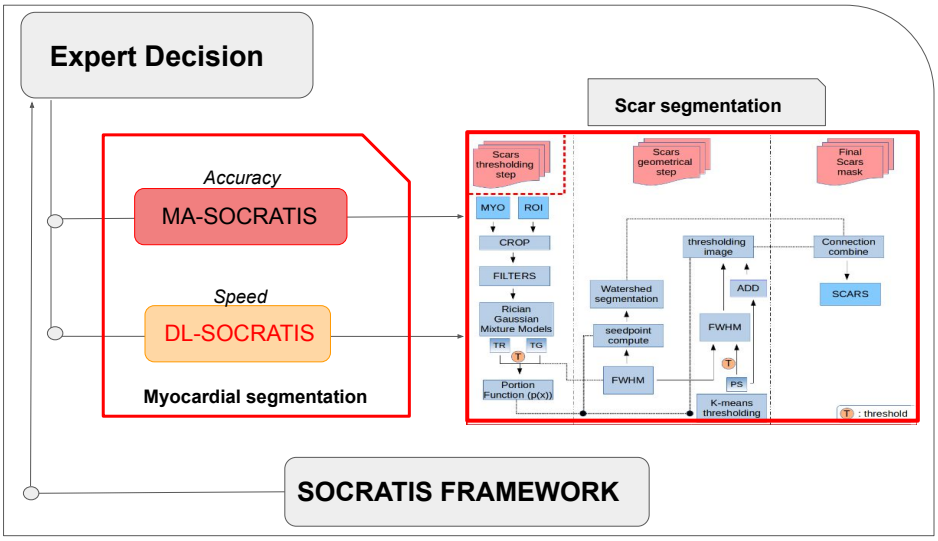

Figure 7.21: The myocardial segmentation part of SOCRATIS framework.

takes 9.22 minutes per patient in total in a Intel(R) Core(TM) i5-4570 CPU @ 3.20GHz with 8GB RAM. For this reason, the DL-SOCRATIS is a quicker and fully unsupervised compare to the MA-SOCRATIS.

7.4.3 The myocardium decision framework

Fig. 7.21 shows the pipeline of the SOCRATIS framework. The decision of the SOCRATIS framework involves the optimal solution that an expert will want to deliver in the problem of the fully automatic segmentation of the LV with scars based on LGE-MRI images. The options are speed, vs, robustness, and accuracy. If the expert chooses the optimal speed, the framework utilizes the DL-SOCRATIS method and delivers 10 to 12 myocardial and scar region slices based on a LGE-MRI short axis image in only 9.22 minutes. If the expert chooses the optimal robustness, the framework uses the MA-SOCRATIS method and delivers 10 to 12 myocardial and scar region slices based on a LGE-MRI short axis image in more than one hour but with high robustness and accuracy results. Lastly, the expert can use a hybrid approach to initially use a quick automatic estimation of the my-

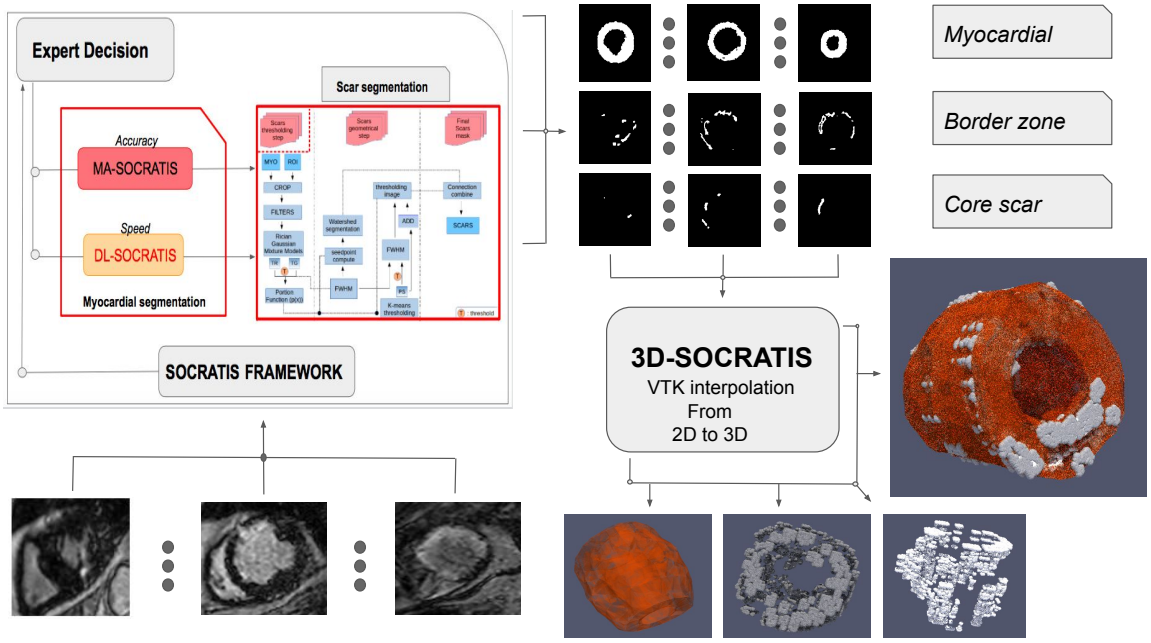

Figure 7.22: The SOCRATIS fully automatically border zone, core scar and LV regions 3D geometries extraction based on expert decisions.

ocardial and scar regions by the DL-SOCRATIS method; and if the results are not acceptable based on clinical global standards, then they can use the more time expensive MA-SOCRATIS method for more robust and accurate results.

7.4.4 The total framework

Figure 7.22 illustrates the fully automatic framework of SOCRATIS method. The expert can decide if they want fast results or robust results regarding the LV myocardium and if they need total scar or BZ and SC regions for the unhealthy 3D mapping tissues of the specific patient.

7.5 Limitations of the study

7.5.1 MA-SOCRATIS pipeline limitations

The main limitation of this study was the computational time required by the MA-SOCRATIS pipeline to deliver the myocardium and scar automatically segmentation (almost an hour per patient), as the registration step of multi-atlas method was time consuming. In addition, in some cases, the detection of the initial position and size of endocardial and epicardial regions was not accurate enough (dice mean value less than 50.0 %). A possible explanation could be the shape variation of multi-atlas process combined with the fact that we used the method for cross-modality application (cine-MRI to LGE-MRI). As a result, the MRI slices of some of the patients were miss-aligned with respect to the multi-atlas slices (registration error). Lastly, the low accuracy of myocardium segmentation in some patients can explain the big difference between AM and MM scar segmentation results (29.5 to 70.5 % respectively).

7.5.2 DL-SOCRATIS pipeline limitations

The detection of the initial position and size of endocardial and epicardial regions was not accurate enough (two DL approaches: DLl, DLc of initial myocardium segmentation step). This is the main limitation of our myocardial pipeline. A possible explanation could be the cross-modality application of the u-net network (cine-MRI training, applied in LGE-MRI validation dataset). Finally, even if we trained the u-net on the LGE-100 dataset or translate the images from cine to LGE images (CGAN-SOCRATIS); when we applied our pipeline in a cross-institute validation (LGE-100 to LGE-20 or LGE-40) the dice value of the results was low. Thus, the position and size of epicardial and endocardial regions was not detected efficiently. Moreover, the STDs of the metrics results of DL-SOCRATIS pipeline are high enough. It is possible that increasing the number of training or

tuning samples of the cross institutes/vendor datasets will solve this problem. As a result, regarding future work, it will be interested in focusing on transfer learning (such as domain adaptation theory) and techniques applied to translation between two modality spaces (cine to LGE) or cross-institute cohorts.

7.5.3 BZ-SOCRATIS pipeline limitations

One of the main limitations of the study is that we cannot claim the intra-observation and inter-observer ground truths of the BZ and CS regions to be manual process. Even if the human interaction and the human error can not be determined in this process (as there is no contour drawing of scar regions); it still exists as the the FWHM needs a human selection of the highest intensity (scar) and the lowest intensity (healthy myocardium) pixels. Therefore, we can claim that in the process of GTs extraction, there exists a inter-observer and intra-variation and uncertainty. Still, this is not the best way of evaluating the pipeline as we observed some cases where our pipeline delivers more highly accurate results than the inter-observer and intra-GTs. This can explain the not so high results of dice and sensitivity scores of BZ and CS regions. Except this limitation, we need to test the cross-institute and cross-vendor robustness of the BZ-SOCRATIS pipeline. Unfortunately, to the authors knowledge, there is not existing any published CMR-LGE cohort to manually label of BZ and CS regions. As a future work, we need to determine a manual contour based BZ and CS region extraction cohort and apply our pipeline in cross-vendor and institute cohorts.

7.6 Future Work

As a future work, we suggest:

- Additional training in cross-vendor and institute cohorts to increase the accuracy of DL-SOCRATIS pipeline.

- A more thorough study of transfer learning techniques to increase the generalization of DL-SOCRATIS pipeline.
- Cross-institute validation of BZ-SOCRATIS pipeline to generalize the use and robustness of pipeline.

Lastly, we suggest the implementation of an artificial intelligence - motion analysis of cardiac region and total infarct scar (AI-MACT). The AI-MACT will be an automatic pipeline to detect the scar position and a portion based on the deformation and strain of ES and ED phases of cine-MRI images. Below is an initial literature review about the aim and possible techniques which can be used.

Zito et al. ([42]) reviews two dimensional speckle-tracking echocardiography (STE) as a way to estimate global longitudinal strain (GLS). They found that the GLS can be used as a marker to diagnose myocardial damage and dysfunction such as myocarditis, systemic disease, neuromuscular disorders, chronic ischemic cardiomyopathy and arrhythmias. In addition, they mention that STE GLS results are almost as accurate as MRI evaluation. Scatteia et al. ([168]) mention that CMR techniques are the best way to identify strains in the ventricle. Thus, they infer the importance of CMR as a reference for the evaluation of bi-ventricular function and morphology. They survey different techniques to estimate myocardial strain.

Acquisition methods track the magnetization of tissue by a specific radio-frequency prepulse spatial modulation of magnetization (SPAMM). Although this operation has good tracking reproducibility and many validation studies, it suffers from low spatial resolution, which is time consuming with low temporal resolution and long acquisition time.

Phase velocity mapping (PVM) uses the myocardial velocity and the phase of signal. Displacement encoding with stimulated echoes (DENSE) encodes the myocardium displacement into the phase of image. Strain encoded (SENC) imaging is a CMR scanner which utilizes quantification of

regional deformation. All three techniques detect a good strain estimation in short acquisition time with quick post-processing, but they suffer from low signal to noise ratio (SNR), fading, the need of clinical experience from the user and the radial strain can not be measured.

Finally, the CMR-FT feature tracking technique computes the deformation of the images with respect to time frames in cine-MRI. It is a postprocessing approach quickly and without the need of further MRI acquisition. The drawbacks are the motion artifacts and the limitation of pixel size within each image. There are CMR-FT studies related to the application of myocardial strain detection in ischemic heart disease.

Some studies in the literature detect pathological and normal heart functionality by utilizing CMR image analysis techniques. Neizel et al. ([133]) uses the SENC technique to compute CMR strain and discriminate different transmurality states in myocardial infraction disease. Neizel et al. ([133]) uses the SENC in healthy subjects too. Finally, Gao et al. ([60]) utilize a b-spline deformable registration as a CMR-FT technique to evaluate the LV myocardial strain and they compare it with DENSE technique. Gao et al. ([60]) conclude that in some cases the b-spline deformable registration overestimates the strain regarding the DENSE. The use of myocardium strain based on cine-MRI can be another way to test and predict the location of scars in the left ventricle ([5]), but reliable automated methods have not yet been developed. This is very interested to the extent of PhD study, so a more robust result of myocardial and scar region segmentation can be delivered, so other MRI modalities and AI techniques take into account.

8

Conclusion

8.1 Overview

In this PhD thesis, we have described and evaluated an automatic segmentation of cardiac regions and total infarct scar (SOCRATIS) pipeline for identifying regions of scar and healthy myocardium from LGE cardiac MRI. In this last chapter, we summarize the importance of each pipeline and of single-modality (LGE-MR) automatic framework.

8.2 Justification of the importance of single-modality (LGE-MR) pipeline

Cardiac MRI by default is multi-parametric imaging. However, the basic CMR protocols can vary between centres. LGE-imaging is included in the majority of ischaemic or non-ischaemic cardiomyopathy assessments as any scar on the myocardium has both diagnostic and prognostic roles ([62,34]). The main reason in this work we preferred to have a single method pipeline is because, in routine clinical practice, different centres will have different acquisition protocols for cines and tissue characterization. This makes running multi-modal pipelines challenging. In addition, a single modality pipeline avoids the issue of spatial miss-registration of data. Moreover, it avoids scenarios where the quality of other modalities is sub-optimum, limiting the automated pipeline segmentation.

8.3 MA-SOCRATIS pipeline

In this part of the thesis, we have presented an automatic unsupervised segmentation pipeline (MA-SOCRATIS) which achieved impressive results when evaluated with two different datasets (STACOM at MICCAI 2019, our collection of 20 patients with challenging scar regions and LV area). Hence, we have tested the adaptability of our pipeline to new cohorts and assessed performance without the need for training, tuning, or changes to the initial hyper-parameters. The robustness and generalization of this pipeline are therefore demonstrated, and our unsupervised method achieves dice metrics comparable to state-of-the-art supervised techniques.

Moreover, we have highlighted the importance of correct detection of myocardium boundaries in LGE-MRI in patients with infarct tissue, in both automatic and manual segmentation. Furthermore, we elucidated the sensitivity of our two pipelines to GT from both repeated observations and different experts. As a result, we clarified the robustness of our pipeline, which produces generalized and accurate results free of intra-observation and inter-observer bias effects. We demonstrated a robust unsupervised combination of Rician-Gaussian mixture model and watershed techniques for

automatic scar segmentation of LV in LGE images. Lastly, we justify the importance of an unbiased ground truth to evaluate an automatic pipeline; thus, no overestimation or underestimation effects of the metrics results can be occurred.

8.4 DL-SOCRATIS pipeline

In this section of the thesis, we studied the automatic segmentation results of the myocardium and scar regions if we changed the initial myocardium segmentation step of MA-SOCRATIS. We decided to change the multi-atlas technique by deep learning (DL) networks. We chose this alternative as: i) the DL networks are less time consuming than the multi-atlas technique ii) the DL structure can be trained based on big data of cardiac shape, size, and external conditions (noise and artifacts from different scans and vendors) and iii) DL does not need any landmark initialization. As we wanted to increase as much as possible the cross-vendor automatic segmentation efficiency of the deep learning pipeline; we trained the u-net network based on a sensitivity, specificity dice-weighted cross-entropy (SS-DCE) cost function and the dice accuracy in the cross-vendor validation of 40 patients increased from 96.0 to 96.7 % for the endocardial and 85.8 to 89.8 % for the epicardial regions compared with the ordinary used dice cross entropy (DCE) cost function.

Because the segmentation of scar and healthy myocardium of LV in LGE-MR images is tricky even for experts, we used a variation of different deep learning techniques to train DL-SOCRATIS's initial segmentation of LV myocardium (a. u-net training only in cine-MRI images and testing in LGE-mages, b. CGAN image to image translation from LGE-MRI to cine-MRI scan modalities and testing in LGE-images, c. u-net training in LGE-images, and testing in LGE-images, d. u-net training in LGE-images, tuning in the cohort of interest and testing in LGE-MR images).

The multi-atlas technique delivers better results than the deep learning (DL) approach. However, the DL-SOCRATIS is less time consuming and does not need any landmark determination as

in multi-atlas method. Thus, the DL-SOCRATIS is still a strong candidate for good quality of fully automatic segmentation results without the need of an expert. The fact that it utilizes only one type of scanning modality (cine-MRI or LGE-MRI) without any need of further multi-modal supervision (bSSFP, T2, LGE images) is positioned in an advantage state contrary to the majority of the state of the art techniques. Moreover DL-SOCRATIS pipeline is flexible enough and it can be used as a quick initial evaluation of scar and myocardial regions in cross-modality analysis tasks; such as when expert radiologists have different or constrained modalities of scan images for a specific patient (only cine-MRI, or only LGE-MRI, or a mixed of these two).

Summarizing, in this study we delivered the SOCRATIS framework; which involves the choice of optimal speed or optimal robustness and accuracy, based on which an expert radiologist wants to solve the problem of the fully automatic segmentation of the LV with scars based on LGE-MRI images.

8.5 BZ-SOCRATIS pipeline

In the last part of the PhD thesis, we have presented an automatic unsupervised scar segmentation pipeline (BZ-SOCRATIS) and an automatic unsupervised scar segmentation pipeline (MA-SOCRATIS) which achieved impressive results. Hence, we have evaluated our pipeline in a new cohort and assessed performance without the need for training, tuning, or changes to the initial hyper-parameters. The robustness of our unsupervised method was compared with established and state-of-the-art supervised techniques and delivered high accuracy.

