

Image Processing Techniques for Analysis of Myocardial
Fibrosis and Related Cardiomyopathies in Cardiac
Magnetic Resonance Imaging

by

Nadia Abdel-Aziz Farrag

A doctoral thesis proposal submitted to the Faculty of Graduate
and Postdoctoral Affairs in partial fulfillment of the requirements
for the degree of

Doctor of Philosophy

in

Biomedical Engineering

Carleton University
Ottawa, Ontario

© 2022, Nadia Abdel-Aziz Farrag

Abstract

Myocardial fibrosis (MF) is a common feature of cardiac disease, characterized by excessive deposition of collagen (i.e., scar tissue) and expansion of the myocardial extracellular volume (ECV). This phenomenon contributes to cardiac dysfunction, promotes further cardiac disease, and has implication in preceding cardiac morbidity and mortality. The extent of myocardial MF can be analyzed globally (across the entire myocardial region) and/or regionally (across the fibrotic area exclusively) using cardiac magnetic resonance (CMR) imaging techniques, such as late gadolinium enhanced imaging or quantitative methods like native T1 and ECV mapping. CMR-based measurements of MF, native T1, and ECV allow for differentiation between various cardiac disease states and are shown to be clinically significant predictors of patient outcomes. However, in order to analyze tissue volumes or classify disease states, clinicians must first perform a manual tracing of the myocardial borders to define an initial region of interest (ROI), while regional MF quantification requires additional manual selection of a reference healthy myocardial tissue region. These manual processes are tedious, user-dependent, and highly prone to operator error, which can significantly confound resultant measures of T1, ECV and quantified MF tissue zones. Thus, alternative, minimally user-dependent techniques for MF, T1 and ECV quantification are appropriate. In this dissertation, several techniques for improving automated quantification of myocardial T1, ECV, and MF regions are presented. The proposed approaches presented in this document incorporate concepts from deep learning and image processing to achieve automated or semi-automated segmentation of the myocardium, MF, T1 and/or ECV in the left ventricle (LV) and left atrium (LA).

Co-authorship

This dissertation was written using the Carleton University guidelines for integrated thesis style as per the Graduate Calendar, Section 12.4, whereby materials from manuscripts that have been published in peer-reviewed journals are included per each chapter. Permission to reprint materials was obtained. All citations and information on whether materials are reproduced in part or in full is referenced per each publication in Chapter 1 of this document.

Chapter 3 presents materials from the publication in *Medical Physics* entitled “Evaluation of fully automated myocardial segmentation techniques in native and contrast-enhanced T1-mapping cardiovascular magnetic resonance images using fully convolutional neural networks” by N. A. Farrag, A. Lochbihler, J. A. White, and E. Ukwatta. I developed and implemented the algorithms, performed experiments, analyzed results, and wrote the manuscript. A. Lochbihler contributed to algorithm development and provided ideas on fully convolutional networks, including coding ideas and hyperparameter changes. J. A. White provided CMR imaging data and provided ideas for clinical evaluation. E. Ukwatta provided ideas for statistical analysis, segmentation evaluation, and parameter adjustments for the Modality Independent Neighborhood Descriptor algorithm. All authors contributed to the preparing of the manuscript.

Chapter 4 presents materials from the publication in *Society of Optics and Photonics Medical Imaging* conference entitled, “Automated myocardial segmentation of extra-cellular volume mapping cardiac magnetic resonance images using fully convolutional neural networks” by N. A. Farrag, S. Bhagavan, D. Sebben, P. Ruwanpura, J. A. White, and E. Ukwatta. I developed and implemented the algorithms, performed

experiments, statistical analysis, analyzed results, performed manual segmentations, and wrote the manuscript. S. Bhagavan contributed to algorithm development and performed experiments, evaluated algorithms, and implemented comparisons across various hyperparameters. D. Sebben and S. Ruwanpura completed manual segmentations. J. A. White provided CMR imaging data and contributed to clinical analysis ideas. E. Ukwatta contributed to algorithm development, segmentation evaluation, statistical analysis, and analysis of results. All authors helped in writing the manuscript.

Chapter 5 presents materials from the publication in *PLOS One* entitled “Assessment of left atrial fibrosis progression in canines following rapid ventricular pacing using 3D late gadolinium enhanced CMR images”, by N. A. Farrag, R. E. Thornhill, F. S. Prato, A. C. Skanes, R. Sullivan, D. Sebben, J. Butler, J. Sykes, B. Wilk, and E. Ukwatta. I developed and implemented the algorithms, performed experiments, performed manual segmentations of CMR images, performed and analyzed statistical tests, and wrote the manuscript. R. E. Thornhill performed manual segmentation of CMR images, contributed toward algorithm development and implemented experiments, performed statistical tests, and analyzed results. A. C. Skanes and F. S. Prato produced the experimental design, project administration, funding acquisition and allocation, and provided ideas for implementation of experiments and statistical testing. R. Sullivan performed histological analysis. J. Butler and J. Sykes acquired CMR images and implemented quality control testing. B. Wilk contributed to the study design. E. Ukwatta supervised the image processing aspect of the project and provided ideas for statistical testing and experimental design. All authors helped in preparing the manuscript.

Chapter 6 presents materials from the publication in Medical Physics entitled “Effect of T1-mapping technique and diminished image resolution on quantification of infarct mass and its ability in predicting appropriate ICD therapy” by N. A. Farrag, V. Ramanan, G. A. Wright, and E. Ukwatta. I implemented algorithms, performed experiments and statistical tests, and wrote the manuscript. R. Ramanan provided ideas for the image processing algorithms and contributed toward study design. G. A. Wright supervised and contributed toward study design. E. Ukwatta supervised the image processing aspect of the project, implemented and developed experiments, and came up with ideas for statistical testing and analysis. All authors contributed toward writing the manuscript.

Chapter 7 is in the process of being prepared for submission to a reputable journal and is not yet published. Several people contributed towards Chapter 7 of the thesis, including myself, P. Ayeegoundarvenkatesan, P. Samal, S. Bhagavan, J. A. White, and E. Ukwatta. I developed and implemented the algorithms, performed experiments, provided ideas on appropriate statistical analysis, analyzed results, and performed manual segmentations. P. Ayeegoundarvenkatesan, P. Samal, and S. Bhagavan contributed to algorithm development, performed experiments, and evaluated algorithms. J. A. White provided ideas on the experimental design and classification algorithms, provided supervision for clinical analysis, and provided CMR imaging data. E. Ukwatta contributed toward experimental design and provided ideas for algorithm implementation and statistical analysis.

Acknowledgements

This dissertation was made possible thanks to the support of many people. I would like to firstly thank my supervisors, Dr. Eranga Ukwatta and Dr. Yuu Ono for their guidance throughout this journey. I am sincerely grateful to Dr. Eranga Ukwatta for giving me the opportunity to study medical image processing with his group in September of 2016. Coming from a clinical background and transitioning to the biomedical engineering program was intimidating, but Dr. Ukwatta's faith in my research abilities and his support, expert knowledge, and mentorship are what got me to the place I am at today. I would also like to express my sincere appreciation for all the research opportunities I was able to experience by being a part of Dr. Ukwatta's lab group, which have undoubtedly contributed to shaping me as an independent researcher. To Dr. Yuu Ono, I am sincerely thankful for the valuable insights and guidance you provided in the latter years of my studies, after I transferred from the Master's program to the PhD program.

I would also like to express sincere gratitude to my collaborators. Dr. James White, who gave me the opportunity to work with unique cardiac MR imaging data and provided me with the valuable medical expertise required to support my research goals. Thank you for all of the insights and feedback, which contributed immensely to making this research possible. Dr. Rebecca Thornhill, who provided me with medical and technical expertise. I am so grateful for having gotten the opportunity to collaborate with Dr. Thornhill and truly appreciative for her positivity, encouragement, and valuable review of my work.

I would like to acknowledge the funding support I received from numerous sources, including Carleton University and the Ontario Graduate Scholarship, which were vital in supporting my productivity and dedication to my research.

Finally, I would like to thank my family and friends for their support throughout these years. My parents, who's motivation and encouragement gave me the determination and tenacity needed to pursue this research. My sisters and brother, who's patience, reassurance, and cheering from the sidelines helped me stay focused and reach the finish line. And lastly, my lab mates and colleagues, who made my graduate studies experience so amazing. The friendships I was able to build throughout this journey are invaluable to me and have positively impacted my life so much. The encouragement and comradery of my lab mates was crucial throughout my years of study, and gave me the determination, perseverance, and confidence needed to complete this research. Thank you all.

Table of Contents

Abstract.....	2
Co-authorship.....	3
Acknowledgements	6
Table of Contents	8
List of Tables	16
List of Figures.....	19
List of Abbreviations	22
Chapter 1: Preface	26
1.1 Motivation.....	26
1.2 Possible Challenges and Limitations.....	27
1.3 Summary of Contributions.....	28
1.3.1 Myocardial Border Segmentation.....	28
1.3.2 Clinical Outcomes and Disease Correlation Based on MF Segmentation Indices and Global Myocardial T1/ECV	30
1.4 Thesis Organization	34
1.5 Chapter 1 References	36
Chapter 2: Background and Literature Review	38
2.1 Clinical Problem: Myocardial Fibrosis	38
2.1.1 Cardiac Anatomy, Structure and Function.....	39
2.1.2 Myocardial Fibrosis Subtypes	41
2.1.3 Assessment of Myocardial Fibrosis	44
2.1.4 Conditions Related to Myocardial Fibrosis	46
2.2 Non-invasive Imaging of Myocardial Fibrosis Using Cardiac Magnetic Resonance	

2.2.1	Late Gadolinium Enhancement (LGE)	47
2.2.2	T1 Mapping	50
2.2.3	ECV Mapping	56
2.2.4	Cinematic Cardiac Magnetic Resonance	58
2.3	Measurement of Measurement of Myocardial Fibrosis in Cardiac Magnetic Resonance Imaging	60
2.3.1	Intensity-based Methods for Segmentation of the Myocardium and MF	61
2.3.1.1	Threshold Segmentation Techniques	61
2.3.1.1.1	Full width at half maximum	62
2.3.1.1.2	Signal threshold versus reference mean	62
2.3.1.1.3	Region growing	62
2.3.1.2	Energy-minimization-based Segmentation Techniques	63
2.3.1.2.1	Active contour model	63
2.3.1.2.2	Level set method	64
2.3.1.2.3	Graph cuts	65
2.3.1.3	Clustering Segmentation Techniques	66
2.3.2	Feature-based Methods of Analysis of MF	66
2.3.2.1	Radiomics analysis	67
2.3.2.1.1	Morphological features	69
2.3.2.1.2	Histogram-based features	69
2.3.2.1.3	Texture features	70
2.3.2.1.4	Transform-based features	70
2.3.2.2	Machine learning classification using radiomic analysis	71
2.3.3	Deep learning-based Methods for Analysis of MF	71
2.3.3.1	Fully convolutional networks	74
2.3.3.2	U-Net	75

2.3.3.3	DenseNet.....	78
2.3.3.4	U-Net++	79
2.3.3.5	Attention U-Net.....	81
2.4	Evaluation of Myocardial Segmentation and Fibrosis Measurements.....	82
2.4.1	Comparison Metrics.....	83
2.4.2	T1 and ECV Mapping Measurements of Myocardial Fibrosis.....	86
2.4.3	Statistical Tests	87
2.5	Chapter 2 References	88
Chapter 3: Fully Automated Segmentation of the Left Ventricular Myocardium in Native and Contrast-enhanced T1-mapping and CINE CMR Images Using Fully Convolutional Neural Networks		
100		
3.1	Introduction.....	100
3.2	Selection of study population and image acquisition.....	101
3.3	Method 1: Direct segmentation of native and contrast-enhanced T1 maps using U-Net based CNN model.....	103
3.3.1	Preprocessing.....	104
3.3.2	Training and testing procedure of U-Net-based CNN model	104
3.4	Method 2: Segmentation of native and contrast-enhanced T1-maps using alternative technique based on registration and propagation of CINE contours	106
3.4.1	Preprocessing.....	107
3.4.2	Automated myocardial segmentation of CINE images.....	108
3.4.3	Propagation of myocardial contours from CINE images to T1-maps using deformable non-rigid registration.....	108
3.5	Computation of native and post-contrast global myocardial T1 values	110
3.6	Results.....	110

3.7	Discussion	119
3.8	Conclusion	121
3.9	Chapter 3 References	122
Chapter 4: Automated myocardial segmentation of extra-cellular volume mapping		
cardiac magnetic resonance images using fully convolutional neural networks..... 125		
4.1	Purpose.....	125
4.2	Methods.....	126
4.2.1	Study Population and CMR Protocol.....	126
4.2.2	Image processing	127
4.2.2.1	Standard U-Net	127
4.2.2.2	U-Net++	128
4.2.2.3	Training procedure	129
4.2.3	Inter-operator variability analysis	129
4.3	Results.....	131
4.3.1	Segmentation Results.....	131
4.3.2	Inter-operator variability analysis	133
4.4	Discussion	135
4.5	Conclusion	137
4.6	Chapter 4 References	137
Chapter 5: Evaluation of the Progression of Left Atrial Fibrosis in Canines		
Following Rapid Ventricular Pacing Using 3D Late Gadolinium Enhanced CMR		
Images		
139		
5.1	Introduction.....	139
5.2	Methods.....	140
5.2.1	Image Acquisition.....	140

5.2.2	Image Processing	142
5.2.2.1	Segmentation of LA myocardium and enhanced regions	142
5.2.2.2	Landmark-based 3D image registration	147
5.2.2.3	Histopathological imaging	148
5.2.3	Method Evaluation and Statistical Analysis	149
5.2.3.1	Observer variability.....	149
5.2.3.2	Statistical Tests	150
5.3	Results.....	150
5.3.1	Segmentation of MF volumes using 3D-FLASH CMR	150
5.3.2	Segmentation of slice-wise MF areas using 3D-FLASH CMR.....	154
5.3.3	Registration of Myocardial Volumes.....	156
5.3.4	Histopathological Validation	158
5.3.5	Observer Variability	159
5.4	Discussion	162
5.5	Conclusion	169
5.6	Chapter 5 References	169

Chapter 6: Effect of T1-mapping Technique and Diminished Image Resolution on Quantification of Infarct Mass and its Ability in Predicting Appropriate ICD

Therapy	172	
6.1	Introduction.....	172
6.2	Methods.....	174
6.2.1	Image Acquisition.....	174
6.2.1.1	Study Populations and CMR Protocol	174
6.2.1.2	CMR Pulse Sequence Parameters	176
6.2.2	Image Processing	178

6.2.2.1	IR-FGRE Images.....	179
6.2.2.2	MOLLI and MCLE Images.....	179
6.2.3	Statistical Analyses.....	181
6.3	Results.....	182
6.4	Discussion.....	185
6.5	Conclusions.....	189
6.6	Chapter 6 References.....	189

Chapter 7: Radiomic analysis-based classification of cardiomyopathy states

manifesting with interstitial myocardial fibrosis using native T1 and ECV mapping

CMR.....	192	
7.1	Introduction.....	192
7.2	Methods.....	194
7.2.1	Patient population and CMR imaging protocol.....	194
7.2.2	Image processing.....	195
7.2.2.1	Preprocessing.....	196
7.2.2.2	Myocardial segmentation.....	196
7.2.3	Cardiomyopathy phenotype classification.....	197
7.2.3.1	RA-based classification method.....	198
7.2.3.1.1	Feature extraction.....	199
7.2.3.1.2	Pre-processing of radiomic features.....	200
7.2.3.1.3	Model training.....	201
7.2.3.1.4	Ensemble approach using stacking with linear support vector classifier	202
7.2.3.2	Deep learning-based classification method.....	203
7.2.3.2.1	CNN training and testing procedure.....	203

7.2.3.2.2	CNN Architecture	204
7.2.4	Performance Evaluation and Statistical Analysis	205
7.3	Results	206
7.3.1	Myocardial Segmentation	206
7.3.2	Binary classification using native T1-mapping images	207
7.3.3	Multiclass classification using ECV mapping images	210
7.3.4	Statistical Analyses	213
7.4	Discussion	215
7.5	Conclusion	218
7.6	Chapter 7 References	218
Chapter 8: Conclusions and Future Work		222
8.1	Summary of Findings	222
8.1.1	Fully automated myocardial segmentation methods for native and post-contrast T1-mapping CMR (Chapter 3)	222
8.1.2	Fully automated myocardial segmentation and observer variability analysis of myocardial segmentation in ECV mapping CMR (Chapter 4)	223
8.1.3	Validation of the IIR quantification technique for analysis of the MF progression in the LA following persistent AF in a canine model (Chapter 5)	224
8.1.4	Analysis of the effects of spatial resolution on quantification of replacement MF volumes in MOLLI and MCLE T1-mapping CMR (Chapter 6)	225
8.1.5	Classification of five cardiomyopathy phenotypes associated with interstitial MF using native T1 and ECV mapping CMR (Chapter 7)	226
8.2	Limitations and Future Work	227
8.2.1	Fully automated myocardial segmentation methods for native and post-contrast T1-mapping CMR (Chapter 3)	227

8.2.2	Validation of the IIR quantification technique for analysis of the evolution of MF volumes in the LA following progression AF in a canine model (Chapter 5)	227
8.2.3	Analysis of the effects of spatial resolution on quantification of replacement MF volumes in MOLLI and MCLE T1-mapping CMR (Chapter 6).....	228
8.2.4	Fully automated pipeline for myocardial segmentation and classification of cardiomyopathy phenotypes associated with interstitial MF using native T1 and ECV mapping CMR (Chapters 4 and 7)	229
8.3	Conclusion	230
8.4	Contributions.....	230
8.5	Chapter 8 References	230
Appendices.....		231
	Appendix A Semi-automated Segmentation of the Left Ventricular Myocardium in Native T1-mapping CMR Using Deformable Non-Rigid Registration of CINE Images.....	231
	A.1 Method.....	231
	A.2 Results.....	233
	A.3 Discussion & Conclusions	235
	A.4 Appendix A References.....	236
	Appendix B Inter/Intra Observer Variability of Left Atrial Myocardial Fibrosis Volumes	237
	B.1 Boundary F1 Score Results	237
	B.2 Precision and Recall Results.....	238
	B.3 Relative Error Index.....	239

List of Tables

Table 2-1 Confusion matrix used to compute accuracy metrics.....	83
Table 3-1 Image acquisition parameters of CINE and shMOLLI pulse sequences.....	102
Table 3-2 Results of CNN-based myocardial segmentation techniques in T1-maps	111
Table 3-3 Results of CINE-registration based segmentation technique in T1-maps.....	113
Table 3-4 Native T1 values computed using various segmentation techniques	116
Table 3-5 Post-contrast T1 values computed using various segmentation techniques...	116
Table 4-1 Segmentation results on ECV maps using U-Net and U-Net++	131
Table 4-2 metrics computed across myocardial masks generated by three operators versus U-Net++ and basic U-Net.	134
Table 4-3 Results of the Kruskal-Wallis (KW) test computed on the full distribution of myocardial ECV values derived by operators vs. U-Net and U-Net++.....	134
Table 4-4 Spearman correlation coefficient ρ computed on global (i.e., mean) ECV values derived by operators vs. operators, U-Net++ and basic U-Net.	135
Table 5-1 Average pulse sequence parameters for 3D LGE-CMR sequence.....	142
Table 5-2 LA MF volumes computed using IIR technique	151
Table 5-3 Canine slice-wise LA MF measurements at different time points.	154
Table 5-4 MF measurements in corresponding histopathological images.....	158
Table 5-5 Observer variability DSC metrics	160
Table 5-6 LA MF volumes measured by different operators	160
Table 6-1 Patient characteristics and LV functional parameters in 27 patients with ICD devices.....	176

Table 6-2 CMR imaging parameters of IR-FGRE, MOLLI and MCLE imaging techniques	177
Table 6-3 Normalized IC and BZ masses quantified by different CMR techniques	182
Table 6-4 Comparison of IC and BZ regions quantified by different CMR techniques.	183
Table 6-5 Comparison of infarct masses found by original vs. downsampled MCLE in patients with and without ICD therapy	185
Table 7-1 Image availability per patient population	195
Table 7-2 Hyperparameters used for GBRT binary and multi-class classification algorithm.....	202
Table 7-3 Hyperparameters used for KNN binary and multi-class classification algorithm	202
Table 7-4 Hyperparameters for binary and multiclass classification of ECV mapping images	204
Table 7-5 Contingency table for the McNemar test.....	206
Table 7-6 Accuracy metrics of myocardial segmentations generated using U-Net and transfer learning	206
Table 7-7 Distribution of binary feature importance across categories.....	208
Table 7-8 Accuracy metrics of binary abnormal vs. normal classification using native T1-mapping images	209
Table 7-9 Distribution of multiclass feature importance across categories.....	210
Table 7-10 Performance metrics computed on algorithms used for multiclass classification of ECV mapping images.....	211

Table 7-11 Spearman correlation coefficients of GBRT, KNN and CNN classification models versus true labels	213
--	-----

List of Figures

Figure 2-1 Position of the human heart in the chest cavity in anterior (frontal) view	40
Figure 2-2 Layers of the heart muscle	41
Figure 2-3 Cellular composition of myocardium in healthy vs. diseased states.....	43
Figure 2-4 Example LGE-CMR image demonstrating enhanced fibrosis.....	49
Figure 2-5 Curve fitting method used to create T1-maps.....	52
Figure 2-6 Example T1-maps demonstrating expansion of ECV.....	55
Figure 2-7 ECV maps of patients with different cardiomyopathies	57
Figure 2-8 Cardiovascular diseases associated with changes in ECV	58
Figure 2-9 CINE images and geometric axes of the ventricles	59
Figure 2-10 Radiomic pipeline for classification.....	69
Figure 2-11 Diagram of basic CNN model.....	74
Figure 2-12 U-Net architecture.	76
Figure 2-13 Architecture of Dense Blocks used in the DenseNet model.	79
Figure 2-14 U-Net++ architecture.	81
Figure 2-15 Attention U-Net architecture.....	82
Figure 2-16 Example ROC curves of two classifiers.....	86
Figure 3-1 Framework of CNN-based automated myocardial segmentation steps	104
Figure 3-2 Framework of CINE registration-based segmentation technique	107
Figure 3-3 Segmentation results U-Net vs. CINE-registration methods in native T1-maps.	114
Figure 3-4 Segmentation results U-Net vs. CINE-registration methods in post-contrast T1-maps.	115

Figure 3-5 Box and whisker plots of DSC metric for U-Net and CINE registration segmentation techniques	116
Figure 3-6 Bland-Altman plots of native myocardial T1 values across segmentation methods.....	118
Figure 3-7 Bland-Altman plots of post-contrast myocardial T1 values across segmentation methods.....	118
Figure 4-1 Framework used for CNN model based on standard U-Net architecture in ECV map segmentation.	128
Figure 4-2 Framework used for CNN model based on U-Net++ for ECV map segmentation	129
Figure 4-3 Segmentation results on ECV maps using U-Net and U-Net++	131
Figure 4-4 Sub-par myocardial segmentations output by U-Net++.	133
Figure 4-5 Manual myocardial delineation by three operators versus U-Net and U-Net++	134
Figure 5-1 Flow-chart of study protocol and 3D LGE-CMR imaging time points.	142
Figure 5-2 Image processing pipeline for MF segmentation of the LA.	146
Figure 5-3 Example workflow of canine LA myocardial segmentation.....	147
Figure 5-4 Images of the left atrium in canines at different time points.....	151
Figure 5-5 Dot and line plot of MF volumes in canines.	154
Figure 5-6 Evolution of LA MF volumes across time points (experimental dogs).....	157
Figure 5-7 Evolution of LA MF volumes across time points (control dogs).....	158
Figure 5-8 Histopathological images of LA samples demonstrating MF regions.	159
Figure 6-1 T1* and steady-state maps for MOLLI, MCLE and downsampled MCLE..	181

Figure 6-2 Segmentation of IC and BZ masses in different CMR techniques for two example patients.....	183
Figure 6-3 T1* and steady-state scatterplots for a series of original and downsampled MCLE slices.....	184
Figure 6-4 Example IC and BZ regions in MCLE and downsampled MCLE images. ..	185
Figure 7-1 Image processing pipeline used for cardiomyopathy disease classification.	196
Figure 7-2 Classification approach used for native T1-mapping and ECV mapping images.	197
Figure 7-3 Stacking ensemble approach for RA-based classification	203
Figure 7-4 CNN architecture used for cardiomyopathy classification.	204
Figure 7-5 Myocardial segmentations derived using U-Net versus manually.....	207
Figure 7-6 Pie chart of filters and transforms with highest MI for binary classification	208
Figure 7-7 Normalized confusion matrices for three binary classification models.	209
Figure 7-8 Binary ROC and AUC for three classification models.	210
Figure 7-9 Pie chart of filters and transforms with highest MI for multiclass classification	210
Figure 7-10 Normalized confusion matrices for multiclass classification algorithms. ..	212
Figure 7-11 Multiclass ROC and AUC for three classification models.	213
Figure 7-12 Contingency tables for three-way comparison of binary classification models	215
Figure 7-13 Contingency tables for three-way comparison of multiclass classification models.....	215

List of Abbreviations

Abbreviations	Meaning
ACC	Animal Care Committee
AF	atrial fibrillation
ANOVA	analysis of variance
ARRIVE	Animal Research: Reporting of In vivo Experiments
AUC	area under the curve
AUP	Animal Use Protocol
AV	atrioventricular node
BF	boundary F1 score
BW	bandwidth
BZ	border zone
CAD	coronary artery disease
CINE	Cinematic
CIROC	Cardiovascular Imaging Registry of Calgary
CMF	continuous max-flow
CMR	Cardiac Magnetic Resonance
CNN	convolutional neural network
CT	Computed Tomography
CVD	cardiovascular disease
DCM	dilated cardiomyopathy
DI	Diagnostic Imaging
DICOM	Digital Imaging and Communications in Medicine
DSC	Dice similarity coefficient
DWI	diffusion weighted imaging
EAM	electroanatomical voltage mapping
ECG	electrocardiogram
ECM	extracellular matrix
ECV	extracellular volume
EDV	end diastolic volume
EF	ejection fraction
ELU	exponential linear unit
EM	expectation maximization
ESV	end systolic volume
FCN	fully convolutional networks
FGRE	fast gradient echo
FLASH	fast low angle shot
FN	false negative
FOV	field of view
FP	false positive

FPR	false positive rate
FWHM	full width at half maximum
GBCA	Gadolinium-based contrast agent
GBRT	gradient boosting regression trees
GE	General Electric
GLCM	gray-level co-occurrence matrix
GLRLM	gray-level run length matrix
GLSZM	gray-level size zone matrix
GPU	graphics processing unit
GRE	gradient echo
HCM	hypertrophic cardiomyopathy
HFPEF	heart failure with preserved ejection fraction
HTN	hypertensive heart disease
IBM	International Business Machines
IC	infarct core
ICC	intraclass correlation coefficient
ICD	implantable cardioverter defibrillator
ICM	ischemic cardiomyopathy
IIR	image intensity ratio
IQR	interquartile range
IR	inversion recovery
IV	intravenous
KNN	K nearest neighbours
LA	left atrial
LASSO	least absolute shrinkage and selection operator
LGE	late gadolinium enhanced
LOG	Laplacian of Gaussian
LU	linear unit
LV	left ventricular
LVEF	left ventricular ejection fraction
LVH	left ventricular hypertrophy
LVM	left ventricular mass
LVSV	left ventricular stroke volume
MATLAB	Matrix Laboratory
MCLE	Multi-contrast late enhancement
MF	myocardial fibrosis
MI	myocardial infarction
MIND	modality independent neighborhood descriptor
MOLLI	Modified Look-Locker
MR	Magnetic Resonance
MRI	Magnetic Resonance Imaging

NEX	number of excitations
NGTDM	neighborhood gray-tone difference matrix
NMR	nuclear magnetic resonance
NVS	number of views per segment
NYHA	New York Heart Association
PCA	principle component analysis
PET	Positron Emission Tomography
PLOS	Public Library of Science
PSIR	phase sensitive inversion recovery
RA	radiomic analysis
RE	relative error
RF	radiofrequency
RGB	red, green, and blue
RMS	root mean square
ROC	receiver operating curve
ROI	region of interest
RV	right ventricular
SA	sinoatrial
SAPPHIRE	saturation pulse prepared heart rate independent inversion recovery
SASHA	saturation recovery single-shot acquisition
SCD	sudden cardiac death
SD	standard deviation
SHARCNET	Shared Hierarchical Research Computing Network
SI	signal intensity
SIBZ	signal intensity border zone
SIIC	signal intensity infarct core
SMOTE	synthetic minority oversampling technique
SNR	signal to noise ratio
SPSS	Statistical Package for Social Sciences
SR	Saturation recovery
SRI	Sunnybrook Research Institute
SSD	sum of square differences
SSFP	steady-state free precession
STRM	signal threshold versus reference mean
SV	stroke volume
SVC	support vector classifier
SVM	support vector machine
TE	time to echo
TI	inversion time
TN	true negative
TP	true positive

TPR	true positive rate
TR	repetition time
VF	ventricular fibrillation
VPS	views per segment
VT	ventricular tachycardia

Chapter 1: Preface

1.1 Motivation

Myocardial fibrosis (MF), also referred to as “scarring” of the heart muscle, is a condition that coexists with and/or succeeds a number of cardiac diseases. This condition is characterized by myocardial tissue remodeling and expansion, both of which jeopardize cardiovascular health and increase the likelihood of developing further cardiovascular disease.

Different patterns of MF are associated with different disease states, thus, non-invasive techniques to characterize deposition of MF are desired for diagnostic purposes. To this aim, cardiac magnetic resonance (CMR) imaging is commonly employed. Moreover, MF predicts major adverse cardiovascular events and death [1]– [3], where higher degrees of MF are universally associated with poorer patient prognoses. Thus, MF measurements are often used to predict patient outcomes, stratify risk, and monitor treatment [4],[5]. Therefore, the development of reliable and accurate measurement techniques for MF quantification is crucial for clinicians and their patients. Several MF quantification techniques have been suggested for this purpose, however, universal agreement on a single method is yet to be established [6],[7]. Although highly prone to operator error and variability, manual planimetry of both the myocardial borders and the MF regions is considered the reference standard [8].

Therefore, the motivation for this thesis was to evaluate the ability of quantitative imaging techniques to diagnose and predict disease states associated with MF using

automated and/or minimally user-dependent techniques for segmentation of the myocardial border and MF. We propose several approaches for this aim, using both conventional image processing-based and deep learning-based techniques.

1.2 Possible Challenges and Limitations

To accurately assess and quantify MF non-invasively, an appropriate modality of imaging should be taken into consideration. CMR imaging using late gadolinium enhancement (LGE) is the front-line imaging technique used to diagnose and assess the degree of MF in clinical practice. In this technique, the patient undergoes CMR imaging after intravenous injection of a gadolinium-based contrast agent (GBCA), which serves to “highlight” or “enhance” regions consistent of fibrosis within the myocardium. The enhanced areas can then be quantified and used toward risk stratification, outcome prediction, and treatment planning [1]– [4]. However, as discussed in Section 2.2.1 of this thesis, the LGE CMR technique poses a number of challenges in measurement accuracy, as the method is not inherently quantitative, but rather, expresses enhancement of myocardial tissue on an arbitrary scale. Therefore, in order to produce more accurate and reliable measurements of MF, we propose the use of quantitative CMR-based imaging techniques.

T1 and extracellular volume (ECV) mapping are two methods that have been used toward this aim. Both methods produce images wherein pixel intensities are directly translatable to an intrinsic quantitative parameter (i.e., the T1 relaxation value and ECV percentage, respectively), which is altered in regions consisting of fibrosis. Both T1-maps and ECV-maps have been shown to correlate highly with histological imaging of the human myocardium [5]– [7] and may therefore be stronger predictors of patient outcomes

as compared with LGE CMR. However, measuring MF can be challenging in these kinds of images. Prior studies implement semi-automated techniques to quantify MF in T1 and ECV maps, after defining an initial myocardial region of interest (ROI) using either manual delineation or manual correction of an estimated ROI [6],[8]– [10]. Our aim is to advance current techniques by developing algorithms for fully automated myocardial and/or MF segmentation that are independent of user input.

There are several conceivable obstacles to this task. Contrast differences between the myocardium, the blood pool and other background structures are not as stark or easily conceivable as compared to LGE CMR and other forms of CMR imaging. This may lead to erroneous inclusion of non-myocardial tissues in the automated myocardial ROI. Spatial resolution is also a limiting factor. T1 and ECV mapping images tend to have coarser spatial resolution, which can lead to partial volume effect, a phenomenon where a single pixel contains intensities from several tissue types and results in an “averaging” effect. This may lead to inaccurate computation of T1 and ECV and further, may confound MF measurements, particularly in areas of sporadic or intermittent MF.

1.3 Summary of Contributions

The research work in this thesis has novelty in both clinical application and in method development. The contributions are classified into two groups, (1) myocardial border (i.e., initial ROI) segmentation methods and (2) clinical outcomes and disease correlation based on MF segmentation indices and/or quantitative features of the myocardium.

1.3.1 Myocardial Border Segmentation

We have developed myocardial border segmentation algorithms for direct application to T1 and ECV maps. The methods are described below.

(1) Fully automated myocardial segmentation of native T1, contrast enhanced T1, and ECV maps using convolutional neural networks

In this method discussed in Chapter 3, we use a convolutional neural network (CNN) based on the renowned U-Net [11] architecture for fully automated segmentation of the myocardium in both native (non-contrast) and contrast enhanced T1-maps in patients with three different forms of myocardial disease (also called cardiomyopathy): ischemic, hypertrophic, and dilated (ICM, HCM and DCM, respectively). To our knowledge, this is the first study of its kind to achieve fully automated myocardial segmentation on both native and contrast enhanced maps across different cardiac disease states – a critical milestone, since MF deposition and placement differs across cardiomyopathies and can significantly influence segmentation results. This work is published in the Journal of Medical Physics. See Section 1.4 for citation and Chapter 3 for full details.

Similarly, in Chapter 4, we achieve fully automated myocardial segmentation of ECV maps using a CNN based on the U-Net [11] and U-Net++ [12] architectures. We demonstrate that either framework can be used to achieve accurate myocardial contours in ECV mapping images with negligible differences in accuracy. This study is the first, to our knowledge, to achieve fully automated segmentation of myocardial contours in ECV maps. Prior studies have implemented semi-automated methods for this purpose. We also performed an inter-observer variability analysis in this study, which showed that myocardial ECV values computed across different operators can vary significantly, but differences for the same operators against both U-Net and U-Net++ are not statistically significant. Thus, this work highlights the impact that operator error can have on myocardial ECV values. This work has been published and presented in the Society for

Optics and Photonics Conference on Medical Imaging: Biomedical Applications in Molecular, Structural and Functional Imaging (see section 1.4, description on Chapter 4 for full details).

(2) Fully automated myocardial segmentation of native and contrast enhanced T1 maps using registration and deformation of prior contours from CINE images

Here, we achieve fully automated segmentation of native and contrast enhanced T1 maps using a framework based on deforming and propagating myocardial contours segmented automatically from CINE CMR images. This method takes advantage of the superior contrast of CINE images, wherein myocardial tissue is starkly distinguishable from the blood pool and background tissues. This is the first study to use CINE images toward segmentation of T1-mapping images. Some meaningful findings were reported with regards to accuracy metrics computed in patients with ICM using this technique versus the direct application of the CNN-based technique (described above). This work is part of the same publication mentioned above (see section 1.4, description of Chapter 3).

1.3.2 Clinical Outcomes and Disease Correlation Based on MF Segmentation

Indices and Global Myocardial T1/ECV

These contributions are centered around using measurements from myocardial pixels to predict patient outcomes and/or correlate to various cardiac disease states. We have implemented methods for MF segmentation on LGE CMR and T1 mapping images (with and without the use of intravenous GBCAs). We have used regional MF measurements and overall (i.e., global) T1 and/or ECV measurements to contribute to the quantitative classification of various myocardial disease types. We have also used MF measurements

toward outcome prediction and correlated MF measurements with secondary cardiac disease.

(1) Validation of an algorithm for semi-automated segmentation of MF from the left atrium (LA) and correlation to LA fibrillation

Atrial fibrillation (AF) and MF are two conditions that notoriously coexist; however, the causality of these conditions is yet to be fully recognized. While there is evidence that AF contributes to the development of MF, there is also evidence that MF can cause and/or perpetuate an AF process [13]– [18]. Thus, the aim of this work was to corroborate the causality of AF to MF. Using LGE CMR images of canines, we validated an algorithm for MF segmentation based on the image intensity ratio (IIR) [19], wherein the manually segmented myocardial region is normalized by the mean of the blood pool chamber and MF pixels are classified as the mean normalized myocardial tissue intensity plus two standard deviations. In this work, we validated the performance of this segmentation technique over conventional segmentation techniques such as full width at half maximum (FWHM) and signal threshold to reference mean (STRM), which are notorious to over-segment MF in the thin-walled LA. The IIR method has been implemented for LA MF segmentation in human studies, however, our work is the first to validate IIR MF measurements to histopathological imaging and correlate the evolution of MF measurements before and after AF is present. This work has been published in the Public Library of Science (PLOS) One Journal. Full details can be found in Chapter 4.

(2) Classification of cardiomyopathy disease types using global T1 values in native and contrast-enhanced T1-mapping images

In patients with various forms of cardiac disease involving fibrosis, myocardial T1 values are altered. Thus, different cardiomyopathies may be diagnosed and distinguished using T1 values. Here, we use our automated myocardial segmentation algorithm to delineate myocardial borders in order to report global myocardial T1 values across HCM, ICM and DCM states using both native and contrast-enhanced T1-maps. This work contributes to improved establishment of the accepted ranges for diseased myocardial T1 in native and contrast-enhanced images, which may facilitate improved diagnostics. This research is part of our work published in the Journal of Medical Physics. The full citation information is described below under “Chapter 3”.

(3) Quantification of MF core and “border” zones in multi-contrast late-enhancement T1-mapping images and the impact of spatial resolution in predicting arrhythmic processes

MF caused by myocardial infarction (MI) may consist of an infarct core (IC) and a heterogeneous, semi-viable border zone (BZ). Patients with chronic MI in the left ventricular (LV) myocardium are at increased risk of developing ventricular arrhythmias and may therefore qualify for implantable cardioverter defibrillator (ICD) therapy. Indices based on MF mass are shown to be sensitive in predicting adverse ventricular arrhythmic events. In this work, we implement quantification of IC and BZ regions using a fuzzy c-means clustering algorithm and compare the computed masses determined by T1-mapping CMR techniques to those of conventional LGE CMR. Specifically, we found that diminished spatial resolution in T1-mapping images led to an increase in estimation of the BZ mass, likely due to partial volume effects, which led to a reduced sensitivity in the prediction of appropriate ICD therapy. This work demonstrates the importance of using

imaging techniques with appropriate spatial resolution, particularly when predicting patient outcomes or planning treatment. This work is published in the Journal of Medical Physics. Citation information and full details regarding the methods implemented in this work are available below under “Chapter 6”.

(4) Classification of cardiomyopathy phenotypes presenting with interstitial MF in native T1 and ECV mapping CMR using radiomic analysis and machine learning

Left ventricular hypertrophy and interstitial MF are key features of several cardiac diseases, including HCM, cardiac amyloidosis, hypertension (HTN) and Fabry’s disease [20]-[24]. These conditions typically require multi-modality imaging, inclusive of echocardiography, nuclear imaging, CMR, Computed Tomography (CT), etc. to support diagnosis [20],[23]. Moreover, the treatment and management of such disease states can vary significantly – from implantable cardioverter defibrillator (ICD) therapy and/or protein stabilizing drugs (amyloidosis) to antihypertensive drugs (HTN), beta and calcium blockers (HCM), or enzyme replacement therapy (Fabry’s disease) [23]. In this work, we aim to achieve CMR-based binary (normal versus abnormal) and multiclass classification of patients with amyloidosis, Fabry’s disease, HCM, HTN, and healthy volunteers using native T1 mapping and ECV mapping imaging, respectively. We compared three methods, categorized into two approaches which were based on (1) radiomic analysis (RA) or (2) deep learning. The two RA-based methods included gradient boosting regression trees (GBRT) and k nearest neighbour (KNN), both of which were applied with ensemble learning, i.e., “stacked” with a linear support vector classifier (SVC). The deep learning-based method was a CNN with four convolutional layers, each followed by max pooling and two dense layers. In this work, we found that RA-based methods outperform CNNs for both binary and multiclass

classification of disease states, however, differences in performance for the binary classification task were negligible across all three methods. On the other hand, for multiclass classification, RA-based methods (particularly, the GBRT and SVC ensemble method) outperformed CNNs across all accuracy metrics. While class imbalance led to deficits in accuracy metrics for minority classes (i.e., HTN and Fabry's disease) across all classification techniques, the CNN was more prone to overfitting and over-predicted the majority class (HCM), leading to low AUC scores of 0.5 across minority groups and healthy volunteers. Statistical significance between predicted labels was not demonstrated for the RA-based methods but was demonstrated for the CNN method against true classification labels, as per the Kruskal-Wallis test ($P=0.0042$). Although additional training data is required to optimize all classification methods used in this work, we were able to demonstrate that RA-based methods may be more appropriate for multiclass classification of cardiomyopathy phenotypes using CMR imaging alone. Further studies and/or the addition of supplemental imaging data could permit improved multiclass classification accuracy. Work on improving classification accuracy is still underway. We intend to acquire more imaging data from minority classes and submit our work to a reputable journal in cardiac magnetic imaging science. Full details regarding the methods implemented in this contribution are available below under "Chapter 7".

1.4 Thesis Organization

This doctoral thesis is written and organized in the "integrated article" style and has been prepared according to the integrated thesis policy on the Carleton University Graduate Calendar, Section 12.4. Chapter 2 of this thesis presents the background information and the literature review applicable to the proposed methods presented in this thesis. Chapters

3 – 7 of this thesis present the methods used toward the aim of automated myocardial boundary and/or fibrosis segmentation. These chapters demonstrate the research contributions made through this doctoral work and indicate the reproduced articles, either in whole or in part, that have been published or are awaiting publication. The full bibliographical details for each article are provided with their corresponding chapter are as follows:

Chapter 3 – reproduced in part, published in Journal of Medical Physics:

Farrag, N. A., Lochbihler, A., White, J. A., & Ukwatta, E. “Evaluation of fully automated myocardial segmentation techniques in native and contrast-enhanced T1-mapping cardiovascular magnetic resonance images using fully convolutional neural networks,” *Medical Physics*, 2020, doi: 10.1002/mp.14574.

Chapter 4 –reproduced in part, presented and published in Society for Optics and Photonics (SPIE) Conference on Medical Imaging in Biomedical, Structural and Functional Imaging:

Farrag, N. A., Bhagavan, S., Sebben, D., Ruwanpura, P., White, J. A. & Ukwatta, E. “Automated myocardial segmentation of extra-cellular volume mapping cardiac magnetic resonance images using fully convolutional neural networks,” in *Proc. SPIE*, Apr. 2022, vol. 12036. doi: 10.1117/12.2626738.

Chapter 5 – reproduced in part, published in Public Library of Science (PLOS) One Journal:

Farrag, N. A., Thornhill, R. E., Prato, F. S., Skanes, A. C., Sullivan, R., Sebben, D., Butler, J., Sykes, J., Wilk, B. & Ukwatta, E., “Assessment of left atrial fibrosis progression in canines following rapid ventricular pacing using 3D late gadolinium enhanced CMR images”, *PLoS ONE* 17(7), 2022, doi: 10.1371/journal.pone.0269592.

Chapter 6 – reproduced in part, published in Journal of Medical Physics:

Farrag, N. A., Ramanan, V., Wright, G. A., & Ukwatta, E. “Effect of T1-mapping technique and diminished image resolution on quantification of infarct mass and its ability in predicting appropriate ICD therapy,” *Medical Physics*, vol. 45, no. 4, pp. 1577–1585, 2018, doi: 10.1002/mp.12840.

Chapter 7 – original work in progress (manuscript currently in preparation)

Farrag, N. A., Ayeegoundarvenkatesan, P., Samal, P., Bhagavan, S., White, J. A. & Ukwatta, E., A fully automated pipeline for classification of five HCM phenotypes using transfer-learning and radiomic analysis of extracellular volume mapping. *In progress*.

1.5 Chapter 1 References

- [1] E. B. Schelbert, L. Y. Hsu, S. A. Anderson, B. D. Mohanty, S. M. Karim, P. Kellman *et al.*, “Late gadolinium-enhancement cardiac magnetic resonance identifies postinfarction myocardial fibrosis and the border zone at the near cellular level in ex vivo rat heart,” *Circulation: Cardiovascular Imaging*, vol. 3, no. 6, pp. 743–752, 2010, doi: 10.1161/CIRCIMAGING.108.835793.
- [2] P. A. Scott, J. M. J. Morgan, N. Carroll, D. C. Murday, P. R. Roberts, C. R. Peebles *et al.*, “The extent of left ventricular scar quantified by late gadolinium enhancement MRI is associated with spontaneous ventricular arrhythmias in patients with coronary,” *Circulation: Arrhythmia and Electrophysiology*, vol. 4, no. 3, pp. 324–330, 2011, doi: 10.1161/CIRCEP.110.959544.
- [3] O. Mesubi, K. Ego-Osuala, J. Jeudy, J. Putilo, S. Synowski, A. Abutaleb *et al.*, “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*, vol. 31, no. 2, pp. 359–368, Feb. 2015, doi: 10.1007/s10554-014-0555-0.
- [4] Y. Yang, K. A. Connelly, T. Zeidan-Shwiri, Y. Lu, G. Paul, I. Roifman *et al.*, “Multi-contrast late enhancement CMR determined gray zone and papillary muscle involvement predict appropriate ICD therapy in patients with ischemic heart disease.,” *J Cardiovasc Magn Reson*, vol. 15, no. 1, p. 57, 2013, doi: 10.1186/1532-429X-15-57.
- [5] E. Dall’Armellina, V. Ferreira, and S. Neubauer, “T1 Mapping in Ischemic Heart Disease,” *Current Cardiovascular Imaging Reports*, vol. 7, no. 4, pp. 1–7, 2014, doi: 10.1007/s12410-014-9262-4.
- [6] P. Kellman and M. S. Hansen, “T1-mapping in the heart: accuracy and precision,” *Journal of Cardiovascular Magnetic Resonance*, vol. 16, no. 2, 2014, doi: 10.1186/1532-429X-16-2.
- [7] A. A. Robinson, K. Chow, and M. Salerno, “Myocardial T1 and ECV Measurement: Underlying Concepts and Technical Considerations,” *JACC: Cardiovascular Imaging*, vol. 12, no. 11P2, pp. 2332–2344, 2019, doi: 10.1016/j.jcmg.2019.06.031.
- [8] J. C. Moon, D. R. Messroghli, P. Kellman, S. K. Piechnik, M. D. Robson, M. Ugander *et al.*, “Myocardial T1 mapping and extracellular volume quantification: a Society for Cardiovascular Magnetic Resonance (SCMR) and CMR Working Group of the European Society of Cardiology consensus statement,” *Journal of Cardiovascular Magnetic Resonance*, vol. 15, no. 1, p. 92, 2013, doi: 10.1186/1532-429X-15-92.
- [9] P. Kellman, J. R. Wilson, H. Xue, M. Ugander, and A. E. Arai, “Extracellular volume fraction mapping in the myocardium, part 1: Evaluation of an automated method,” *Journal of Cardiovascular Magnetic Resonance*, vol. 14, no. 1, pp. 1–11, 2012, doi: 10.1186/1532-429X-14-63.
- [10] R. J. Perea, J. T. Ortiz-perez, and M. Sole, “T1 mapping: characterisation of myocardial interstitial space,” pp. 189–202, 2015, doi: 10.1007/s13244-014-0366-9.

- [11] O. Ronneberger, P. Fischer, and T. Brox, “U-Net: Convolutional Networks for Biomedical Image Segmentation,” pp. 1–8.
- [12] Z. Zhou, M. M. Rahman Siddiquee, N. Tajbakhsh, and J. Liang, “Unet++: A nested u-net architecture for medical image segmentation,” *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 11045 LNCS, pp. 3–11, 2018, doi: 10.1007/978-3-030-00889-5_1.
- [13] H. Cochet, R. Dubois, S. Yamashita, A. Jefairi, and B. Berte, “Relationship Between Fibrosis Detected on Late Gadolinium-Enhanced Cardiac Magnetic Resonance and Re-Entrant Activity Assessed with Electrocardiographic Imaging in Human Persistent Atrial Fibrillation Relationship Between Fibrosis Detected on Late Gadolin,” *JACC: Clinical Electrophysiology*, vol. 4, no. 1, 2018, doi: 10.1016/j.jacep.2017.07.019.
- [14] B. Burstein and S. Nattel, “Atrial Fibrosis: Mechanisms and Clinical Relevance in Atrial Fibrillation,” vol. 51, no. 8, 2008, doi: 10.1016/j.jacc.2007.09.064.
- [15] S. Nattel, B. Burstein, and D. Dobrev, “Atrial Remodelling and Atrial Fibrillation: Mechanisms and Implications,” *Circulation: Arrhythmia and Electrophysiology*, vol. 1, no. 1, pp. 62–73, 2008, doi: 10.1161/CIRCEP.107.754564.
- [16] I. V Kazbanov, K. H. W. J. Tusscher, and A. V Panfilov, “Effects of Heterogeneous Diffuse Fibrosis on Arrhythmia Dynamics and Mechanism,” *Nature Publishing Group*, no. December 2015, pp. 1–14, 2016, doi: 10.1038/srep20835.
- [17] P. G. Platonov, “Atrial fibrosis: an obligatory component of arrhythmia mechanisms in atrial fibrillation?” pp. 233–237, 2017, doi: 10.11909/j.issn.1671-5411.2017.04.008.
- [18] M. Pop *et al.*, “Quantification of fibrosis in infarcted swine hearts by *ex vivo* late gadolinium-enhancement and diffusion-weighted MRI methods,” *Physics in Medicine and Biology*, vol. 58, no. 15, pp. 5009–5028, 2013, doi: 10.1088/0031-9155/58/15/5009.
- [19] I. M. Khurram, R. Beinart, V. Zipunnikov, J. Dewire, H. Yarmohammadi, T. Sasaki *et al.*, “Magnetic resonance image intensity ratio, a normalized measure to enable interpatient comparability of left atrial fibrosis,” *Heart Rhythm*, vol. 11, no. 1, pp. 85–92, 2014, doi: 10.1016/j.hrthm.2013.10.007.
- [20] E. Monda, G. Imiero, M. Lioncino, M. Rubino, A. Cirillo, A. Fusco *et al.*, “Multimodality Imaging in Cardiomyopathies with Hypertrophic Phenotypes,” *Journal of Clinical Medicine*, vol. 11, no. 3, p. 868, Feb. 2022, doi: 10.3390/JCM11030868.
- [21] B. P. Halliday and S. K. Prasad, “The Interstitium in the Hypertrophied Heart,” *JACC Cardiovasc Imaging*, vol. 12, no. 11 Pt 2, pp. 2357–2368, Nov. 2019, doi: 10.1016/J.JCMG.2019.05.033.
- [22] W. S. Aronow, “Hypertension and left ventricular hypertrophy”, doi: 10.21037/atm.2017.06.14.
- [23] R. Sankaranarayanan, E. J. Fleming, and C. J. Garratt, “Mimics of Hypertrophic Cardiomyopathy – Diagnostic Clues to Aid Early Identification of Phenocopies,” *Arrhythmia & Electrophysiology Review*, vol. 2, no. 1, p. 36, 2013, doi: 10.15420/AER.2013.2.1.36.
- [24] Z. Yousef, P. M. Elliott, F. Cecchi, B. Escoubet, A. Linhart, L. Monserrat *et al.*, “Left ventricular hypertrophy in Fabry disease: a practical approach to diagnosis,” *European Heart Journal*, vol. 34, no. 11, p. 802, Mar. 2013, doi: 10.1093/EURHEARTJ/EHS166.

Chapter 2: Background and Literature Review

2.1 Clinical Problem: Myocardial Fibrosis

Cardiovascular disease (CVD) is an umbrella-term used to describe a group of disorders that impact the heart and cardiovascular system, which are responsible for up to 31% of deaths worldwide [1]. Heart disease, a component of CVD, is the second leading cause of death in Canada. About 1 in 12 Canadian adults over the age of 20 live with heart disease and every hour, about 12 Canadian adults die of this disease (Statistics Canada, 2012-2013 Public Health Agency of Canada's Canadian Chronic Disease Surveillance System). Myocardial fibrosis (MF) is a common feature of nearly all forms of heart disease [2], characterized by excessive deposition of collagen (i.e., scar tissue) and expansion of the extracellular volume (ECV) in the heart muscle (i.e., the myocardium). This phenomenon is often referred to as “cardiac remodeling”, wherein a variety of quantitative and qualitative changes occur within the interstitial myocardial collagen network that lead to cardiac stiffness and dysfunction [2]. As a result, several complications may occur, including heart failure, arrhythmia, and sudden cardiac death (SCD), giving MF a 5-year mortality rate of up to 20% [3]. Thus, MF has implication in preceding cardiac morbidity and mortality and may therefore be an appropriate biomarker to predict cardiac outcomes. MF of the left atrium (LA) is correlated with atrial fibrillation (AF), a condition characterized by a quivering, irregular heartbeat that is associated with increased risk of secondary CVDs such as stroke and heart failure [4]. However, quantification of left atrial MF via non-invasive imaging is challenging due to low image resolution relative to the LA wall, which is merely 1 mm in thickness [5]. MF of the left ventricular (LV) myocardium has been found to correlate with incidence of both atrial and ventricular arrhythmias and

adverse cardiac outcomes [6], [7]. Moreover, the LV is the thickest-walled heart chamber in the body, making left ventricular MF naturally more conspicuous on non-invasive imaging compared to the LA [8]. Thus, MF of the LV is most predominantly assessed for diagnostics and risk stratification due to ease of visualization and quantification using non-invasive imaging and/or histopathological imaging. As such, MF measurements from the LV may provide an improved means of risk stratification for heart failure, arrhythmia and SCD compared to the left ventricular ejection fraction (LVEF), the current risk metric, which is considered outdated and unreliable.

2.1.1 Cardiac Anatomy, Structure and Function

The mammalian heart is a mechanical and electrical organ that is positioned obliquely in the chest cavity, typically between the level of the third costal cartilage and the diaphragm. The great veins, including the superior and inferior vena cava, aortic arch and pulmonary artery are situated above the superior (i.e., the upper-most) surface of the heart, called the heart base. The inferior (lower-most) tip of the heart, called the apex, typically lies to the left side. The right side of the heart is tipped anteriorly while the left side is tipped posteriorly [9]. Figure 2-1 demonstrates the typical position of the heart in the human chest cavity.

The mammalian heart is made up of a several layers of muscle cells, namely the endocardium, myocardium, epicardium and pericardium, demonstrated in Figure 2-2. The innermost muscular layer, the endocardium, lines the inner wall of the heart. The myocardium sits above the endocardium and is the thickest layer of the heart, chiefly responsible for the contraction motion that occurs during the heart's beating cycle. The outermost muscular layer is the epicardium, which serves as an additional layer of

protection and contained the coronary blood vessels that oxygenate the tissues of the heart. The pericardium acts as an additional protection barrier and lubricant to reduce friction between the heart and surrounding tissues[10], [11].

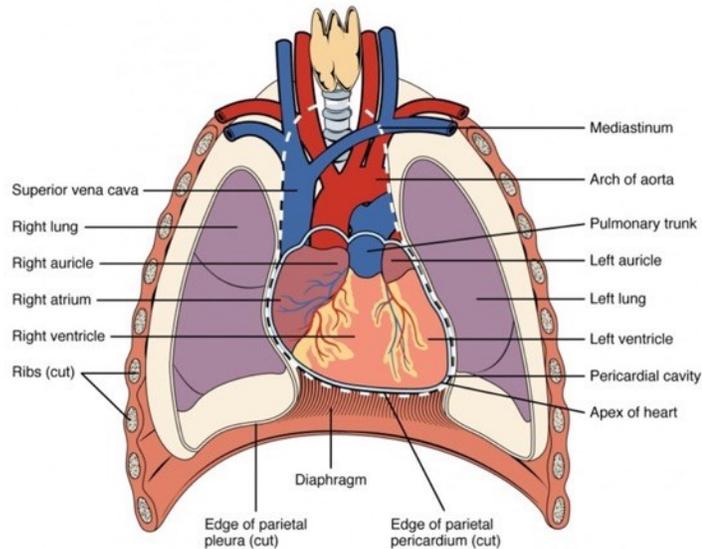


Figure 2-1 Position of the human heart in the chest cavity in anterior (frontal) view. The heart base is situated superiorly (i.e., top-most) while the heart apex is situated inferiorly (bottom-most) and tipped left. This image was provided freely <https://courses.lumenlearning.com/ap2/chapter/heart-anatomy/>.

In each of the aforementioned muscular layers of the heart, an extracellular matrix (ECM) exists, which is a molecular network of macromolecules that is interleaved between the cardiac muscle cells, called myocytes. The ECM consists of macromolecules such as collagen, elastic fibers and glycoproteins. This network is essential in providing biochemical and structural support to the cardiac tissue [12].

Structurally, the heart is made up of four chambers, which are the right and left atria and the right and left ventricles. The right atrium receives deoxygenated blood from the body, which is pumped into the right ventricle via the tricuspid valve and onward to the lungs via the pulmonary artery for oxygenation. Oxygenated blood returns to the left atrium via the pulmonary veins and is pumped through the mitral valve to the left ventricle, the

largest and strongest heart chamber, which can then be pumped through the aortic valve to the aorta, thereby providing oxygen to all the body's organs/tissues through the systemic circulation system. This process is repeated for every heartbeat. The pumping of the blood through the each of the cardiac valves, arteries and veins is a function of the myocardium, facilitated and regulated by the heart's electrical conduction system. Each heartbeat is stimulated by the sinoatrial (SA) node, the main pacemaker of the heart, which releases stimuli causing the left and right atria to contract, thereby pumping blood into the respective ventricles. Following this, the stimulus from the SA node eventually reaches the atrioventricular (AV) node, which briefly slows down electrical signal to allow ventricles time to receive blood from the atria. After the electrical signal leaves the AV node, it travels to a bundle of conduction cells called the Bundle of His, which branch and divide into a system of conducting fibers in the left and right ventricles, causing ventricular contraction.