It is essential to mention that this is a novel complete study regarding automatic and semi-automatic segmentation techniques of both border zone and core scar regions in LV with infarct LGE-MR images. Moreover, we have highlighted the importance of correct detection of myocardium boundaries in LGE-MRI in patients with infarct tissue, in both automatic and manual segmentation. We

demonstrated a robust unsupervised combination of Rician-Gaussian mixture model and watershed techniques for automatic total scar and border zone core scar region segmentation of LV in LGE images. We justified the importance of an unbiased fully manually ground truth (contour extraction based) to evaluate an automatic pipeline; thus no overestimation or underestimation effects of metrics scores can be occurred. Lastly, we delivered a 3D geometric representation/mapping of the border zone, core scar, and LV fully automatically and share an automatic system framework for expert radiologist interaction in a 3D space.

A

Appendix

Appendix A1. Automatic segmentation thoroughly background

A.0.1 Overview

In this section of Appendix, we would like to extend the literature review of automatic segmentation techniques. Reconstruction of anatomical models is related to segmentation of cardiac CT, MRI images into the labeled region. Manual segmentation of medical images is a time consuming

and tricky operation. Valid threshold values and contours of the cardiac region are not always easy determinable due to the variation of the cardiac anatomy, thus an expert clinician is needed. There are studies of manual segmentation of heart ([59] ,[75] ,[113]). Segmentation is recommended by clinicians to be automated due to time consuming and subjective characteristics of the manual process. Methods for automatic segmentation is a common case study in many literature studies ([3], [103], [58]). The main challenges of automatic CMR segmentation methods are the non-clear localization of cavity borders due to blood flow, the artifacts and the noise of images, and the presence of papillary muscles in the LV and trabeculations in the RV (have the same gray level as the myocardium) which have to exclude from myocardium segmentation, as the high complexity of patient-specific shape of the RV. Because isolated RV pathology is not so common, most research effort has focused on the LV ([173]).

There are several techniques for automatic segmentation of the cardiac regions and the whole heart. In the literature the automatic segmentation algorithms are divided mainly in four categories ([85],[71]) boundary-driven, region-based, and statistical shape models. This section is organised as follow: A.02 automatic segmentation background of our pipelines, A.03 boundary-driven, A.04 region-based, A.05 statistical shape models (model-based), and A.06 hybrid techniques.

A.0.2 Automatic segmentation background of our pipelines

Probabilistic Atlas

Atlases have been used in many model fitting application of medical imaging. Frangi et al. ([59]) outlined an automatic construction of atlas from MRI and CT images (Fig. A.1). These images were segmented in regions and aligned to a reference model. A basic atlas model was then created by an iterative algorithm. Via a shape based blending ([59]) an average atlas model was created. The testing images and the average atlas model were transformed to natural coordinate systems. Landmarks in

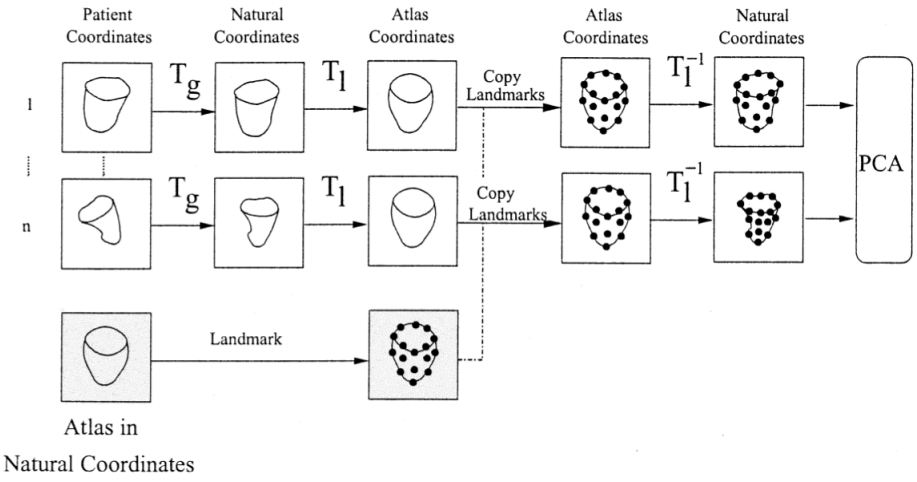

Figure A.1: Pipeline of anatomical cardiac atlas

the testing images were generated. Via a non-rigid transformation matrix the nature coordinate systems of testing images were modified on atlas nature coordinates system. Then landmark points of atlas model, were copied and applied to them. These models were converted back to their initial nature coordinates through a inverse local nonrigid transformation matrix. A Principal Component Analysis (PCA) system was used to produced statistical models ([71]). Multi-atlas structures involve a similar process, iterated through multi-vendors samples.

Deep Learning

Deep learning algorithms are state-of-the-art methodology for cardiac automatic segmentation and region of interest (ROI) detection. Avendi et al. ([13]) detected the ROI of left ventricle by using a convolution network. The network combined from four layers; two convoluted features:

$$C_l[i,j] = f(Z_l[i,j]) \tag{A.1}$$

$$Z_l[i,j] = \sum_{k_1=1}^{11} \sum_{k_2=1}^{11} F_l[k_1,k_2] I[i+k_1-1, J+k_2-1] + b_o[l] \tag{A.2}$$

where i and j the index of pixel location, and two average pooling features:

$$P_l[i_1,j_1] = \frac{1}{6} \sum_{i=(6i_1-5)}^{6i_1} \sum_{j=(6j_1-5)}^{6j_1} C_l[i,j] \tag{A.3}$$

where I(x,y) defines image intensity, F_l convolution filter and b_o biased weight.

A.0.3 K-Means:

An unsupervised region-based technique is k-means clustering. Bayes rule be determined by $p(x|y) \propto p(y|x)p(x)$, given an image y and its segmented ROI x. Let assume that there are k region of interest to segment. Moreover, let assume that distributions p(x) and p(y|x) follow $exp(-\sum_C(V_c(x)))$ and $exp(-\sum_s(\frac{1}{2\sigma_s^2}(y_s-\mu_s)^2))$ respectively. Then $exp(-\sum_s(\frac{1}{2\sigma_s^2}(y_s-\mu_s)^2) - \sum_C(V_c(x)))$ where μ_s is the mean and σ_s the variance at a pixel locations (posteriori probability). Clique potential $V_c(x)$ was equal with a constant β if $x_s = x_t$ or - β if $x_s \neq x_t$ with s,t are belong to a clique C and t be the pixels of a define neighborhood of s ([31]). Clique is a set of sites where all the pixels are neighborhood. The k-means problem solved by optimization algorithms.

Active Contour:

Active contour is a main methodology for detecting edges and lines. Kass et al. ([92]) called them 'snake'. The basic snake equilibrium involves internal and external image constraint ('forces'). The external forces are functions of constraints that they have to set the snake near the region of interest. The internal forces is related to piecewise smoothness constraint of curves. The image 'forces' are equations that used to detect the edges, lines and subjective contours of image topology. The active

contour theory follows an optimization-minimization problem of these 'forces', which guide the snake to the ROI contour. The main equations are ([92]):

$$E^*_{snake} = \int_0^1 E_{snake}(v(s))ds = \int_0^1 E_{int}(v(s)) + E_{image}(v(s)) + E_{con}(v(s))ds \tag{A.4}$$

The internal energy is given from:

$$E_{int} = (\alpha(s)|v_s(s)|^2 + \beta(s)|v_{ss}(s)|^2)/2 \tag{A.5}$$

where the $\alpha(s)$ is a first order term related to the relative importance of the membrane and the $\beta(s)$ controls the thin-plate terms (v(s) is the parametrical position of snake $v(s) = (x(s), y(s))$ and v_s is the first derivative with respect of s). The image energy is given from:

$$E_{image} = w_{line}E_{line} + w_{edge}E_{edge} + w_{term}E_{term} \tag{A.6}$$

where the E_{line} is the intensity of image, $E_{edge} = -|\nabla I(x, y)|^2$, and the $E_{term} = \frac{d^2C/dn^2_\perp}{dC/dn}$. The $C(x, y) = G_\sigma(x, y)I(x, y)$ and is a slightly version of image (G_σ is a Gaussian of standard deviation of "σ), $n = (cos\theta, sin\theta)$, $n_\perp = (cos\theta, -sin\theta)$ and $\theta = tan^{-1}(C_y/C_x)$.

Geodesic active contour is a combination of active contour theory and minimal distance curves theory in Riemannian space ([28]). More analytically, the basic function of snake energy including an edge regularization order is given by:

$$E(C) = \alpha \int_0^1 |C'(q)|^2 dq + \beta \int_0^1 |C''(q)|^2 dq + \lambda \int_0^1 |\nabla I(C(q))|dq \tag{A.7}$$

Caselles et al. ([28]) justified that the curve smoothing will be obtained when $\beta = 0$, so internal energy of equation (3.5) take into account only the first regularization order. Moreover, they assume

a restricted decreasing function g(r), where when it approaches the infinite the r goes to zero. Thus the edge regularization part can become:

$$E(C) = \alpha \int_0^1 |C^{'}(q)|^2 dq + \lambda \int_0^1 g(|\nabla I(C(q))|)^2 dq \tag{A.8}$$

The last part is the external energy of the curve. The equation (3.8) is the geodesic active contour characteristic equation.

The equation (3.5) needs to minimized from a specific curve C(q). Caselles et al. ([28]) shown that a geodesic curve in a Riemannian space, is the solution. This justified from the theorem of Maupertuis' Principle which defines when an energy problem have a geodesic curve solution in a Riemannian space ([14]). The geodesic curve solution in a Riemannian space are given from the functions:

$$U(C) = -\lambda g(|\nabla I(C)|)^2, \tag{A.9}$$

$$E(C) = \int_0^1 L|C(q)|dq, \tag{A.10}$$

$$L(C) = m/2|C^{'}|^2 - U(C), \tag{A.11}$$

$$H(C) = p^2/2m + U(C), \tag{A.12}$$

$$g_{ij} = 2m(E_o - U(C))\delta_{ij} \tag{A.13}$$

where $\alpha = m/2, p = mC^{'}$, H is Hamiltonian anf L Lagrangian equations. Caselles et al. ([28])

solve the minimization problem by the use of the light rays problem ([14], [25]) by applying the equation above. By using the theorem of Fermat's principle they defined the light waves metric in Riemannian space ([14]) given by:

$$g_{ij} = \frac{1}{c^2(X)}\delta_{ij} \tag{A.14}$$

where c(X) is the speed light at X point, and $c(X) = 1/g(X)$. The high values of speed light are edges, and the low are non-edges. Casseles et al. ([28]) used the above equation and solved the minimization problem, by applying optic theory and minimizing instead of equation (3.5) the:

$$\int_0^1 g(|\nabla I(C(q))|)|C^{'}(q)|dq \tag{A.15}$$

with $E_o = 0$.

The distance transformation that they use in Riemannian space is the Euclidean ([66]). This transformation used to define the Euclidean length curve in Riemannian space (where the Euclidean length is L). The $C_t = \kappa N$ is the flow that gives the fastest way to reduce the length equation. Base on that the final Euclidean length curve is $L_R = \int_0^L (C)g(|\nabla I(C(q))|)ds$ where $L = \int ds$. This transformation determine the geodesic curve flow that is the principle of geodesic active contour methodology. Finally the minimization problem is just an optimization of the L_R function, by using the main curve evolution flow:

$$\frac{dC(t)}{dt} = g(I)\kappa N - (\nabla g N)N \tag{A.16}$$

where the κ is the Euclidean curvature and N is the unit inward normal. This equation is delivered based on the snake generalized equation, and the use of theorem of Maupertuis' Principle.

Generally, an edge-detector can be defined as a decrease function g(z) (when z approaches infinite

the g(z) is zero) and it is depend on the gradient of image I. Then the gradient define as:

$$g(|\nabla I(x,y)|) = \frac{1}{1+|\nabla G_\sigma(x,y) * I(x,y)|^p}|, \quad p \geq 1 \tag{A.17}$$

where G_σ a Gaussian function and $G_\sigma(x,y) * I(x,y)$ a smoother version of I ([29]).

Using the level set formulation ([142]) the derivation of (4.13) becomes:

$$\frac{du}{dt} = g(I)|\nabla u|\kappa + \nabla g(I)\nabla u \tag{A.18}$$

where the C represents by a level set of u ($u_t = |\nabla u|$) and $\kappa = div(\frac{\nabla u}{|\nabla u|})$.

The equation (3.12) is the minimization equation has to be solved, detecting the contour of interest. Caselles et al. ([28]) use the level set formulation and solve the geodesic problem of equation (3.12) by the formulation :

$$\frac{du}{dt} = |\nabla u| \quad \text{div}(g(I)\frac{\nabla u}{|\nabla u|} + cg(I)|\nabla u|, \tag{A.19}$$

where $cg(I)|\nabla u|$ is an area constraint of (4.12) minimization problem (c Lagrange multiplier). Based on the use of mean curvature of Osher and Sethian ([143]) the (4.16) become the final geodesic active contour model with level set formulation:

$$\begin{cases} \frac{du}{dt} = |\nabla u| div(g(I)\frac{\nabla u}{|\nabla u|}) + cg(I)|\nabla u|, & \text{in} \quad [0,\infty] \times \Re^2 \\ u(t,x,y) = u_o(x,y), & \text{in} \quad \Re^2 \quad \text{for} \quad t = 0 \end{cases} \tag{A.20}$$

An interesting application of a prior statistical shape curve (shape and pose) and geodesic active contour is the one of Leventon ([107]). They used as an initial surface curve, a statistic shape model based on principal component analysis method (PCA). A prior knowledge of a curve of interest (COI) $\tilde{u}$ can be estimate by:

$$\tilde{u} = U_k \alpha + \mu \tag{A.21}$$

where $\alpha = U_k^T(u - \mu)$ and U_k is the k first columns of U and are the k-first principal components of shape model. In the majority of cases the α follows an Gaussian distribution. Leventon et al . ([107]) estimated pose, and shape of the specific patient image , and solves the optimization problem of equation (3.12) by using the maximum α posteriori (MAP) approach:

$$< \alpha_M AP, p_M AP >= argmax_{a,p} P(\alpha, p|u, \nabla I) \tag{A.22}$$

where α is the shape parameter given from PCA prior shape and p is the rigid pose parameters of shape (translation and rotation with respect of PCA training scheme). To compute the maximum posteriori final curve they used Bayes' Rule and solve the:

$$P(\alpha, p|u, \nabla I) = \frac{P(u|\alpha, p)P(\nabla I|\alpha, p, u)P(\alpha)P(p)}{P(u, \nabla I)} \tag{A.23}$$

where $P(u|a, p) = exp(-V_{outside})$ ($V_{outside}$ the volume outside of the estimated compute edge (u^*) at t+1 time) and $P(\nabla I|\alpha, p, u) = exp(-h(\alpha, p) - \nabla I^2)$ with $h(\alpha, p)$ the best fit of Gaussian distribution to the sample (u^*, ∇I).

Lastly, they used as evolve surface the equation:

$$u(t+1) = u(t) + \lambda_1(g(c+\kappa)|\nabla u(t)| + \nabla u(t)\nabla g) + \lambda_2(u(t)^* - u(t)) \tag{A.24}$$

where $u(t)^*$ is the final estimation of curve in t: time and λ_1 and λ_2 are coefficients between [0,1]. These coefficients used to variate the influence of the shape model and the gradient-curvature model respectively.

A.0.4 Boundary-driven techniques

In boundary-driven segmentation the most common algorithm is the active contour or "snake". The "snake" uses the external and internal image forces and a minimization problem is solved. Kass ([92]), was the first who applied the deformable models (snakes) for image segmentation. The active contour algorithm has some limitations regarding the contour estimation. The curve can agglutinates in a local minimum near the initial counter or passes out of the boundaries of the fields of view. Moreover the accurate of algorithms is limited to the type of the minimization technique. The geodesic active contour is an extension of "snake" algorithm where the level set formulation is utilized. In the majority of the active contour methods a edge-detector filter has to be applied, so the contour of interest can be detected.

A.0.5 Region-based techniques

The region-based techniques involve the selection of a global model so the estimation of the region of interest (ROI) could be accomplished. Kang et al. ([85]) classified this techniques in three subcategories the parametric representation algorithms, level-set and clustering. In parametric representation field is dominated by the **Mumford-Shah functional** and the Bayesian, maximum α posteriori

(MAP) estimators. The Mumford-Shah function ([48]) utilizes a piecewise smooth model given by:

$$E(f, C) = \mu \text{Length}(C) + \lambda \int \int_R (f(x,y) - I(x,y))^2 \, dx \, dy$$

$$+ \int \int_{R-C} ||\nabla f(x,y)||^2 \, dx \, dy + \mu |C| \quad \text{(A.25)}$$

where λ and μ are positive parameters C is the boundary length, R is a domain I the image and f the piecewise smooth function which estimate the image of interest. The computational cost and the no robustness in the face of strong noise, enact the use of a fuzzy algorithm with Bayesian and MAP estimators. Chan et al. ([29]) implemented an active counter based on techniques of Mumford-Shah segmentation and level sets methods. The novelty was that the model does not utilize an edge-function to converge the detection of the desired boundary, there is no need of smoothing the initial image, and the model can detect object without the use of gradient definition. Finally, this method can determine interior curve automatically, starting by only one initial curve.