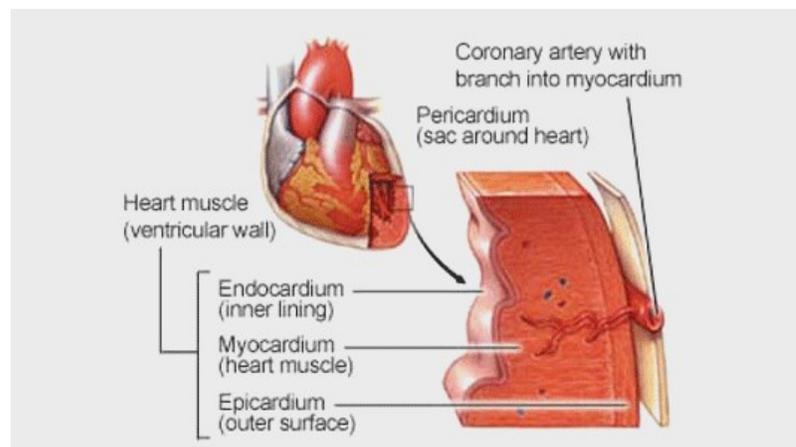


Figure 2-2 Layers of the heart muscle

The layers of the heart muscle are shown from innermost (left) to outermost (right). The endocardium is the innermost layer, followed by myocardium, epicardium and pericardium. This image was downloaded freely from '<http://www.vascularinnovations.com/heart-basics.html>'.

2.1.2 Myocardial Fibrosis Subtypes

MF is classified into two main subtypes, replacement MF and interstitial MF, as demonstrated in Figure 2-3. Replacement MF is also referred to as scarring fibrosis,

wherein a dense, focal region of the myocardium endures significant tissue damage or necrosis (i.e., tissue death). This phenomenon most frequently occurs after the patient experiences an ischemic event, i.e., myocardial infarction (MI) [13]. MI is most typical caused by coronary artery disease (CAD), wherein a coronary artery that provides blood to the myocardium is blocked due to buildup of fatty plaque deposits. The obstruction of blood flow in the artery prevents oxygen and essential nutrients within the blood from reaching the myocardial tissue, eventually causing myocardial tissue death in the region of the blocked artery. The pathophysiological mechanism leading to MI-related fibrosis is acute, wherein an intense inflammatory response is induced, leading to degradation of the normal interstitial matrix and rapid replacement of dead cells with collagen type-I scar [1]. This results in a dense, localized region of irreversible scarring. Aside from MI-related ischemic injury, replacement MF may also occur from other cardiac diseases, such as hypertrophic cardiomyopathy, myocarditis or sarcoidosis [13].

Interstitial, often referred to as “diffuse” MF, differs from replacement MF in its distribution within the myocardium and its pathophysiological mechanism. Interstitial MF is characterized by a dispersed deposition of collagen and/or proteins within the extracellular matrix, occurring in the absence of cardiac necrosis [14]. The mechanism leading to interstitial MF tends to be progressive rather than acute as in the case of replacement MF. To further differentiate from replacement MF, interstitial MF may potentially be reversible, contingent on early detection and appropriate treatment measures [12]–[14].

Interstitial MF is further sub-divided into two types: reactive or infiltrative. In reactive interstitial MF, myofibroblasts are stimulated to increase collagen production, and

collagen is deposited in the extracellular matrix. This phenomenon is often present in common conditions such as aging, hypertension, diabetes mellitus or dilated cardiomyopathy [12], [13]. Infiltrative interstitial MF is caused by progressive deposition of insoluble proteins or glycosphingolipids in the interstitial myocardial space. This condition is much rarer and may be seen in conditions such as amyloidosis or Anderson-Fabry disease [12], [13]. Both reactive and infiltrative interstitial MF can eventually lead to cardiomyocyte apoptosis (i.e., programmed cell death of the myocardial cells), which ultimately results in irreversible replacement MF [15]–[17]. Thus, early detection and treatment intervention is of utmost importance in order prevent further deterioration of interstitial MF to irreversible replacement MF.

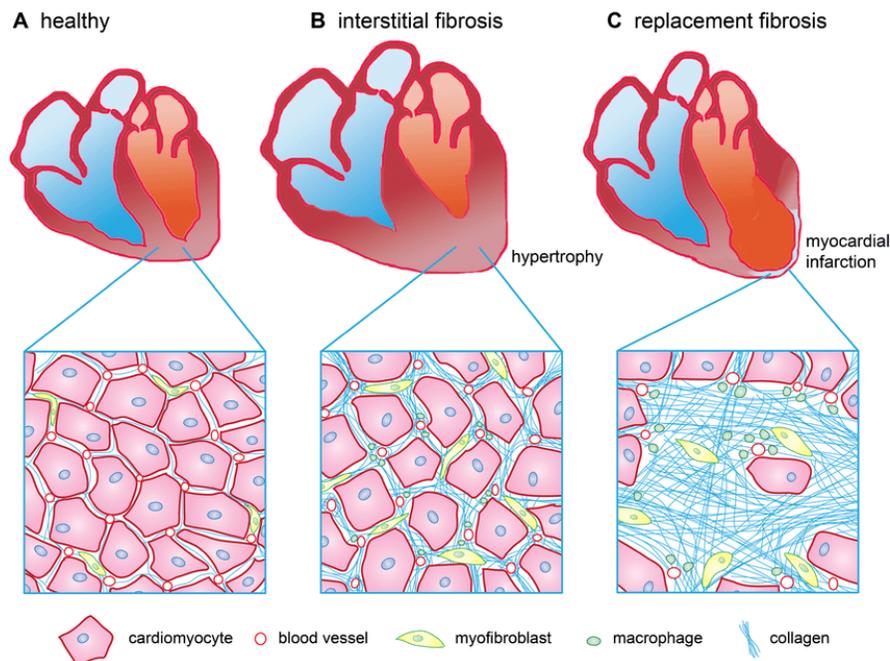


Figure 2-3 Cellular composition of myocardium in healthy vs. diseased states
 Healthy myocardium (A), interstitial myocardial fibrosis due to hypertrophy (B) and replacement myocardial fibrosis due to MI (C) are demonstrated from left to right. Collagen (scar) is shown in blue color between the cardiomyocyte cells (red) (i.e., within the ECM). Myofibroblasts, responsible for collagen synthesis and deposition, are demonstrated within the ECM in green. In interstitial MF (B), scar is distributed diffusely and sporadically within the ECM. In replacement MF, scar is condensed and distributed focally to one region. This image is referenced from de Graaf et al.'s 2014 manuscript [18], referenced in the bibliography.

2.1.3 Assessment of Myocardial Fibrosis

MF can be assessed in a number of ways. Until recently, the gold standard method for MF assessment has been endomyocardial tissue biopsy, an invasive method wherein a small sample is taken from the right ventricular (RV) side of the interventricular septum that separates the RV from the LV [12]. In this approach, a transthoracic needle is inserted into either the femoral or internal jugular vein and a sheath is introduced into the vein such that a biopsy forceps can be advanced to the region of interest [19], [20]. This procedure is facilitated under x-ray fluoroscopic guidance. This sample is typical $<1\text{mm}^3$ in size and is assessed via histopathological imaging using Masson trichome staining, allowing for absolute quantitative assessment of the collagen volume within the sampled region [12]. Although the endomyocardial biopsy technique is accurate, it also has several downsides. Firstly, the procedure itself is quite invasive and therefore carries a risk of complications, including carotid puncture, perforation of coronary artery and/or myocardium, prolonged bleeding, vasovagal reactions, arrhythmic events or conduction abnormalities [20]. Moreover, this method does not allow for assessment of the entire LV nor the entire heart, and furthermore carries the risk of sampling error, which restricts accuracy of results [12].

Recently, non-invasive diagnostic imaging (DI) techniques have been established as safer alternatives to endomyocardial biopsy for MF assessment. Cardiac magnetic resonance (CMR) and computed tomography (CT) are the two most commonly used DI methods for diagnosis, assessment and quantification of MF.

CT imaging is an ionizing radiation modality that has been increasingly explored as a clinical tool for evaluation of CAD, allowing for evaluation of coronary artery plaques and assessment of artery stenosis [12]. This allows for precise localization of the

intervention area for treatment. Patients with CAD also frequently have large, focal regions of replacement MF, identifiable on CT scans acquired using an extravascular iodinated contrast agent, which lingers and enhances in extracellular regions where scar is present thereby allowing clinicians to discern MF regions from healthy myocardium [15], [21]. However, from a tissue characterization standpoint, CT imaging is less appropriate, due to low contrast resolution and poor signal-to-noise ratio [12], [14], [21]. This is particularly problematic for the assessment of interstitial MF, which is diffusely distributed in the myocardium. In addition, this technique requires the use of ionizing radiation, which is harmful and particularly problematic for patients requiring serial imaging. Thus, CT imaging has the potential to assess MF, however, widespread application of this technique to clinical practice is unlikely, as the drawbacks to this technique may be addressed by using CMR [12], [14], [21].

CMR imaging has recently been increasingly explored for the assessment of MF. This technique allows for non-invasive, comprehensive assessment of the anatomical structure and function of the heart with high resolution, high image contrast, and without the use of ionizing radiation [22]–[24]. There are three main techniques used to evaluate MF in CMR: late-gadolinium enhancement (LGE), T1-mapping CMR and ECV-mapping CMR. In addition, the cinematic (CINE) CMR imaging technique is used for assessment of cardiac function. While the CINE technique does not explicitly allow for MF tissue characterization, this method may assist in the diagnosis and treatment planning for patients with MF by allowing measurement of various functional parameters that are altered by various cardiac diseases. CMR imaging methods used in diagnosis of MF are discussed in detail in Section 2.2.

2.1.4 Conditions Related to Myocardial Fibrosis

MF is a common end point of a number of pathological processes that affect the myocardium and is therefore a prominent feature of several forms of myocardial disease. Myocardial infarction (MI) is one of the most common conditions that causes MF, however, several other conditions can bring about MF, including hypertrophic and dilated cardiomyopathies (HCM and DCM, respectively), aortic stenosis, myocarditis, sarcoidosis, hypertensive heart disease, among many others [1], [14], [25].

Furthermore, MF is also precursor of heart failure and cardiac conduction disorders, including ventricular and atrial arrhythmias and sudden cardiac death (SCD). The changes in the ECM associated with the pathological mechanism of MF may create pathways for abnormal cardiac conduction that can cause or perpetuate arrhythmic events. Thus, MF has been proposed as a risk-stratification method for arrhythmic cardiac events and SCD [26]–[30]. In patients with secondary arrhythmic conditions such as AF, MF segmentation is an essential step toward creating computational models of the heart for treatment with catheter ablation using electroanatomical voltage mapping systems (EAMs) [31], [32].

2.2 Non-invasive Imaging of Myocardial Fibrosis Using Cardiac Magnetic Resonance

Myocardial tissue characterization and MF quantification is permissible using cardiac magnetic resonance (CMR) imaging. CMR is a non-invasive imaging technique that uses a strong magnetic field and radiofrequency (RF) waves to generate diagnostic images with superior contrast and temporal resolution [33].

Magnetic resonance imaging (MRI) of all body regions is based on the physics of nuclear magnetic resonance (NMR) of hydrogen protons, which are prominent components

of water, fat and other body tissues [34]. Hydrogen atoms have magnetic properties due to their electron motion, which acts as a current loop. Under the influence of a strong, static magnetic field (i.e., the field of the magnetic resonance imaging machine, typically 1.5 or 3T), hydrogen protons in the body experience a bulk magnetic effect and will align parallel with the static magnetic field direction in the longitudinal (z) plane. Influenced by radiofrequency waves tuned to the hydrogen protons' precession frequency (i.e., the Larmour frequency), the proton magnetization vector will excite and "flip" from the z-plane into the transverse (x-y) plane. Excited hydrogen protons will eventually relax back to equilibrium in the z-plane in a time-varying fashion, dependent on each tissue's composition of hydrogen protons [34], [35]. During this relaxation process, a magnetic resonance signal is induced and collected in a nearby RF receiver coil to create the image. The time with which the excited magnetization vector takes to recover back to 63% of its maximum value in the longitudinal (z) plane is referred to as T1 recovery. The time by which the transverse (x-y) magnetization component decays back to approximately 37% of its initial value is referred to as T2 relaxation [34]–[36]. Magnetic resonance images can be created on the basis of T1 recovery (called T1-weighting), T2 relaxation (called T2-weighting), the density of hydrogen protons (called proton density weighting), among many other weightings that allow clinicians to visualize multiple variations of contrast. The differences in the respective recovery or relaxation times between tissues contributes to image contrast [36].

2.2.1 Late Gadolinium Enhancement (LGE)

LGE CMR has been established as the preferred technique for diagnosis and quantification of MF. This technique is based on the concept that the expanded

extracellular space in fibrotic regions allows for increased pooling of fluids and other molecules compared to the normal, unexpanded ECM belonging to healthy myocardial tissue. Moreover, the molecules that accumulate into expanded ECM areas take longer to wash out, due to decreased capillary density and efficiency in the region [13].

To acquire LGE CMR images, a cardiac gated T1-weighted gradient echo sequence (GRE) is used after application of an inversion pulse, which nulls the signal of healthy myocardium. Prior to image acquisition, an intravenous (IV) gadolinium-based contrast agent (GBCA) is introduced into the patient's body, which acts to facilitate T1-shortening in the regions at which the gadolinium resides. After being introduced into the body intravenously, the GBCA will flow through the patient's circulation system and eventually reach the heart, at which point it will infiltrate the myocardial tissue and distribute into the extracellular space. Due to the ECV expansion and reduced capillary function that is characteristic of MF regions, the GBCA will concentrate into the MF areas and wash out at a slower rate as compared to surrounding healthy tissue. This results in a visual difference in signal in fibrotic regions, which appear brighter due to the shorter T1-relaxation time associated with the GBCA. Moreover, the nulling of normal myocardial signal allows for increased contrast between the regions, wherein healthy myocardium appears darker and MF appears brighter [13], [14], [37]. This phenomenon is demonstrated in Figure 2-4.

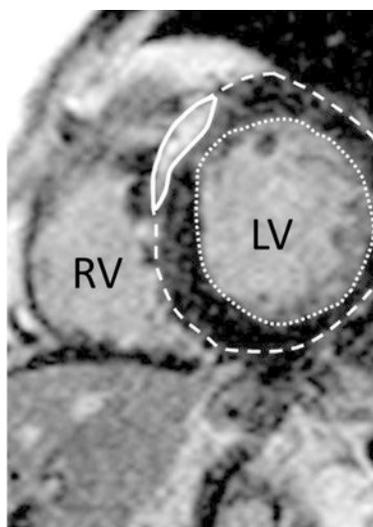


Figure 2-4 Example LGE-CMR image demonstrating enhanced fibrosis

Late gadolinium enhanced (LGE) CMR image in short-axis view of a patient with hypertrophic cardiomyopathy. Left ventricle (LV) and right ventricle (RV) blood pool chambers are labeled. Normal LV myocardial tissue is demonstrated as the dark, oval-shaped region surrounding the LV chamber, delineated by the dotted lines. A large region of MF is demonstrated as the bright region within the normal myocardial segmentation, delineated with a solid line. This image is referenced from Fluchter et al.'s 2010 manuscript [38], cited in the bibliography.

While LGE CMR imaging represents a major milestone in diagnosis and management of MF, several drawbacks are associated with this technique. From a technical standpoint, LGE CMR is associated with many variables that make reproducibility challenging. For example, healthy myocardial null point is chosen manually at the time of acquisition, which is subject to operator error [14], [39]. Likewise, image contrast is subject to timing of gadolinium washout, which varies between patients and even in sequential imaging of the same patient due to metabolism differences. Similarly, since the image acquisition is gated by the patient's heart rate, some acquisitions may permit more gadolinium washout, diminishing the contrast between MF and healthy surrounding regions [12].

LGE CMR has several drawbacks from a clinical standpoint, as well. While stark, regional contrast differences may be prominent in patients with replacement MF, visualization and quantification of interstitial fibrosis is more problematic in LGE-CMR

[40]. Sporadic distribution of MF may make it impossible for the user to define a region of clearly unaffected healthy myocardial tissue to null in the LGE CMR acquisition, and as such, these images may demonstrate nulled signal in areas of fibrosis, which would obscure results [41], [42]. Thus, diffuse/interstitial MF may go undetected on LGE CMR [12], [41]–[44]. Histopathological assessment of diffuse fibrosis has found LGE to correlate poorly with areas of increased collagen volume computed using endomyocardial biopsy [42], [45].

Finally, the LGE CMR technique presents challenges in the quantification of MF due to subjectivity of enhancement measurements. As mentioned, the enhancement of myocardial signal is expressed on an arbitrary scale, posing challenges when it comes to comparison MF measurements in serial imaging. Moreover, the threshold used to define healthy tissue versus MF is also subjective. Several image thresholding techniques exist to quantify MF in LGE CMR images (discussed in Section 2.3.2), wherein a threshold is arbitrarily picked by the user during post-processing [12]. This poses an issue when it comes to comparison of MF metrics from one imaging study to the next. Lastly, in LGE-CMR images, direct MF quantification is not possible, since pixel intensities are a function of GBCA retention and are not directly representative of a quantitative parameter [46], [47].

The LGE CMR method remains heavily relied on for clinical evaluation of MF; however, the aforementioned drawbacks have led to exploration of alternative imaging approaches to better assess and quantify this disease.

2.2.2 T1 Mapping

T1 mapping is a novel CMR imaging technique that addresses some of the drawbacks of LGE-CMR by allowing for quantitative characterization of myocardial

tissue. T1 mapping CMR differs from T1-weighted CMR in that each pixel within the image is directly representative of the corresponding tissue's T1-relaxation parameter, measured in milliseconds, which is intrinsic and specific to every tissue in the body [48]–[50]. The principle of T1-mapping is based on the principle of T1 magnetization recovery, a fundamental concept of magnetic resonance imaging. In the presence of a large magnetic field, tissues in the body experience a small magnetic moment due to hydrogen content in the tissues [34], [35]. Hydrogen atoms have magnetic properties due to their electron motion, which acts as a current loop. The magnetic moment of the body tissues aligns with a large static magnetic field (i.e., the field of the magnetic resonance imaging machine, typically 1.5 or 3 Tesla) in the z-plane. Following application of a radiofrequency (RF) pulse, the body tissues' magnetization vectors are rotated into the x-y plane [34]–[36]. The magnetic moments will eventually relax back to equilibrium in the z-plane, the speed of which is called the “T1 relaxation” or “longitudinal recovery” time and is dependent on the tissue type. This longitudinal recovery speed is modeled by an exponential function [50], [51]. By acquiring data at varying time points after RF excitation, it is possible to model the T1 recovery function by measuring T1 vs. time, as demonstrated in Figure 2-5. Mathematical curve fitting can then be used to define the T1-relaxation parameter for each pixel, the values of which are then used to reconstruct the final T1 map [13], [52]. Each raw image pixel is sampled at every time point to model an exponential function, which can be fit to determine a relaxation time parameter. This parameter is the pixel intensity of the resultant T1-map.

The T1-mapping acquisition process is therefore based on three steps, as follows:
(1) a T1 magnetization preparation pulse of either 90° or 180° (i.e., saturation recovery

(SR) or inversion recovery (IR), respectively) is applied, (2) a single image is readout after a variable delay time and (3) the process is repeated several times during T1 recovery to generate multiple raw, T1-weighted images [51]–[53]. The T1-map is a reconstructed image that is created by sampling all of the raw images generated by the T1-mapping pulse sequence and computing a T1-recovery parameter for each pixel based on mathematical curve fitting [13], [41], [50], [54]. Motion compensation algorithms are applied when reconstructing the T1-map to account for respiratory and/or cardiac motion differences across each raw image. Each raw image pixel is sampled at every time point to model an exponential function, which can be fit to determine a relaxation time parameter. This parameter is the pixel intensity of the resultant T1-map, expressed in milliseconds (ms). T1-maps are two-dimensional images that are often generated using color-coding, however, pixel intensities may also be mapped to grayscale.

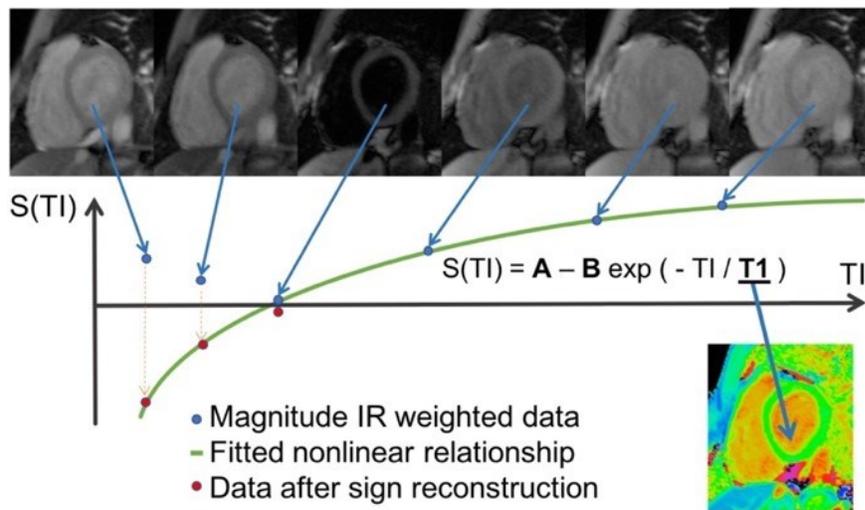


Figure 2-5 Curve fitting method used to create T1-maps

Creation of a T1-map using six T1-weighted images (top row) acquired at varying inversion times (TIs). Signal intensity ($S(TI)$) of a particular pixel is graphed on the y-axis, against TI on the x-axis. The signal intensity models an exponential function vs. time, as demonstrated by the green curve. The respective pixel's T1 parameter can be computed using the equation $S(TI) = A - B e^{-\frac{t}{T1}}$, where A and B are fitting parameters related to the equilibrium magnetization and type of preparation pulse, t is inversion time (TI) or time after the saturation pulse, and $T1$ is the T1 relaxation time of the selected pixel. This process is repeated for all pixels to create the final T1-map (demonstrated in color scale in bottom right). This image is referenced from Werys et al.'s 2020 manuscript, referenced in the bibliography [55].

There are several techniques that can be used to acquire T1-maps. Most methods are acquired during cardiac gating in the diastole phase while the patient holds their breath, in order to reduce motion artifact. Several repeated breath holds are required for acquisition of a single pulse sequence. The most commonly used T1 mapping technique was established by Messroghli et al. called the Modified Look-Locker Inversion-recovery (MOLLI) method [56]. This IR-based technique provides high resolution T1-maps in a single breath hold over 17 heartbeats. A variation of this technique is called the shortened-Modified Look-Locker Inversion-recovery (shMOLLI) technique, which uses the same methodology with reduced breath hold times, acquiring high-resolution images over 9 heartbeats [57]. Other T1-mapping techniques based on SR have been proposed, namely, saturation recovery single-shot acquisition (SASHA) and saturation pulse prepared heart rate independent inversion recovery (SAPPHIRE), but are less commonly used in clinical practice due to poorer precision as compared to MOLLI and shMOLLI sequences [58].

T1-mapping can be performed with or without the use of GBCAs, called native T1-mapping and contrast-enhanced/post-contrast T1-mapping, respectively. In either case, tissue characterization is permitted by direct differentiation of the T1-relaxation parameter, which is altered in certain disease processes. Native T1 relaxation times increase in tissue regions with myocyte injury and/or expansion of the ECV, such is the case of MF [23], [54], [59], [60]. T1-mapping following the administration of GBCAs leads to a proportional reduction in T1-relaxation relative to the GBCA's retention within the ECV, and thus, reduced T1 relaxation times in fibrotic regions [23], [54], [61], [62]. The signal coming from native T1-maps is predominantly reflective of the signal from myocytes and best demonstrate areas of replacement MF [61], [63]. Post-contrast T1-values are

predominantly representative of signal coming from the ECV space [64]. Hence, native T1-maps may be used to characterize replacement fibrosis, while post-contrast T1-maps may better characterize interstitial fibrosis [65]. Native and post-contrast T1-maps are created in grayscale but are often displayed using color mapping in order to better differentiate myocardial contrast. Figure 2-6 demonstrates an example of native and post-contrast T1-maps in grayscale and using color mapping.

Myocardial T1 values are dependent on magnetic field strength, scanner manufacturer, the T1-mapping pulse sequence and other acquisition parameters and post-processing [60]. Thus, various magnetic resonance manufacturers are making efforts to standardize T1 measurements between different vendors [51]. Several studies have aimed at reporting ranges for normal, healthy myocardial T1 values among different magnetic resonance vendors, field strengths and using different T1-mapping pulse sequences. Healthy human native (i.e., non-contrast) myocardial T1 values have been recorded between 950 to 1050 ms using the MOLLI sequence across Philips, Siemens and General Electric vendors at 1.5 Tesla field strength [60], [66]– [69]. T1 is inherently increased at higher field strengths, thus healthy native myocardial values of 1100 to 1300 have been reported in humans at 3 T field strength across the aforementioned vendors [59], [60], [66], [70]– [72]. Post-contrast T1 values vary on the time of imaging acquisition after injection of the GBCA, contrast agent dosing and renal function [48], but have been reported in the range of 402 – 504 ms between 10 to 20 minutes after introduction of the GBCA [59], [73]. Due to confounding factors effecting the measured T1 value after injection of a GBCA, T1 values reported from contrast enhanced T1 maps may have higher measures of variation and may therefore be less robust compared to native T1 and ECV [73].

In patients with various forms of cardiac disease involving fibrosis, myocardial T1 values are altered. Global (i.e., overall) myocardial T1 values in patients with HCM have been recorded in the range of 1230 – 1254 ms in native images and 300 – 430 ms in post-contrast images at 3 T. Global myocardial T1 values in patients with DCM have been recorded at 1264 – 1364 ms in native T1 maps and 444 – 549 ms in contrast enhanced T1 maps at 3T [59]. Although several studies have demonstrated characteristic changes in myocardial T1 in different forms of cardiac disease, further studies are required to better establish accepted ranges for normal and diseased myocardium among different field strengths.

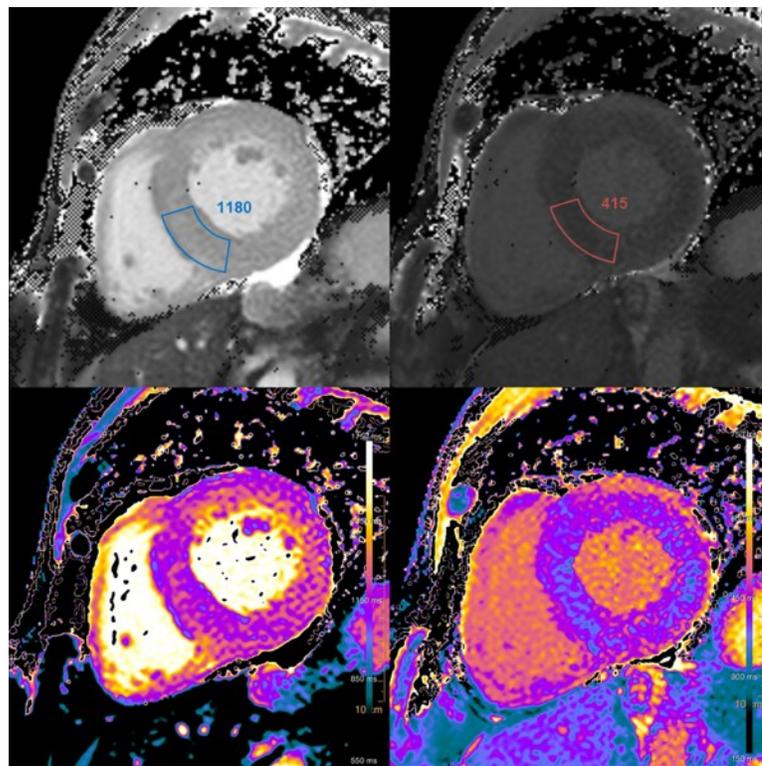


Figure 2-6 Example T1-maps demonstrating expansion of ECV
Native (left) and post-contrast (right) T1-maps acquired using the Look-Locker pulse sequence are shown in grayscale (top row) and color mapping (bottom row). The images shown are of a patient with amyloidosis, a condition wherein an insoluble protein called amyloid is deposited in excess in the extracellular space of the myocardium. This condition is associated with an elevation of native T1 values and decrease in post-contrast T1 values relative to normal myocardium (see delineated regions in the grayscale native and post-contrast T1-maps). This image was downloaded freely by the courtesy of Dr Joachim Feger via Radiopaedia.org, rID: 74008 (<https://radiopaedia.org/articles/myocardial-mapping>).

2.2.3 ECV Mapping

ECV mapping is a quantitative imaging technique wherein pixel intensities are reflective of an estimation of each tissue's ECV. Changes to the myocardial ECV is a diagnostic marker of myocardial tissue remodeling [49], [50], [54], [61]. Expansion of the ECV due to collagen deposition (in the absence of amyloid or edema) is a hallmark indicator of MF due to deposition of collagen [49], [50], [60]. Therefore, ECV mapping may be particularly useful in quantitation of MF in cardiomyopathies exhibiting interstitial fibrosis or cardiac diseases that are not associated with infarction.

The ECV mapping technique is based on measuring the ratio of change in T1 between myocardium and blood before and after the administration of a GBCA and is expressed as the percent of tissue comprising the extracellular space. This technique offers some unique advantages over other CMR techniques since it is a physiologically intuitive unit of measurement that is independent of magnetic resonance field [52], [61], [74]. Moreover, the ECV measurement is unaffected by confounding factors such as obesity, renal function, time of image acquisition post-injection of CBGA, and GBCA characteristics such as water exchange rate, which effect both the post-contrast T1 values measured by contrast-enhanced T1-mapping and signal intensity measured by LGE CMR [48], [61], [73], [74]. Thus, ECV mapping provides less variability between different scanners at different field strengths and could be of significant benefit for monitoring interval changes and/or assessing treatment effect in patients with cardiac disease [61], [73], [74] and has been shown to correlate highly with histological imaging of human myocardium [52], [62], [75].

ECV maps are created using native and post-contrast T1-mapping data and are estimated using the following equation,

$$ECV = (1 - Hct) \frac{\frac{1}{\text{post contrast T1 myocardium}} \frac{1}{\text{native T1 myocardium}}}{\frac{1}{\text{post contrast T1 blood}} \frac{1}{\text{native T1 blood}}} \quad \text{Equation 2-1}$$

where *Hct* is blood hematocrit, a measurement of the cellular fraction of blood [54], [61].