A level set formulation of a curve C based on a $\varphi(x,y)$ Lipschitz function is defined by ([29]), by:

$$\begin{cases} C = d\omega = (x,y) \in \Omega : \varphi(x,y) = 0 \\ \text{inside}(C) = \omega = (x,y) \in \Omega : \varphi(x,y) > 0 \\ \text{outside}(C) = \Omega / \omega = (x,y) \in \Omega : \varphi(x,y) < 0 \end{cases} \quad \text{(A.26)}$$

where Ω is the pixels of image and x,y the index of a pixel (Figure A.2). [29] utilize the energy equation regarding the optimization problem:

$$F(c_1, c_2, C) = \mu \text{Length}(C) + \nu \text{Area(inside(C))}$$

$$+ \text{"}\lambda_1 \int_\Omega (u_o(x,y) - c_1)^2 \, dx \, dy$$

$$+ \text{"}\lambda_2 \int_\Omega (u_o(x,y) - c_2)^2 \, dx \, dy$$

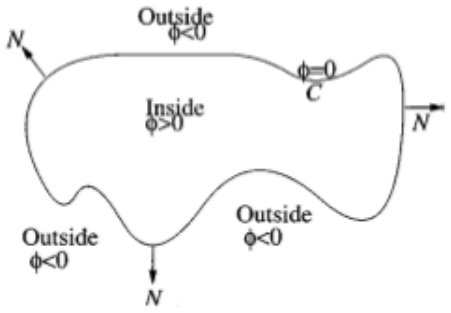

Figure A.2: curve C propagation in N direction

(A.27)

Based on the above equation, the Mumford-Shah function ([48]), and the Heaviside function H, Chan et al. established the following energy equation:

$$F(c_1, c_2, \varphi) = \mu \int_{\Omega} \delta_{\varepsilon}(\varphi(x,y))|\nabla\varphi(x,y)|\,dx\,dy$$

$$+ \nu \int_{\Omega} H_{\varepsilon}(\varphi(x,y))\,dx\,dy$$

$$+ \lambda_1 \int_{\Omega} (u_o(x,y) - c_1)^2 H_{\varepsilon}(\varphi(x,y))\,dx\,dy$$

$$+ \lambda_2 \int_{\Omega} (u_o(x,y) - c_2)^2 (1 - H_{\varepsilon}(\varphi(x,y)))\,dx\,dy$$

(A.28)

where H_{ε} is a regularization form of H given by:

$$H_\varepsilon(z) = \begin{cases} 1, & \text{if } z > \varepsilon \\ 0, & \text{if } z < -\varepsilon \\ \frac{1}{2}[1 + \frac{z}{\varepsilon} + \frac{1}{\pi}\sin(\frac{\pi z}{\varepsilon})], & \text{if } |z| \leq \varepsilon \end{cases} \tag{A.29}$$

and $\delta_\varepsilon = \frac{dH_{epsilon}(z)}{dz}$. Lastly, they use Euler-Lagrance equation to update the $\varphi(t,x,y)$:

$$\begin{cases} \frac{d\varphi}{dt} = \delta_{epsilon}[\mu div(\frac{\nabla\varphi}{|\nabla\varphi|}) - \nu - \lambda_1(u_o - c_1)^2 + \lambda_2(u_o - c_2)^2] = 0 \in (0,\infty) \times \Omega, \\ \varphi(0,x,y) = \varphi_o(x,y) \in \Omega \quad \text{(initial contour)} \\ \frac{\delta_\varepsilon(\varphi)}{|\nabla\varphi|}\frac{d\varphi}{d\vec{n}} = 0 \in d\Omega \end{cases} \tag{A.30}$$

where *vecn* external normal to the boundary $d\Omega$.

An other region-based techniques is the **k-means clustering**. Given an image y and the segmented region of interest x the Bayes rule is $p(x|y) \propto p(y|x)p(x)$. Let assume that there are k region of interest for segmentation. Then by assuming that the distribution p(x) and p(y|x) follows $exp(-\sum_C(V_c(x)))$ and $exp(-\sum_s(\frac{1}{2\sigma_s^2}(y_s - \mu_s)^2))$ respectively. The final posteriori probability can define as $exp(-\sum_s(\frac{1}{2\sigma_s^2}(y_s - \mu_s)^2) - \sum_C(V_c(x)))$ where μ_s is the mean and σ_s the variance at a pixel location s, clique potential $V_c(x)$ is equal with a constant β if $x_s = x_t$ or - β if $x_s \neq x_t$ with s,t are belong to a clique C (clique is a set of sites where all the pixels are neighborhood) with t be the pixels of a define neighborhood of s ([31]). The k-means implement using optimization techniques such us MAP.

Graph-cut techniques

Graph cut techniques uses optimization algorithms to detect the regions with maximum flow for binary images. The most common optimization algorithm is the MAP.

A.0.6 Model-based techniques

An other common automatic segmentation technique is the statistical shape algorithms. These algorithms use a predefined statistical shape model to fit in a specific patient image region of interest (ROI). We will present bellow the mathematics of some basic model based techniques.

To create a statistical shape model, each shape in the training data determine from points (landmarks) $x = (x_1, y_1, z_1, ...x_n, y_n, z_n)^T$). From this land marks points a mean shape model can extract as $\mu = 1/n \sum x_i$ and a co-variance matrix S $S = 1/(n-1) \sum (x_i - \mu)(x_i - \mu)^T$ ([43]). Moreover, model-based techniques uses the PCA formulation to compute the eigenvector and eigenvalues of S matrix, and the model parameter vector α.

Active shape model (ASM) is a technique which initially is used to define and locate rigid objects. Cootes et al. ([43])outlined the method of ASM. Shape model is a statistically based technique, rely on representing the object of a flexible objects via determine labeled points. Through the statistic study of the position and shape of the objects a point distribution model (PDM) is used by the authors. Cootes et al. ([43]) demonstrates the ASM method focusing in the statistics of a set of shapes, and the fitting techniques of the training sample. To begin with, the chosen parts of the images have to be labeled, and a set of training shapes must be defined. The alignment of the shape by scaling rotating and translating transformation with respect to a set of axes is needed (aligning the training set). A minimization problem between the landmarks of each training sample is solved. Moreover, the mean shape **x**, co-variance matrix (S) of the shapes are computed and the eigenvector p of S, eigenvalue λ are generated as:

$$Sp = \lambda p \tag{A.31}$$

$$x = \mathbf{x} + Pb \tag{A.32}$$

where P is the matrix of the eigenvectors and b the vector of weights. If a Gaussian distribution is

assumed to solve the statistic problem of the shape model, then b can be obtained by solving the Mahalanobis distance minimization (D_m).

$$g = \mathbf{g} + Pb \tag{A.33}$$

Regarding the gray level appearance modeling, the utilize of (4) is needed (g is the gray level of the shapes). In this method, the measure of how well the model fits the profile is used to minimize the distance from the object with gray scale intensity. A minimization problem is used based on the Mahalanobis distance and the sum of the squares of differences R^2 given by:

$$R^2 = (g - \mathbf{g})^T(g - \mathbf{g}) - b^T b \tag{A.34}$$

Based on the flexible shape model above, Cootes et al. ([43]) define the PDM and used the ASM for image search. To define the position, orientation, and scale of the target shape of the image, create an instance X by solving the equation:

$$X = M(s, \theta)[x] + X_c \tag{A.35}$$

where s is the scale of the image and θ the rotating angle. The approach the authors used was to calculate the displacement of the point required to move it to a better location, adjustment the pose and scale of the shape parameters basing the point distribution method (PDM), and update the model parameters. By this way to the test samples are remaining similar to those of the training set.

Deformable models is based on the elastic and deformation theory. Mcinerney et al. ([123]) elucidates the deformable models, a technique to assisted medical image analysis. Deformable model is an elastic model which deforms by 'forces' and restricted by constraints. Thus, an external potential energy defines, based on the data of interest, to apply 'forces' in the model to fit at the data. The most

public deformable model is the snakes ([134]). The snakes is a parametric template initialized inside the image I(x,y). The template is given by:

$$v(s) = (x(s), y(s))^T, s \in [0, 1] \tag{A.36}$$

where x,y are the coordinate functions of the 2D image and s is the parametric domain. The shape of the contour subject to an image is given by:

$$E(v) = S(v) + P(v) \tag{A.37}$$

S is the internal deformation energy given by:

$$S(v) = \int_0^1 w_1(s)|\frac{dv}{ds}|^2 + w_2(s)|\frac{d^2v}{ds^2}|^2 \, ds \tag{A.38}$$

The w1(s) controls the tension and w2(s) the rigidity of the snake.

The P(x,y) is defined by:

$$P(v) = \int_0^1 P(v(s)) \, ds \tag{A.39}$$

which is a scalar potential function to align the snake with the image. For instance this can be done choosing a potential $P(x, y) = -c|\nabla[G_s * I(x, y)]|$ to attract the intensity extrema of the Image (where Gs is a Gaussian smoothing filter convoluted with the image). The extrema of the external potential P(v) are designed to concur with the edges boundaries of the region, and other image features of interest. Regarding the use of this the minimization of the energy E(v) is needed. The equilibrium equation that the minimization problem must be satisfied is the Euler-Lagrange equation. The utilization of this model in dynamic phenomena such as anatomical structures which are

deformable. Thus, nonrigid motion (heart cycle, etc.) could be delivered via dynamic deformable models. For example, the dynamic snake is taking into account the inertia and damping forces (mass density m(s), and damping density g(s) ([123])). To solve the optimization problem, there are two popular methods a discretization of the energy E(v), and a probabilistic framework approach. The discretization of the energy can be delivered by using finite elements, finite differences, geometric splines (local basis functions), and Fourier bases (global basis functions). The time derivatives of an dynamic snake model are approximated by finite difference and explicit or implicit numerical time integration methods. The discrete form of energy is written as:

$$M\ddot{u} + C\dot{u} + Ku = f \tag{A.40}$$

where M is the mass matrix, C is the damping matrix and $f = -\nabla P$ with P the external potential. Concerning the probabilistic framework, a maximization of a posteriori probability is used p(u|I) based on Bayes' theorem ($p(u|I) = \frac{p(I|u)p(u)}{p(I)}$). The u represents the deformable model shape parameters and I the image. A Boltzmann or Gibbs distribution is used. The prior probability is defined by:

$$p(u) = \frac{1}{Z_s}\exp(-S(u)) \tag{A.41}$$

and the likelihood probability by:

$$p(I|u) = \frac{1}{Z_i}\exp(-P(u)) \tag{A.42}$$

The probabilistic framework can be applied in dynamic models or in system models by resulting in Kalman filter ([179]). Mcinerney et al. ([123]) review the application of the deformable model related to image segmentation with deformable curves, volume image segmentation with deformable sur-

faces, matching, motion tracking analysis, and incorporating an initial knowledge of the geometry (pose and shape).

Bilinear models are based on the point distribution analysis theory (PDA). The main framework to solve a two factor task using bilinear models is described by Tenenbaum et al. ([178]). The bi-linearity is defined as the effect of the observations to be linear with respect to the variable factor when the other factor remains constant (property of separability). Tenenbaum et al. ([178]) employ the terms of style and content for the variable and invariable independent factors, to subsentence a set of observations. Contrary to the hierarchical factorial model, the fitting model techniques are more efficient for the study of linear models (singular value decomposition SVD and the expectation-maximization algorithm EM), due to the minimization of the stochastic approximations. Tenenbaum et al. ([178]), illustrates the classification, extrapolation, and translation problems of a bilinear face image problem, where the style factor was the different pose of persons and the content was the different persons. In the study of Hoogendoorn et al. ([76]) a bilinear model is created to reconstruct the shape of the heart at discrete points of the heart cycle. A style factor was determined based on the shape variability of inter-subject variation, and it contents the intrinsic dynamics of the human heart. Bilinear models separated in symmetric and asymmetric models. Fitting of a bilinear model to an array of observations of content and style classes during the training and testing phase is recommended ([178]). The aim of the model during the training (and approximately in testing) phase is to minimize the squared error over the training observations of the sample for symmetric and asymmetric models (or to maximize the likelihood estimation of the style and content parameters given the training data). The symmetric models are given by the equation:

$$y_k^{sc} = a^{s^T} W_k b^c \tag{A.43}$$

where a^s and b^c denote a dimensionality I and J respectively, and the y_{sc} denote a K-dimensional

observation of style s and content class c. The K matrices W_k describe a bilinear map from style and content vector spaces to the K-dimensional observation space. The asymmetric model is given by the equation:

$$y_k^{sc} = \sum_j a_{jk}^s b_j^c \tag{A.44}$$

where

$$\alpha_{jk}^s = \sum_{j=1}^{I} w_{ijk}^s a_i^s \tag{A.45}$$

Let assume that y(t) is the training observations (y=1...,T) and let the indicator variable of equation (4.36) be h^{sc}(t)=1 if y(t) is in style or content class, otherwise 0. Then the total square error for the asymmetric model fitting ([178]) is given by :

$$E = \sum_{t=1}^{T} \sum_{s=1}^{S} \sum_{c=1}^{C} h^{sc}(t) ||y(t) - A^s b^c||^2 \tag{A.46}$$

where the part $A^s b^c$ is compute by solving the SVD of y of matrix estimation Y. Lastly, the total square error for the symmetric model fitting ([178]) is given by :

$$E = \sum_{t=1}^{N} \sum_{s=1}^{S} \sum_{c=1}^{C} \sum_{k=1}^{K} h^{sc}(t)(y(t) - a^{sT} W_k b^c)^2 \tag{A.47}$$

A.0.7 Hybrid techniques

Image-driven or Hybrid techniques are a combination of computer vision, image processing, and segmentation techniques. The segmentation of dynamic medical (cardiac) images, such as steady state free procession (SSFP), and 3D+time short axis MRI can be performed automatically. Cocosco et al. ([35]) proposed a novel automatic image-driven segmentation method to deal with 3D+time short axis MRI (cine cardiac MRI) and the handle of both left and right ventricles of the heart.

This automatic method uses simple prior knowledge, segments heart's ventricles dynamically, decreases the time consuming, and contents an automatic method for computing the region of interest (ROI). The pipeline of this study is summarized to extract ROI by a fully automatic method. It uses an automatic unsupervised binary voxel classification from the intensity histogram of the 4D image within the ROI. It is applying a binary thresholding method with no adjustable parameters ([144]), and it uses to the foreground of the 4D image thinning operation to label the region of the 3D+time binary image. It uses a region growing technique based on the previous classification to compute the difference between the extrema of the spatial volume (X-Y-Z) for each connected component, and it uses a final region growing in the Z direction. With this process, ([35]) create an segmented 3D+time image with labeled voxels of the MRI.

Table A.1: Summary of initial parameters of four RG models.

first case: Fix means values				
Initial Parameters	RG	RGG	RRG	RRGG
portion first Rician model	0.5	0.5	0.25	0.25
mean value first Rician model	0	0	0	0
variance first Rician model	250	250	250	250
portion second Rician model	-	-	0.25	0.25
mean value second Rician model	-	-	50	50
variance second Rician model	-	-	250	250
portion first Gaussian model	0.5	0.25	0.5	0.25
mean value first Gaussian model	200	200	200	200
variance first Gaussian model	250	250	250	250
portion second Gaussian model	-	0.25	-	0.25
mean value second Gaussian model	-	100	-	100
variance second Gaussian model	-	250	-	250

Appendix B1. Additional details of Rician Gaussian mixture model

Threshold detection component: We described the use of a Rician-Gaussian mixture model ([175] and [164]) to estimate the intensity threshold of healthy and unhealthy myocardium. We defined a threshold value of scar in each myocardium region. We used four different mixture models of Rician-Gaussian (RG) distributions (RG, RGG, RRG, and RRGG), and these are shown in Figure 4.9. For each LGE-MRI image, we tested each combination of RG mixture model. Then, we sorted the mean values of each RG model from maximum to minimum. The maximum value was defined as Gaussian threshold (TG) and the minimum as Rician threshold (TR). We assumed two different cases of mixture model initialisation; fixed mean values (pixel inetnsity of: 0, 60, 100, 200) and initial mean values based on Hennemuth et al. ([72]) study.