Normal ranges for hematocrit are 40 – 54% in men and 36 – 48% in women [76], [77].

ECV maps are typically generated in color-scale, as demonstrated in Figure 2-7.

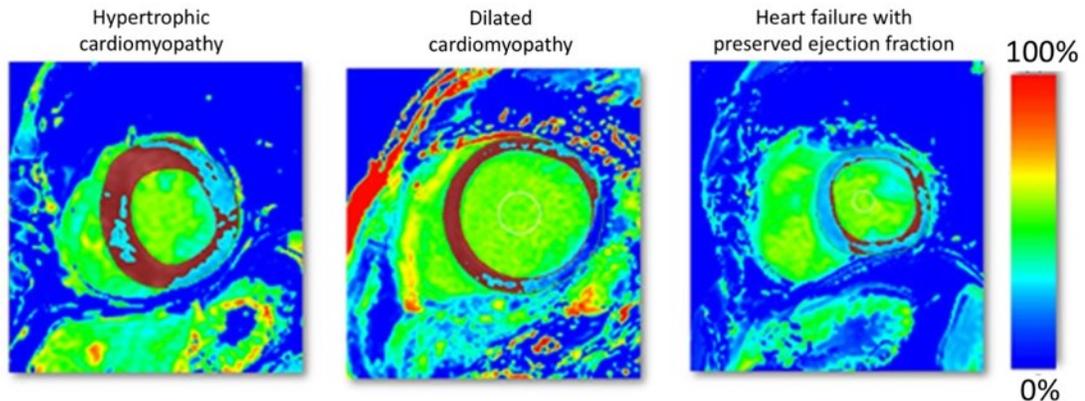


Figure 2-7 ECV maps of patients with different cardiomyopathies

ECV maps were created using MOLLI T1 data at 1.5 T for patients with HCM, DCM and heart failure with preserved ejection fraction (HFPEF). Images are demonstrated in colour scale ranging from 0 – 100%, as demonstrated by the color bar on the right. In these particular ECV maps, red areas demonstrate regions of ECV > 30%. This image is referenced from Haaf et al.'s 2016 manuscript [74], cited in the bibliography.

ECV values are increased in myocardial regions that exhibit MF due to collagen deposition and expansion that is inherent to this disease, as demonstrated in Figure 2-8.

Normal myocardial ECV ranges have been reported around the range of 23 – 28% at both 1.5 T and 3T [60], [61], [66], [68]. Significant increases in ECV are seen in diseases such

as HCM, DCM, amyloid and MI [53], [60], [62]. Several studies have reported ECV ranges for patients with the aforementioned cardiac diseases, however, a standard diagnostic

criterion is yet to be concluded. Currently, most clinical imaging centers use ECV > 28 or

30% as a diagnostic marker for fibrosis [61], [73], [75], [79]. However, the diagnostic range of ECV values is variable between diffuse and replacement fibrosis types. ECV values >40% have been reported in areas of dense replacement while interstitial fibrosis and myocardial areas at risk are reported with ECV values > 30% [61], [79]. Thus, further studies are required to establish standardization of ECV ranges for diagnostic characterization of various cardiac diseases.

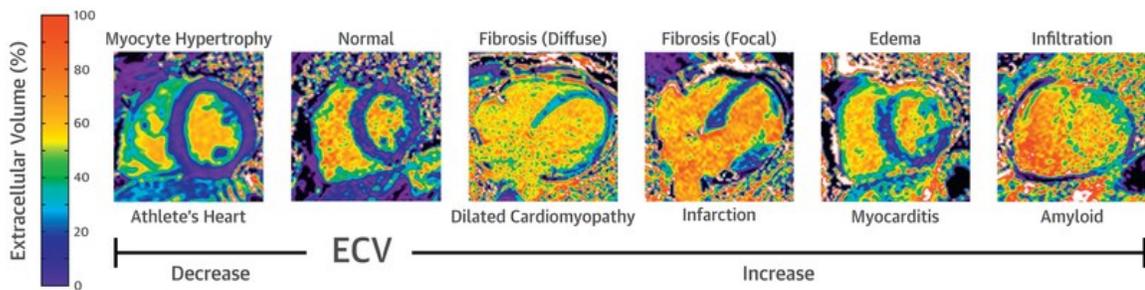


Figure 2-8 Cardiovascular diseases associated with changes in ECV

Various diseases associated with altered ECV values are demonstrated. Notably, both interstitial (diffuse) and replacement (focal) MF are associated with increased ECV. This image is referenced from Robison et al.'s (2019) manuscript [80], cited in the bibliography.

2.2.4 Cinematic Cardiac Magnetic Resonance

Cinematic (CINE) CMR is used to assess the condition and function of the heart in patients with various cardiac diseases and is included by default in virtually any CMR imaging examination [39], [81]. The CINE technique is unique in that it allows clinicians to assess the functionality of the heart by acquiring “movies” demonstrating the heart’s motion. CINE images are obtained by repeatedly imaging the area of interest at a single slice location throughout the course of the cardiac beating cycle, which is recorded during CINE imaging using MRI-compatible electrocardiogram (ECG) technology. Typically, 10 – 30 image frames are acquired at each slice location in synchronization with ECG data using steady-state free precession (SSFP) gradient echo (GRE) imaging sequences with very short echo times (TEs). The short TE times associated with SSFP GRE imaging allows

for multiple lines of k-space (i.e., multiple frames) to be acquired at each slice location during a single heartbeat [22], [39], [82]. When multiple lines of k-space are collected during each heartbeat (called the R-R interval), the total imaging time decreases by a factor of the number of views per segment (NVS) [83]. Figure 2-9 demonstrates example of image frames acquired in the CINE pulse sequence in short axis view.

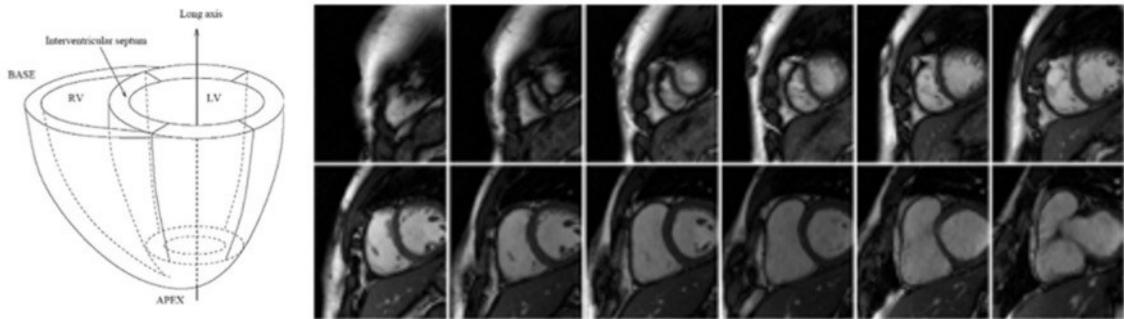


Figure 2-9 CINE images and geometric axes of the ventricles

A typical geometric model of the long axis of the LV and RV is shown on the left-most side. The right-most side demonstrates a sequence of twelve CINE images, acquired in short axis (i.e., images are acquired perpendicular to the long axis geometry that is shown) across slice locations from the apex (top left) to the base (bottom right) of the LV. Only a single frame per image location is demonstrated, however, typically between 10-30 frames are acquired per slice location across the cardiac beating cycle in a full CINE imaging acquisition. This image was downloaded from Petitjean, C. & Dacher, J-N.'s article (2010) and is referenced in the bibliography [79].

CINE imaging is typically used to evaluate right and left ventricular wall motion, contractile function, volume and mass measurements, blood flow as well as the viability of cardiac valves [22], [39], [81], [85]. Typically, several CINE sequences are acquired at various cardiac imaging planes in order to get an overall assessment of the aforementioned parameters. Common cardiac imaging planes include short axis, horizontal long axis (four chamber view) and vertical long axis (two chamber view), which are prescribed along a line extending from the apex of the heart to center of the mitral valve [86]. Valve viability and blood flow are typically assessed qualitatively through visual assessment of CINE images across various imaging planes by a cardiologist or a radiologist specializing in

cardiac anatomy. Volume and mass parameters, such as ejection fraction (EF), left/right ventricular mass, stroke volume (SV), end systolic volume and end diastolic volume (ESV and EDV, respectively) are measured quantitatively by measuring the differences in the size of endocardial and epicardial contours between systolic and diastolic phases of the cardiac cycle [87].

The left ventricular ejection fraction (LVEF) is a particularly important parameter measured using CINE CMR and/or echocardiography which is used to assess prognosis and severity of functional damage caused by various cardiac diseases. LVEF is the percentage of blood that the LV pumps out with each contraction, normally ranging between 50 – 70% [88]. Reduced LVEF (<40%) has been correlated with high levels of morbidity and mortality [89], [90] and is therefore widely recognized as a risk stratification measure for arrhythmias, heart failure and SCD [91], [92]. Thus, LVEF is a critical parameter used in decision making for implantation of implantable cardioverter-defibrillator (ICD) devices following myocardial infarction (MI) [91].

An important distinction between CINE CMR compared to LGE, T1-mapping and ECV-mapping CMR is the inability to distinguish fibrotic myocardial tissue from healthy myocardial tissue. While CINE imaging permits measurement of critical volumetric and functional parameters that are associated with cardiac disease, tissue characterization is not possible [93], [94].

2.3 Measurement of Measurement of Myocardial Fibrosis in Cardiac Magnetic Resonance Imaging

Analysis and measurement of MF is a critical step in determining prognosis and treatment planning for patients with various forms of cardiac disease. In order to assess the

severity of MF, it is necessary to first isolate the myocardium as an initial region-of-interest (ROI). This is done by delineating the inner (endocardial) and outer (epicardial) borders, a process typically performed by a trained clinician, by manually tracing these regions using commercial image processing software [93], [94]. This process is tedious, time consuming and highly prone to inter- and intra-operator error. Erroneous inclusion of blood pool and/or pericardial tissue regions in the myocardial ROI can significantly confound MF measurements, which may affect prognosis of treatment planning. Ideally, automating this process would eliminate the potential for operator error while improving diagnostic efficiency. Automated myocardial segmentation has been achieved in CMR images using both conventional image-processing algorithms and deep learning-based techniques. Following isolation of the myocardial region, MF may be analyzed visually or quantitatively using a variety of conventional and/or machine learning algorithms.

2.3.1 Intensity-based Methods for Segmentation of the Myocardium and MF

Both the myocardial ROI and the fibrotic regions within the myocardium can be quantified using various segmentation techniques that are based on the intensity differences existing between the relevant region and its surroundings. These methods include threshold, energy-minimization, and clustering segmentation techniques.

2.3.1.1 Threshold Segmentation Techniques

Intensity-based segmentation techniques are often used for quantification of MF, specifically in LGE-CMR, after a prior delineation (either automated or manual) of the myocardial ROI. Intensity-based segmentation techniques are based on the definition of a threshold value, with which pixel intensities above the value are defined as MF and below the value are defined as healthy myocardium. Widely used intensity-based thresholding

methods include Full Width at Half Maximum (FWHM), signal threshold versus reference mean (STRM), and region growing.

2.3.1.1.1 Full width at half maximum

The FWHM technique is typically used in LGE-CMR to quantify MF and defines the fibrotic threshold point as any pixel intensity that exceeds half of the maximum intensity of enhanced area. Thus, this technique requires that an expert user manually define an initial ROI inclusive of fibrosis, being careful not to erroneously include any remote (healthy) myocardium and/or non-myocardial tissues in the manually selected region [95].

2.3.1.1.2 Signal threshold versus reference mean

The STRM method is also used for MF quantification, defining the fibrotic threshold point as any pixel intensity exceeding x standard deviations above the mean intensity of a manually selected region of healthy myocardial tissue. Typically, between 2 to 6 standard deviations (i.e., STRM 2-SD – STRM 6-SD) are used to define MF regions, such that several MF measurements can be expressed [95]. There is currently no standard accepted number of standard deviations above mean myocardium used in clinical practice to quantify MF. Typically, the segmentation providing the best agreement with visual assessment is accepted.

2.3.1.1.3 Region growing

The region growing technique is an iterative intensity-based segmentation approach, typically used in LGE-CMR, that defines the MF segmentation based on the similarities of neighboring pixels with a manually selected “seed point” that is placed by

the user in an enhanced tissue region. The segmentation is created by examining pixels adjacent to the initial seed point in a small neighborhood and adding them to the segmentation based on some inclusion criteria, such as grayscale range, texture, or variance until reaching a convergence point. If separate enhanced regions exist throughout the myocardium, additional seed points may be required [96].

The FWHM, STRM and region growing algorithms are commonly used techniques to segment MF in LGE-CMR, however, since these techniques require user-interaction in defining ROIs and/or seed points, they are therefore prone to inter- and intra- operator error. Thus, several alternative segmentation approaches have been established, which may provide more robustness than the aforementioned techniques.

2.3.1.2 Energy-minimization-based Segmentation Techniques

Energy-minimization algorithms may be used to segment the entire myocardium and/or the MF regions within the myocardium using an optimization-based approach, which allow for improved results by incorporating shape, smoothness and location into the desired segmentation [97]. Active contours, level sets, and graph cuts are commonly used energy-minimization segmentation algorithms [98].

2.3.1.2.1 Active contour model

Active contour models, also called deformable models (3D) or snakes, are commonly used algorithms for automated or semi-automated myocardial segmentation. The active contour model uses an energy minimizing spline function that is influenced by image forces to drive it towards the edges of an object, generating the segmentation [94]. The active contour is a set of points $(X(s), Y(s))$ that are parameterized by $s \in [0,1]$.

These contour points are found by minimizing an energy function $E = E_{internal} + E_{external}$, where $E_{internal}$ and $E_{external}$ are the internal and external energy, respectively.

The internal energy function controls the contour smoothness and is defined as follows:

$$E_{internal} = \frac{1}{2} \int_0^1 \{ \alpha [\left(\frac{dX(s)}{ds} \right)^2 + \left(\frac{dY(s)}{ds} \right)^2] + \beta [\left(\frac{d^2X(s)}{ds^2} \right)^2 + \left(\frac{d^2Y(s)}{ds^2} \right)^2] \} ds \quad \text{Equation 2-2}$$

where α and β are non-negative weighting parameters, respectively expressing the degree of contour resistance and bending. The external energy function finds the edges of the image I as follows,

$$E_{external} = \int_0^1 f(X(s), Y(s)) ds \quad \text{Equation 2-3}$$

where $f(x, y) = |\nabla I(x, y)|^2$, the image gradient [95], [96].

Active contour models have been used to segment the myocardium in LGE, CINE and T1-mapping CMR [102]. Dharanibai et al. (2018) achieved fully automated myocardial segmentation of the LV using an active contour-based model driven by a local Gaussian distribution fitting (LGDF) energy and an auxiliary global intensity fitting energy [103]. Wei et al. (2011) achieved myocardial segmentation of LGE-CMR images using registration and propagation of CINE contours derived by an active contour model [104].

2.3.1.2.2 Level set method

The level set algorithm is an energy-minimization function that has the ability to adapt to changes in topology in images. In this technique, contours are represented as the zero-level set of a function that is defined in a higher dimensional hyper-surface (referred to as the level set function), one level higher than the object dimension to be segmented [97], [105]. The segmentation is iteratively refined by optimizing the level set function as follows:

$$\frac{\partial \Phi(x)}{\partial t} + F|\nabla \Phi(x)| = 0 \quad \text{Equation 2-4}$$

where F represents the speed function and $\Phi(x)$ the level set function, defined as $\Phi(x): \Omega \rightarrow \mathfrak{R}$. The speed function F depends on image intensity and surface characteristics, such as curvature and normal direction.

The level set algorithm allows for improved refinement of segmented region at the cost of computational time [106]. This technique is most frequently used in segmentation of the myocardial ROI in CINE-CMR, but has also been used to quantify MF in LGE-CMR [98].

2.3.1.2.3 Graph cuts

The graph cuts technique is an energy-minimization segmentation method wherein object segmentation is achieved by representing the image as a graph, $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ consistent of nodes \mathcal{V} and edges \mathcal{E} . The nodes set, $\mathcal{V} = s, t \cup \mathcal{P}$ contains terminal nodes called the source s and the sink t , and non-terminal nodes \mathcal{P} . Segmentation is achieved by separating foreground and background by partitioning a cut \mathcal{C} that separates the graph into subsets S , containing the source nodes, and T , containing the source sink nodes. Graph edges are assigned a cost $w(p, q)$, and the cost of a cut $\mathcal{C} = S, T$ separating the source and sink nodes is equal to the sum of the weights of edges (p, q) such that $p \in S$ and $q \in T$. Ideal segmentation is achieved by finding the minimum cut, i.e., the cut that has the minimum cost of summed edge weights [107]. Graph cuts have been used for myocardial segmentation in CINE and LGE-CMR and MF segmentation in LGE- [98], [102], [108], [109]. Similarly, the graph cuts method is often used in combination with other approaches to refine a prior myocardial or fibrosis segmentation. For example, Oghli et al. (2017) used a combination of region growing with graph cuts to segment the left ventricular

myocardium in CINE images [105]. Lu et al. (2019) integrated the graph cuts method with multi-scale feature learning to refine right ventricular myocardial segmentations in CINE images [106]. Ukwatta et al. (2014) used a convex optimization-based algorithm for segmenting MF associated with myocardial infarction based on min-cut using continuous max-flow (CMF).

2.3.1.3 Clustering Segmentation Techniques

Clustering segmentation algorithms are unsupervised machine learning techniques that achieve segmentation by classifying pixels into groups based on some similarity metric. Support vector machines (SVM), k-means clustering, and fuzzy c-means clustering are some commonly used algorithms used for cardiac segmentation.

SVM achieve segmentation by finding a hyperplane in an N-dimensional space that classifies pixels in to foreground and background classes. In k-means clustering, pixels are separated into k clusters based on similarity of features, such that the within-cluster sum of squared differences is minimized. Similarly, fuzzy c-means clustering also assigns pixels into clusters based on the distance between the cluster center and the data point, however, each pixel is also assigned probability score which defines the likeliness of each pixel belonging to certain clusters. Clustering-based techniques are typically used to segment MF and semi-viable interstitial fibrotic tissue zones, called border-zone or gray-zone, in LGE-CMR and T1-maps [26], [98], [112], [113].

2.3.2 Feature-based Methods of Analysis of MF

Qualitative and quantitative features of CMR images can be used to characterize cardiac disease phenotypes manifesting in MF. For example, radiologists may use qualitative image descriptors, such as the shape, size, or relative intensity of a region to

support diagnostic conclusions pertaining to a particular disease state. For example, HCM is a cardiac disease state that is characterized by severe LV hypertrophy, i.e., significant enlargement of the LV myocardium [38]. Thus, the size of the LV myocardium may be used as a descriptor to support the diagnosis of HCM. Similarly, the size, shape, and distribution of MF allow radiologists to discern between interstitial versus replacement - type fibrosis, which may discriminate ischemic from nonischemic disease phenotypes [114].

However, diagnostic markers in medical images are not always visually appreciable. Subtle differences in image contrast, texture or shape may go unnoticed and/or bring about subjectivity in image interpretation. Thus, the field of radiomics emerged, which aims to quantify features through mathematical extraction of the intensity patterns, spatial distribution, and interrelationships of pixels within medical images [115]– [117].

2.3.2.1 Radiomics analysis

Figure 2-10 demonstrates the typical pipeline used to derive radiomic features from medical images. Medical images typically go through some pre-processing to begin the pipeline. Such steps may include intensity normalization, intensity transformation, image filtering, and/or denoising. Next, the relevant region or lesion is segmented using either manual, automated or semi-automated techniques. Features from a specific ROI may then be derived using the segmentation mask, thereby providing textural, intensity or shape differences that are specific to that region.

During the feature extraction step, several parameters/options may be adjusted in order to improve the robustness of the extracted features. These options may include, but are not limited to, (1) pixel interpolation, which allows for feature sets to become

rotationally invariant; (2) intensity outlier filtering, which removes pixels from the segmented region that fall outside of a specified range, and (3) discretization, which groups the original ROI pixel intensity values within a specified range of intervals (called bins) and may be advantageous for imaging modalities with arbitrary intensity units [112]. These parameters are adjusted as needed to finally extract several radiomic features from within the segmented ROI. These features can broadly be divided into four categories: morphological (i.e., shape), histogram-based, texture and filter/transform -based features [115]– [118].

Once features have been derived from the ROI, feature selection is performed using various feature reduction and/or dimensionality reduction techniques to remove redundant, non-reproducible and non-relevant features from the dataset [112]. To this aim, correlation-based methods or principal component analysis (PCA) are often employed. Lastly, the remaining features can be used to train a model for a classification task, approached using various machine learning methods [115]– [117], [119], [120].

Prior works have used machine learning-based pipelines to achieve radiomic analysis (RA) -based classification of various cardiomyopathies using native T1 mapping, LGE and CINE CMR imaging [121]– [128]. Recently, HCM phenotypes were successfully differentiated from alternative disease states presenting with similar imaging signs, such as LV enlargement and interstitial MF. For example, Antonopoulos et al.'s (2021) study differentiated between cardiac amyloid and left ventricular hypertrophy (LVH) vs. normal [122], and Wang et al.'s (2020) work differentiated between HCM genotypes caused by MYH7 and MYBPC3 genetic mutations [121], wherein both works used manually delineated myocardial contours and native T1 mapping images for RA classification. As

of now, the majority of RA-based methods for disease classification rely on manual or semi-manual myocardial contours [125].

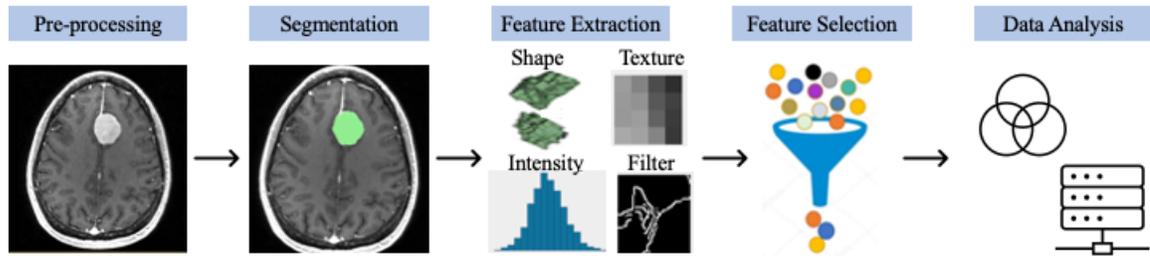


Figure 2-10 Radiomic pipeline for classification.

Radiomic studies begin with pre-processing, such as intensity normalization and/or noise reduction, followed by segmentation (automated or manual) in order to derive features specific to a certain ROI. Morphological (shape), texture, intensity-based and transform or filter -based radiomic features can then be quantified from the ROI. Feature selection or dimensionality reduction is then performed in order to remove redundant features and/or reduce the number of features used in the final data analysis step. The remaining features may then be used for various analyses, including model building for classification tasks.

2.3.2.1.1 Morphological features

Morphological features describe the shape and geometry of the region of interest.

Examples of shape features include area or volume, maximum diameter, bounding box, sphericity, maximum surface, and compactness [116].

2.3.2.1.2 Histogram-based features

Histogram-based features are also often called first-order statistical features. These are based on the global image histogram and do not consider spatial relationships between pixels [116]. Examples of histogram-based features are the mean, median, minimum, and maximum pixel intensity across the image. Other first-level statistical features include skewness, kurtosis, energy and entropy [116], [119]. The image skewness is the degree of asymmetry of the pixel distribution. Kurtosis refers to the flatness of the distribution relative to that of the Gaussian distribution. Energy refers to the image uniformity and is

larger when most pixels in the image are of similar intensities [129]. Image entropy is a measure of uncertainty and refers to the randomness of the pixel distribution [116].

2.3.2.1.3 Texture features

Texture features are also called second-order statistical features and differ from histogram-based features in that they take into account the statistical inter-relationships between neighboring pixels [116], [117]. The image gradient (i.e., the degree of abruptness of intensity fluctuations between neighboring pixels) can be used to derive several features specific to the gradient, including the gradient mean, variance, skewness, and kurtosis [116].

Other texture features are commonly derived from a gray-level co-occurrence matrix (GLCM), which is a matrix of co-occurring grayscale values computed at a given offset over an image [130]. The GLCM models intensity pairs and the frequency with which those pairs exist within the neighborhood [117]. GLCM features include contrast, correlation, angular momentum (also known as joint energy), joint entropy, difference average, maximum probability, among many others [116], [119], [130]. Similarly, texture features can also be derived from the gray-level size zone matrix (GLSZM), gray-level run length matrix (GLRLM), or neighborhood gray-tone difference matrix (NGTDM) [119], [131].

2.3.2.1.4 Transform-based features

Transform-based features are also called higher-order statistical features and are derived using first and second -order statistical methods after applying mathematical transforms or filters to the image. By applying various transforms to the image, a new

image is created with which patterns can be detected. Examples of commonly employed filters for higher-order statistical feature extraction include wavelet, Laplacian of Gaussian (LoG), logarithm, and exponential filters, among many others [116], [117], [119], [131].

2.3.2.2 Machine learning classification using radiomic analysis

Radiomic analysis permits numeric features to be extracted from medical images, from which classification can be performed. Binary classification distinguishes two classes from one another (e.g., healthy versus abnormal cardiac phenotype) and multiclass classification distinguishes three or more classes from one another (e.g., different types of cardiomyopathies). Various machine learning algorithms can be used to achieve classification. Some popular classifiers include decision trees/random forests, k nearest neighbors, and SVMs.

2.3.3 Deep learning-based Methods for Analysis of MF

In the last decade, deep learning algorithms have been implemented in a wide variety of cardiac segmentation problems. Deep learning-based algorithms offer advantages over traditional segmentation techniques based on edge detection and/or energy minimization, which are sensitive to noise and inhomogeneities in image intensity. Many of the traditional techniques used commercially for myocardial and/or MF segmentation either require refinement by a clinician and/or a user-defined initialization step [132]. Thus, these methods may be used to assist clinicians and save time, however, user interaction is still required. On the contrary, deep learning-based algorithms allow for completely automated segmentation, mitigating the need for user interaction and thereby eliminating inter and intra operator error. Deep learning methods are considered highly robust,

generalizable and scalable to several applications [133]. Moreover, these methods have been shown to outperform traditional methods in a variety of segmentation tasks across multiple imaging modalities.

Convolutional neural networks (CNNs) are one of the most used deep learning algorithms for segmentation tasks in cardiac imaging. CNNs achieve segmentation by extracting image features through a set of hierarchical layers and pooling operations. CNNs may contain tens or hundreds of layers, each functioning to learn and detect crucial features that can be used to define regions of an image. A CNN contains an input layer, output layer and several hidden layers. The output from each hidden layer is used as input to the subsequent layer. Hidden layers perform various operations, such as convolution or pooling. The convolution operation applies filters to each training image at different resolutions across each layer of the CNN in order to identify various image features. Convolution layers can be applied with a rectified linear unit (ReLU) function in a single layer, wherein the ReLU is an activation function that maps negative values to zero in order to overcome the vanishing gradient problem, allowing for increased training efficiency and improved accuracy. Exponential linear unit (ELU) is an alternative activation function that can be used in CNNs. Unlike ReLUs, ELUs can produce negative values, thereby converging cost to zero faster with improved accuracy. The pooling layer is a down-sampling operation, typically applied after a convolution layer, used to simplify output from a previous layer by reducing the dimensionality of features [134]. Max pooling is often used, whereby the maximum value for patches of a feature map is computed and used to create a down sampled feature map. A basic CNN model is demonstrated in Figure 2-11.

After learning features through several layers in the CNN, the model aims to generate the desired output. If segmentation is the desired output of the CNN, a binary segmentation mask of the same dimensions as the input image will be output, wherein the relevant object of interest is mapped to pixel values of 1 and irrelevant structures and/or background pixels are mapped to pixel values of 0. If the desired output is classification rather than segmentation (e.g., classification of a disease state), the CNN model will output a single value that is mapped to a certain class. Whether segmentation or classification is desired, the output is usually achieved through a fully connected layer which is normally placed as the second-to-last layer of the model. This layer outputs a vector of k dimensions, wherein k is the number of classes the network will predict. Following the fully connected layer, a final classification layer is applied, such as a softmax function. This function transforms the unnormalized output of the preceding fully connected layer to a vector of probabilities between 0 and 1 (i.e., a probability distribution), which can be mapped to a discrete value based on a pre-defined threshold to provide the final output segmentation or classification [133]– [135].

There have been several developments to the basic CNN architecture that have been used to achieve cardiac segmentation in CMR, such as fully convolutional networks (FCNs) [133], DenseNets [134], U-Net [135] and U-Net++ [136], among several others [137].