Additional details of Rician Gaussian mixture model configurations

The main description of RGMM framework is presented in chapter 4.2. Table A.A.2 shows the initial parameters of GR models based on Hennemuth et al. ([72]) assumption extended to three and

Table A.2: Summary of Hennemuth et al. ([72]) study initial parameters of four RG models.

second case: Initial parameters based on[72]				
Initial Parameters	RG	RGG	RRG	RRGG
(ar) portion first Rician model	$freq_{max}$	$freq_{max}$	$freq_{max}$	$freq_{max}$
(mr) mean value first Rician model	0	0	0	0
(sr) variance first Rician model	$observ_{max}$	$observ_{max}$	$observ_{max}$	$observ_{max}$
(ar_2) portion second Rician model	-	-	ar * 0.5	ar * 0.5
(mr_2) mean value second Rician model	-	-	Eq. A.48	Eq. A.48
(sr_2) variance second Rician model	-	-	Eq. A.48	Eq. A.48
(ag) portion first Gaussian model	1-ar	1- ar- 0.25	1- (ar * 1.5)	0.25
(mg) mean value first Gaussian model	Eq. A.48	Eq. A.48	Eq. A.48	Eq. A.48
(sg) variance first Gaussian model	Eq. A.48	Eq. A.48	Eq. A.48	Eq. A.48
(ag_2)portion second Gaussian model	-	0.25	-	0.15
(mg_2) mean value second Gaussian model	-	Eq. A.48	-	Eq. A.48
(sg_2) variance second Gaussian model	-	Eq. A.48	-	Eq. A.48

four mixture models (RGG, RRG, RRGG). The $freq_{max}$ is the maximum frequency of each image histogram, and $observ_{max}$ the maximum observation of a specific pixel intensity.

$$\begin{cases} mg = freq_{max} + (\frac{I_{max} - observ_{max}}{2}) \\ sg = sqrt(\frac{I_{max} - mg}{3}) \\ mg_2 = (mg + \frac{mg}{2}) - 1 \\ sg_2 = sg * 0.8 \\ mr_2 = (mg - \frac{mg}{2}) - 1 \\ sr_2 = sr * 0.8 \end{cases} \tag{A.48}$$

Lastly, we defined a threshold of the scar region based on the FWHM method of maximum (Gaussian mean value) and minimum values (Rician mean value). The thresholds of Rician (TR) and Gaussian (TG) distributions were used to define a linear partial volume model by:

$$p(x) = \begin{cases} 0, & \text{if } x \leq TR \\ \frac{x - TR}{TG - TR}, & \text{if } TR < x \leq TG \\ 1, & \text{if } x < TG \end{cases} \tag{A.49}$$

Bibliography

[1] Abramson, H.G., Popescu, D.M., Yu, R., Lai, C., Shade, J.K., Wu, K.C., Maggioni, M., Trayanova, N.A.: Anatomically-informed deep learning on contrast-enhanced cardiac mri for scar segmentation and clinical feature extraction (2021), https://arxiv.org/abs/2010.11081

[2] Al-Khatib, S.M., Stevenson, W.G., Ackerman, M.J., Bryant, W.J., Callans, D.J., Curtis, A.B., Deal, B.J., Dickfeld, T., Field, M.E., Fonarow, G.C., Gillis, A.M., Hlatky, M.A., Granger, C.B., Hammill, S.C., Joglar, J.A., Kay, G.N., Matlock, D.D., Myerburg, R.J., Page, R.L.: 2017 aha/acc/hrs guideline for management of patients with ventricular arrhythmias and the prevention of sudden cardiac death: Executive summary: A report of the american college of cardiology/american heart association task force on clinical practice guidelines and the heart rhythm society. Journal of the American College of Cardiology (2017). https://doi.org/https://doi.org/10.1016/j.jacc.2017.10.053, http://www.sciencedirect.com/science/article/pii/S0735109717413052

[3] Albà, X., i Ventura, R.M., Lekadir, K., Frangi, A.F.: Healthy and Scar Myocardial Tissue Classification in DE-MRI. In: Camara, O., Mansi, T., Pop, M., Rhode, K., Sermesant, M.,

Young, A. (eds.) Statistical Atlases and Computational Models of the Heart. Imaging and Modelling Challenges. pp. 62–70. Springer Berlin Heidelberg, Berlin, Heidelberg (2013)

[4] Amado, L.C., Gerber, B.L., Gupta, S.N., Rettmann, D.W., Szarf, G., Schock, R., Nasir, K., Kraitchman, D.L., Lima, J.A.: Accurate and objective infarct sizing by contrast-enhanced magnetic resonance imaging in a canine myocardial infarction model. Journal of the American College of Cardiology **44**(12), 2383 – 2389 (2004). https://doi.org/https://doi.org/10.1016/j.jacc.2004.09.020, http://www.sciencedirect.com/science/article/pii/S0735109704018376

[5] Amzulescu, M.S., De Craene, M., Langet, H., Pasquet, A., Vancraeynest, D., Pouleur, A.C., Vanoverschelde, J.L., Gerber, B.L.: Myocardial strain imaging: review of general principles, validation, and sources of discrepancies. European Heart Journal - Cardiovascular Imaging **20**(6), 605–619 (03 2019). https://doi.org/10.1093/ehjci/jez041, https://doi.org/10.1093/ehjci/jez041

[6] Andreopoulos, A., Tsotsos, J.K.: Efficient and generalizable statistical models of shape and appearance for analysis of cardiac mri. Medical Image Analysis **12**(3), 335 – 357 (2008). https://doi.org/https://doi.org/10.1016/j.media.2007.12.003, http://www.sciencedirect.com/science/article/pii/S1361841508000029

[7] Andreu, D., Berruezo, A., Ortiz-Pérez, J.T., Silva, E., Mont, L., Borràs, R., de Caralt, T.M., Perea, R.J., Fernández-Armenta, J., Zeljko, H., Brugada, J.: Integration of 3d electroanatomic maps and magnetic resonance scar characterization into the navigation system to guide ventricular tachycardia ablation. Circulation: Arrhythmia and Electrophysiology **4**(5), 674–683 (2011). https://doi.org/10.1161/CIRCEP.111.961946, https://www.ahajournals.org/doi/abs/10.1161/CIRCEP.111.961946

[8] Antzelevitch, C., Burashnikov, A.: Overview of basic mechanisms of cardiac arrhythmia. Cardiac Electrophysiology Clinics **3**(1), 23 – 45 (2011). https://doi.org/https://doi.org/10.1016/j.ccep.2010.10.012, http://www.sciencedirect.com/science/article/pii/S1877918210001425, basic Science for the Clinical Electrophysiologist

[9] Arheden, H., Saeed, M., Higgins, C.B., Gao, D.W., Ursell, P.C., Bremerich, J., Wyttenbach, R., Dae, M.W., Wendland, M.F.: Reperfused rat myocardium subjected to various durations of ischemia: Estimation of the distribution volume of contrast material with echo-planar mr imaging. Radiology **215**(2), 520–528 (2000). https://doi.org/10.1148/radiology.215.2.rooma38520, https://doi.org/10.1148/radiology.215.2.r00ma38520, pMID: 10796935

[10] Arlot, S., Celisse, A.: A survey of cross-validation procedures for model selection. Statistics Surveys **4**(none), 40 – 79 (2010). https://doi.org/10.1214/09-SS054, https://doi.org/10.1214/09-SS054

[11] Ascha, M., Renapurkar, R., Tonelli, A.: A review of imaging modalities in pulmonary hypertension. Annals of Thoracic Medicine **12**(2), 61–73 (2017). https://doi.org/10.4103/1817-1737.203742, http://www.thoracicmedicine.org/article.asp?issn=1817-1737;year=2017;volume=12;issue=2;spage=61;epage=73;aulast=Ascha;t=6

[12] Ashikaga, H., Sasano, T., Dong, J., Zviman, M.M., Evers, R., Hopenfeld, B., Castro, V., Helm, R.H., Dickfeld, T., Nazarian, S., Donahue, J.K., Berger, R.D., Calkins, H., Abraham, M.R., Marbán, E., Lardo, A.C., McVeigh, E.R., Halperin, H.R.: Magnetic resonance–based anatomical analysis of scar-related ventricular tachycardia. Circulation

Research **101**(9), 939–947 (2007). https://doi.org/10.1161/CIRCRESAHA.107.158980, https://www.ahajournals.org/doi/abs/10.1161/CIRCRESAHA.107.158980

[13] Avendi, M., Kheradvar, A., Jafarkhani, H.: A combined deep-learning and deformable-model approach to fully automatic segmentation of the left ventricle in cardiac mri. Medical Image Analysis **30**, 108 – 119 (2016). https://doi.org/https://doi.org/10.1016/j.media.2016.01.005, http://www.sciencedirect.com/science/article/pii/S1361841516000128

[14] B. A. Dubrovin, A. T. Fomenko, S.P.N.a.: Modern Geometry — Methods and Applications: Part I. The Geometry of Surfaces, Transformation Groups, and Fields, Graduate Texts in Mathematics 93, vol. Part 1. Springer New York, 2nd ed edn. (1984), http://gen.lib.rus.ec/book/index.php?md5=326D5A29E5EC933283E648C81A5DBE37

[15] Bai, J., Li, P., Wang, K.: Automatic whole heart segmentation based on watershed and active contour model in ct images. In: 2016 5th International Conference on Computer Science and Network Technology (ICCSNT). pp. 741–744 (Dec 2016). https://doi.org/10.1109/ICCSNT.2016.8070256

[16] Bai, W., Shi, W., O'Regan, D.P., Tong, T., Wang, H., Jamil-Copley, S., Peters, N.S., Rueckert, D.: A probabilistic patch-based label fusion model for multi-atlas segmentation with registration refinement: Application to cardiac mr images. IEEE Transactions on Medical Imaging **32**(7), 1302–1315 (July 2013). https://doi.org/10.1109/TMI.2013.2256922

[17] Bai, W., Shi, W., de Marvao, A., Dawes, T.J., O'Regan, D.P., Cook, S.A., Rueckert, D.: A bi-ventricular cardiac atlas built from 1000+ high resolution mr images of healthy subjects and an analysis of shape and motion. Medical Image Analysis **26**(1), 133 – 145 (2015).

https://doi.org/https://doi.org/10.1016/j.media.2015.08.009, http://www.sciencedirect.com/science/article/pii/S1361841515001346

[18] Beare R., L.G.: The watershed transform in itk - discussion and new developments (2006), http://hdl.handle.net/1926/202

[19] Ben-naceur, M., Saouli, R., Akil, M., Kachouri, R.: Fully automatic brain tumor segmentation using end-to-end incremental deep neural networks in mri images. Computer Methods and Programs in Biomedicine **166**, 39–49 (2018). https://doi.org/https://doi.org/10.1016/j.cmpb.2018.09.007, https://www.sciencedirect.com/science/article/pii/S0169260718309015

[20] Beucher, S., Lantuéjoul, C.: Use of watersheds in contour detection. workshop published (Sep 1979), http://cmm.ensmp.fr/~beucher/publi/watershed.pdf

[21] Biglands John D., Radjenovic Aleksandra, R.J.P.: Cardiovascular magnetic resonance physics for clinicians: part ii. Journal of Cardiovascular Magnetic Resonance **14**, 66 (2012). https://doi.org/https://doi.org/10.1186/1532-429X-14-66, individual Subject Prediction

[22] Bilchick, K.C.: Integration of cmr scar imaging and electroanatomic mapping: The future of vt ablation?**editorials published in jacc: Cardiovascular imaging reflect the views of the authors and do not necessarily represent the views of jacc: Cardiovascular imaging or the american college of cardiology. JACC: Cardiovascular Imaging **5**(2), 211 – 213 (2012). https://doi.org/https://doi.org/10.1016/j.jcmg.2011.09.021, http://www.sciencedirect.com/science/article/pii/S1936878X11008837

[23] BioDigital: MS Windows NT © 2018 biodigital, inc. all rights reserved. (2018), https://www.biodigital.com/

[24] Bizino, M.B., Tao, Q., Amersfoort, J., Siebelink, H.M.J., van den Bogaard, P.J., van der Geest, R.J., Lamb, H.J.: High spatial resolution free-breathing 3d late gadolinium enhancement cardiac magnetic resonance imaging in ischaemic and non-ischaemic cardiomyopathy: quantitative assessment of scar mass and image quality. European Radiology **28**(9), 4027–4035 (Sep 2018). https://doi.org/10.1007/s00330-018-5361-y, https://doi.org/10.1007/s00330-018-5361-y

[25] Born, M., Wolf, E.: Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light (7th Edition). Cambridge University Press, 7th edn. (1999)

[26] Calkins, H., Epstein, A., Packer, D., Arria, A.M., Hummel, J., Gilligan, D.M., Trusso, J., Carlson, M., Luceri, R., Kopelman, H., Wilber, D., Wharton, J., Stevenson, W., : Catheter ablation of ventricular tachycardia in patients with structural heart disease using cooled radiofrequency energy. Journal of the American College of Cardiology **35**(7), 1905–1914 (2000). https://doi.org/10.1016/S0735-1097(00)00615-X, http://www.onlinejacc.org/content/35/7/1905

[27] Campello V.M., Martín-Isla C., I.C.P.S.B.M.L.K.: Combining multi-sequence and synthetic images for improved segmentation of late gadolinium enhancement cardiac mri. Statistical Atlases and Computational Models of the Heart. Multi-Sequence CMR Segmentation, CRT-EPiggy and LV Full Quantification Challenges. STACOM 2019 (2020). https://doi.org/https://doi.org/10.1002/mp.14022

[28] Caselles, V., Kimmel, R., Sapiro, G.: Geodesic active contours. International Journal of Computer Vision **22**(1), 61–79 (Feb 1997)

[29] Chan, T.F., Vese, L.A.: Active contours without edges. IEEE Transactions on Image Processing **10**(2), 266–277 (feb 2001). https://doi.org/10.1109/83.902291

[30] Chartsias, A., Joyce, T., Dharmakumar, R., Tsaftaris, S.A.: Adversarial image synthesis for unpaired multi-modal cardiac data. In: Tsaftaris, S.A., Gooya, A., Frangi, A.F., Prince, J.L. (eds.) Simulation and Synthesis in Medical Imaging. pp. 3–13. Springer International Publishing, Cham (2017)

[31] Chen, C.W., Luo, J., Parker, K.J.: Image segmentation via adaptive k-mean clustering and knowledge-based morphological operations with biomedical applications. IEEE Transactions on Image Processing 7(12), 1673–1683 (Dec 1998). https://doi.org/10.1109/83.730379

[32] Chen, C., Ouyang, C., Tarroni, G., Schlemper, J., Qiu, H., Bai, W., Rueckert, D.: Unsupervised Multi-modal Style Transfer for Cardiac MR Segmentation. arXiv e-prints arXiv:1908.07344 (Aug 2019)

[33] Chen, J., Li, H., Zhang, J., Menze, B.: Adversarial convolutional networks with weak domain-transfer for multi-sequence cardiac mr images segmentation. In: Pop, M., Sermesant, M., Camara, O., Zhuang, X., Li, S., Young, A., Mansi, T., Suinesiaputra, A. (eds.) Statistical Atlases and Computational Models of the Heart. Multi-Sequence CMR Segmentation, CRT-EPiggy and LV Full Quantification Challenges. pp. 317–325. Springer International Publishing, Cham (2020)

[34] Chew, P.G., Dobson, L.E., Garg, P., Fairbairn, T.A., Musa, T.A., Uddin, A., Swoboda, P.P., Foley, J.R., Fent, G.J., Brown, L.A.E., Onciul, S., Plein, S., Blackman, D.J., Greenwood, J.P.: Cmr quantitation of change in mitral regurgitation following transcatheter aortic valve replacement (tavr): impact on left ventricular reverse remodeling and outcome. The International Journal of Cardiovascular Imaging 35(1), 161–170 (2019). https://doi.org/10.1007/s10554-018-1441-y, https://doi.org/10.1007/s10554-018-1441-y

[35] Chris, A., J., N.W., Thomas, N., P.A., V.E., Gunnar, L., Alexander, S., A., V.M.: Automatic image-driven segmentation of the ventricles in cardiac cine MRI. Journal of Magnetic Resonance Imaging **28**(2), 366–374. https://doi.org/10.1002/jmri.21451, https://onlinelibrary.wiley.com/doi/abs/10.1002/jmri.21451

[36] Chuang, M.L., Gona, P., Hautvast, G.L., Salton, C.J., Blease, S.J., Yeon, S.B., Breeuwer, M., O'Donnell, C.J., Manning, W.J.: Correlation of trabeculae and papillary muscles with clinical and cardiac characteristics and impact on cmr measures of lv anatomy and function. JACC: Cardiovascular Imaging **5**(11), 1115 – 1123 (2012). https://doi.org/https://doi.org/10.1016/j.jcmg.2012.05.015, http://www.sciencedirect.com/science/article/pii/S1936878X12007486

[37] Chubb, H., Karim, R., Roujol, S., Nuñez-Garcia, M., Williams, S.E., Whitaker, J., Harrison, J., Butakoff, C., Camara, O., Chiribiri, A., Schaeffter, T., Wright, M., O'Neill, M., Razavi, R.: The reproducibility of late gadolinium enhancement cardiovascular magnetic resonance imaging of post-ablation atrial scar: a cross-over study. Journal of Cardiovascular Magnetic Resonance **20**(1), 21 (Mar 2018). https://doi.org/10.1186/s12968-018-0438-y, https://doi.org/10.1186/s12968-018-0438-y

[38] Clark, M.C., Hall, L.O., Goldgof, D.B., Velthuizen, R., Murtagh, F.R., Silbiger, M.S.: Automatic tumor segmentation using knowledge-based techniques. IEEE Transactions on Medical Imaging **17**(2), 187–201 (1998). https://doi.org/10.1109/42.700731

[39] Clayton, R., Bernus, O., Cherry, E., Dierckx, H., Fenton, F., Mirabella, L., Panfilov, A., Sachse, F., Seemann, G., Zhang, H.: Models of cardiac tissue electrophysiology: Progress, challenges and open questions. Progress in Biophysics and Molecular Biology **104**(1), 22 – 48 (2011). https://doi.org/https://doi.org/10.1016/j.pbiomolbio.2010.05.008,

http://www.sciencedirect.com/science/article/pii/S0079610710000362, cardiac Physiome project: Mathematical and Modelling Foundations

[40] Cocosco, C.A., Niessen, W.J., Netsch, T., Vonken, E.j.P., Lund, G., Stork, A., Viergever, M.A.: Automatic image-driven segmentation of the ventricles in cardiac cine mri. Journal of Magnetic Resonance Imaging **28**(2), 366–374 (2008). https://doi.org/10.1002/jmri.21451, https://onlinelibrary.wiley.com/doi/abs/10.1002/jmri.21451

[41] Colli-Franzone, P., Gionti, V., Pavarino, L., Scacchi, S., Storti, C.: Role of infarct scar dimensions, border zone repolarization properties and anisotropy in the origin and maintenance of cardiac reentry. Mathematical Biosciences **315**, 108228 (2019). https://doi.org/https://doi.org/10.1016/j.mbs.2019.108228, https://www.sciencedirect.com/science/article/pii/S002555641830590X

[42] Concetta Zito, Luca Longobardo, R.C.e.a.: Ten years of 2d longitudinal strain for early myocardial dysfunction detection: A clinical overview. BioMed Research International (Jul 2018)

[43] Cootes, T.F., Hill, A., Taylor, C.J., Haslam, J.: Use of active shape models for locating structures in medical images. Image and Vision Computing **12**(6), 355–365 (1994). https://doi.org/https://doi.org/10.1016/0262-8856(94)90060-4, http://www.sciencedirect.com/science/article/pii/0262885694900604

[44] Cootes, T., Taylor, C., Cooper, D., Graham, J.: Active shape models-their training and application. Computer Vision and Image Understanding **61**(1), 38 – 59 (1995). https://doi.org/https://doi.org/10.1006/cviu.1995.1004, http://www.sciencedirect.com/science/article/pii/S1077314285710041

[45] Couprie, C., Grady, L., Najman, L., Talbot, H.: Power watershed: A unifying graph-based optimization framework. IEEE Transactions on Pattern Analysis and Machine Intelligence **33**(7), 1384–1399 (July 2011). https://doi.org/10.1109/TPAMI.2010.200

[46] Daniel Perry, Alan Morris, N.B.C.M.R.M.J.C.: Automatic classification of scar tissue in late gadolinium enhancement cardiac mri for the assessment of left-atrial wall injury after radiofrequency ablation. vol. 8315, pp. 8315 – 8315 – 9 (2012). https://doi.org/10.1117/12.910833, https://doi.org/10.1117/12.910833

[47] Das, S., Dey, A., Pal, A.: Applications of artificial intelligence in machine learning : Review and prospect (2015)

[48] David, M., Jayant, S.: Optimal approximations by piecewise smooth functions and associated variational problems. Communications on Pure and Applied Mathematics **42**(5), 577–685. https://doi.org/10.1002/cpa.3160420503, https://onlinelibrary.wiley.com/doi/abs/10.1002/cpa.3160420503

[49] Desjardins, B., Crawford, T., Good, E., Oral, H., Chugh, A., Pelosi, F., Morady, F., Bogun, F.: Infarct architecture and characteristics on delayed enhanced magnetic resonance imaging and electroanatomic mapping in patients with postinfarction ventricular arrhythmia. Heart Rhythm **6**(5), 644–651 (2009). https://doi.org/https://doi.org/10.1016/j.hrthm.2009.02.018, https://www.sciencedirect.com/science/article/pii/S1547527109001672

[50] Detsky, J.S., Paul, G., Dick, A.J., Wright, G.A.: Reproducible classification of infarct heterogeneity using fuzzy clustering on multicontrast delayed enhancement magnetic resonance images. IEEE Transactions on Medical Imaging **28**(10), 1606–1614 (Oct 2009). https://doi.org/10.1109/TMI.2009.2023515

[51] Ecabert, O., Peters, J., Schramm, H., Lorenz, C., von Berg, J., Walker, M.J., Vembar, M., Olszewski, M.E., Subramanyan, K., Lavi, G., Weese, J.: Automatic model-based segmentation of the heart in ct images. IEEE Transactions on Medical Imaging **27**(9), 1189–1201 (Sep 2008). https://doi.org/10.1109/TMI.2008.918330

[52] Ecabert, O., Peters, J., Walker, M.J., Ivanc, T., Lorenz, C., von Berg, J., Lessick, J., Vembar, M., Weese, J.: Segmentation of the heart and great vessels in ct images using a model-based adaptation framework. Medical Image Analysis **15**(6), 863 – 876 (2011). https://doi.org/https://doi.org/10.1016/j.media.2011.06.004, http://www.sciencedirect.com/science/article/pii/S1361841511000910

[53] El Berbari, R., Bloch, I., Redheuil, A., Angelini, E.D., Mousseaux, E., Frouin, F., Herment, A.: Automated segmentation of the left ventricle including papillary muscles in cardiac magnetic resonance images. In: Sachse, F.B., Seemann, G. (eds.) Functional Imaging and Modeling of the Heart. pp. 453–462. Springer Berlin Heidelberg, Berlin, Heidelberg (2007)

[54] Estner, H.L., Zviman, M.M., Herzka, D., Miller, F., Castro, V., Nazarian, S., Ashikaga, H., Dori, Y., Berger, R.D., Calkins, H., Lardo, A.C., Halperin, H.R.: The critical isthmus sites of ischemic ventricular tachycardia are in zones of tissue heterogeneity, visualized by magnetic resonance imaging. Heart Rhythm **8**(12), 1942 – 1949 (2011). https://doi.org/https://doi.org/10.1016/j.hrthm.2011.07.027, http://www.sciencedirect.com/science/article/pii/S1547527111008472

[55] Estner, H.L., Zviman, M.M., Herzka, D., Miller, F., Castro, V., Nazarian, S., Ashikaga, H., Dori, Y., Berger, R.D., Calkins, H., Lardo, A.C., Halperin, H.R.: The critical isthmus sites of ischemic ventricular tachycardia are in zones of tissue heterogeneity, visualized by magnetic resonance imaging. Heart Rhythm **8**(12), 1942–1949

(2011). https://doi.org/https://doi.org/10.1016/j.hrthm.2011.07.027, https://www.sciencedirect.com/science/article/pii/S1547527111008472

[56] Fatma Usta, Wail Gueaieb, J.A.W.C.M.E.U.: Comparison of myocardial scar geometries from 2d and 3d lge-mri (2018). https://doi.org/10.1117/12.2293961, https://doi.org/10.1117/12.2293961

[57] Feng, Y., Guo, Z., Dong, Z., Zhou, X.Y., Kwok, K.W., Ernst, S., Lee, S.L.: An efficient cardiac mapping strategy for radiofrequency catheter ablation with active learning. International Journal of Computer Assisted Radiology and Surgery **12**(7), 1199–1207 (Jul 2017). https://doi.org/10.1007/s11548-017-1587-4, https://doi.org/10.1007/s11548-017-1587-4

[58] Fonseca, C.G., Backhaus, M., Bluemke, D.A., Britten, R.D., Chung, J.D., Cowan, B.R., Dinov, I.D., Finn, J.P., Hunter, P.J., Kadish, A.H., Lee, D.C., Lima, J.A.C., Medrano Gracia, P., Shivkumar, K., Suinesiaputra, A., Tao, W., Young, A.A.: The Cardiac Atlas Project—an imaging database for computational modeling and statistical atlases of the heart. Bioinformatics **27**(16), 2288–2295 (2011). https://doi.org/10.1093/bioinformatics/btr360, http://dx.doi.org/10.1093/bioinformatics/btr360

[59] Frangi, A.F., Rueckert, D., Schnabel, J.A., Niessen, W.J.: Automatic construction of multiple-object three-dimensional statistical shape models: application to cardiac modeling. IEEE Transactions on Medical Imaging **21**(9), 1151–1166 (sep 2002). https://doi.org/10.1109/TMI.2002.804426

[60] Gao, H., Allan, A., McComb, C., Luo, X., Berry, C.: Left ventricular strain and its pattern estimated from cine CMR and validation with DENSE. Physics in Medicine and Biology

59(13), 3637–3656 (jun 2014). https://doi.org/10.1088/0031-9155/59/13/3637, https://doi.org/10.1088%2F0031-9155%2F59%2F13%2F3637

[61] Gao, Y., Gholami, B., MacLeod, R.S., Blauer, J., Haddad, W.M., Tannenbaum, A.R.: Segmentation of the endocardial wall of the left atrium using local region-based active contours and statistical shape learning. In: Dawant, B.M., Haynor, D.R. (eds.) Medical Imaging 2010: Image Processing. vol. 7623, pp. 1542 – 1549. International Society for Optics and Photonics, SPIE (2010). https://doi.org/10.1117/12.844321, https://doi.org/10.1117/12.844321

[62] Garg, P., Assadi, H., Jones, R., Chan, W.B., Metherall, P., Thomas, R., van der Geest, R., Swift, A.J., Al-Mohammad, A.: Left ventricular fibrosis and hypertrophy are associated with mortality in heart failure with preserved ejection fraction. Scientific Reports **11**(1), 617 (2021). https://doi.org/10.1038/s41598-020-79729-6, https://doi.org/10.1038/s41598-020-79729-6

[63] Geman, S., Geman, D.: Stochastic relaxation, gibbs distributions, and the bayesian restoration of images. IEEE Transactions on Pattern Analysis and Machine Intelligence **PAMI-6**(6), 721–741 (Nov 1984). https://doi.org/10.1109/TPAMI.1984.4767596

[64] Goetti, R., Kozerke, S., Donati, O.F., Sürder, D., Stolzmann, P., Kaufmann, P.A., Lüscher, T.F., Corti, R., Manka, R.: Acute, subacute, and chronic myocardial infarction: Quantitative comparison of 2d and 3d late gadolinium enhancement mr imaging. Radiology **259**(3), 704–711 (2011). https://doi.org/10.1148/radiol.11102216, https://doi.org/10.1148/radiol.11102216, pMID: 21467254

[65] Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., Bengio, Y.: Generative adversarial nets. In: Ghahramani, Z., Welling, M., Cortes, C.,

Lawrence, N.D., Weinberger, K.Q. (eds.) Advances in Neural Information Processing Systems 27, pp. 2672–2680. Curran Associates, Inc. (2014), http://papers.nips.cc/paper/5423-generative-adversarial-nets.pdf

[66] Grayson, M.A.: The heat equation shrinks embedded plane curves to round points. J. Differential Geom. **26**(2), 285–314 (1987). https://doi.org/10.4310/jdg/1214441371, https://doi.org/10.4310/jdg/1214441371

[67] Greig, D.M., Porteous, B.T., Seheult, A.H.: Exact maximum a posteriori estimation for binary images. Journal of the Royal Statistical Society: Series B (Methodological) **51**(2), 271–279 (jan 1989)

[68] Gucuk Ipek, E., Nazarian, S.: Cardiac magnetic resonance for prediction of arrhythmogenic areas. Trends in Cardiovascular Medicine **25**(7), 635–642 (2015). https://doi.org/https://doi.org/10.1016/j.tcm.2015.02.012, https://www.sciencedirect.com/science/article/pii/S1050173815000638

[69] Hamelmann, P., Vullings, R., Kolen, A.F., Bergmans, J.W.M., van Laar, J.O.E.H., Tortoli, P., Mischi, M.: Doppler ultrasound technology for fetal heart rate monitoring: A review. IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control **67**(2), 226–238 (2020). https://doi.org/10.1109/TUFFC.2019.2943626

[70] Harris, P., Lysitsas, D.: Ventricular arrhythmias and sudden cardiac death. BJA Education **16**(7), 221–229 (2016). https://doi.org/10.1093/bjaed/mkv056, http://dx.doi.org/10.1093/bjaed/mkv056

[71] Heimann, T., Meinzer, H.P.: Statistical shape models for 3d medical image segmentation: A review. Medical Image Analysis **13**(4), 543 – 563 (2009).

https://doi.org/https://doi.org/10.1016/j.media.2009.05.004, http://www.sciencedirect.com/science/article/pii/S1361841509000425

[72] Hennemuth, A., Seeger, A., Friman, O., Miller, S., Klumpp, B., Oeltze, S., Peitgen, H.O.: A comprehensive approach to the analysis of contrast enhanced cardiac mr images. IEEE Transactions on Medical Imaging **27**(11), 1592–1610 (Nov 2008). https://doi.org/10.1109/TMI.2008.2006512

[73] Hertzmann, A., Jacobs, C.E., Oliver, N., Curless, B., Salesin, D.H.: Image analogies. Association for Computing Machinery, New York, NY, USA (2001). https://doi.org/10.1145/383259.383295, https://doi.org/10.1145/383259.383295

[74] Hoffman, J., Tzeng, E., Park, T., Zhu, J., Isola, P., Saenko, K., Efros, A.A., Darrell, T.: Cycada: Cycle-consistent adversarial domain adaptation. CoRR **abs/1711.03213** (2017), http://arxiv.org/abs/1711.03213

[75] Hoogendoorn, C., Duchateau, N., Sánchez-Quintana, D., Whitmarsh, T., Sukno, F., Craene, M.D., Lekadir, K., Frangi, A.F.: A High-Resolution Atlas and Statistical Model of the Human Heart From Multislice CT. IEEE Transactions on Medical Imaging **32**, 28–44 (2013)

[76] Hoogendoorn, C., Sukno, F.M., Ordás, S., Frangi, A.F.: Bilinear Models for Spatio-Temporal Point Distribution Analysis. International Journal of Computer Vision **85**(3), 237–252 (dec 2009). https://doi.org/10.1007/s11263-009-0212-6, https://doi.org/10.1007/s11263-009-0212-6

[77] Huang, S., Liu, J., Lee, L.C., Venkatesh, S.K., Teo, L.L.S., Au, C., Nowinski, W.L.: An image-based comprehensive approach for automatic segmentation of left ventricle from cardiac short axis cine mr images. Journal of Digital Imaging **24**(4), 598–608

(Aug 2011). https://doi.org/10.1007/s10278-010-9315-4, https://doi.org/10.1007/s10278-010-9315-4

[78] HáKon, G., Samuel, P.: The rician distribution of noisy mri data. Magnetic Resonance in Medicine **34**(6), 910–914 (1995). https://doi.org/10.1002/mrm.1910340618, https://onlinelibrary.wiley.com/doi/abs/10.1002/mrm.1910340618

[79] Isola, P., Zhu, J., Zhou, T., Efros, A.A.: Image-to-image translation with conditional adversarial networks. CoRR **abs/1611.07004** (2016), http://arxiv.org/abs/1611.07004

[80] Jerosch-Herold, M., Kwong, R.Y.: Magnetic resonance imaging in the assessment of ventricular remodeling and viability. Current Heart Failure Reports **5**(1), 5–10 (Mar 2008). https://doi.org/10.1007/s11897-008-0002-4, https://doi.org/10.1007/s11897-008-0002-4

[81] Jiang, F., Jiang, Y., Zhi, H., Dong, Y., Li, H., Ma, S., Wang, Y., Dong, Q., Shen, H., Wang, Y.: Artificial intelligence in healthcare: past, present and future. Stroke and Vascular Neurology **2**(4), 230–243 (2017). https://doi.org/10.1136/svn-2017-000101, https://svn.bmj.com/content/2/4/230

[82] Johnson, J., Alahi, A., Li, F.: Perceptual losses for real-time style transfer and super-resolution. CoRR **abs/1603.08155** (2016), http://arxiv.org/abs/1603.08155

[83] Johnson, K.W., Torres Soto, J., Glicksberg, B.S., Shameer, K., Miotto, R., Ali, M., Ashley, E., Dudley, J.T.: Artificial intelligence in cardiology. Journal of the American College of Cardiology **71**(23), 2668–2679 (2018). https://doi.org/10.1016/j.jacc.2018.03.521, http://www.onlinejacc.org/content/71/23/2668

[84] Kamnitsas, K., Baumgartner, C.F., Ledig, C., Newcombe, V.F.J., Simpson, J.P., Kane, A.D., Menon, D.K., Nori, A.V., Criminisi, A., Rueckert, D., Glocker, B.: Unsupervised domain adaptation in brain lesion segmentation with adversarial networks. In: IPMI (2017)

[85] Kang, D., Woo, J., Kuo, C.C.J., Slomka, P.J., Dey, D., Germano, G.: Heart chambers and whole heart segmentation techniques: review **21**, 10901–10917 (2012), https://doi.org/10.1117/1.JEI.21.1.010901