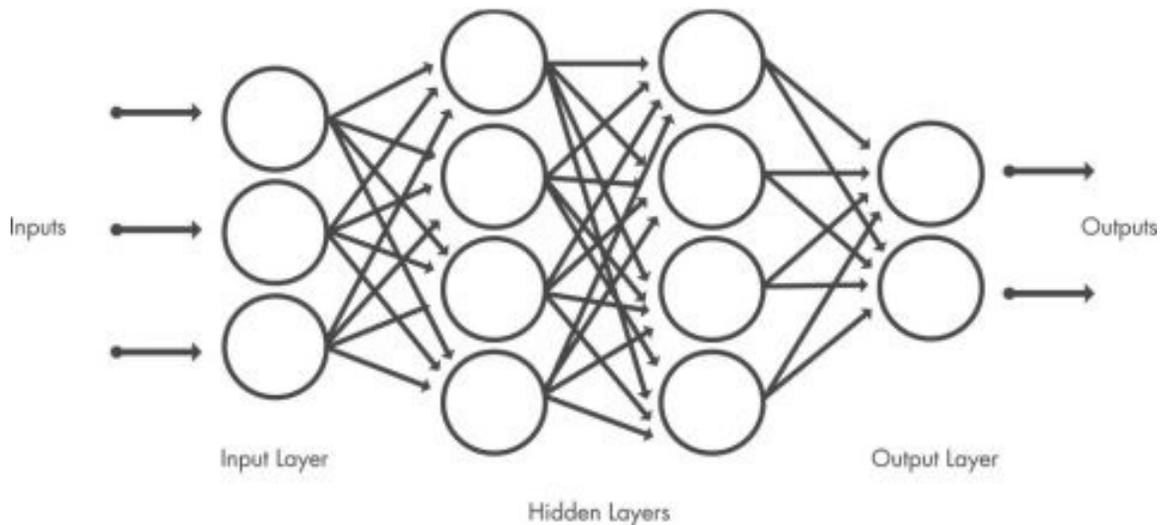


Figure 2-11 Diagram of basic CNN model

Input, output and hidden layers are demonstrated. The image is fed into the input layer of the model (left-most layer) and features are learned through convolution, pooling and rectified linear unit (ReLU) activation function layers. The classification is provided at the output layer (right-most layer). This image was freely downloaded courtesy of Mathworks®, via <https://www.mathworks.com/discovery/convolutional-neural-network-matlab.html>.

2.3.3.1 Fully convolutional networks

Conventional CNNs contain fully connected layers, meaning that all subsequent layers are connected to their preceding layer. FCNs differ from conventional CNNs in that they do not contain any fully connected layers. Instead, subsequent layers are only connected to locally connected layers through operations such as convolution, pooling, or up sampling [138]. The lack of fully connected layers reduces the number of hyperparameters required for tuning, making FCNs simpler and faster to train [133], [139]. FCN-based methods have been shown to achieve automated myocardial and/or MF segmentation with high accuracy; however, these methods require large datasets for training, which is a potential drawback, as labeled medical imaging data is often scarce, particularly in segmentation tasks evaluating different disease states [132], [140]. Nevertheless, this problem can be mitigated using transfer learning and/or data augmentation. Several studies have implemented techniques based on FCNs for both left

and right ventricular segmentation in CINE images [94], [141]– [143]. Likewise, both myocardial and MF segmentation have been achieved in LGE-CMR using techniques based on FCNs [144]– [146]. On the other hand, myocardial segmentation of T1 mapping images using FCNs has been achieved in just one study. Fahmy et al. reported automated myocardial segmentation of native T1-mapping images acquired using the slice interleaved (STONE) pulse sequence using FCNs in 2019 [147]. While this is a critical milestone, there is still a need for further development of segmentation algorithms for myocardial segmentation of T1-mapping and ECV images. For example, FCN-based myocardial segmentation algorithms have not been implemented in alternative, more widely used T1 mapping pulse sequences, such as MOLLI and shMOLLI. Moreover, FCN-based segmentation methods have not yet been applied to post-contrast T1 or ECV mapping images.

2.3.3.2 U-Net

The U-Net is a renowned FCN that was developed for biomedical image segmentation in 2015 by Ronneberger et al. [134]. U-Net differs from conventional FCNs in that it has a symmetric architecture composed of an encoder (down sampling) and decoder (up sampling) path and skip connections between the encoder and decoder paths, as shown in Figure 2-12. U-Net offers some advantages over alternative deep learning-based segmentation algorithms in that it is robust across numerous types of medical images and requires very little annotated images [138].

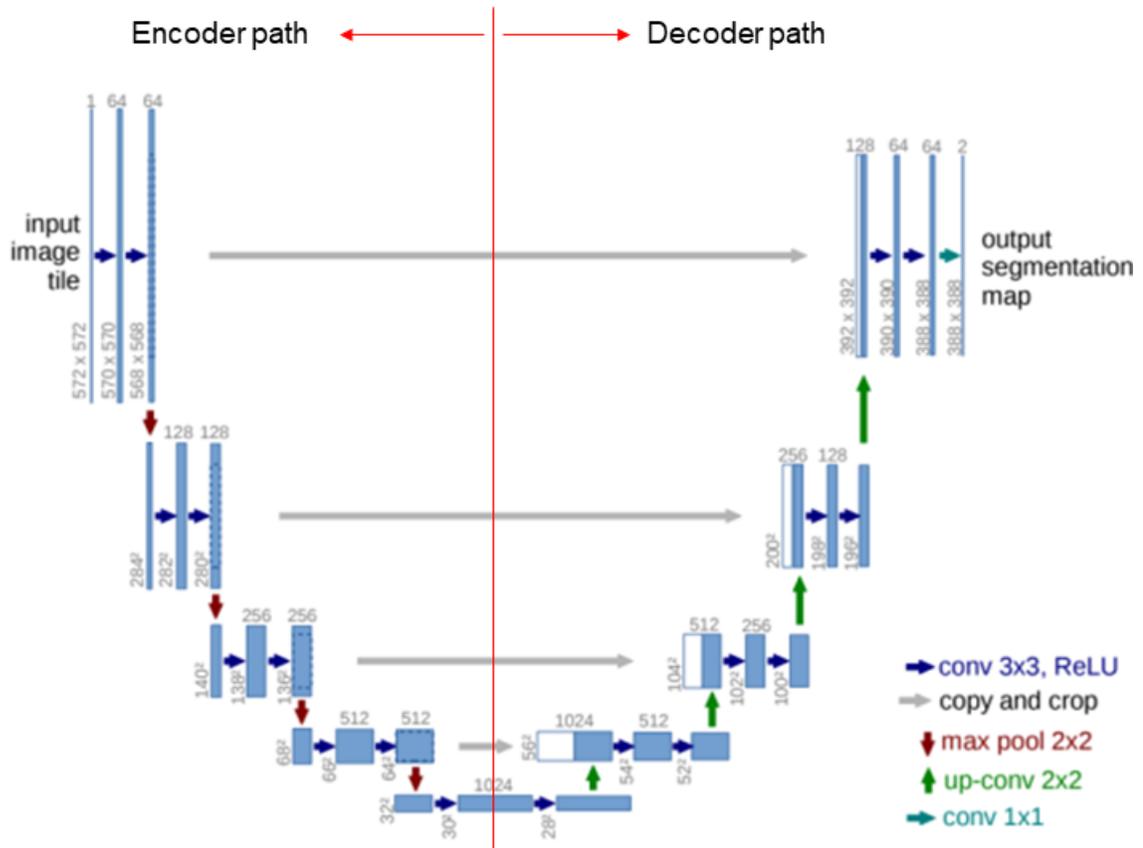


Figure 2-12 U-Net architecture.

This figure is referenced from Ronneberger et al.’s (2015) manuscript [134], cited in the bibliography. The blue boxes represent multi-channel feature maps. The number of channels for each feature map is denoted on top of the respective blue box. The input image is down sampled and reduced in dimension as it passes through each convolution, wherein the x-y dimensions are denoted per each feature map. White boxes represent copied feature maps. Blue right arrows denote 3x3 convolution and ReLU operations, grey right arrows represent copy and crop operations, and light blue right arrows represent 1x1 convolution. Red down arrows represent 2x2 max pooling operations and green up arrows represent 2x2 up-convolution. Note that the red line dividing encoder and decoder paths is for visualization/understanding purposes only.

The first half of the U-Net architecture is the encoder path, which follows the typical architecture of an FCN. The encoder path consists of several repeated “steps” whereby two 3x3 convolution operations are applied each followed by a ReLU (blue right arrows in Figure 2-12), after which, 2x2 max pooling is applied for down sampling (red down arrows in Figure 2-12). “Step one” of the encoder path can be visualized at the left-most side of Figure 2-12. The first convolution operation is initially applied to the input image and changes the number of channels (i.e., the “depth”) of the image from 1 to 64

(i.e., this operation consists of 64 filters of size 3×3 and 1 channel per filter). The second convolution operation contains 64 filters of size 3×3 whereby each filter consists of 64 channels. Thus, the output of the second convolution at “step one” is 64 feature maps. Output from “step one” is then max pooled to the next “step”, which doubles the number of feature channels and halves the image dimensions. This process is repeated a total of four times. Two border pixels are lost at every convolution operation; therefore, the dimensions of the image are reduced by four in length and width at each “step” in the encoder path. The number of feature channels for each convolution and max pooling operation with corresponding image dimensions can be visualized across the encoder path can be seen in Figure 2-12.

The decoder path, which is the second half of the architecture, consists of several “steps”, whereby 2×2 “up-convolution” is performed for up sampling of the feature map (green up arrows in Figure 2-12), followed by concatenation, whereby the feature map of corresponding dimensions from the encoder path is copied to the decoder path and cropped (grey right arrows and white boxes in Figure 2-12), and lastly, two 3×3 convolution operations, each followed by a ReLU. This process is repeated four times. The 2×2 “up-convolution” operation halves the number of feature channels and doubles the dimensions of the image. For the concatenation operation, cropping is required because border pixels are lost at every 3×3 convolution layer; thus, cropping ensures the output segmentation map will have the same dimensionality as the original input image. At the final “step” in the decoder path, a 1×1 convolution layer is applied, which reduces the depth of the preceding convolution layer and maps the 64-component feature vector to the desired

number of classes, i.e., two classes, whereby class 1 is the segmented object and class 0 is the background and other irrelevant structures within the image.

The U-Net model has been implemented in a number of segmentation problems in CMR. Zabihollahy et al. (2019) proposed a U-Net based technique for fully automated segmentation of the LV myocardium in 3D LGE-CMR [150]. Fahmy et al. (2018) also used a U-Net based architecture to segment the LV myocardium and MF regions in LGE-CMR in patients with hypertrophic cardiomyopathy [151].

2.3.3.3 DenseNet

DenseNet is a type of CNN that uses dense connections and Dense Blocks, such that all layers are directly connected with each other [135]. Dense connections (i.e., fully connected layers) use a linear operation to connect inputs to subsequent layers using a weighting factor such that each layer gets input from all preceding layers and passes feature maps to all subsequent layers. Dense Blocks consist of several layers that are densely connected with matching feature map sizes to all other layers. Features are concatenated before being passed to the next layer (in contrast to ResNets [152], whereby features are instead combined through summation). In this way, the l^{th} layer has l inputs, consisting of feature maps from all preceding convolutional layers [135], [153]. An example of a Dense Block architecture can be seen in Figure 2-13. The original DenseNet architecture consists of three Dense Blocks, wherein convolution and pooling layers are placed in between (called “transition layers”) to change feature map sizes. DenseNets offer the advantage of finding more refined features which may increase accuracy in image segmentation, however, computational time is higher due to the dense connections.

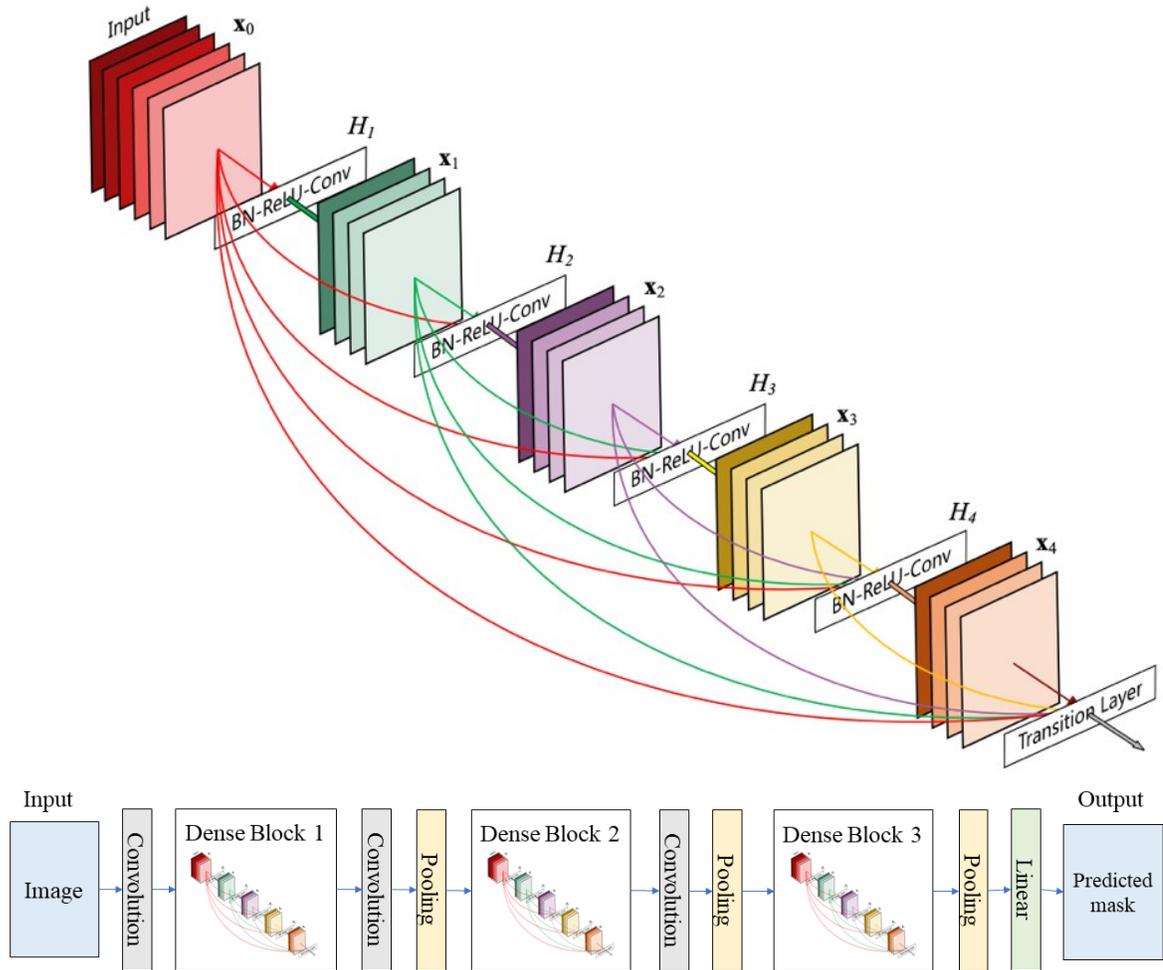


Figure 2-13 Architecture of Dense Blocks used in the DenseNet model (top) and DenseNet network with embedded dense blocks (below). Each layer within the Dense Block takes all preceding feature maps as input. This image is referenced from Huang et al.'s (2017) manuscript [135], referenced in the bibliography.

2.3.3.4 U-Net++

U-Net++ is an FCN based on the original U-Net architecture which incorporates ideas from DenseNets [135] to improve accuracy [136]. U-Net++ differs from U-Net in that convolution layers are applied on skip-pathways to bridge the gap between encoder to decoder feature maps. Additionally, dense connections are applied on skip pathways, which improves gradient flow. The architecture of the U-Net++ model is demonstrated in Figure 2-14. The U-Net++ architecture contains the same encoder path, called the

backbone, and decoder path as original U-Net (both denoted by black arrows in Figure 2-14). However, the original skip connections of U-Net are replaced by re-designed skip connections in U-Net++ (shown in blue and green in Figure 2-14), wherein every two adjacent nodes are connected. The skip pathways consist of dense convolution blocks consisting of three convolutional layers. Concatenation is performed before each convolutional layer, which fuses the output from the preceding convolutional layer of the same dense block. Another notable difference between original U-Net and U-Net++ is deep supervision, whereby pruning may be performed to accelerate segmentation speed. U-Net++'s skip pathways allow for full resolution feature maps to be output at four "branches", resulting in four output binary segmentation masks. Thus, deep supervision allows for the user to define how many branches will be used toward the final output segmentation. In "accurate mode", the output segmentation mask is generated via averaging of all four branch outputs. In "fast mode", only the segmentation map from the last branch is selected.

U-Net++ combines multiscale features from all preceding nodes at the same resolution while also integrating multiscale features of differing resolutions to subsequent nodes. While this architecture offers improved accuracy over conventional U-Net, it is more computational expensive and memory intensive [136], [138].

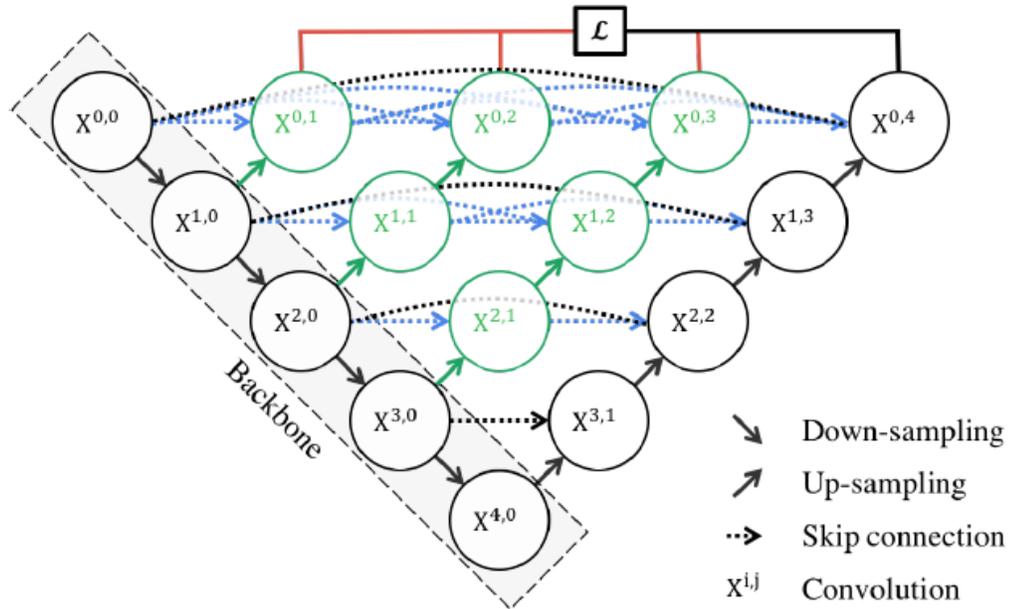


Figure 2-14 U-Net++ architecture.

The U-Net++ encoder (backbone) and decoder pathways are shown by black nodes (circles marked by $X^{i,j}$) and black arrows. Each node represents a convolution operation. Down arrows represent down sampling and up arrows represent up sampling. Skip pathways connecting each adjacent node are shown in blue and green. The red lines represent deep supervision, wherein “accurate mode” implements averaging of all outputs from nodes $X^{0,1}$, $X^{0,2}$, $X^{0,3}$ and $X^{0,4}$ while “fast mode” uses only the output from node $X^{0,4}$ toward the final binary segmentation mask. This image is referenced from Zhou et al.’s (2018) manuscript [136], referenced in the bibliography.

2.3.3.5 Attention U-Net

The Attention U-Net is an adaptation to the original U-Net model that incorporates attention gates into the skip connections, which serve to filter and suppress irrelevant features while highlighting salient features [139]. In the basic U-Net model, skip connections are used to combine spatial information from the encoder path with the decoder path in order to combat the risk of producing imprecise spatial information. However, this process may introduce several redundant, low level feature extractions to the model, since feature maps are coarser along the encoder path. The Attention U-Net counteracts this problem by adding attention gates to each “step” in the decoder path of the original U-Net architecture, as seen in Figure 2-15.

The attention gate operation scales input features from a convolutional layer (green up arrows in Figure 2-15), whereby the feature maps go through strided convolution in order to reduce dimensions to the corresponding coarser dimensions of the proceeding layer in the model. The proceeding layer undergoes a 1x1 convolution, such that it can be summed elementwise with the preceding feature map of matching dimensions.

By implementing attention gates into the original U-Net architecture, only important features are preserved while redundant features are suppressed, which may improve prediction accuracy without a significant increase in computing cost [139].

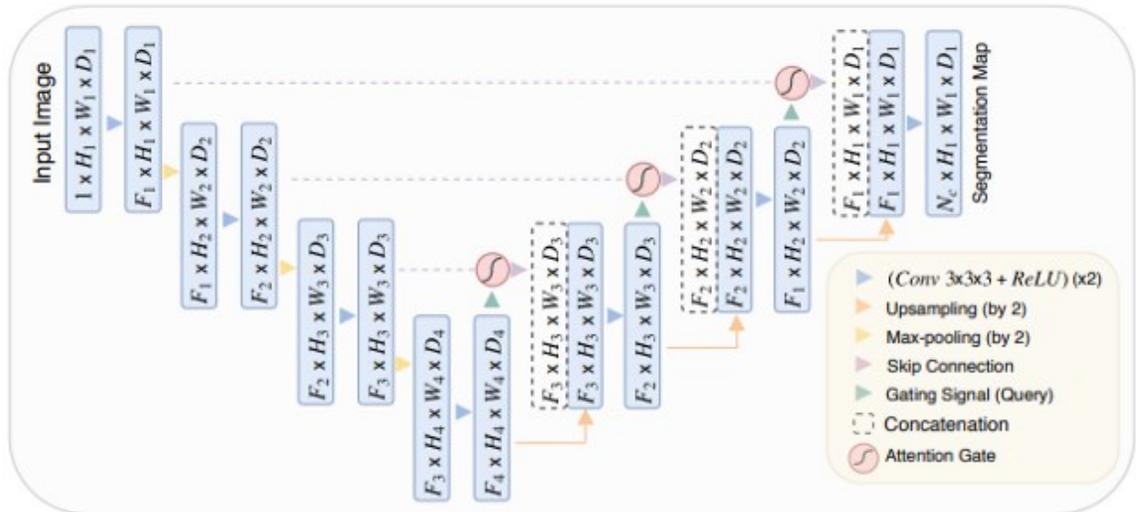


Figure 2-15 Attention U-Net architecture.

The Attention U-Net model architecture contains the same encoder and decoder paths as conventional U-Net; however, attention gates are added to each step in the decoder path. The attention gates filter features propagated through the skip connections in order to reduce redundancy. This figure is referenced from Oktay et al.'s (2018) article [139], referenced in the bibliography.

2.4 Evaluation of Myocardial Segmentation and Fibrosis Measurements

In order to translate the use of segmentation and classification algorithms to clinical practice, it is essential to evaluate method performance against a reference gold standard. Typically, segmentation algorithms are evaluated against a manual segmentation of the region delineated by an expert, while classification algorithms are evaluated against true

class labels from a pre-defined subset of testing data. Several attributes are used for this purpose, including accuracy/similarity metrics, tissue characterization and statistical tests.

2.4.1 Comparison Metrics

Common comparison metrics used for segmentation evaluation include the dice similarity coefficient (DSC) and error metrics, such as absolute and relative mass or volume error. Other metrics, such as precision, recall, F1 score, and accuracy can be computed for evaluation of both segmentation and classification algorithms and may be computed using true positive (TP), true negative (TN), false positive (FP) and false negative (FN) metrics provided by a confusion matrix, as demonstrated in Table 1. Lastly, the receiver operating curve (ROC) and area under the curve (AUC) are common metrics used to evaluate classification algorithms.

Table 2-1 Confusion matrix used to compute accuracy metrics

Actual value	A	TP	FN
	B	FP	TN
		A	B
		Predicted value	

The DSC is a measure of the spatial overlap between the binary segmentation generated manually, A and the binary segmentation generated by the algorithm, B . The DSC metric is given as follows:

$$DSC = \frac{2(A \cap B)}{|A| + |B|} \quad \text{Equation 2-5}$$

where A and B represent two sets, $A \cap B$ is the intersection of regions A and B , $|A|$ and $|B|$ is the cardinality of region A and region B (i.e., the number of elements in the set) respectively. The DSC may also be expressed in terms of TP, FP, and FN as:

$$DSC = \frac{2TP}{2(TP+FP+FN)} \quad \text{Equation 2-6}$$

where a TP instance is a pixel that is found in both segmentation A and segmentation B . A FP is a pixel that is found in segmentation A but not in segmentation B . A FN is a pixel that is not included in segmentation A but is included in segmentation B . The DSC metric is given on a scale of 0 – 1 or can be expressed as a percentage, where values closer to 1 or 100% denote greater similarity between segmentations and a DSC = 1 or 100% denotes identity [144].

Error metrics, including absolute error and relative error indices, are measures used to evaluate the difference in area or volume of regions segmented manually versus algorithmically. The absolute and relative error indices are given as follows:

$$Absolute\ Error = |A - B| \quad \text{Equation 2-7}$$

$$Relative\ Error = \frac{|A-B|}{|A|} \quad \text{Equation 2-8}$$

where $|A - B|$ is the absolute difference of the manually segmented area or volume A and the area or volume segmented by the algorithm, B . The relative error index normalizes this difference.

Precision, recall, accuracy and F1 score are computed using confusion matrix indices and can be computed to evaluate both segmentation and classification algorithms. Precision is a measure of the number of TP instances among the amount of observed positive instances. Recall (also called sensitivity) is the measure of the number of TP instances compared to the actual number of positive instances. Accuracy is a measure of the number of correct predictions (TPs and true negatives (i.e., TNs)) over the total number

of predictions. The F1 score is the weighted average of precision and recall. Precision, recall and accuracy are computed as follows:

$$\textit{Precision} = \frac{TP}{TP+FP} \quad \text{Equation 2-9}$$

$$\textit{Recall} = \frac{TP}{TP+FN} \quad \text{Equation 2-10}$$

$$\textit{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad \text{Equation 2-11}$$

$$\textit{F1 Score} = 2 \cdot \frac{\textit{precision} \cdot \textit{recall}}{\textit{precision} + \textit{recall}} \quad \text{Equation 2-12}$$

The ROC and AUC metrics, both visual and numerical means to evaluate a classification model, are used to summarize the performance of an algorithm at different threshold values. The ROC curve is generated by plotting the true positive rate (TPR, also called sensitivity) on the x-axis against the false positive rate (FPR) (1 – specificity) on the y-axis. Each point in the ROC curve corresponds to the classification that can be obtained using a given decision threshold, thereby demonstrating the tradeoff between TPR and FPR [145], [146]. The ROC curve demonstrates the performance of the classifier in comparison to chance, which is referenced on the graph as a diagonal “chance line”. Equations for TPR and FPR are given and an example ROC of two classifiers is demonstrated in Figure 2-16.

$$\textit{TPR} = \frac{TP}{TP+FN} \quad \text{Equation 2-13}$$

$$\textit{FPR} = 1 - \textit{TPR} = \frac{FP}{TN+FP} \quad \text{Equation 2-14}$$

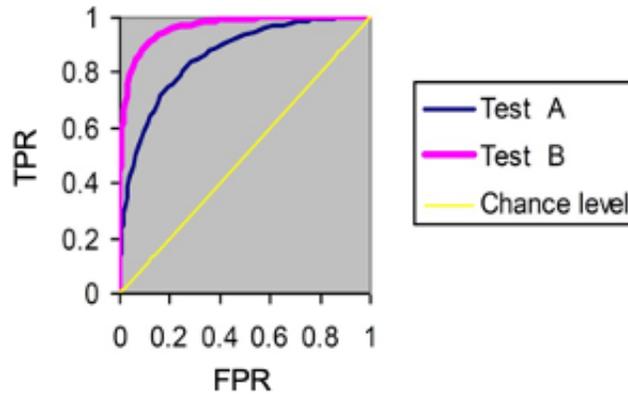


Figure 2-16 Example ROC curves of two classifiers

The AUC summarizes the entire ROC and aggregates the performance of the classifier at all threshold values. It is computed using integral calculus and is expressed as a value between 0 and 1, wherein the closer to 1, the better the performance of the classifier. In Figure 2-16, Classifier B has better performance than Classifier A, as demonstrated by the larger area under the ROC.

2.4.2 T1 and ECV Mapping Measurements of Myocardial Fibrosis

In addition to the aforementioned evaluation metrics, it is also common to evaluate the differences between quantitative pixel values of T1 maps and/or ECV maps reported by manual segmentation versus automated segmentation. Since pixel values in T1-mapping and ECV-mapping images are representative of intrinsic tissue parameters, the similarity of these pixel values is indicative of segmentation accuracy.

Tissue characterization comparison is achieved by evaluating basic statistical parameters of the global (i.e., overall) pixel intensity measured by the manual segmentation versus the global pixel intensity measured by the automated algorithm. The mean and median pixel intensity, standard deviation, and intensity range are typically measured

within the manual and automated segmented regions. These values can be compared qualitatively, by comparison using box and whisker plots, and quantitatively using Bland-Altman analysis and/or correlation coefficients, such as Pearson's r or Spearman's ρ . Tissue characterization comparison between manual and automated segmentation is particularly useful for the analysis of the efficacy of automated algorithms in images demonstrating various forms of cardiac disease with which T1 and ECV values fall within a certain known range.

2.4.3 Statistical Tests

Several statistical tests may be used for evaluation of one or more automated segmentation techniques versus manual segmentation. The student's t-test is frequently used for this purpose, wherein the goal is to determine if the difference between the mean value of two sets of observations is significantly different. The null hypothesis assumes that the difference of the mean value between the two sets of observations is zero, and alternative hypothesis assumes that the mean hypothesis is either greater than or less than zero (two-tailed t-test), greater than zero (right-tailed t-test), or smaller than zero (left-tailed t-test). This statistical test assumes that the data follows a normal distribution.

In datasets where normality cannot be assumed, non-parametric statistical tests such as the Wilcoxon signed rank test may be used as an alternative to the student's t-test. The Wilcoxon signed rank test is a paired difference test that is used to compare two related observations (such as matched observations generated by two different methods) to assess whether their population mean ranks differ. This method can be used to evaluate the performance of two algorithms across a particular accuracy metric and is typically

evaluated at 5% level of confidence (P), wherein $P < 0.05$ indicates significant differences between the two measurements.