[86] Kang, H.C., Kim, B., Lee, J., Shin, J., Shin, Y.G.: Automatic left and right heart segmentation using power watershed and active contour model without edge. Biomedical Engineering Letters **4**(4), 355–361 (Dec 2014). https://doi.org/10.1007/s13534-014-0164-9, https://doi.org/10.1007/s13534-014-0164-9

[87] Karim, R., Arujuna, A., Housden, R.J., Gill, J., Cliffe, H., Matharu, K., Gill, J., Rindaldi, C.A., O'Neill, M., Rueckert, D., Razavi, R., Schaeffter, T., Rhode, K.: A method to standardize quantification of left atrial scar from delayed-enhancement mr images. IEEE Journal of Translational Engineering in Health and Medicine **2**, 1–15 (2014). https://doi.org/10.1109/JTEHM.2014.2312191

[88] Karim, R., Bhagirath, P., Claus, P., Housden, R.J., Chen, Z., Karimaghaloo, Z., Sohn, H.M., Rodríguez, L.L., Vera, S., Albà, X., Hennemuth, A., Peitgen, H.O., Arbel, T., Ballester, M.A.G., Frangi, A.F., Götte, M., Razavi, R., Schaeffter, T., Rhode, K.: Evaluation of state-of-the-art segmentation algorithms for left ventricle infarct from late gadolinium enhancement mr images. Medical Image Analysis **30**, 95 – 107 (2016). https://doi.org/https://doi.org/10.1016/j.media.2016.01.004, http://www.sciencedirect.com/science/article/pii/S1361841516000050

[89] Karim, R., Bhagirath, P., Claus, P., Housden, R.J., Chen, Z., Karimaghaloo, Z., Sohn, H.M., Rodríguez, L.L., Vera, S., Albà, X., Hennemuth, A., Peitgen, H.O., Arbel, T., Ballester, M.A.G., Frangi, A.F., Götte, M., Razavi, R., Schaeffter, T., Rhode, K.: Evaluation of state-of-the-art segmentation algorithms for left ventricle infarct from late gadolinium enhancement mr images. Medical Image Analysis **30**, 95 – 107 (2016). https://doi.org/https://doi.org/10.1016/j.media.2016.01.004, http://www.sciencedirect.com/science/article/pii/S1361841516000050

[90] Karim, R., Housden, R.J., Balasubramaniam, M., Chen, Z., Perry, D., Uddin, A., Al-Beyatti, Y., Palkhi, E., Acheampong, P., Obom, S., Hennemuth, A., Lu, Y., Bai, W., Shi, W., Gao, Y., Peitgen, H.O., Radau, P., Razavi, R., Tannenbaum, A., Rueckert, D., Cates, J., Schaeffter, T., Peters, D., MacLeod, R., Rhode, K.: Evaluation of current algorithms for segmentation of scar tissue from late gadolinium enhancement cardiovascular magnetic resonance of the left atrium: an open-access grand challenge. Journal of Cardiovascular Magnetic Resonance **15**(1), 105 (Dec 2013). https://doi.org/10.1186/1532-429X-15-105, https://doi.org/10.1186/1532-429X-15-105

[91] Karimaghaloo, Z., Shah, M., Francis, S.J., Arnold, D.L., Collins, D.L., Arbel, T.: Automatic detection of gadolinium-enhancing multiple sclerosis lesions in brain mri using conditional random fields. IEEE Transactions on Medical Imaging **31**(6), 1181–1194 (June 2012). https://doi.org/10.1109/TMI.2012.2186639

[92] Kass, M., Witkin, A., Terzopoulos, D.: Snakes: Active contour models. International Journal of Computer Vision **1**(4), 321–331 (jan 1988). https://doi.org/10.1007/BF00133570, https://doi.org/10.1007/BF00133570

[93] Kellman, P., Arai, A.E.: Cardiac imaging techniques for physicians: Late en-

hancement. Journal of Magnetic Resonance Imaging **36**(3), 529–542 (2012). https://doi.org/10.1002/jmri.23605, https://onlinelibrary.wiley.com/doi/abs/10.1002/jmri.23605

[94] Kido, T., Kido, T., Nakamura, M., Kawaguchi, N., Nishiyama, Y., Ogimoto, A., Miyagawa, M., Mochizuki, T.: Three-dimensional phase-sensitive inversion recovery sequencing in the evaluation of left ventricular myocardial scars in ischemic and non-ischemic cardiomyopathy: Comparison to three-dimensional inversion recovery sequencing. European Journal of Radiology **83**(12), 2159 – 2166 (2014). https://doi.org/https://doi.org/10.1016/j.ejrad.2014.09.014, http://www.sciencedirect.com/science/article/pii/S0720048X14004574

[95] Kim, J., Kim, M., Kang, H., Lee, K.: U-GAT-IT: unsupervised generative attentional networks with adaptive layer-instance normalization for image-to-image translation. CoRR **abs/1907.10830** (2019), http://arxiv.org/abs/1907.10830

[96] Kingma, D.P., Ba, J.: Adam: A method for stochastic optimization. CoRR **abs/1412.6980** (2015)

[97] Kolipaka, A., Chatzimavroudis, G.P., White, R.D., O'Donnell, T.P., Setser, R.M.: Segmentation of non-viable myocardium in delayed enhancement magnetic resonance images. The International Journal of Cardiovascular Imaging **21**(2), 303–311 (Apr 2005). https://doi.org/10.1007/s10554-004-5806-z, https://doi.org/10.1007/s10554-004-5806-z

[98] Koplan, B.A., Stevenson, W.G.: Ventricular tachycardia and sudden cardiac death. Mayo Clinic Proceedings **84**(3), 289 – 297 (2009).

https://doi.org/https://doi.org/10.4065/84.3.289, http://www.sciencedirect.com/science/article/pii/S002561961161149X

[99] Krähenbühl, P., Koltun, V.: Efficient inference in fully connected crfs with gaussian edge potentials. In: Shawe-Taylor, J., Zemel, R.S., Bartlett, P.L., Pereira, F., Weinberger, K.Q. (eds.) Advances in Neural Information Processing Systems 24, pp. 109–117. Curran Associates, Inc. (2011), http://papers.nips.cc/paper/4296-efficient-inference-in-fully-connected-crfs-with-gaussian-edge-potentials.pdf

[100] Krizhevsky, A.: Learning multiple layers of features from tiny images. Master's thesis, University of Tront (2009), https://ci.nii.ac.jp/naid/20001706980/en/

[101] Kurzendorfer, T., Breininger, K., Steidl, S., Brost, A., Forman, C., Maier, A.: Myocardial scar segmentation in lge-mri using fractal analysis and random forest classification. In: 2018 24th International Conference on Pattern Recognition (ICPR). pp. 3168–3173 (2018). https://doi.org/10.1109/ICPR.2018.8545636

[102] Kurzendorfer, T., Forman, C., Schmidt, M., Tillmanns, C., Maier, A., Brost, A.: Fully automatic segmentation of left ventricular anatomy in 3-d lge-mri. Computerized Medical Imaging and Graphics **59**, 13 – 27 (2017). https://doi.org/https://doi.org/10.1016/j.compmedimag.2017.05.001, http://www.sciencedirect.com/science/article/pii/S0895611117300447

[103] Lara, L., Vera, S., Perez, F., Lanconelli, N., Morisi, R., Donini, B., Turco, D., Corsi, C., Lamberti, C., Gavidia, G., Bordone, M., Soudah, E., Curzen, N., Rosengarten, J., Morgan, J., Herrero, J., González Ballester, M.A.: Supervised Learning Modelization and Segmentation of Cardiac Scar in Delayed Enhanced MRI. In: Camara, O., Mansi, T., Pop, M., Rhode, K.,

Sermesant, M., Young, A. (eds.) Statistical Atlases and Computational Models of the Heart. Imaging and Modelling Challenges. pp. 53–61. Springer Berlin Heidelberg, Berlin, Heidelberg (2013)

[104] Ledig, C., Theis, L., Huszar, F., Caballero, J., Cunningham, A., Acosta, A., Aitken, A., Tejani, A., Totz, J., Wang, Z., Shi, W.: Photo-realistic single image super-resolution using a generative adversarial network (2016), https://arxiv.org/abs/1609.04802v5

[105] Lee, H., Codella, N.C.F., Cham, M.D., Weinsaft, J.W., Wang, Y.: Automatic left ventricle segmentation using iterative thresholding and an active contour model with adaptation on short-axis cardiac mri. IEEE Transactions on Biomedical Engineering **57**(4), 905–913 (April 2010). https://doi.org/10.1109/TBME.2009.2014545

[106] Lehmann, G.: Label object representation and manipulation with itk. http://hdl.handle.net/1926/584 (08 2007)

[107] Leventon, M.E., Grimson, W.E.L., Faugeras, O.: Statistical shape influence in geodesic active contours. In: 5th IEEE EMBS International Summer School on Biomedical Imaging, 2002. pp. 8 pp.– (June 2002). https://doi.org/10.1109/SSBI.2002.1233989

[108] Li, C., Wand, M.: Precomputed real-time texture synthesis with markovian generative adversarial networks. CoRR **abs/1604.04382** (2016), http://arxiv.org/abs/1604.04382

[109] Li, X., Aldridge, B., Fisher, R., Rees, J.: Estimating the ground truth from multiple individual segmentations incorporating prior pattern analysis with application to skin lesion segmentation. In: 2011 IEEE International Symposium on Biomedical Imaging: From Nano to Macro. pp. 1438–1441 (March 2011). https://doi.org/10.1109/ISBI.2011.5872670

[110] Liu, M., Breuel, T., Kautz, J.: Unsupervised image-to-image translation networks. CoRR **abs/1703.00848** (2017), http://arxiv.org/abs/1703.00848

[111] Liu, T., Tian, Y., Zhao, S., Huang, X., Xu, Y., Jiang, G., Wang, Q.: Pseudo-3d network for multi-sequence cardiac mr segmentation. In: Pop, M., Sermesant, M., Camara, O., Zhuang, X., Li, S., Young, A., Mansi, T., Suinesiaputra, A. (eds.) Statistical Atlases and Computational Models of the Heart. Multi-Sequence CMR Segmentation, CRT-EPiggy and LV Full Quantification Challenges. pp. 237–245. Springer International Publishing, Cham (2020)

[112] Liu, Y., Wang, W., Wang, K., Ye, C., Luo, Gongning", e.M., Sermesant, M., Camara, O., Zhuang, X., Li, S., Young, A., Mansi, T., Suinesiaputra, A.: An automatic cardiac segmentation framework based on multi-sequence mr image pp. 220–227 (2020)

[113] Lorenz, C., von Berg, J.: A comprehensive shape model of the heart. Medical Image Analysis **10**(4), 657–670 (2006). https://doi.org/https://doi.org/10.1016/j.media.2006.03.004, http://www.sciencedirect.com/science/article/pii/S1361841506000168

[114] Lorenzo-Valdés, M., Sanchez-Ortiz, G.I., Elkington, A.G., Mohiaddin, R.H., Rueckert, D.: Segmentation of 4d cardiac mr images using a probabilistic atlas and the em algorithm. Medical Image Analysis **8**(3), 255 – 265 (2004). https://doi.org/https://doi.org/10.1016/j.media.2004.06.005, http://www.sciencedirect.com/science/article/pii/S1361841504000271, medical Image Computing and Computer-Assisted Intervention - MICCAI 2003

[115] Lu, Y., Connelly, K.A., Dick, A.J., Wright, G.A., Radau, P.E.: Watershed segmentation of basal left ventricle for quantitation of cine cardiac mri function. Journal of Cardiovascular Magnetic Resonance **13**(1), P4 (Feb 2011). https://doi.org/10.1186/1532-429X-13-S1-P4, https://doi.org/10.1186/1532-429X-13-S1-P4

[116] Lu, Y., Yang, Y., Connelly, K.A., Wright, G.A., Radau, P.E.: Automated quantification of myocardial infarction using graph cuts on contrast delayed enhanced magnetic resonance

images. Quantitative imaging in medicine and surgery **2 2**, 81–6 (2012)

[117] de Luna, A.B., Coumel, P., Leclercq, J.F.: Ambulatory sudden cardiac death: Mechanisms of production of fatal arrhythmia on the basis of data from 157 cases. American Heart Journal **117**(1), 151 – 159 (1989). https://doi.org/https://doi.org/10.1016/0002-8703(89)90670-4, http://www.sciencedirect.com/science/article/pii/0002870389906704

[118] Lötjönen, J., Kivistö, S., Koikkalainen, J., Smutek, D., Lauerma, K.: Statistical shape model of atria, ventricles and epicardium from short- and long-axis mr images. Medical Image Analysis **8**(3), 371 – 386 (2004). https://doi.org/https://doi.org/10.1016/j.media.2004.06.013, http://www.sciencedirect.com/science/article/pii/S1361841504000374, medical Image Computing and Computer-Assisted Intervention - MICCAI 2003

[119] Mamalakis, M., Garg, P., Nelson, T., Lee, J., Wild, J.M., Clayton, R.H.: Masocratis: An automatic pipeline for robust segmentation of the left ventricle and scar. Computerized Medical Imaging and Graphics p. 101982 (2021). https://doi.org/https://doi.org/10.1016/j.compmedimag.2021.101982, https://www.sciencedirect.com/science/article/pii/S0895611121001312

[120] Manniesing, R., Viergever, M.A., Niessen, W.J.: Vessel enhancing diffusion: A scale space representation of vessel structures. Medical Image Analysis **10**(6), 815–825 (2006). https://doi.org/https://doi.org/10.1016/j.media.2006.06.003, https://www.sciencedirect.com/science/article/pii/S1361841506000442

[121] Martin, R., Maury, P., Bisceglia, C., Wong, T., Estner, H., Meyer, C., Dallet, C., Martin, C.A., Shi, R., Takigawa, M., Rollin, A., Frontera, A., Thompson, N., Kitamura, T., Vlachos, K., Wolf, M., Cheniti, G., Duchâteau, J., Massoulié, G., Pambrun, T., Denis, A., Derval, N., Hocini, M., Bella, P.D., Haïssaguerre, M., Jaïs, P., Dubois, R., Sacher, F.:

Characteristics of scar-related ventricular tachycardia circuits using ultra-high-density mapping. Circulation: Arrhythmia and Electrophysiology **11**(10), e006569 (2018). https://doi.org/10.1161/CIRCEP.118.006569

[122] McAlindon, E., Pufulete, M., Lawton, C., Angelini, G.D., Bucciarelli-Ducci, C.: Quantification of infarct size and myocardium at risk: evaluation of different techniques and its implications. European Heart Journal - Cardiovascular Imaging **16**(7), 738–746 (03 2015). https://doi.org/10.1093/ehjci/jev001, https://doi.org/10.1093/ehjci/jev001

[123] McInerney, T., Terzopoulos, D.: Deformable models in medical image analysis: a survey. Medical Image Analysis **1**(2), 91–108 (1996). https://doi.org/https://doi.org/10.1016/S1361-8415(96)80007-7, http://www.sciencedirect.com/science/article/pii/S1361841596800077

[124] Menze, B.H., Jakab, A., Bauer, S., Kalpathy-Cramer, J., Farahani, K., Kirby, J., Burren, Y., Porz, N., Slotboom, J., Wiest, R., Lanczi, L., Gerstner, E., Weber, M., Arbel, T., Avants, B.B., Ayache, N., Buendia, P., Collins, D.L., Cordier, N., Corso, J.J., Criminisi, A., Das, T., Delingette, H., Demiralp, C., Durst, C.R., Dojat, M., Doyle, S., Festa, J., Forbes, F., Geremia, E., Glocker, B., Golland, P., Guo, X., Hamamci, A., Iftekharuddin, K.M., Jena, R., John, N.M., Konukoglu, E., Lashkari, D., Mariz, J.A., Meier, R., Pereira, S., Precup, D., Price, S.J., Raviv, T.R., Reza, S.M.S., Ryan, M., Sarikaya, D., Schwartz, L., Shin, H., Shotton, J., Silva, C.A., Sousa, N., Subbanna, N.K., Szekely, G., Taylor, T.J., Thomas, O.M., Tustison, N.J., Unal, G., Vasseur, F., Wintermark, M., Ye, D.H., Zhao, L., Zhao, B., Zikic, D., Prastawa, M., Reyes, M., Van Leemput, K.: The multimodal brain tumor image segmentation benchmark (brats). IEEE Transactions on Medical Imaging **34**(10), 1993–2024 (2015). https://doi.org/10.1109/TMI.2014.2377694

[125] Mesubi, O., Ego-Osuala, K., Jeudy, J., Purtilo, J., Synowski, S., Abutaleb, A., Niekoop, M., Abdulghani, M., Asoglu, R., See, V., Saliaris, A., Shorofsky, S., Dickfeld, T.: Differences in quantitative assessment of myocardial scar and gray zone by lge-cmr imaging using established gray zone protocols. The International Journal of Cardiovascular Imaging **31**(2), 359–368 (Feb 2015). https://doi.org/10.1007/s10554-014-0555-0, https://doi.org/10.1007/s10554-014-0555-0

[126] Meyer, F., Beucher, S.: Morphological segmentation. Journal of Visual Communication and Image Representation **1**(1), 21 – 46 (1990)

[127] Meyer, F.: Mathematical Morphology in Image Processing:The morphological approach to segmentation: the watershed transformation. Imprint CRC Press (1993)

[128] Mines, G.R.: On circulating excitations in heart muscles and their possible relation to tachycardia and fibrillation. Transactions of the Royal Society of Canada. 4. section **1914**(4), 43 – 52 (2010)

[129] Mitchell, S.C., Bosch, J.G., Lelieveldt, B.P.F., van der Geest, R.J., Reiber, J.H.C., Sonka, M.: 3-d active appearance models: segmentation of cardiac mr and ultrasound images. IEEE Transactions on Medical Imaging **21**(9), 1167–1178 (Sep 2002). https://doi.org/10.1109/TMI.2002.804425

[130] Moccia, S., Banali, R., Martini, C., Moscogiuri, G., Pontone, G., Pepi, M., Caiani, E.G.: Automated scar segmentation from cmr-lge images using a deep learning approach. In: 2018 Computing in Cardiology Conference (CinC). vol. 45, pp. 1–4 (Sep 2018). https://doi.org/10.22489/CinC.2018.278

[131] Moccia, S., Banali, R., Martini, C., Muscogiuri, G., Pontone, G., Pepi, M., Caiani, E.G.: Development and testing of a deep learning-based strategy for scar segmentation on cmr-

lge images. Magnetic Resonance Materials in Physics, Biology and Medicine **32**(2), 187–195 (2019). https://doi.org/10.1007/s10334-018-0718-4, https://doi.org/10.1007/s10334-018-0718-4

[132] Monserrat, L., Elliott, P.M., Gimeno, J.R., Sharma, S., Penas-Lado, M., McKenna, W.J.: Non-sustained ventricular tachycardia in hypertrophic cardiomyopathy: an independent marker of sudden death risk in young patients. Journal of the American College of Cardiology **42**(5), 873 – 879 (2003). https://doi.org/https://doi.org/10.1016/S0735-1097(03)00827-1, http://www.sciencedirect.com/science/article/pii/S0735109703008271

[133] Neizel, M., Lossnitzer, D., Korosoglou, G., Schäufele, T., Peykarjou, H., Steen, H., Ocklenburg, C., Giannitsis, E., Katus, H.A., Osman, N.F.: Strain-encoded mri for evaluation of left ventricular function and transmurality in acute myocardial infarction. Circulation: Cardiovascular Imaging **2**(2), 116–122 (2009). https://doi.org/10.1161/CIRCIMAGING.108.789032

[134] Nerbonne, J.M., Kass, R.S.: Molecular Physiology of Cardiac Repolarization. Physiological Reviews **85**(4), 1205–1253 (2005). https://doi.org/10.1152/physrev.00002.2005, https://doi.org/10.1152/physrev.00002.2005

[135] Ng, H.P., Ong, S.H., Foong, K.W.C., Goh, P.S., Nowinski, W.L.: Medical image segmentation using k-means clustering and improved watershed algorithm. In: 2006 IEEE Southwest Symposium on Image Analysis and Interpretation. pp. 61–65 (March 2006). https://doi.org/10.1109/SSIAI.2006.1633722

[136] Ngo, T.A., Lu, Z., Carneiro, G.: Combining deep learning and level set for the automated segmentation of the left ventricle of the heart from car-

diac cine magnetic resonance. Medical Image Analysis **35**, 159 – 171 (2017). https://doi.org/https://doi.org/10.1016/j.media.2016.05.009, http://www.sciencedirect.com/science/article/pii/S136184151630038X

[137] Oakes, R.S., Badger, T.J., Kholmovski, E.G., Akoum, N., Burgon, N.S., Fish, E.N., Blauer, J.J., Rao, S.N., DiBella, E.V., Segerson, N.M., Daccarett, M., Windfelder, J., McGann, C.J., Parker, D., MacLeod, R.S., Marrouche, N.F.: Detection and quantification of left atrial structural remodeling with delayed-enhancement magnetic resonance imaging in patients with atrial fibrillation. Circulation **119**(13), 1758–1767 (2009). https://doi.org/10.1161/CIRCULATIONAHA.108.811877, http://circ.ahajournals.org/content/119/13/1758

[138] Oktay, O., Bai, W., Guerrero, R., Rajchl, M., de Marvao, A., O'Regan, D.P., Cook, S.A., Heinrich, M.P., Glocker, B., Rueckert, D.: Stratified decision forests for accurate anatomical landmark localization in cardiac images. IEEE Transactions on Medical Imaging **36**(1), 332–342 (Jan 2017). https://doi.org/10.1109/TMI.2016.2597270

[139] Olivier Ecabert, Jochen Peters, J.W.: Modeling shape variability for full heart segmentation in cardiac computed-tomography images (2006). https://doi.org/10.1117/12.652105, https://doi.org/10.1117/12.652105

[140] Olshausen, K.V., Witt, T., Pop, T., Treese, N., Bethge, K.P., Meyer, J.: Sudden cardiac death while wearing a holter monitor. The American Journal of Cardiology **67**(5), 381 – 386 (1991). https://doi.org/https://doi.org/10.1016/0002-9149(91)90046-N, http://www.sciencedirect.com/science/article/pii/000291499190046N

[141] Ordovas, K.G., Higgins, C.B.: Delayed contrast enhancement on mr images of myocardium: Past, present, future. Radiology **261**(2), 358–374 (2011).

https://doi.org/10.1148/radiol.11091882, https://doi.org/10.1148/radiol.11091882, pMID: 22012903

[142] Osher, S., Sethian, J.A.: Fronts propagating with curvature-dependent speed: Algorithms based on hamilton-jacobi formulations. Journal of Computational Physics **79**(1), 12 – 49 (1988). https://doi.org/https://doi.org/10.1016/0021-9991(88)90002-2, http://www.sciencedirect.com/science/article/pii/0021999188900022

[143] Osher, S., Sethian, J.A.: Fronts propagating with curvature-dependent speed: Algorithms based on hamilton-jacobi formulations. Journal of Computational Physics **79**(1), 12 – 49 (1988). https://doi.org/https://doi.org/10.1016/0021-9991(88)90002-2, http://www.sciencedirect.com/science/article/pii/0021999188900022

[144] Otsu, N.: A Threshold Selection Method from Gray-Level Histograms. IEEE Transactions on Systems, Man, and Cybernetics **9**(1), 62–66 (jan 1979). https://doi.org/10.1109/TSMC.1979.4310076

[145] Paragios, N.: A level set approach for shape-driven segmentation and tracking of the left ventricle. IEEE Transactions on Medical Imaging **22**(6), 773–776 (June 2003). https://doi.org/10.1109/TMI.2003.814785

[146] Park, K., Montillo, A., Metaxas, D., Axel, L.: Volumetric heart modeling and analysis. Commun. ACM **48**(2), 43–48 (Feb 2005). https://doi.org/10.1145/1042091.1042118, http://doi.acm.org/10.1145/1042091.1042118

[147] Pashakhanloo, F., Herzka, D.A., Halperin, H., McVeigh, E.R., Trayanova, N.A.: Role of 3-dimensional architecture of scar and surviving tissue in ventricular tachycardia. Circulation: Arrhythmia and Electrophysiology **11**(6), e006131 (2018). https://doi.org/10.1161/CIRCEP.117.006131

[148] Perez-David, E., Ángel Arenal, Rubio-Guivernau, J.L., del Castillo, R., Atea, L., Arbelo, E., Caballero, E., Celorrio, V., Datino, T., Gonzalez-Torrecilla, E., Atienza, F., Ledesma-Carbayo, M.J., Bermejo, J., Medina, A., Fernández-Avilés, F.: Noninvasive identification of ventricular tachycardia-related conducting channels using contrast-enhanced magnetic resonance imaging in patients with chronic myocardial infarction: Comparison of signal intensity scar mapping and endocardial voltage mapping. Journal of the American College of Cardiology **57**(2), 184–194 (2011). https://doi.org/https://doi.org/10.1016/j.jacc.2010.07.043, https://www.sciencedirect.com/science/article/pii/S0735109710043032

[149] Piehler, K.M., Wong, T.C., Puntil, K.S., Zareba, K.M., Lin, K., Harris, D.M., Deible, C.R., Lacomis, J.M., Czeyda-Pommersheim, F., Cook, S.C., Kellman, P., Schelbert, E.B.: Free-breathing, motion-corrected late gadolinium enhancement is robust and extends risk stratification to vulnerable patients. Circulation: Cardiovascular Imaging **6**(3), 423–432 (2013). https://doi.org/10.1161/CIRCIMAGING.112.000022, https://www.ahajournals.org/doi/abs/10.1161/CIRCIMAGING.112.000022

[150] Piers, S.R., Tao, Q., de Riva Silva, M., Siebelink, H.M., Schalij, M.J., van der Geest, R.J., Zeppenfeld, K.: Cmr–based identification of critical isthmus sites of ischemic and nonischemic&#xao;ventricular tachycardia. JACC: Cardiovascular Imaging **7**(8), 774–784 (2014). https://doi.org/10.1016/j.jcmg.2014.03.013, https://www.jacc.org/doi/abs/10.1016/j.jcmg.2014.03.013

[151] Pluempitiwiriyawej, C., Moura, J.M.F., and: Stacs: new active contour scheme for cardiac mr image segmentation. IEEE Transactions on Medical Imaging **24**(5), 593–603 (May 2005). https://doi.org/10.1109/TMI.2005.843740

[152] Popović, Z.B., Thomas, J.D.: Assessing observer variability: a user's guide. Cardiovascular Diagnosis and Therapy **7**(3) (2017), http://cdt.amegroups.com/article/view/14271

[153] Quammen, C.: A very important contribution for microscopy application. %url-http://hdl.handle.net/10380/3207 (10 2007)

[154] Ranganath, S.: Contour extraction from cardiac mri studies using snakes. IEEE Transactions on Medical Imaging **14**(2), 328–338 (June 1995). https://doi.org/10.1109/42.387714

[155] Reddy, V.Y., Malchano, Z.J., Holmvang, G., Schmidt, E.J., D'Avila, A., Houghtaling, C., Chan, R.C., Ruskin, J.N.: Integration of cardiac magnetic resonance imaging with three-dimensional electroanatomic mapping to guide left ventricular catheter manipulation: Feasibility in a porcine model of healed myocardial infarction. Journal of the American College of Cardiology **44**(11), 2202–2213 (2004). https://doi.org/https://doi.org/10.1016/j.jacc.2004.08.063, http://www.sciencedirect.com/science/article/pii/S0735109704017553

[156] Ren, J., Sun, H., Huang, Y., Gao, H.: Knowledge-based multi-sequence mr segmentation via deep learning with a hybrid u-net++ model. In: Pop, M., Sermesant, M., Camara, O., Zhuang, X., Li, S., Young, A., Mansi, T., Suinesiaputra, A. (eds.) Statistical Atlases and Computational Models of the Heart. Multi-Sequence CMR Segmentation, CRT-EPiggy and LV Full Quantification Challenges. pp. 280–289. Springer International Publishing, Cham (2020)

[157] Rezaee, M.R., van der Zwet, P.M.J., Lelieveldt, B.P.E., van der Geest, R.J., Reiber, J.H.C.: A multiresolution image segmentation technique based on pyramidal segmentation and fuzzy clustering. IEEE Transactions on Image Processing **9**(7), 1238–1248 (July 2000). https://doi.org/10.1109/83.847836

[158] Ridgway, J.P.: Cardiovascular magnetic resonance physics for clinicians: part i. Journal of Cardiovascular Magnetic Resonance **12**, 71 (2010). https://doi.org/doi:10.1186/1532-429X-12-71, individual Subject Prediction

[159] van Rikxoort, E.M., Isgum, I., Arzhaeva, Y., Staring, M., Klein, S., Viergever, M.A., Pluim, J.P., van Ginneken, B.: Adaptive local multi-atlas segmentation: Application to the heart and the caudate nucleus. Medical Image Analysis **14**(1), 39 – 49 (2010). https://doi.org/https://doi.org/10.1016/j.media.2009.10.001, http://www.sciencedirect.com/science/article/pii/S1361841509000887

[160] Roes, S.D., Borleffs, C.J.W., van der Geest, R.J., Westenberg, J.J., Marsan, N.A., Kaandorp, T.A., Reiber, J.H., Zeppenfeld, K., Lamb, H.J., de Roos, A., Schalij, M.J., Bax, J.J.: Infarct tissue heterogeneity assessed with contrast-enhanced mri predicts spontaneous ventricular arrhythmia in patients with ischemic cardiomyopathy and implantable cardioverter-defibrillator. Circulation: Cardiovascular Imaging **2**(3), 183–190 (2009). https://doi.org/10.1161/CIRCIMAGING.108.826529, https://www.ahajournals.org/doi/abs/10.1161/CIRCIMAGING.108.826529

[161] Roes, S.D., Borleffs, C.J.W., van der Geest, R.J., Westenberg, J.J., Marsan, N.A., Kaandorp, T.A., Reiber, J.H., Zeppenfeld, K., Lamb, H.J., de Roos, A., Schalij, M.J., Bax, J.J.: Infarct tissue heterogeneity assessed with contrast-enhanced mri predicts spontaneous ventricular arrhythmia in patients with ischemic cardiomyopathy and implantable cardioverter-defibrillator. Circulation: Cardiovascular Imaging **2**(3), 183–190 (2009). https://doi.org/10.1161/CIRCIMAGING.108.826529

[162] Ronneberger, O., P.Fischer, Brox, T.: U-net: Convolutional networks for biomedical image segmentation. In: Medical Image Computing and Computer-Assisted Intervention

(MICCAI). LNCS, vol. 9351, pp. 234–241. Springer (2015), http://lmb.informatik.uni-freiburg.de/Publications/2015/RFB15a, (available on arXiv:1505.04597 [cs.CV])

[163] Roth, H., Zhu, W., Yang, D., Xu, Z., Xu, D.: Cardiac segmentation of lge mri with noisy labels. In: Pop, M., Sermesant, M., Camara, O., Zhuang, X., Li, S., Young, A., Mansi, T., Suinesiaputra, A. (eds.) Statistical Atlases and Computational Models of the Heart. Multi-Sequence CMR Segmentation, CRT-EPiggy and LV Full Quantification Challenges. pp. 228–236. Springer International Publishing, Cham (2020)

[164] Roy, S., Carass, A., Bazin, P.L., Resnick, S., Prince, J.L.: Consistent segmentation using a rician classifier. Medical Image Analysis **16**(2), 524 – 535 (2012). https://doi.org/https://doi.org/10.1016/j.media.2011.12.001, http://www.sciencedirect.com/science/article/pii/S136184151100171X

[165] Rueckert, D., Burger, P.: Shape-based segmentation and tracking in 4d cardiac mr images. In: Troccaz, J., Grimson, E., Mösges, R. (eds.) CVRMed-MRCAS'97. pp. 43–52. Springer Berlin Heidelberg, Berlin, Heidelberg (1997)

[166] Sacher, F., Martin, R.: Risk factors for early mortality after vt ablation. Journal of the American College of Cardiology **69**(17), 2116–2118 (2017). https://doi.org/10.1016/j.jacc.2017.03.537, http://www.onlinejacc.org/content/69/17/2116

[167] Sasaki, T., Miller, C.F., Hansford, R., Yang, J., Caffo, B.S., Zviman, M.M., Henrikson, C.A., Marine, J.E., Spragg, D., Cheng, A., Tandri, H., Sinha, S., Kolandaivelu, A., Zimmerman, S.L., Bluemke, D.A., Tomaselli, G.F., Berger, R.D., Calkins, H., Halperin, H.R., Nazarian, S.: Myocardial structural associations with local electrograms. Circulation: Arrhythmia and

Electrophysiology **5**(6), 1081–1090 (2012). https://doi.org/10.1161/CIRCEP.112.970699, https://www.ahajournals.org/doi/abs/10.1161/CIRCEP.112.970699

[168] Scatteia, A., Baritussio, A., Bucciarelli-Ducci, C.: Strain imaging using cardiac magnetic resonance. Heart Failure Reviews **22**(4), 465–476 (Jul 2017). https://doi.org/10.1007/s10741-017-9621-8, https://doi.org/10.1007/s10741-017-9621-8

[169] van de Schoot, R., Depaoli, S., King, R., Kramer, B., Märtens, K., Tadesse, M.G., Vannucci, M., Gelman, A., Veen, D., Willemsen, J., Yau, C.: Bayesian statistics and modelling. Nature Reviews Methods Primers **1**(1), 1 (2021). https://doi.org/10.1038/s43586-020-00001-2, https://doi.org/10.1038/s43586-020-00001-2

[170] Shameer, K., Johnson, K.W., Glicksberg, B.S., Dudley, J.T., Sengupta, P.P.: Machine learning in cardiovascular medicine: are we there yet? Heart **104**(14), 1156–1164 (2018). https://doi.org/10.1136/heartjnl-2017-311198, https://heart.bmj.com/content/104/14/1156