The Mann-Whitney U test is an alternative non-parametric statistical test that may be used to evaluate differences in two independent groups, such as patients with no cardiac disease versus patients with cardiac disease. This statistical test is typically used as an alternative to the student's t-test in very small datasets where no relationship exists between the recorded measurements (i.e., measurements are not paired). This test is also evaluated at a 5% level of confidence, wherein $P < 0.05$ indicates significant differences between the groups.

Similarly, the Kruskal-Wallis test is frequently used to evaluate whether observations from two datasets originate from the same distribution. This test is a non-parametric rank-based method to compare the means of two or more groups. This test is the non-parametric alternative to the one-way analysis of variance (ANOVA) test and is an extension of the Mann-Whitney U Test, which is only used to compare two groups. This test is evaluated at a 5% level of confidence, wherein $P < 0.05$ indicates significant differences between the groups.

2.5 Chapter 2 References

- [1] S. Hinderer and K. Schenke-Layland, "Cardiac fibrosis – A short review of causes and therapeutic strategies," *Adv Drug Deliv Rev*, vol. 146, pp. 77–82, 2019, doi: 10.1016/j.addr.2019.05.011.
- [2] J. G. Travers, F. A. Kamal, J. Robbins, K. E. Yutzey, and B. C. Blaxall, "Cardiac fibrosis: The fibroblast awakens," *Circ Res*, vol. 118, no. 6, pp. 1021–1040, 2016, doi: 10.1161/CIRCRESAHA.115.306565.
- [3] P. G. Masci *et al.*, "Incremental Prognostic Value of Myocardial Fibrosis in Patients with Non – Ischemic Cardiomyopathy Without Congestive Heart Failure," pp. 448–456, 2014, doi: 10.1161/CIRCHEARTFAILURE.113.000996.
- [4] A. Alonso and F. L. Norby, "Predicting atrial fibrillation and its complications," *Circulation*, vol. 80, no. 5, pp. 1061–1066, 2016, doi: 10.1253/circj. CJ-16-0239.

- [5] A. G. Almeida, “Searching for the key to foreseeing left atrial fibrosis in atrial fibrillation: Could this be simple and reliable with MRI?” *Radiology*, vol. 292, no. 3, pp. 583–584, 2019, doi: 10.1148/radiol.2019191221.
- [6] F. R. van de Schoor *et al.*, “Myocardial Fibrosis in Athletes,” *Mayo Clin Proc*, vol. 91, no. 11, pp. 1617–1631, 2016, doi: 10.1016/j.mayocp.2016.07.012.
- [7] L. Zhao *et al.*, “Prognostic Significance of Left Ventricular Fibrosis Assessed by T1 Mapping in Patients with Atrial Fibrillation and Heart Failure,” *Sci Rep*, vol. 9, no. 1, pp. 1–9, 2019, doi: 10.1038/s41598-019-49793-8.
- [8] M. N. Berman, C. Tupper, and A. Bhardwaj, “Physiology, Left Ventricular Function,” *StatPearls*, Sep. 2020.
- [9] J. K. Volpe and A. N. Makaryus, “Anatomy, Thorax, Heart and Pericardial Cavity,” *StatPearls*, Jul. 2021.
- [10] Y. M. Carter, C. J. Wehrle, and K. Mahajan, “Anatomy, Thorax, Heart Anomalous Left Coronary Artery,” Aug. 2021.
- [11] I. Rehman and A. Rehman, “Anatomy, Thorax, Heart,” *StatPearls*, Dec. 2020.
- [12] D. A. Bluemke and P. Pattanayak, “Tissue Characterization of the Myocardium: State of the art characterization by magnetic resonance and computed tomography imaging,” *Radiol Clin North Am*, vol. 53, no. 2, pp. 413–423, 2015, doi: 10.1016/j.rcl.2014.11.005.
- [13] N. Mewton, C. Y. Liu, P. Croisille, D. Bluemke, and J. A. C. Lima, “Assessment of myocardial fibrosis with cardiovascular magnetic resonance,” *J Am Coll Cardiol*, vol. 57, no. 8, pp. 891–903, 2011, doi: 10.1016/j.jacc.2010.11.013.
- [14] R. Bing and M. R. Dweck, “Myocardial fibrosis: Why image, how to image and clinical implications,” *Heart*, vol. 105, no. 23, pp. 1832–1840, 2019, doi: 10.1136/heartjnl-2019-315560.
- [15] T. Liu *et al.*, “Current understanding of the pathophysiology of myocardial fibrosis and its quantitative assessment in heart failure,” *Front Physiol*, vol. 8, no. APR, pp. 1–13, 2017, doi: 10.3389/fphys.2017.00238.
- [16] F. Weidemann *et al.*, “Impact of myocardial fibrosis in patients with symptomatic severe aortic stenosis,” *Circulation*, vol. 120, no. 7, pp. 577–584, 2009, doi: 10.1161/CIRCULATIONAHA.108.847772.
- [17] S. Herrmann *et al.*, “Myocardial Fibrosis Predicts 10-Year Survival in Patients Undergoing Aortic Valve Replacement,” *Circ Cardiovasc Imaging*, vol. 11, no. 8, p. e007131, 2018, doi: 10.1161/CIRCIMAGING.117.007131.
- [18] W. L. de Graaf, K. Vandoorne, F. Arslan, K. Nicolay, and G. J. Strijkers, “Contrast-Enhanced T1-Mapping MRI for the Assessment of Myocardial Fibrosis,” *Curr Cardiovasc Imaging Rep*, vol. 7, no. 4, pp. 1–12, 2014, doi: 10.1007/s12410-014-9260-6.
- [19] M. Gyöngyösi *et al.*, “Myocardial fibrosis: biomedical research from bench to bedside,” *Eur J Heart Fail*, vol. 19, no. 2, pp. 177–191, 2017, doi: 10.1002/ejhf.696.
- [20] F. Saraiva, V. Matos, L. Goncalves, M. Antunes, and L. A. Providência, “Complications of endomyocardial biopsy in heart transplant patients: A retrospective study of 2117 consecutive procedures,” *Transplant Proc*, vol. 43, no. 5, pp. 1908–1912, 2011, doi: 10.1016/j.transproceed.2011.03.010.
- [21] P. R. Scully, G. Bastarrika, J. C. Moon, and T. A. Treibel, “Myocardial Extracellular Volume Quantification by Cardiovascular Magnetic Resonance and Computed Tomography,” 2018.

- [22] L. Scholtz, A. Sarkin, and Z. Lockhat, “Current clinical applications of cardiovascular magnetic resonance imaging.,” *Cardiovasc J Afr*, vol. 25, no. 4, pp. 185–90, 2014, doi: 10.5830/CVJA-2014-021.
- [23] J. R. Burt, S. L. Zimmerman, I. R. Kamel, M. Halushka, and D. A. Bluemke, “Myocardial T1 Mapping: Techniques and Potential Applications,” *RadioGraphics*, vol. 34, no. 2, pp. 377–395, Mar. 2014, doi: 10.1148/rg.342125121.
- [24] E. B. Schelbert and T. C. Wong, “Clinical Applications of T1 Mapping,” *American College of Cardiology*, 2015. https://www.acc.org/latest-in-cardiology/articles/2015/02/13/09/24/clinical-applications-of-t1-mapping?w_nav=TI
- [25] E. Braunwald, “Cardiomyopathy Compendium Cardiomyopathies An Overview,” pp. 711–721, 2017, doi: 10.1161/CIRCRESAHA.117.311812.
- [26] Y. Yang *et al.*, “Multi-contrast late enhancement CMR determined gray zone and papillary muscle involvement predict appropriate ICD therapy in patients with ischemic heart disease.,” *J Cardiovasc Magn Reson*, vol. 15, no. 1, p. 57, 2013, doi: 10.1186/1532-429X-15-57.
- [27] N. A. Farrag, V. Ramanan, G. A. Wright, and E. Ukwatta, “Effect of T1-mapping technique and diminished image resolution on quantification of infarct mass and its ability in predicting appropriate ICD therapy,” *Med Phys*, vol. 45, no. 4, pp. 1577–1585, 2018, doi: 10.1002/mp.12840.
- [28] M. Pop *et al.*, “Quantification of fibrosis in infarcted swine hearts by *ex vivo* late gadolinium-enhancement and diffusion-weighted MRI methods,” *Phys Med Biol*, vol. 58, no. 15, pp. 5009–5028, 2013, doi: 10.1088/0031-9155/58/15/5009.
- [29] G. I. Fishman *et al.*, “Sudden Cardiac Death Prediction and Prevention: Report From a National Heart, Lung, and Blood Institute and Heart Rhythm Society Workshop,” *Circulation*, vol. 122, no. 22, pp. 2335–2348, Nov. 2010, doi: 10.1161/CIRCULATIONAHA.110.976092.
- [30] M. Pop, S. Oduneye, L. Zhang, S. Newbigging, and G. Wright, “In Vivo Parametric T1 Maps Correlate with Structural and Molecular Characteristics of Focal Fibrosis BT - Functional Imaging and Modelling of the Heart: 9th International Conference, FIMH 2017, Toronto, ON, Canada, June 11-13, 2017, Proceedings,” M. Pop and G. A. Wright, Eds. Cham: Springer International Publishing, 2017, pp. 13–22. doi: 10.1007/978-3-319-59448-4_2.
- [31] E. Ukwatta *et al.*, “Image-based reconstruction of three-dimensional myocardial infarct geometry for patient-specific modeling of cardiac electrophysiology,” *Med Phys*, vol. 42, no. 8, pp. 4579–4590, 2015, doi: 10.1118/1.4926428.
- [32] T. Skála and M. Táborský, “Electromechanical mapping in electrophysiology and beyond,” *Cor Vasa*, vol. 57, no. 6, pp. e470–e482, 2015, doi: 10.1016/j.crvasa.2015.10.002.
- [33] E. Lin and A. Alessio, “What are the basic concepts of temporal, contrast, and spatial resolution in cardiac CT?” *J Cardiovasc Comput Tomogr*, vol. 3, no. 6, pp. 403–408, 2009, doi: doi:10.1016/j.jcct.2009.07.003.
- [34] D. B. Plewes and W. Kucharczyk, “Physics of MRI: A primer,” *Journal of Magnetic Resonance Imaging*, vol. 35, no. 5, pp. 1038–1054, 2012, doi: 10.1002/jmri.23642.
- [35] V. S. Vassiliou, D. Cameron, S. K. Prasad, and P. D. Gatehouse, “Magnetic resonance imaging: Physics basics for the cardiologist,” *JRSM Cardiovasc Dis*, vol. 7, p. 204800401877223, 2018, doi: 10.1177/2048004018772237.

- [36] V. P. B. Grover, J. M. Tognarelli, M. M. E. Crossey, I. J. Cox, S. D. Taylor-Robinson, and M. J. W. McPhail, "Magnetic Resonance Imaging: Principles and Techniques: Lessons for Clinicians," *J Clin Exp Hepatol*, vol. 5, no. 3, pp. 246–255, 2015, doi: 10.1016/j.jceh.2015.08.001.
- [37] M. S. Dzeshka, G. Y. H. Lip, V. Snezhitskiy, and E. Shantsila, "Cardiac Fibrosis in Patients with Atrial Fibrillation: Mechanisms and Clinical Implications," *J Am Coll Cardiol*, vol. 66, no. 8, 2015, doi: 10.1016/j.jacc.2015.06.1313.
- [38] S. Fluechter *et al.*, "Extent of late gadolinium enhancement detected by cardiovascular magnetic resonance correlates with the inducibility of ventricular tachyarrhythmia in hypertrophic cardiomyopathy," *Journal of Cardiovascular Magnetic Resonance*, vol. 12, no. 1, 2010, doi: 10.1186/1532-429X-12-30.
- [39] C. M. Kramer, J. Barkhausen, S. D. Flamm, R. J. Kim, and E. Nagel, "Standardized cardiovascular magnetic resonance (CMR) protocols 2013 update," pp. 1–10, 2013.
- [40] V. O. Puntmann, E. Peker, Y. Chandrashekhar, and E. Nagel, "T1 Mapping in Characterizing Myocardial Disease," *Circ Res*, vol. 119, no. 2, 2016.
- [41] V. O. Puntmann, E. Peker, Y. Chandrashekhar, and E. Nagel, "T1 Mapping in Characterizing Myocardial Disease: A Comprehensive Review," *Circ Res*, vol. 119, no. 2, pp. 277–299, 2016, doi: 10.1161/CIRCRESAHA.116.307974.
- [42] P. Haaf, P. Garg, D. R. Messroghli, D. A. Broadbent, J. P. Greenwood, and S. Plein, "Cardiac T1 Mapping and Extracellular Volume (ECV) in clinical practice: A comprehensive review," *Journal of Cardiovascular Magnetic Resonance*, vol. 18, no. 1, pp. 1–12, 2016, doi: 10.1186/s12968-016-0308-4.
- [43] V. O. Puntmann *et al.*, "Native T1 mapping in differentiation of normal myocardium from diffuse disease in hypertrophic and dilated cardiomyopathy," *JACC Cardiovasc Imaging*, vol. 6, no. 4, pp. 475–484, 2013, doi: 10.1016/j.jcmg.2012.08.019.
- [44] L. Altabella, M. Petrini, M. Ali, P. M. Cannao, F. Secchi, and F. Sardanelli, "Novel cardiac magnetic resonance biomarkers: native T1 and extracellular volume myocardial mapping," vol. 18, pp. 64–71, 2016, doi: 10.1093/eurheartj/suw022.
- [45] C. Parsai, R. Ohanlon, S. K. Prasad, and R. H. Mohiaddin, "Diagnostic and prognostic value of cardiovascular magnetic resonance in non-ischaemic cardiomyopathies," *Journal of Cardiovascular Magnetic Resonance*, vol. 14, no. 1, p. 1, 2012, doi: 10.1186/1532-429X-14-54.
- [46] P. Garg, L. C. Saunders, A. J. Swift, J. M. Wild, and S. Plein, "Role of cardiac T1 mapping and extracellular volume in the assessment of myocardial infarction," vol. i, no. 0, pp. 404–411, 2018, doi: 10.14744/AnatolJCardiol.2018.39586.
- [47] P. Garg *et al.*, "Is signal intensity of late gadolinium enhancement a substitute for extracellular volume mapping in acute myocardial infarction?" *Journal of Cardiovascular Magnetic Resonance*, vol. 17, no. S1, pp. 1–2, 2015, doi: 10.1186/1532-429x-17-s1-p156.
- [48] E. B. Schelbert and D. R. Messroghli, "State of the Art: Clinical Applications of Cardiac T1 Mapping 1," vol. 278, no. 3, 2016.
- [49] F. M. Caballeros *et al.*, "T1 mapping of the myocardium: what a resident needs to know," *European Cardiology Review*, pp. 1–17, 2016, doi: 10.1594/ecr2016/C-0385.
- [50] J. C. Moon *et al.*, "Myocardial T1 mapping and extracellular volume quantification: a Society for Cardiovascular Magnetic Resonance (SCMR) and CMR Working Group of

- the European Society of Cardiology consensus statement,” *Journal of Cardiovascular Magnetic Resonance*, vol. 15, no. 1, p. 92, 2013, doi: 10.1186/1532-429X-15-92.
- [51] A. J. Taylor, M. Salerno, R. Dharmakumar, and M. Jerosch-Herold, “T1 Mapping Basic Techniques and Clinical Applications,” *JACC Cardiovasc Imaging*, vol. 9, no. 1, pp. 67–81, 2016, doi: 10.1016/j.jcmg.2015.11.005.
- [52] P. Kellman and M. S. Hansen, “T1-mapping in the heart: accuracy and precision,” *Journal of Cardiovascular Magnetic Resonance*, vol. 16, no. 2, 2014, doi: 10.1186/1532-429X-16-2.
- [53] P. K. Kim *et al.*, “Myocardial T1 and T2 Mapping: Techniques and Clinical Applications,” *Korean J Radiol*, vol. 18, no. 1, p. 113, 2017, doi: 10.3348/kjr.2017.18.1.113.
- [54] D. Radenkovic, S. Weingärtner, L. Ricketts, J. C. Moon, and G. Captur, “T 1 mapping in cardiac MRI,” pp. 415–430, 2017, doi: 10.1007/s10741-017-9627-2.
- [55] K. Werys *et al.*, “Total Mapping Toolbox (TOMATO): An open source library for cardiac magnetic resonance parametric mapping,” *SoftwareX*, vol. 11, Jan. 2020, doi: 10.1016/j.softx.2019.100369.
- [56] D. R. Messroghli, A. Radjenovic, S. Kozerke, D. M. Higgins, M. U. Sivananthan, and J. P. Ridgway, “Modified Look-Locker inversion recovery (MOLLI) for high-resolution T1 mapping of the heart,” *Magn Reson Med*, vol. 52, no. 1, pp. 141–146, Jul. 2004, doi: 10.1002/mrm.20110.
- [57] S. K. Piechnik *et al.*, “Shortened Modified Look-Locker Inversion recovery (ShMOLLI) for clinical myocardial T1- mapping at 1.5 and 3 T within a 9 heartbeat breathhold,” pp. 1–11, 2010.
- [58] S. Roujol *et al.*, “Accuracy, precision, and reproducibility of four T1 mapping sequences: a head-to-head comparison of MOLLI, ShMOLLI, SASHA, and SAPHIRE.” *Radiology*, vol. 272, no. 3, 2014, doi: 10.1148/radiol.14140296.
- [59] G. Carr-white, R. Razavi, T. Schaeffter, and E. Nagel, “Native T1 Mapping in Differentiation of Normal Myocardium from Diffuse Disease in Hypertrophic and Dilated Cardiomyopathy,” *JCMG*, vol. 6, no. 4, pp. 475–484, 2013, doi: 10.1016/j.jcmg.2012.08.019.
- [60] R. J. Perea, J. T. Ortiz-perez, and M. Sole, “T1 mapping: characterisation of myocardial interstitial space,” pp. 189–202, 2015, doi: 10.1007/s13244-014-0366-9.
- [61] P. Haaf, P. Garg, D. R. Messroghli, D. A. Broadbent, J. P. Greenwood, and S. Plein, “Cardiac T1 Mapping and Extracellular Volume (ECV) in clinical practice: A comprehensive review,” *Journal of Cardiovascular Magnetic Resonance*, vol. 18, no. 1, pp. 1–12, 2016, doi: 10.1186/s12968-016-0308-4.
- [62] E. Dall’Armellina, V. Ferreira, and S. Neubauer, “T1 Mapping in Ischemic Heart Disease,” *Curr Cardiovasc Imaging Rep*, vol. 7, no. 4, pp. 1–7, 2014, doi: 10.1007/s12410-014-9262-4.
- [63] C. R. Hamilton-craig, “T 1 Mapping for Myocardial Fibrosis by Cardiac Magnetic Resonance Relaxometry — A Comprehensive,” vol. 3, no. March, pp. 1–11, 2017, doi: 10.3389/fcvm.2016.00049.
- [64] C. R. Hamilton-craig, “T 1 Mapping for Myocardial Fibrosis by Cardiac Magnetic Resonance Relaxometry — A Comprehensive,” vol. 3, no. March, pp. 1–11, 2017, doi: 10.3389/fcvm.2016.00049.

- [65] D. Dabir *et al.*, “Reference values for healthy human myocardium using a T1 mapping methodology: results from the International T1 Multicenter cardiovascular magnetic resonance study.,” *J Cardiovasc Magn Reson*, vol. 16, 2014, doi: 10.1186/s12968-014-0069-x.
- [66] D. Dabir *et al.*, “Reference values for healthy human myocardium using a T1 mapping methodology: results from the International T1 Multicenter cardiovascular magnetic resonance study,” *Journal of Cardiovascular Magnetic Resonance*, vol. 16, no. 1, p. 69, 2014, doi: 10.1186/s12968-014-0069-x.
- [67] S. K. Piechnik *et al.*, “Shortened Modified Look-Locker Inversion recovery (ShMOLLI) for clinical myocardial T1- mapping at 1.5 and 3 T within a 9 heartbeat breathhold,” pp. 1–11, 2010.
- [68] M. Fontana *et al.*, “Comparison of T1 mapping techniques for ECV quantification. Histological validation and reproducibility of ShMOLLI versus multibreath-hold T1 quantification equilibrium contrast CMR.,” *J Cardiovasc Magn Reson*, vol. 14, no. 1, p. 88, 2012, doi: 10.1186/1532-429X-14-88.
- [69] A. Meloni *et al.*, “Myocardial T1 Values at 1.5 T: Normal Values for General Electric Scanners and Sex-Related Differences,” *Journal of Magnetic Resonance Imaging*, pp. 1–15, 2021, doi: 10.1002/jmri.27639.
- [70] J. J. Lee *et al.*, “Myocardial T1 and extracellular volume fraction mapping at 3 tesla,” *Journal of Cardiovascular Magnetic Resonance*, vol. 13, no. 1, pp. 1–10, 2011, doi: 10.1186/1532-429X-13-75.
- [71] V. O. Puntmann *et al.*, “Native T1 mapping in differentiation of normal myocardium from diffuse disease in hypertrophic and dilated cardiomyopathy,” *JACC Cardiovasc Imaging*, vol. 6, no. 4, pp. 475–484, 2013, doi: 10.1016/j.jcmg.2012.08.019.
- [72] D. Liu *et al.*, “CMR Native T1 Mapping Allows Differentiation of Reversible Versus Irreversible Myocardial Damage in ST-Segment-Elevation Myocardial Infarction: An OxAMI Study (Oxford Acute Myocardial Infarction),” *Circ Cardiovasc Imaging*, vol. 10, no. 8, pp. 2–11, 2017, doi: 10.1161/CIRCIMAGING.116.005986.
- [73] F. aus dem Siepen *et al.*, “T1 mapping in dilated cardiomyopathy with cardiac magnetic resonance: quantification of diffuse myocardial fibrosis and comparison with endomyocardial biopsy,” *Eur Heart J Cardiovasc Imaging*, vol. 16, no. 2, 2015, doi: 10.1093/ehjci/jeu183.
- [74] C. L. Jellis and D. H. Kwon, “Myocardial T1 mapping: modalities and clinical applications,” vol. 4, no. 2, pp. 126–137, 2014, doi: 10.3978/j.issn.2223-3652.2013.09.03.
- [75] A. A. Robinson, K. Chow, and M. Salerno, “Myocardial T1 and ECV Measurement: Underlying Concepts and Technical Considerations,” *JACC Cardiovasc Imaging*, vol. 12, no. 11P2, pp. 2332–2344, 2019, doi: 10.1016/j.jcmg.2019.06.031.
- [76] H. Mondal and S. Lotfollahzadeh, “Hematocrit,” *StatPearls*, Jan. 2022, Accessed: Sep. 26, 2022. [Online]. Available: <https://www.ncbi.nlm.nih.gov/books/NBK542276/>
- [77] H. H. Billett, “Hemoglobin and Hematocrit,” *Anesthesiology*, vol. 28, no. 4, pp. 763–763, Jul. 1990, doi: 10.1097/00000542-196707000-00028.
- [78] P. Haaf, P. Garg, D. R. Messroghli, D. A. Broadbent, J. P. Greenwood, and S. Plein, “Cardiac T1 Mapping and Extracellular Volume (ECV) in clinical practice: A comprehensive review,” *Journal of Cardiovascular Magnetic Resonance*, vol. 18, no. 1, pp. 1–12, 2016, doi: 10.1186/s12968-016-0308-4.

- [79] E. Y. Yang *et al.*, “Myocardial Extracellular Volume Fraction Adds Prognostic Information beyond Myocardial Replacement Fibrosis,” *Circ Cardiovasc Imaging*, vol. 12, no. 12, pp. 1–14, 2019, doi: 10.1161/CIRCIMAGING.119.009535.
- [80] A. A. Robinson, K. Chow, and M. Salerno, “Myocardial T1 and ECV Measurement: Underlying Concepts and Technical Considerations,” *JACC Cardiovasc Imaging*, vol. 12, no. 11 Pt 2, p. 2332, Nov. 2019, doi: 10.1016/J.JCMG.2019.06.031.
- [81] E.-S. H. Ibrahim *et al.*, “Value CMR: Towards a Comprehensive, Rapid, Cost-Effective Cardiovascular Magnetic Resonance Imaging,” *ISSN 2502-3632 (Online) ISSN 2356-0304 (Paper) Jurnal Online Internasional & Nasional Vol. 7 No.1, Januari – Juni 2019 Universitas 17 Agustus 1945 Jakarta*, vol. 53, no. 9, pp. 1689–1699, 2019.
- [82] J. C. Carr, O. Simonetti, J. Bundy, D. Li, S. Pereles, and J. P. Finn, “Cine MR angiography of the heart with segmented true fast imaging with steady-state precession,” *Radiology*, vol. 219, no. 3, pp. 828–834, 2001, doi: 10.1148/radiology.219.3.r01jn44828.
- [83] D. A. Bluemke, J. L. Boxerman, E. Atalar, and E. R. McVeigh, “Segmented K-space cine breath-hold cardiovascular MR imaging: Part 1. Principles and technique,” *American Journal of Roentgenology*, vol. 169, no. 2, pp. 395–400, 1997, doi: 10.2214/ajr.169.2.9242742.
- [84] C. Petitjean and J. Dacher, “A review of segmentation methods in short axis cardiac MR images,” *Med Image Anal*, vol. 15, no. 2, pp. 169–184, 2011, doi: 10.1016/j.media.2010.12.004.
- [85] T. Lockie, E. Nagel, S. Redwood, and S. Plein, “Use of cardiovascular magnetic resonance imaging in acute coronary syndromes,” *Circulation*, vol. 119, no. 12, pp. 1671–1681, 2009, doi: 10.1161/CIRCULATIONAHA.108.816512.
- [86] D. T. Ginat, M. W. Fong, D. J. Tuttle, S. K. Hobbs, and R. C. Vyas, “Cardiac imaging: Part 1, MR pulse sequences, imaging planes, and basic anatomy,” *American Journal of Roentgenology*, vol. 197, no. 4, pp. 808–815, 2011, doi: 10.2214/AJR.10.7231.
- [87] P. Thunberg, K. Emilsson, P. Rask, and A. Kähäri, “Estimation of ejection fraction and stroke volume using single- and biplane magnetic resonance imaging of the left cardiac ventricle,” *Acta radiol*, vol. 49, no. 9, pp. 1016–1023, 2008, doi: 10.1080/02841850802415544.
- [88] M. Fiechter *et al.*, “Age-related normal structural and functional ventricular values in cardiac function assessed by magnetic resonance,” *BMC Med Imaging*, vol. 13, no. 1, 2013, doi: 10.1186/1471-2342-13-6.
- [89] B. J. Lerman, R. A. Popat, T. L. Assimes, P. A. Heidenreich, and S. M. Wren, “Association of Left Ventricular Ejection Fraction and Symptoms with Mortality after Elective Noncardiac Surgery among Patients with Heart Failure,” *JAMA - Journal of the American Medical Association*, vol. 321, no. 6, pp. 572–579, 2019, doi: 10.1001/jama.2019.0156.
- [90] P. Angaran *et al.*, “Association of Left Ventricular Ejection Fraction with Mortality and Hospitalizations,” *Journal of the American Society of Echocardiography*, vol. 33, no. 7, pp. 802–811.e6, 2020, doi: 10.1016/j.echo.2019.12.016.
- [91] N. Dagres and G. Hindricks, “Risk stratification after myocardial infarction: Is left ventricular ejection fraction enough to prevent sudden cardiac death?” *Eur Heart J*, vol. 34, no. 26, pp. 1964–1971, 2013, doi: 10.1093/eurheartj/eh109.