[171] Singh, Y., Deepa, D., Wu, S.Y., Friebe, M., Tavares, J.M.R.S., Hu, W.: Cardiac electrophysiology studies based on image and machine learning. Journal of Biomedical Engineering and Technology **6**(1), 1–6 (2018). https://doi.org/10.12691/jbet-6-1-1, http://pubs.sciepub.com/jbet/6/1/1

[172] Sinop, A.K., Grady, L.: A seeded image segmentation framework unifying graph cuts and random walker which yields a new algorithm. In: 2007 IEEE 11th International Conference on Computer Vision. pp. 1–8 (Oct 2007). https://doi.org/10.1109/ICCV.2007.4408927

[173] Slomka, P.J., Dey, D., Sitek, A., Motwani, M., Berman, D.S., Germano, G.: Cardiac imaging: working towards fully-automated machine analysis & interpretation. Expert Review of

Medical Devices **14**(3), 197–212 (2017). https://doi.org/10.1080/17434440.2017.1300057, https://doi.org/10.1080/17434440.2017.1300057, pMID: 28277804

[174] Solis-Lemus, J.A., Razeghi, O., Roney, C., Sim, I., Mukherjee, R., Williams, S., O'Neill, M., Niederer, S.: Software framework to quantify pulmonary vein isolation atrium scar tissue. In: 2020 Computing in Cardiology. pp. 1–4 (2020). https://doi.org/10.22489/CinC.2020.052

[175] Tao, Q., Milles, J., Zeppenfeld, K., Lamb, H.J., Bax, J.J., Reiber, J.H., van der Geest, R.J.: Automated segmentation of myocardial scar in late enhancement mri using combined intensity and spatial information. Magnetic Resonance in Medicine **64**(2), 586–594 (2010). https://doi.org/10.1002/mrm.22422, https://onlinelibrary.wiley.com/doi/abs/10.1002/mrm.22422

[176] Tao, Q., Yan, W., Wang, Y., Paiman, E.H.M., Shamonin, D.P., Garg, P., Plein, S., Huang, L., Xia, L., Sramko, M., Tintera, J., de Roos, A., Lamb, H.J., van der Geest, R.J.: Deep learning–based method for fully automatic quantification of left ventricle function from cine mr images: A multivendor, multicenter study. Radiology **290**(1), 81–88 (2019)

[177] Tao, X., Wei, H., Xue, W., Ni, D.: Segmentation of multimodal myocardial images using shape-transfer gan. In: Pop, M., Sermesant, M., Camara, O., Zhuang, X., Li, S., Young, A., Mansi, T., Suinesiaputra, A. (eds.) Statistical Atlases and Computational Models of the Heart. Multi-Sequence CMR Segmentation, CRT-EPiggy and LV Full Quantification Challenges. pp. 271–279. Springer International Publishing, Cham (2020)

[178] Tenenbaum, J.B., Freeman, W.T.: Separating Style and Content with Bilinear Models. Neural Computation **12**(6), 1247–1283 (jun 2000). https://doi.org/10.1162/089976600300015349

[179] Terzopoulos, D., Szeliski, R.: Active Vision. chap. Tracking w, pp. 3–20. MIT Press, Cambridge, MA, USA (1993), http://dl.acm.org/citation.cfm?id=163213.163214

[180] Tom Nelson, Garg P, R.C.J.L.: The role of cardiac mri in the management of ventricular arrhythmias in ischaemic and non-ischaemic dilated cardiomyopathy. Arrhythm Electrophysiol Rev **8**(3), 191–201 (2019). https://doi.org/10.15420/aer.2019.5.1

[181] Top, A., Hamarneh, G., Abugharbieh, R.: Active learning for interactive 3d image segmentation. In: Fichtinger, G., Martel, A., Peters, T. (eds.) Medical Image Computing and Computer-Assisted Intervention – MICCAI 2011. pp. 603–610. Springer Berlin Heidelberg, Berlin, Heidelberg (2011)

[182] Tusscher, K.T., Panfilov, A.: Modelling of the ventricular conduction system. Progress in Biophysics and Molecular Biology **96**(1), 152 – 170 (2008). https://doi.org/https://doi.org/10.1016/j.pbiomolbio.2007.07.026, http://www.sciencedirect.com/science/article/pii/S007961070700065X, cardiovascular Physiome

[183] Tzikas, D.G., Likas, A.C., Galatsanos, N.P.: The variational approximation for bayesian inference. IEEE Signal Processing Magazine **25**(6), 131–146 (November 2008). https://doi.org/10.1109/MSP.2008.929620

[184] Ukwatta, E., Arevalo, H., Li, K., Yuan, J., Qiu, W., Malamas, P., Wu, K.C., Trayanova, N.A., Vadakkumpadan, F.: Myocardial infarct segmentation from magnetic resonance images for personalized modeling of cardiac electrophysiology. IEEE Transactions on Medical Imaging **35**(6), 1408–1419 (June 2016). https://doi.org/10.1109/TMI.2015.2512711

[185] van Assen, H.C., Danilouchkine, M.G., Dirksen, M.S., Reiber, J.H.C., Lelieveldt, B.P.F.: A 3-d active shape model driven by fuzzy inference: Application to cardiac ct and mr. IEEE

Transactions on Information Technology in Biomedicine **12**(5), 595–605 (Sep 2008). https://doi.org/10.1109/TITB.2008.926477

[186] Varoquaux, G., Raamana, P.R., Engemann, D.A., Hoyos-Idrobo, A., Schwartz, Y., Thirion, B.: Assessing and tuning brain decoders: Cross-validation, caveats, and guidelines. NeuroImage **145**, 166–179 (2017). https://doi.org/https://doi.org/10.1016/j.neuroimage.2016.10.038, https://www.sciencedirect.com/science/article/pii/S105381191630595X, individual Subject Prediction

[187] Vassiliou, V.S., Cameron, D., Prasad, S.K., Gatehouse, P.D.: Magnetic resonance imaging: Physics basics for the cardiologist. JRSM Cardiovascular Disease **7**, 2048004018772237 (2018). https://doi.org/10.1177/2048004018772237, https://doi.org/10.1177/2048004018772237, pMID: 30128147

[188] VERMA, A., MARROUCHE, N.F., SCHWEIKERT, R.A., SALIBA, W., WAZNI, O., CUMMINGS, J., ABDUL-KARIM, A., BHARGAVA, M., BURKHARDT, J.D., KILICASLAN, F., MARTIN, D.O., NATALE, A.: Relationship between successful ablation sites and the scar border zone defined by substrate mapping for ventricular tachycardia post-myocardial infarction. Journal of Cardiovascular Electrophysiology **16**(5), 465–471. https://doi.org/https://doi.org/10.1046/j.1540-8167.2005.40443.x, https://onlinelibrary.wiley.com/doi/abs/10.1046/j.1540-8167.2005.40443.x

[189] Vesal, S., Ravikumar, N., Maier, A.: Automated multi-sequence cardiac mri segmentation using supervised domain adaptation. In: Pop, M., Sermesant, M., Camara, O., Zhuang, X., Li, S., Young, A., Mansi, T., Suinesiaputra, A. (eds.) Statistical Atlases and Computational

Models of the Heart. Multi-Sequence CMR Segmentation, CRT-EPiggy and LV Full Quantification Challenges. pp. 300–308. Springer International Publishing, Cham (2020)

[190] Wang, G., Li, W., Ourselin, S., Vercauteren, T.: Automatic brain tumor segmentation based on cascaded convolutional neural networks with uncertainty estimation. Frontiers in Computational Neuroscience **13**, 56 (2019). https://doi.org/10.3389/fncom.2019.00056, https://www.frontiersin.org/article/10.3389/fncom.2019.00056

[191] Wang, J., Huang, H., Chen, C., Ma, W., Huang, Y., Ding, X.: Multi-sequence cardiac mr segmentation with adversarial domain adaptation network. In: Pop, M., Sermesant, M., Camara, O., Zhuang, X., Li, S., Young, A., Mansi, T., Suinesiaputra, A. (eds.) Statistical Atlases and Computational Models of the Heart. Multi-Sequence CMR Segmentation, CRT-EPiggy and LV Full Quantification Challenges. pp. 254–262. Springer International Publishing, Cham (2020)

[192] Wang, X., Yang, S., Tang, M., Wei, Y., Han, X., He, L., Zhang, J.: Sk-unet: An improved u-net model with selective kernel for the segmentation of multi-sequence cardiac mr. In: Pop, M., Sermesant, M., Camara, O., Zhuang, X., Li, S., Young, A., Mansi, T., Suinesiaputra, A. (eds.) Statistical Atlases and Computational Models of the Heart. Multi-Sequence CMR Segmentation, CRT-EPiggy and LV Full Quantification Challenges. pp. 246–253. Springer International Publishing, Cham (2020)

[193] Welander, P., Karlsson, S., Eklund, A.: Generative adversarial networks for image-to-image translation on multi-contrast mr images - a comparison of cyclegan and unit (2018), https://arxiv.org/abs/1806.07777

[194] Xiong, Z., Xia, Q., Hu, Z., Huang, N., Bian, C., Zheng, Y., Vesal, S., Ravikumar, N., Maier, A., Yang, X., Heng, P.A., Ni, D., Li, C., Tong, Q., Si, W., Puybareau, E., Khoudli, Y., Géraud,

T., Chen, C., Bai, W., Rueckert, D., Xu, L., Zhuang, X., Luo, X., Jia, S., Sermesant, M., Liu, Y., Wang, K., Borra, D., Masci, A., Corsi, C., de Vente, C., Veta, M., Karim, R., Preetha, C.J., Engelhardt, S., Qiao, M., Wang, Y., Tao, Q., Nuñez-Garcia, M., Camara, O., Savioli, N., Lamata, P., Zhao, J.: A global benchmark of algorithms for segmenting the left atrium from late gadolinium-enhanced cardiac magnetic resonance imaging. Medical Image Analysis **67**, 101832 (2021). https://doi.org/https://doi.org/10.1016/j.media.2020.101832, https://www.sciencedirect.com/science/article/pii/S1361841520301961

[195] Xu, H., Xu, Z., Gu, W., Zhang, Q.: A two-stage fully automatic segmentation scheme using both 2d and 3d u-net for multi-sequence cardiac mr. In: Pop, M., Sermesant, M., Camara, O., Zhuang, X., Li, S., Young, A., Mansi, T., Suinesiaputra, A. (eds.) Statistical Atlases and Computational Models of the Heart. Multi-Sequence CMR Segmentation, CRT-EPiggy and LV Full Quantification Challenges. pp. 309–316. Springer International Publishing, Cham (2020)

[196] Yan, A.T., Shayne, A.J., Brown, K.A., Gupta, S.N., Chan, C.W., Luu, T.M., Carli, M.F.D., Reynolds, H.G., Stevenson, W.G., Kwong, R.Y.: Characterization of the peri-infarct zone by contrast-enhanced cardiac magnetic resonance imaging is a powerful predictor of post–myocardial infarction mortality. Circulation **114**(1), 32–39 (2006). https://doi.org/10.1161/CIRCULATIONAHA.106.613414, https://www.ahajournals.org/doi/abs/10.1161/CIRCULATIONAHA.106.613414

[197] Yang, G., Zhuang, X., Khan, H., Haldar, S., Nyktari, E., Ye, X., Slabaugh, G., Wong, T., Mohiaddin, R., Keegan, J., Firmin, D.: A fully automatic deep learning method for atrial scarring segmentation from late gadolinium-enhanced mri images. In: 2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017). pp. 844–848 (2017). https://doi.org/10.1109/ISBI.2017.7950649

[198] Yang, G., Zhuang, X., Khan, H., Haldar, S., Nyktari, E., Li, L., Ye, X., Slabaugh, G., Wong, T., Mohiaddin, R., Keegan, J., Firmin, D.: Multi-atlas propagation based left atrium segmentation coupled with super-voxel based pulmonary veins delineation in late gadolinium-enhanced cardiac MRI. In: Styner, M.A., Angelini, E.D. (eds.) Medical Imaging 2017: Image Processing. vol. 10133, pp. 312 – 318. International Society for Optics and Photonics, SPIE (2017). https://doi.org/10.1117/12.2250926, https://doi.org/10.1117/12.2250926

[199] Yogananda, C.G.B., Nalawade, S.S., Murugesan, G.K., Wagner, B., Pinho, M.C., Fei, B., Madhuranthakam, A.J., Maldjian, J.A.: Fully automated brain tumor segmentation and survival prediction of gliomas using deep learning and mri. bioRxiv (2019). https://doi.org/10.1101/760157, https://www.biorxiv.org/content/early/2019/09/08/760157

[200] Yushkevich, P.A., Piven, J., Hazlett, H.C., Smith, R.G., Ho, S., Gee, J.C., Gerig, G.: User-guided 3d active contour segmentation of anatomical structures: Significantly improved efficiency and reliability. NeuroImage **31**(3), 1116 – 1128 (2006). https://doi.org/https://doi.org/10.1016/j.neuroimage.2006.01.015, http://www.sciencedirect.com/science/article/pii/S1053811906000632

[201] Zabihollahy, F., Rajchl, M., White, J.A., Ukwatta, E.: Fully automated segmentation of left ventricular scar from 3d late gadolinium enhancement magnetic resonance imaging using a cascaded multi-planar u-net (cmpu-net). Medical Physics **47**(4), 1645–1655 (2020). https://doi.org/https://doi.org/10.1002/mp.14022, https://aapm.onlinelibrary.wiley.com/doi/abs/10.1002/mp.14022

[202] Zeineldin, R.A., Karar, M.E., Coburger, J., Wirtz, C.R., Burgert, O.: Deepseg: deep neural network framework for automatic brain tumor segmentation using magnetic resonance flair

images. International Journal of Computer Assisted Radiology and Surgery **15**(6), 909–920 (2020). https://doi.org/10.1007/s11548-020-02186-z, https://doi.org/10.1007/s11548-020-02186-z

[203] Zhang, Y., Brady, M., Smith, S.: Segmentation of brain mr images through a hidden markov random field model and the expectation-maximization algorithm. IEEE Transactions on Medical Imaging **20**(1), 45–57 (Jan 2001). https://doi.org/10.1109/42.906424

[204] Zhao, H.K., Chan, T., Merriman, B., Osher, S.: A variational level set approach to multiphase motion. Journal of Computational Physics **127**(1), 179 – 195 (1996). https://doi.org/https://doi.org/10.1006/jcph.1996.0167, http://www.sciencedirect.com/science/article/pii/S0021999196901679

[205] Zheng, Q., Delingette, H., Duchateau, N., Ayache, N.: 3-d consistent and robust segmentation of cardiac images by deep learning with spatial propagation. IEEE Transactions on Medical Imaging **37**, 2137–2148 (2018)

[206] Zheng, R., Zhao, X., Zhao, X., Wang, H.: Deep learning based multi-modal cardiac mr image segmentation. In: Pop, M., Sermesant, M., Camara, O., Zhuang, X., Li, S., Young, A., Mansi, T., Suinesiaputra, A. (eds.) Statistical Atlases and Computational Models of the Heart. Multi-Sequence CMR Segmentation, CRT-EPiggy and LV Full Quantification Challenges. pp. 263–270. Springer International Publishing, Cham (2020)

[207] Zhu, J.Y., Park, T., Isola, P., Efros, A.A.: Unpaired image-to-image translation using cycle-consistent adversarial networks. 2017 IEEE International Conference on Computer Vision (ICCV) pp. 2242–2251 (2017)

[208] Zipes, D.P., Camm, A.J., Borggrefe, M., Buxton, A.E., Chaitman, B., Fromer, M., Gregoratos, G., Klein, G., Myerburg, R.J., Quinones, M.A., Roden, D.M., Silka, M.J., Tracy,

C., Smith, S.C., Jacobs, A.K., Adams, C.D., Antman, E.M., Anderson, J.L., Hunt, S.A., Halperin, J.L., Nishimura, R., Ornato, J.P., Page, R.L., Riegel, B., Priori, S.G., Moss, A.J., Priori, S.G., Blanc, J.J., Budaj, A., Camm, A.J., Dean, V., Deckers, J.W., Despres, C., Dickstein, K., Lekakis, J., McGregor, K., Metra, M., Morais, J., Osterspey, A., Tamargo, J.L., Zamorano, J.L.: Acc/aha/esc 2006 guidelines for management of patients with ventricular arrhythmias and the prevention of sudden cardiac death: A report of the american college of cardiology/american heart association task force and the european society of cardiology committee for practice guidelines (writing committee to develop guidelines for management of patients with ventricular arrhythmias and the prevention of sudden cardiac death). Journal of the American College of Cardiology **48**(5), e247 – e346 (2006). https://doi.org/https://doi.org/10.1016/j.jacc.2006.07.010, http://www.sciencedirect.com/science/article/pii/S0735109706018171

Acknowledgments

We acknowledge the use of facilities of the Research Software Engineering Sheffield (RSE) UK and the JADE Tier 2 HPC UK system specification. This work was funded by University of Sheffield Department of Computer Science Scholarships. We thank Gavin Bainbridge, Caroline Richmond, Margaret Saysell, and Petra Bijsterveld for their invaluable assistance in recruiting and collecting data for this study. The author declare that they have no competing interests.