- [92] B. M. L. Rocha *et al.*, “Risk stratification in hf with mid-range LVEF: the role of cardiopulmonary exercise testing,” *Eur Heart J*, vol. 41, no. Supplement_2, p. 2020, 2020, doi: 10.1093/ehjci/ehaa946.0940.
- [93] N. A. Farrag, J. A. White, and E. Ukwatta, “Semi-automated myocardial segmentation in native T1-mapping CMR using deformable non-rigid registration of CINE images,” *Proc SPIE 10953*, vol. Biomedical, no. March 2019, doi: 10.1117/12.2513054.
- [94] N. A. Farrag, A. Lochbihler, J. A. White, and E. Ukwatta, “Evaluation of fully automated myocardial segmentation techniques in native and contrast-enhanced T1-mapping cardiovascular magnetic resonance images using fully convolutional neural networks,” *Med Phys*, 2020, doi: 10.1002/mp.14574.
- [95] A. S. Flett *et al.*, “Evaluation of techniques for the quantification of myocardial scar of differing etiology using cardiac magnetic resonance,” *JACC Cardiovasc Imaging*, vol. 4, no. 2, pp. 150–156, 2011, doi: 10.1016/j.jcmg.2010.11.015.
- [96] N. R. Pal and S. K. Pal, “A Review on Image Segmentation Techniques,” *Pattern Recognit*, vol. 26, no. 9, pp. 1277–1294, 1993, doi: [https://doi.org/10.1016/0031-3203\(93\)90135-J](https://doi.org/10.1016/0031-3203(93)90135-J).
- [97] F. Usta, “Image Processing Methods for Myocardial Scar Analysis from 3D Late-Gadolinium Enhanced Cardiac Magnetic Resonance Images,” 2018.
- [98] R. Karim *et al.*, “Evaluation of state-of-the-art segmentation algorithms for left ventricle infarct from late Gadolinium enhancement MR images,” *Med Image Anal*, vol. 30, pp. 95–107, 2016, doi: 10.1016/j.media.2016.01.004.
- [99] V. Caselles, R. Kimmel, and G. Sapiro, “Geodesic active contours,” *Int J Comput Vis*, vol. 22, no. 1, pp. 67–79, 1997, doi: 10.1007/3-540-48236-9_4.
- [100] M. Kass, A. Witkin, and D. Terzopoulos, “Snakes: Active contour models,” *Int J Comput Vis*, vol. 1, no. 4, pp. 321–331, 1988, doi: 10.1007/BF00133570.
- [101] T. F. Chan and L. A. Vese, “Adaptive active contours without edges,” *Math Comput Model*, vol. 55, no. 5–6, pp. 1705–1721, 2001, doi: 10.1016/j.mcm.2011.11.014.
- [102] P. Peng, K. Lekadir, A. Gooya, L. Shao, S. E. Petersen, and A. F. Frangi, “A review of heart chamber segmentation for structural and functional analysis using cardiac magnetic resonance imaging,” *Magnetic Resonance Materials in Physics, Biology and Medicine*, vol. 29, no. 2, pp. 155–195, 2016, doi: 10.1007/s10334-015-0521-4.
- [103] G. Dharanibai, A. Chandrasekharan, and Z. C. Alex, “Automated Segmentation of Left Ventricle Using Local and Global Intensity Based Active Contour and Dynamic Programming,” *International Journal of Automation and Computing*, vol. 15, no. 6, pp. 673–688, 2018, doi: 10.1007/s11633-018-1112-4.
- [104] D. Wei, Y. Sun, P. Chai, A. Low, and S. H. Ong, “Myocardial segmentation of late gadolinium enhanced MR images by propagation of contours from cine MR images,” *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 6893 LNCS, no. PART 3, pp. 428–435, 2011, doi: 10.1007/978-3-642-23626-6_53.
- [105] S. Osher and J. A. Sethian, “Fronts Propagating with Curvature Dependent Speed: Algorithms Based on Hamilton-Jacobi Formulations,” *J Comput Phys*, vol. 79, no. 1988, pp. 12–49, 1988, doi: 10.1016/0021-9991(88)90002-3.
- [106] M. Maciejewski, W. Surtel, B. Maciejewska, and T. Małeczka-Massalska, “Level-set image processing methods in medical image segmentation,” *Bio-Algorithms and Med-Systems*, vol. 11, no. 1, pp. 47–51, 2015, doi: 10.1515/bams-2014-0017.

- [107] Y. Boykov and O. Veksler, “Graph Cuts in Vision and Graphics: Theories and Applications BT - Handbook of Mathematical Models in Computer Vision,” *Handbook of Mathematical Models in Computer Vision*, no. Chapter 5, pp. 79–96, 2006.
- [108] E. Ukwatta *et al.*, “Myocardial Infarct Segmentation from Magnetic Resonance Images for Personalized Modeling of Cardiac Electrophysiology,” *IEEE Trans Med Imaging*, vol. 35, no. 6, pp. 1408–1419, 2016, doi: 10.1109/TMI.2015.2512711.
- [109] E. Ukwatta, J. Yuan, W. Qiu, K. C. Wu, N. Trayanova, and F. Vadakkumpadan, “Myocardial infarct segmentation and reconstruction from 2D late-gadolinium enhanced magnetic resonance images,” *Med Image Comput Assist Interv*, vol. 17, pp. 554–561, 2014.
- [110] M. Ghelich Oghli, M. Mohammadzadeh, V. Mohammadzadeh, S. Kadivar, and A. M. Zadeh, “Left ventricle segmentation using a combination of region growing and graph-based method,” *Iranian Journal of Radiology*, vol. 14, no. 2, 2017, doi: 10.5812/iranjradiol.42272.
- [111] X. Lu, X. Chen, W. Li, and Y. Qiao, “Graph cut segmentation of the right ventricle in cardiac MRI using multi-scale feature learning,” *ACM International Conference Proceeding Series*, pp. 231–235, 2019, doi: 10.1145/3309074.3309117.
- [112] J. S. Detsky, G. Paul, A. J. Dick, and G. A. Wright, “Reproducible classification of infarct heterogeneity using fuzzy clustering on multicontrast delayed enhancement magnetic resonance images,” *IEEE Trans Med Imaging*, vol. 28, no. 10, pp. 1606–1614, 2009, doi: 10.1109/TMI.2009.2023515.
- [113] N. A. Farrag and G. A. Wright, “Effect of T1-mapping technique and diminished image resolution on quantification of infarct mass and its ability in predicting appropriate ICD therapy”.
- [114] J. Díez, A. González, and J. C. Kovacic, “Myocardial Interstitial Fibrosis in Nonischemic Heart Disease, Part 3/4: JACC Focus Seminar,” *J Am Coll Cardiol*, vol. 75, no. 17, pp. 2204–2218, May 2020, doi: 10.1016/J.JACC.2020.03.019.
- [115] J. E. van Timmeren, D. Cester, S. Tanadini-Lang, H. Alkadhi, and B. Baessler, “Radiomics in medical imaging— ‘how-to’ guide and critical reflection,” *Insights Imaging*, vol. 11, no. 1, Dec. 2020, doi: 10.1186/S13244-020-00887-2.
- [116] S. Rizzo *et al.*, “Radiomics: the facts and the challenges of image analysis,” *Eur Radiol Exp*, vol. 2, no. 1, Dec. 2018, doi: 10.1186/S41747-018-0068-Z.
- [117] S. H. Lee, H. Park, and E. S. Ko, “Radiomics in Breast Imaging from Techniques to Clinical Applications: A Review,” *Korean J Radiol*, vol. 21, no. 7, p. 779, Jul. 2020, doi: 10.3348/KJR.2019.0855.
- [118] S. Li, Y. Q. Deng, Z. L. Zhu, H. L. Hua, and Z. Z. Tao, “A Comprehensive Review on Radiomics and Deep Learning for Nasopharyngeal Carcinoma Imaging,” *Diagnostics 2021, Vol. 11, Page 1523*, vol. 11, no. 9, p. 1523, Aug. 2021, doi: 10.3390/DIAGNOSTICS11091523.
- [119] M. E. Mayerhoefer *et al.*, “Introduction to Radiomics,” *Journal of Nuclear Medicine*, vol. 61, no. 4, pp. 488–495, Apr. 2020, doi: 10.2967/JNUMED.118.222893.
- [120] S. Li, Y. Q. Deng, Z. L. Zhu, H. L. Hua, and Z. Z. Tao, “A Comprehensive Review on Radiomics and Deep Learning for Nasopharyngeal Carcinoma Imaging,” *Diagnostics 2021, Vol. 11, Page 1523*, vol. 11, no. 9, p. 1523, Aug. 2021, doi: 10.3390/DIAGNOSTICS11091523.

- [121] J. Wang *et al.*, “Radiomic Analysis of Native T1 Mapping Images Discriminates Between MYH7 and MYBPC3-Related Hypertrophic Cardiomyopathy,” 2020, doi: 10.1002/jmri.27209.
- [122] A. S. Antonopoulos *et al.*, “Machine learning of native T1 mapping radiomics for classification of hypertrophic cardiomyopathy phenotypes,” *Scientific Reports 2021 11:1*, vol. 11, no. 1, pp. 1–11, Dec. 2021, doi: 10.1038/s41598-021-02971-z.
- [123] J. Son *et al.*, “Radiomics Feature Analysis Using Native T1 Mapping for Discriminating Between Cardiac Tumors and Thrombi,” *Acad Radiol*, vol. 29, pp. S1–S8, Apr. 2022, doi: 10.1016/J.ACRA.2020.12.009.
- [124] U. Neisius, H. El-Rewaidy, S. Nakamori, J. Rodriguez, W. J. Manning, and R. Nezafat, “Radiomic Analysis of Myocardial Native T1 Imaging Discriminates Between Hypertensive Heart Disease and Hypertrophic Cardiomyopathy,” *JACC Cardiovasc Imaging*, vol. 12, no. 10, pp. 1946–1954, Oct. 2019, doi: 10.1016/J.JCMG.2018.11.024/SUPPL_FILE/MMC1.DOCX.
- [125] S. Chang, K. Han, Y. J. Suh, and B. W. Choi, “Quality of science and reporting for radiomics in cardiac magnetic resonance imaging studies: a systematic review,” *Eur Radiol*, pp. 1–13, Mar. 2022, doi: 10.1007/S00330-022-08587-9/TABLES/5.
- [126] C. Izquierdo *et al.*, “Radiomics-Based Classification of Left Ventricular Non-compaction, Hypertrophic Cardiomyopathy, and Dilated Cardiomyopathy in Cardiovascular Magnetic Resonance,” *Front Cardiovasc Med*, vol. 0, p. 1502, Oct. 2021, doi: 10.3389/FCVM.2021.764312.
- [127] E. Avard *et al.*, “Non-contrast Cine Cardiac Magnetic Resonance image radiomics features and machine learning algorithms for myocardial infarction detection,” *Comput Biol Med*, vol. 141, Feb. 2022, doi: 10.1016/J.COMPBIOMED.2021.105145.
- [128] T. di Noto *et al.*, “Radiomics for distinguishing myocardial infarction from myocarditis at late gadolinium enhancement at mri: Comparison with subjective visual analysis,” *Radiol Cardiothorac Imaging*, vol. 1, no. 5, Dec. 2019, doi: 10.1148/RYCT.2019180026/SUPPL_FILE/RYCT180026SUPPA1.PDF.
- [129] H. A. Leopold, A. Singh, S. Sengupta, and V. Lakshminarayanan, “Deep learning for ophthalmology using optical coherence tomography,” *State of the Art in Neural Networks and Their Applications: Volume 1*, pp. 239–269, Jan. 2021, doi: 10.1016/B978-0-12-819740-0.00012-7.
- [130] R. M. Haralick, I. Dinstein, and K. Shanmugam, “Textural Features for Image Classification,” *IEEE Trans Syst Man Cybern*, vol. SMC-3, no. 6, pp. 610–621, 1973, doi: 10.1109/TSMC.1973.4309314.
- [131] J. J. M. van Griethuysen *et al.*, “Focus on Computer Resources Computational Radiomics System to Decode the Radiographic Phenotype”, doi: 10.1158/0008-5472.CAN-17-0339.
- [132] A. H. Curiale, F. D. Colavecchia, P. Kaluza, R. A. Isoardi, and G. Mato, “Automatic myocardial segmentation by using a deep learning network in cardiac MRI,” *2017 43rd Latin American Computer Conference, CLEI 2017*, vol. 2017-Janua, pp. 1–6, 2017, doi: 10.1109/CLEI.2017.8226420.
- [133] E. Shelhamer, J. Long, and T. Darrell, “Fully Convolutional Networks for Semantic Segmentation,” *IEEE Trans Pattern Anal Mach Intell*, vol. 39, no. 4, pp. 640–651, 2017, doi: 10.1109/TPAMI.2016.2572683.
- [134] O. Ronneberger, P. Fischer, and T. Brox, “U-Net: Convolutional Networks for Biomedical Image Segmentation,” pp. 1–8.

- [135] G. Huang and K. Q. Weinberger, “Densely Connected Convolutional Networks”.
- [136] Z. Zhou, M. M. Rahman Siddiquee, N. Tajbakhsh, and J. Liang, “Unet++: A nested u-net architecture for medical image segmentation,” *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 11045 LNCS, pp. 3–11, 2018, doi: 10.1007/978-3-030-00889-5_1.
- [137] F. Sultana, A. Sufian, and P. Dutta, “Evolution of Image Segmentation using Deep Convolutional Neural Network: A Survey,” *Knowl Based Syst*, vol. 201–202, pp. 1–38, 2020, doi: 10.1016/j.knosys.2020.106062.
- [138] N. S. Punn and S. Agarwal, “Modality specific U-Net variants for biomedical image segmentation: a survey,” *Artif Intell Rev*, vol. 55, no. 7, p. 5845, Oct. 2022, doi: 10.1007/S10462-022-10152-1.
- [139] O. Oktay *et al.*, “Attention U-Net: Learning Where to Look for the Pancreas,” no. Midl, 2018.
- [140] J. Long, E. Shelhamer, and T. Darrell, “Fully Convolutional Networks for Semantic Segmentation”.
- [141] D. Nie, L. Wang, Y. Gao, and D. Sken, “FULLY CONVOLUTIONAL NETWORKS FOR MULTI-MODALITY ISOINTENSE INFANT BRAIN IMAGE SEGMENTATION,” *Proceedings / IEEE International Symposium on Biomedical Imaging: from nano to macro. IEEE International Symposium on Biomedical Imaging*, vol. 2016, p. 1342, Jun. 2016, doi: 10.1109/ISBI.2016.7493515.
- [142] H. Seo *et al.*, “Machine learning techniques for biomedical image segmentation: An overview of technical aspects and introduction to state-of-art applications,” *ArXiv*, pp. 1–35, 2019.
- [143] L. K. Tan, R. A. Mclaughlin, E. Lim, Y. Faridah, A. Aziz, and Y. M. Liew, “Fully Automated Segmentation of the Left Ventricle in Cine Cardiac MRI Using Neural Network Regression,” 2018, doi: 10.1002/jmri.25932.
- [144] T. Anh, Z. Lu, and G. Carneiro, “Combining deep learning and level set for the automated segmentation of the left ventricle of the heart from cardiac cine magnetic resonance,” vol. 35, pp. 159–171, 2017, doi: 10.1016/j.media.2016.05.009.
- [145] P. V. Tran, “A Fully Convolutional Neural Network for Cardiac Segmentation in Short-Axis MRI,” pp. 1–21, 2016.
- [146] F. Zabihollahy, J. A. White, and E. Ukwatta, “Fully automated segmentation of left ventricular myocardium from 3D late gadolinium enhancement magnetic resonance images using a U-net convolutional neural network-based model,” *Proc. SPIE 10950*, no. Computer-Aided Diagnosis, 2019, doi: 10.1117/12.2512146.
- [147] E. de la Rosa, D. Sidibé, T. Decourselle, T. Leclercq, A. Cochet, and A. Lalande, “Myocardial infarction quantification from late gadolinium enhancement mri using top-hat transforms and neural networks,” *ArXiv*, 2019.
- [148] A. S. Fahmy *et al.*, “Three-dimensional deep convolutional neural networks for automated myocardial scar quantification in hypertrophic cardiomyopathy: A multicenter multivendor study,” *Radiology*, vol. 294, no. 1, pp. 52–60, 2020, doi: 10.1148/radiol.2019190737.
- [149] A. S. Fahmy, H. El-rewaidy, M. Nezafat, S. Nakamori, and R. Nezafat, “Automated analysis of cardiovascular magnetic resonance myocardial native T 1 mapping images using fully convolutional neural networks,” vol. 1, pp. 1–12, 2019.

- [150] F. Zabihollahy, J. A. White, and E. Ukwatta, “Fully automated segmentation of left ventricular myocardium from 3D late gadolinium enhancement magnetic resonance images using a U-net convolutional neural network-based model,” *Proc. SPIE 10950*, no. Computer-Aided Diagnosis, 2019, doi: 10.1117/12.2512146.
- [151] A. S. Fahmy *et al.*, “Automated Cardiac MR Scar Quantification in Hypertrophic Cardiomyopathy Using Deep Convolutional Neural Networks,” *JACC Cardiovasc Imaging*, vol. 11, no. 12, pp. 1917–1918, 2018, doi: 10.1016/j.jcmg.2018.04.030.
- [152] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 2016-Decem, pp. 770–778, 2016, doi: 10.1109/CVPR.2016.90.
- [153] Z. Luo, A. Jara, and W. Ou, “Image Segmentation using U-Net, DenseNet and CapsuleNet”.
- [154] L. R. Dice, “Measures of the Amount of Ecologic Association Between Species Author (s): Lee R. Dice Published by: Wiley on behalf of the Ecological Society of America Stable URL: <https://www.jstor.org/stable/1932409>,” *Ecology*, vol. 26, no. 3, pp. 297–302, 1945.
- [155] P. Sonego and S. Ndor Pongor, “ROC analysis: applications to the classification of biological sequences and 3D structures”, doi: 10.1093/bib/bbm064.
- [156] K. Hajian-Tilaki, “Receiver Operating Characteristic (ROC) Curve Analysis for Medical Diagnostic Test Evaluation,” 2013.

Chapter 3: Fully Automated Segmentation of the Left Ventricular Myocardium in Native and Contrast-enhanced T1-mapping and CINE CMR Images Using Fully Convolutional Neural Networks

3.1 Introduction

MF presents itself differently in various forms of cardiomyopathy in both native and post-contrast T1-mapping CMR images. As such, it is possible to evaluate and distinguish between different disease types based on native and post-contrast T1 values. Isolating the myocardium as a whole is an essential first step to this aim in order for parametric values and measurements to be derived from all myocardial pixels. However, myocardial isolation can be a particularly challenging task in T1-mapping CMR images wherein contrast differences between the myocardium and outside structures are not as easily appreciable as compared to other imaging types. In this work, we aim to mitigate the challenges of myocardial isolation in both native and post-contrast enhanced T1-mapping CMR images acquired using the shMOLLI imaging sequence by establishing a technique for automated myocardial segmentation using a fully convolutional neural network based on the renowned U-Net architecture. We compare this technique to an alternative registration-based automated segmentation method which eludes the issue of poor contrast differentiation, wherein spatially corresponding CINE images are segmented automatically using U-Net, and a non-rigid registration technique transforms and propagates CINE contours to the myocardial regions of T1-maps.

3.2 Selection of study population and image acquisition

The Cardiovascular Imaging Registry of Calgary (CIROC) was used to retrospectively select sixty (60) patients with known MF, visually confirmed by LGE CMR imaging. Participants were prospectively recruited using written informed consent at the time of referral to the Stephenson Cardiac Imaging Centre, Libin Cardiovascular Institute, in Calgary, AB, Canada. The study was approved by the Conjoint Health Research Ethics Board at University of Calgary.

Subjects with both a confirmed clinical diagnosis and cardiac MRI-based diagnosis of a cardiomyopathy state commonly associated with MF were selected for this study. Twenty subjects, each with ischemic, dilated and hypertrophic cardiomyopathy (ICM, DCM and HCM, respectively) were retrospectively identified on the basis of clearly visualized signal enhancement of the myocardium on LGE imaging in ≥ 2 contiguous slices. ICM was defined as presence of systolic dysfunction (LVEF $< 50\%$) with known obstructive ($\geq 70\%$ lesion in ≥ 1 vessel) coronary artery disease (CAD) by invasive catheterization. DCM was defined as systolic dysfunction with dilation (LV volume $\geq 2SD$ reference value) with no CAD by catheterization. All HCM subjects had unexplained LV wall thickness $\geq 15\text{mm}$ with positive family history or positive genetic markers. Since this study was performed retrospectively, we were not able to include any normal group in this work for comparison against the HCM, DCM and ICM groups for this particular T1 mapping pulse sequence. All patients underwent a standardized imaging protocol using a Siemens 3T magnet (Prisma or Skyra, Siemens Healthineers, Erlangen), inclusive of short and long-axis CINE and LGE imaging using SSFP and inversion-recovery gradient echo (IR-GRE) pulse sequences, respectively. LGE imaging was performed 10 minutes

following the intravenous administration of 0.1 mmol/kg Gadovist (Bayer Inc., Canada). Short axis native (i.e. pre-contrast) and post-contrast T1-mapping was performed in short axis using the shMOLLI [57] technique at the basal, mid and apex positions of the left ventricle. Native and contrast-enhanced T1-maps were generated from 6 or 7 raw shMOLLI images (obtained at varying inversion times) using commercially available software cvi42 (Circle Cardiovascular Imaging, Calgary, Canada). The average image parameters for CINE and shMOLLI acquisitions are demonstrated in Table 3-1.

Table 3-1 Image acquisition parameters of CINE and shMOLLI pulse sequences

Table 3-1		
Average image acquisition parameters for CINE and shMOLLI sequences.		
	CINE	shMOLLI
Slice thickness (mm)	8	8
TR (ms)	39.00	388.65
TE (ms)	1.46	1.07
Flip Angle (°)	51.78	35
Percent FOV (%)	83.89	81.30
Percent Sampling (%)	58.62	99.44

The number of acquired native and contrast-enhanced T1-maps was not standard between patients, i.e., between one and three maps were acquired at the apical, basal and/or mid-level of the LV, per patient. Six native T1-maps were excluded from the dataset due to motion and/or respiration (1 mid-level DCM; 1 mid-level HCM), image artifact (1 mid-level ICM), or excessive noise (1 mid-level HCM and 2 apical HCM), totaling 125 native T1-maps overall. Eleven patients did not undergo contrast-enhanced T1-mapping imaging (4 ICM, 7 DCM), and one contrast-enhanced T1-map was removed from the dataset due

to excessive noise (1 apical HCM). In total, 100 contrast-enhanced T1-maps were available across all patient groups.

To account for the small data set size in this study, training and testing of our methodologies was performed using five-fold cross validation. In this approach, native and contrast enhanced T1 maps were divided into five subsets, wherein four of those subsets were used for training and the remaining subset used for testing. This process was repeated five times, each with a different subset for testing, such that the entire dataset could be used in evaluation of the methodologies.

3.3 Method 1: Direct segmentation of native and contrast-enhanced T1 maps using U-Net based CNN model

Myocardial segmentation of native and contrast-enhanced T1-maps was achieved using a CNN based on the U-Net architecture.[112] A series of pre-processing steps were implemented prior to dividing the dataset into five subsets, as previously described. Following preprocessing of native and contrast-enhanced T1-maps, the U-Net-based CNN model was trained separately for native T1-maps and contrast-enhanced T1-maps. The output myocardial masks for the testing subsets were validated through comparison against contours manually delineated by an expert by using conventional accuracy metrics, i.e., Dice similarity coefficient (DSC), relative error index, precision, recall and boundary F1 score. Output myocardial masks were also used to compute global (average) myocardial T1-values, which were compared to global T1-values generated from manual contours. Global T1-values were validated using Pearson correlation and Bland-Altman analysis. A diagram of the myocardial segmentation framework is demonstrated in Figure 3-1.

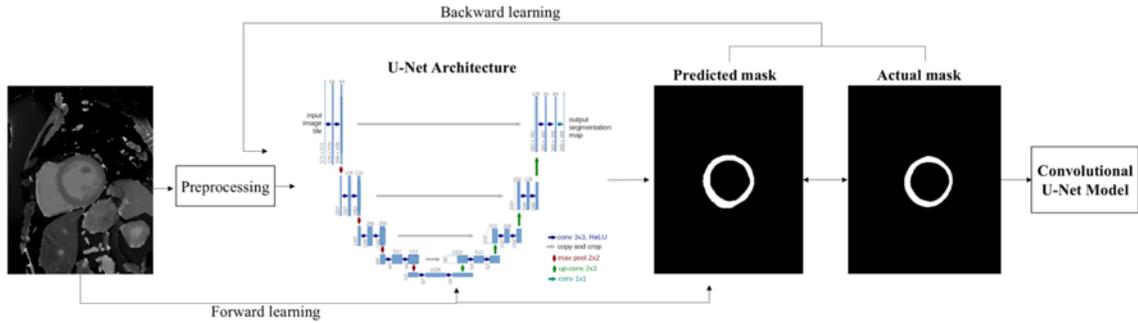


Figure 3-1 Framework of CNN-based automated myocardial segmentation steps

3.3.1 Preprocessing

Dimensions of native and contrast-enhanced T1-maps were not standardized between patients. The dimensions of images were required to be a factor of 16 in both length and width in order to be compatible with the U-Net architecture. Therefore, the dimensions of native and contrast-enhanced T1-maps were cropped and/or padded to a standard size prior to implementing the training procedure. Dimensions of all images were 336x272 in length and width, respectively, after cropping and padding. Automated resizing of T1-maps was implemented using MATLAB r2020a.

3.3.2 Training and testing procedure of U-Net-based CNN model

Myocardial segmentation of native and contrast enhanced T1-maps was achieved using two separate CNN models based on the standard U-Net architecture [112]. Hyperparameters were chosen based on the results of a random grid search in order to optimize the model. The kernel size of convolution layers in the standard model was size 3x3. Filter sizes were increased at 64, 128, 256, 512, and 1024, and then decreased at the same interval back to 64 across convolution layers. The convolutional layers used an ELU activation function. In the shrinking path, max pooling of size 2x2 was used in order to reduce the dimensions of the image and allow for the collection of the learned features. The expanding path contained up-sampling layers in order to bring the image back up to

its original dimensions while applying the learned features along the path [112]. The final output of the model was produced using a Sigmoid function that created a probability map from 0-1. A batch normalization layer was added before each convolutional layer to regularize the data and force new feature learning [124]. Data augmentation was also implemented, wherein training images were flipped, rotated, and shifted to create 1000 unique images per epoch. Once the model was optimized, it was trained by passing in the respective images and corresponding manual myocardial segmentation mask, as demonstrated in Figure 3-1. The model was trained for 100 epochs allowing for the loss metric (i.e., Dice similarity coefficient) to plateau and achieve its optimal score. The training procedure took 12 hours to complete. Training of the CNN model was implemented using the Python libraries Keras and Tensorflow, and then run on a Tesla K80 graphics processing unit (GPU) on the Shared Hierarchical Research Computing Network (SHARCNET).

We compared the results of the base U-Net model described above to two alternative state-of-the-art CNN-based segmentation methods, Attention Net [125] and Dense Net [126]. Attention gates are integrated into the base U-Net model in the Attention Net, such that the algorithm can learn to focus on target structures and suppress irrelevant regions in a given input space [125]. The Dense U-Net concatenates the outputs of all preceding layers for the input of every following layer [126]. Therefore, a “collective knowledge” is being passed on to all future layers which may improve output. The Dense Net and Attention Net were trained using the same optimizer and learning rate for native T1-maps, contrast enhanced T1-maps and CINE images. Through experimental testing, it was found that the base U-Net yielded significantly higher DSC than the Dense U-Net and

Attention U-Net in all image types and was therefore the chosen method for further analyses on global T1 values.

Following the training procedure, the optimized base U-Net-based CNN model was used to segment the myocardium on the testing datasets of the native T1-maps and contrast-enhanced T1-maps. The U-Net-based model produced a binary mask representing the predicted myocardial region, with which accuracy metrics could be evaluated with respect to corresponding manual myocardial contours.

3.4 Method 2: Segmentation of native and contrast-enhanced T1-maps using alternative technique based on registration and propagation of CINE contours

To further validate the U-Net-based myocardial segmentation technique for native and contrast-enhanced T1-maps, we also implemented an alternative myocardial segmentation technique based on prior CINE data for comparison. In this technique, automated myocardial contours from CINE images are registered and propagated to T1-maps based on a Modality Independent Neighborhood Descriptor (MIND).[127] This CINE registration-based pipeline is demonstrated in Figure 3-2.

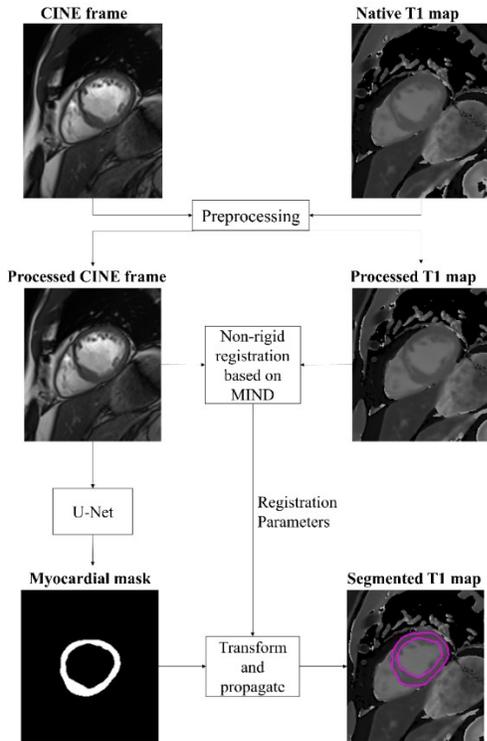


Figure 3-2 Framework of CINE registration-based segmentation technique
 A CNN based on the U-Net architecture is used to segment the myocardium in CINE images. Contours generated using this technique are then transformed and propagated to native (pictured) and post-contrast T1-mapping images.

3.4.1 Preprocessing

The native and contrast-enhanced T1-maps were limited to a single slice per acquisition in short-axis. Thus, in order to compare the U-Net-based methodology to the CINE registration-based myocardial segmentation technique, a single, end-diastolic frame of the CINE sequence was chosen in short-axis at the slice location corresponding to each T1-map for preprocessing. Preprocessing steps included CINE interpolation, in order to match the pixel size of T1-maps, followed by an automated resizing procedure such that dimensions were a factor of 16 in order to be standardized to the dimensions of each T1-map and inherently be compatible with the U-Net architecture. Interpolation of CINE images and automatic image resizing procedures were implemented using MATLAB r2020a.

3.4.2 Automated myocardial segmentation of CINE images

Automated myocardial segmentation in CINE images was achieved using the same U-Net-based CNN model described in Section 2.2. The CINE images were divided for training and testing using five-fold cross validation, as previously described.

The CINE and shMOLLI pulse sequences were acquired within the same examination, but at different points in time. Therefore, in order to take into account inadvertent patient motion from one imaging sequence to the next, we implement a non-rigid image registration technique to align CINE contours to corresponding (slice-matched) T1-maps.

3.4.3 Propagation of myocardial contours from CINE images to T1-maps using deformable non-rigid registration

The MIND [127] technique was used to register T1-maps to corresponding (slice-matched) CINE frames. This technique is based on the concept of self-similarity and aims to find an image descriptor, based on image patches, which represents the distinctive image structure in a local neighborhood. The MIND is defined generally by a distance Dp , variance estimate V and spatial search region R as follows:

$$MIND(I, x, r) = \frac{1}{n} \exp\left(\frac{-Dp(I, x, x+r)}{V(I, x)}\right) r \hat{1} R \quad \text{Equation 3-1}$$

where n is a normalization constant and $r \hat{1} R$ defines the search region. The MIND algorithm represents images as a vector of the size $|R|$ at each location x . The distance metric Dp , or distance between patches, is defined as the point-wise sum of squared distances (SSD) of all voxels between two patches, which is computed in a computationally inexpensive manner as follows:

$$Dp(I, x, x + r) = C * (I - I'(r))^2 \quad \text{Equation 3-2}$$

where C is a convolution filter of size $(2p + 1)^d$ and $I'(r)$ is a copy of the image I translated by r . The convolution filter C to the difference between the original image I and the image translated by r on a pixel-wise basis (sum of squares). Computation of the distance metric is repeated for all positions in \hat{R} . The variance estimate $V(I, x)$ represents the average SSD in a small neighborhood N around pixel x as follows:

$$V(I, x) = \frac{1}{\text{size}(N)} \sum_{n \in N} Dp(I, x, x + n) \quad \text{Equation 3-3}$$

In the framework, a six-neighborhood $n \in N$ and patch size of 3 is used.

In our implementation of MIND registration, we chose to register the shMOLLI map directly with the CINE image, as opposed to registering CINE with a single phase of the raw shMOLLI data as previously [88]. We define the shMOLLI map as the target and CINE as the moving image.

The MIND algorithm creates a matrix of vectors of the size $|R|$ at each pixel location, which demonstrates the deformation of the CINE image relative to the structure of the shMOLLI image. The MIND registration approach was implemented individually for native T1-maps and contrast enhanced T1-maps of the same patient, since these maps were acquired at different points in time. Thus, a different matrix of vectors was created following MIND registration for native and contrast-enhanced T1-maps, respectively, and separate deformation/propagation of CINE contours was performed for native and contrast-enhanced T1-maps based on these matrices.

After deformation of CINE contours to the T1-maps, a morphological closing procedure was implemented to smooth the myocardial contours. A disk-shaped structuring element with a radius of 9 pixels was used for this application. MIND registration,

propagation of CINE contours, and closing morphological operations were implemented using MATLAB r2019a.

An alternative approach to this technique may be considered, wherein a conventional image processing algorithm, such as active contours, can be used in place of the U-Net for the CINE segmentation step. This approach is discussed in Appendix A.

3.5 Computation of native and post-contrast global myocardial T1 values

As the clinical goal of T1-mapping segmentation is directed toward accurate identification and quantification of myocardial fibrosis, we aimed to demonstrate this application by computing the global T1-values within the myocardial regions segmented by (a) direct application of the U-Net-based CNN on T1-maps, (b) MIND-based propagation of CINE contours, and (c) manual delineation. After masking the native and contrast-enhanced T1-maps by myocardial contours generated using the above three methodologies, the mean pixel value was computed within each mask to express a global T1 value.

3.6 Results

Output myocardial masks were compared to manual masks delineated by an expert. As mentioned in Section 2.2, the Dense U-Net and Attention U-Net were also employed for comparison to the base U-Net model. Our results indicated that the base U-Net model outperformed these adapted models for myocardial segmentation in both native and contrast enhanced T1-maps. Likewise, the base U-Net model also outperformed Dense Net and Attention Net in myocardial segmentation of CINE images. While Dense Net and Attention Net have been found to perform well in several biomedical image segmentation problems, in our particular dataset, the desired segmentation region is small and a relatively simple shape. Thus, dense skip connections are likely not required as the desired output is

not complex. On the other hand, since the desired output segmentation is not complex, Attention Net may have over-simplified the feature space, reducing accuracy.

We demonstrate accuracy metrics of the base U-Net, Dense Nets and Attention Nets in direct myocardial segmentation of native T1-maps, contrast enhanced T1-maps and CINE images in Table 3-2. The U-Net based segmentation model had the highest performance in CINE images and lowest in contrast enhanced T1-maps. The U-Net model failed in segmentation of 7 contrast enhanced T1-maps and 1 native T1-map and was successful in segmentation of all CINE images. The Dense Net model failed (i.e., DSC < 10%) in segmentation of 2 native T1-maps, 3 contrast enhanced T1-maps and 6 CINE images. The Attention Net model failed in 1 native T1-map, 5 contrast enhanced T1-maps and 2 CINE images.

Our comparative technique, based on deformation of CINE contours from a MIND[127] registration algorithm, utilized prior myocardial segmentations from CINE images. Due to the suboptimal performance of CINE myocardial segmentation using Dense Nets and Attention Nets as demonstrated in Table 3-2, MIND registration of CINE contours derived using these techniques was not performed for the analysis of our alternative registration-based myocardial segmentation method.

Table 3-2 Results of CNN-based myocardial segmentation techniques in T1-maps

Performance of automated myocardial segmentations generated by base U-Net, Dense Net and Attention Net in native T1-maps, contrast enhanced T1-maps and CINE images across different patient populations. The values are expressed as average \pm standard deviation across all images. Myocardial segmentations generated by the CNN models are compared to contours manually delineated by an expert.

Native T1-maps						
Patient group	CNN model	DSC (%)	Rel. Error (%)	Precision (%)	Recall (%)	BF Score (%)
DCM group (n=20)	U-Net	83.8 ± 5.8	14.7 ± 13	92.7 ± 9.0	94.0 ± 8.2	93.3 ± 8.3
	Dense Net	64.5 ± 28	34.9 ± 34	67.3 ± 34	91.9 ± 10	76.5 ± 25
	Attention Net	73.7 ± 18	25.3 ± 25	78.7 ± 25	88.9 ± 12	82.1 ± 21
HCM group (n=20)	U-Net	81.1 ± 17	16.25 ± 21	83.00 ± 22	85.93 ± 18	84.02 ± 20
	Dense Net	59.5 ± 27	63.2 ± 57	61.9 ± 31	59.3 ± 22	59.5 ± 23
	Attention Net	70.5 ± 24	31.9 ± 28	70.6 ± 28	81.4 ± 18	75.1 ± 23
ICM group (n=20)	U-Net	83.8 ± 4.9	9.69 ± 7.9	91.23 ± 11	91.83 ± 9.3	91.41 ± 9.6
	Dense Net	62.2 ± 31	35.5 ± 35	63.1 ± 35	82.6 ± 21	68.1 ± 31
	Attention Net	60.3 ± 26	41.3 ± 32	59.1 ± 31	83.7 ± 15	67.1 ± 27
All groups (n=60)	U-Net	82.7 ± 12	13.79 ± 16	88.38 ± 16	90.1 ± 13	89.0 ± 15
	Dense Net	61.8 ± 28	46.4 ± 47	63.9 ± 33	76.1 ± 23	67.2 ± 27
	Attention Net	68.5 ± 23	32.8 ± 29	69.28 ± 29	84.4 ± 16	74.5 ± 24
Contrast enhanced T1-maps						
Patient group	CNN model	DSC (%)	Rel. Error (%)	Precision (%)	Recall (%)	BF Score (%)
DCM group (n=20)	U-Net	77.5 ± 13	19.7 ± 18	81.2 ± 18	85.5 ± 13	82.7 ± 15
	Dense Net	69.7 ± 20	37.2 ± 29	65.9 ± 22	73.4 ± 14	70.5 ± 15
	Attention Net	71.1 ± 22	34.5 ± 30	72.1 ± 26	79.5 ± 17	76.5 ± 19
HCM group (n=20)	U-Net	78.9 ± 20	20.2 ± 25	74.95 ± 24	79.2 ± 16	75.1 ± 23
	Dense Net	69.8 ± 22	32.9 ± 26	62.2 ± 25	68.9 ± 19	63.6 ± 22
	Attention Net	75.5 ± 22	33.1 ± 28	64.3 ± 21	62.6 ± 17	63.5 ± 17
ICM group (n=20)	U-Net	65.5 ± 18	31.7 ± 25	60.50 ± 23	71.5 ± 16	64.13 ± 20
	Dense Net	57.0 ± 27	38.8 ± 32	53.5 ± 27	67.3 ± 15	59.7 ± 20
	Attention Net	50.0 ± 30	45.8 ± 32	45.6 ± 28	64.0 ± 22	55.4 ± 24
All groups (n=60)	U-Net	74.1 ± 18	23.79 ± 23	72.11 ± 23	78.6 ± 16	73.8 ± 21
	Dense Net	66.3 ± 23	35.6 ± 28	60.5 ± 25	69.5 ± 17	64.1 ± 20
	Attention Net	65.7 ± 24	37.0 ± 29	60.7 ± 26	66.8 ± 20	64.3 ± 21
CINE images						
Patient group	CNN model	DSC (%)	Rel. Error (%)	Precision (%)	Recall (%)	BF Score (%)
DCM group (n=20)	U-Net	81.6 ± 4.9	9.00 ± 7.3	93.1 ± 6.9	93.8 ± 5.4	93.4 ± 5.8
	Dense Net	57.3 ± 36	39.1 ± 40	65.8 ± 41	89.3 ± 25	82.2 ± 29
	Attention Net	71.8 ± 25	23.9 ± 25	82.7 ± 25	94.1 ± 8.0	85.7 ± 21
HCM group (n=20)	U-Net	84.2 ± 7.0	11.1 ± 8.7	89.9 ± 8.5	90.9 ± 6.5	90.3 ± 7.2
	Dense Net	66.7 ± 32	31.0 ± 36	68.2 ± 36	84.9 ± 20	76.5 ± 28
	Attention Net	78.8 ± 19	19.2 ± 22	81.5 ± 22	89.0 ± 10	86.5 ± 12

ICM group (n=20)	U-Net	81.9 ± 7.0	9.19 ± 7.8	93.4 ± 9.5	95.8 ± 7.5	55.4 ± 24
	Dense Net	68.6 ± 19	30.2 ± 26	75.9 ± 24	95.0 ± 7.9	82.2 ± 18
	Attention Net	62.4 ± 24	35.7 ± 31	68.7 ± 28	92.1 ± 11	75.0 ± 24
All groups (n=60)	U-Net	83.42 ± 5.1	9.51 ± 7.4	92.6 ± 7.3	93.4 ± 6.2	92.9 ± 6.5
	Dense Net	64.6 ± 30	33.1 ± 34	69.9 ± 34	89.4 ± 19	79.9 ± 26
	Attention Net	71.7 ± 23	25.6 ± 27	77.9 ± 26	91.4 ± 9.9	82.7 ± 20

The MIND registration procedure and propagation of CINE contours derived from the base U-Net model took approximately 2.5 s per image in native and contrast-enhanced T1-maps. Resultant CINE-registered contours were compared to manual contours delineated by an expert. Table 3-3 demonstrates average accuracy metrics generated by the CINE registration-based myocardial segmentation technique in native and contrast-enhanced T1-maps.

Table 3-3 Results of CINE-registration based segmentation technique in T1-maps

Myocardial segmentations generated by the CINE-registration approach are compared to respective manual contours of the T1-maps.

Patient group	DSC (%)	Rel. Error (%)	Precision (%)	Recall (%)	BF Score (%)
Native T1-maps					
DCM group (n=20)	79.4 ± 6.3	13.8 ± 11	88.8 ± 9.8	88.1 ± 9.5	88.4 ± 9.5
HCM group (n=20)	83.6 ± 6.8	12.0 ± 9.0	84.6 ± 11	84.0 ± 11	84.3 ± 11
ICM group (n=20)	80.3 ± 7.0	13.44 ± 9.0	88.6 ± 10	89.1 ± 9.1	88.7 ± 9.3
All groups (n=60)	81.4 ± 6.9	12.96 ± 9.5	87.0 ± 10	86.8 ± 10	86.8 ± 10.2
Contrast enhanced T1-maps					
DCM group (n=20)	78.4 ± 7.2	16.2 ± 9.3	85.6 ± 11	85.6 ± 11	85.5 ± 11
HCM group (n=20)	78.5 ± 11	15.8 ± 11	74.2 ± 16	73.6 ± 16	73.8 ± 16
ICM group (n=20)	73.2 ± 7.3	23.0 ± 13	72.4 ± 11	74.6 ± 11	73.4 ± 11
All groups (n=60)	77.0 ± 9.6	19.3 ± 12	76.4 ± 14	76.6 ± 14	76.4 ± 14

As demonstrated in Tables 3-2 and 3-3, the CINE-registration segmentation method performs closely to the base U-Net in native T1-maps and outperforms U-Net in contrast enhanced T1-maps.

To further establish validity of the U-Net based myocardial segmentation technique, we performed a Wilcoxon signed rank test with respect to paired DSC values generated using the direct U-Net-based approach and the CINE-registration approach. A statistically significant difference between paired DSC values computed using these techniques was demonstrated in both native T1-maps ($P < 0.0001$) and contrast enhanced T1-maps ($P = 0.0014$). Figures 3-3 and 3-4 demonstrate resultant myocardial segmentations using the direct U-Net-based approach and the CINE-registration approach in native and contrast-enhanced T1-maps, respectively. The distribution of DSC metrics computed from the U-Net based method and CINE-registration based method for native T1-maps and contrast enhanced T1-maps are demonstrated in Figure 3-5.

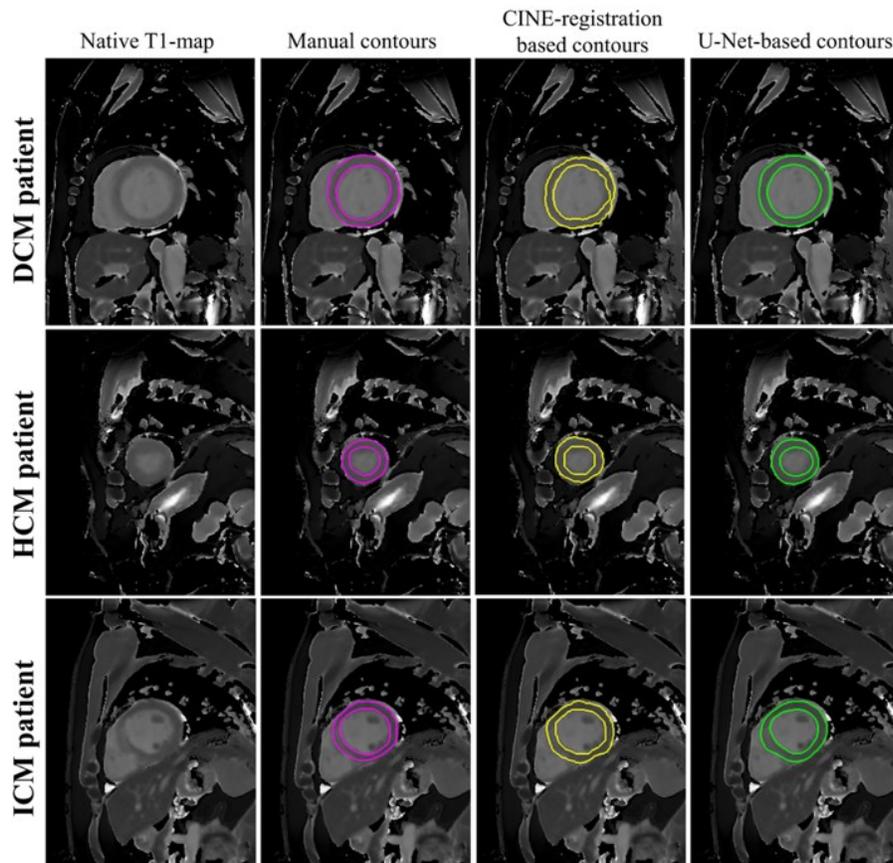


Figure 3-3 Segmentation results U-Net vs. CINE-registration methods in native T1-maps. Myocardial contours of native T1-maps generated using automated techniques in patients with DCM, HCM and ICM. Original native T1-maps are shown in the first column. Manual contours delineated by an expert

are shown in magenta in column 2. The third column demonstrates contours generated by the CINE-registration based approach in yellow, while the last column demonstrates contours generated by the U-Net-based algorithm, shown in green.

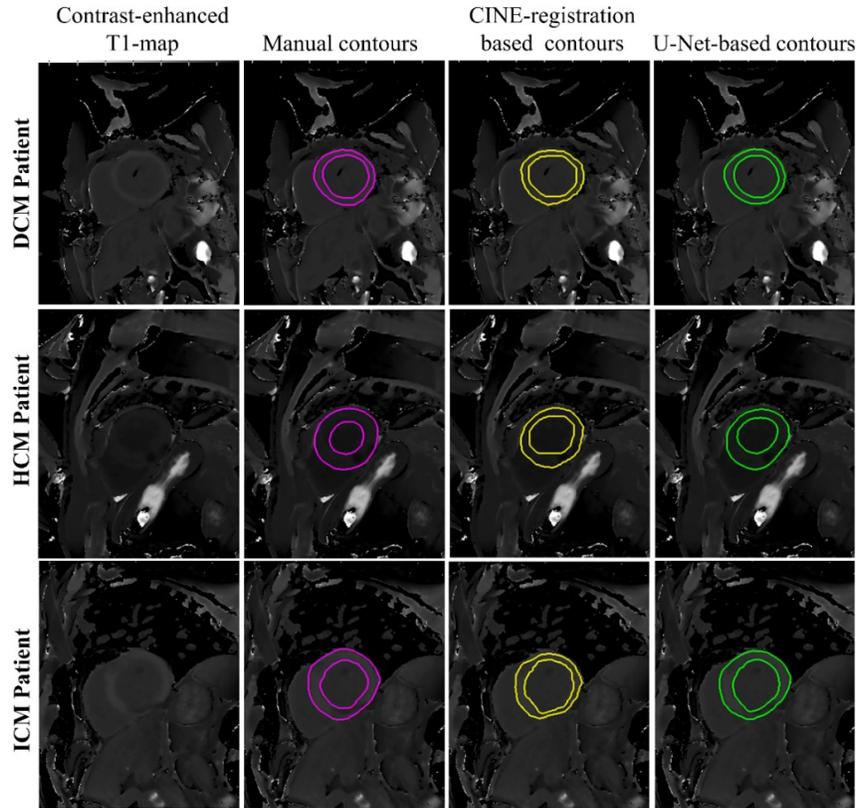


Figure 3-4 Segmentation results U-Net vs. CINE-registration methods in post-contrast T1-maps. Myocardial contours of post-contrast T1-maps generated using automated techniques in patients with DCM, HCM and ICM. Original contrast-enhanced T1-maps are shown in the first column. Manual contours delineated by an expert are shown in magenta in column 2. The third column demonstrates contours generated by the CINE-registration based approach in yellow, while the last column demonstrates contours generated by the U-Net-based algorithm, shown in green.

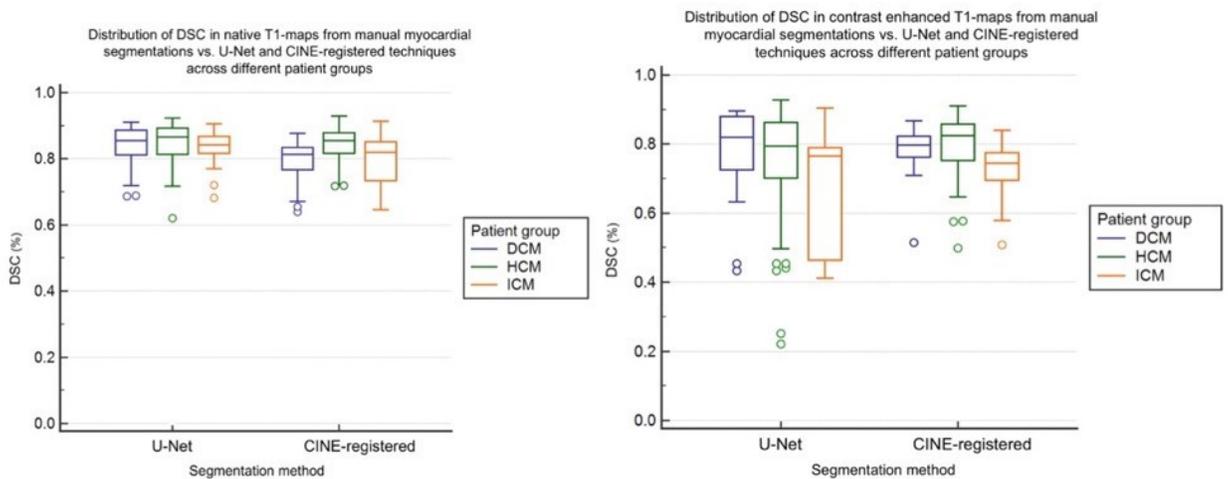


Figure 3-5 Box and whisker plots of DSC metric for U-Net and CINE registration segmentation techniques. Plots of the DSC metric generated by U-Net-based and CINE registration-based methods in native T1-maps (left) and contrast-enhanced T1-maps (right) are shown, separated by patient groups DCM (blue), HCM (green) and ICM (orange).

Global myocardial T1-values were computed by taking the mean pixel intensity of all pixels within the myocardial region after masking native and contrast-enhanced T1-maps by manual contours, contours generated directly by U-Net, and by our comparative technique using propagation of CINE contours by MIND registration parameters. Tables 3-4 and 3-5 demonstrate global T1-values averaged over the myocardial region using the three aforementioned approaches in native and contrast-enhanced T1-maps, respectively. Figure 3-5 demonstrates the distribution of computed global T1 values using U-Net and CINE-registration based methods in native T1-maps and contrast enhanced T1-maps, respectively.

Table 3-4 Native T1 values computed using various segmentation techniques

Average native global T1 values are demonstrated in ms after masking using manual myocardial contours (first column), direct U-Net-generated contours (middle column) and CINE-registered contours (last column). Global T1 values are computed in each T1-map by taking the average pixel intensity within the myocardial region masked by the three aforementioned techniques. The average global T1 value is then computed across all patients for each segmentation technique. The values are expressed as the average global T1 value \pm standard deviation across all images

Global T1 using manual contours (ms)	Global T1 using direct U-Net-based contours (ms)	Global T1 using CINE-registration based contours (ms)
DCM patient group (n=20)		
1256 \pm 83.5	1264 \pm 97.9	1258 \pm 99.0
HCM patient group (n=20)		
1235 \pm 59.8	1230 \pm 59.5	1239 \pm 68.1
ICM patient group (n=20)		
1216 \pm 198	1244 \pm 77.9	1240 \pm 89.6
All patient group (n=60)		
1236 \pm 124	1244 \pm 78.6	1245 \pm 84.3

Table 3-5 Post-contrast T1 values computed using various segmentation techniques

Average contrast-enhanced global T1 values are demonstrated in ms after masking using manual myocardial contours (first column), direct U-Net generated contours (middle column) and CINE-registered contours (last column). Global T1 values are computed in each T1-map by taking the average pixel intensity within the

myocardial region masked by the three aforementioned techniques. The average global T1 value is then computed across all patients for each segmentation technique. The values are expressed as the average global T1 value \pm standard deviation across all images.

Global T1 using manual contours (ms)	Global T1 using direct U-Net-based contours (ms)	Global T1 using CINE-registration based contours (ms)
DCM patient group (n=20)		
550 \pm 54.3	549 \pm 52.3	555 \pm 47.5
HCM patient group (n=20)		
324 \pm 48.2	338 \pm 36.4	325 \pm 49.6
ICM patient group (n=20)		
540 \pm 99.0	554 \pm 112	541 \pm 97.5
All patient group (n=60)		
440 \pm 130	477 \pm 128	442 \pm 130

The Pearson correlation coefficient r was computed to measure the linear agreement between global myocardial T1-values in native and contrast-enhanced T1-maps calculated using manual contours vs. direct U-Net contours and CINE-registered contours. In native T1-maps, global T1 values generated by both the U-Net and CINE-registration methods showed high correlation with those generated by manual delineation ($r=0.93$, $P<0.0001$ and $r= 0.87$, $P<0.0001$, respectively). Likewise, high correlation was demonstrated for global myocardial T1 values in contrast-enhanced T1-maps generated by the direct U-Net approach and CINE-registered approach ($r=0.93$, $P<0.0001$ and $r=0.98$, $P<0.0001$, respectively).

Bland-Altman analysis was used to evaluate the agreement between global myocardial T1-values generated after manual myocardial masking vs. masking with contours generated by the U-Net-based CNN and those generated using MIND registration of CINE contours in both native and contrast-enhanced T1-maps. Bland-Altman plots for global myocardial T1-values using manual contouring vs. direct U-Net and CINE-

registered approaches are demonstrated in Figure 3-6 for native T1-maps and Figure 3-7 for contrast-enhanced T1-maps. Good agreement is demonstrated between manual T1-values and those generated by both the direct U-Net approach and CINE-registration approach. Separation of points in the contrast-enhanced Bland Altman plots is attributed to the different disease states. For example, many points on the contrast-enhanced Bland Altman plot lie toward the 200 – 400 ms range, which likely reflects the HCM disease group.

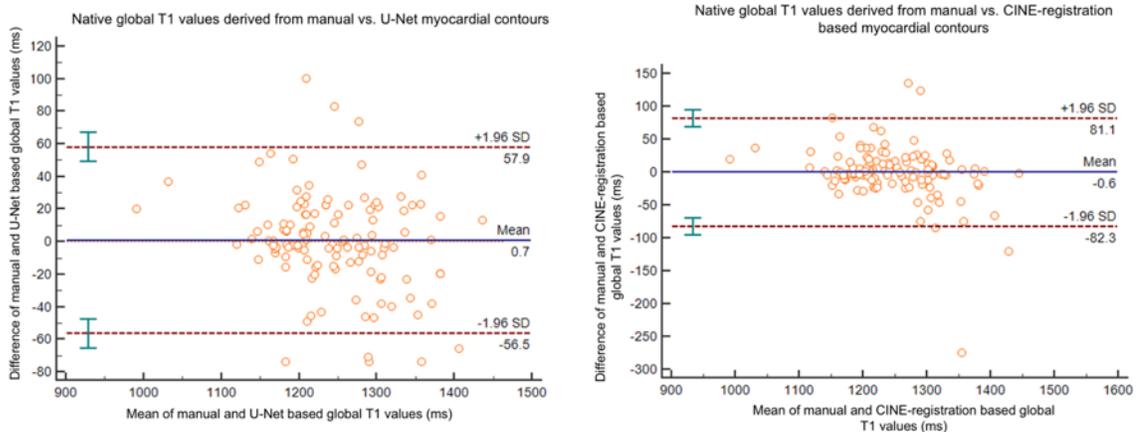


Figure 3-6 Bland-Altman plots of native myocardial T1 values across segmentation methods
Bland-Altman plots of native myocardial T1 values computed using U-Net based contours (left) and CINE-registration based contours (right) vs. myocardial T1 values computed manually.

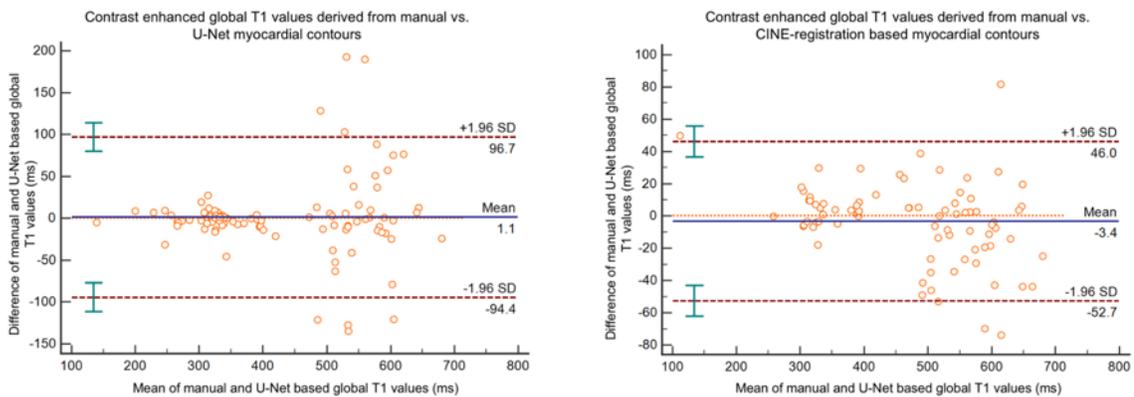


Figure 3-7 Bland-Altman plots of post-contrast myocardial T1 values across segmentation methods
Bland-Altman plots of contrast-enhanced myocardial T1 values computed using U-Net-based (left) and CINE registration-based (right) vs. myocardial T1 values computed manually.

3.7 Discussion

In this study, we present a pipeline for automated segmentation of the myocardium in native and post-contrast T1 maps acquired by the shMOLLI technique using a CNN model based on the U-Net architecture. We also developed an alternative automated myocardial segmentation technique that uses slice-matched CINE contours, also generated from a U-Net-based CNN model, and propagates those contours to native or contrast-enhanced T1-maps based on deformation metrics derived by a modality independent neighborhood descriptor (MIND) [127]. This is a development of our previous study [88] wherein we achieved myocardial segmentation of native T1-maps using a semi-automated approach from CINE contours generated by a deformable model.

Both the U-Net-based approach and the CINE registration-based approach were found to produce accurate myocardial segmentations of both native and contrast-enhanced T1-maps; with both techniques significantly outperforming Dense Nets and Attention Nets in all image types. In native T1-maps, the U-Net-based approach was shown to produce segmentations with higher accuracy in DSC in comparison to the CINE registration-based approach (i.e., $P < 0.0001$ by Wilcoxon signed rank test). However, the CINE-registration method outperformed the direct U-Net in myocardial segmentation of contrast-enhanced T1-maps (i.e., $P = 0.0014$ by Wilcoxon signed rank test). The direct U-Net based approach performed the poorest in contrast enhanced T1-maps in the ICM group, as demonstrated by the vast distribution and the average reported values of DSC metrics (see Figure 3-6 and Table 3-2). This phenomenon was likely due to the inherent low contrast and poor spatial gradient between the myocardium, blood pool, and dense fibrotic regions in these images. In such cases, our alternative proposed technique based on CINE-registration may be more

appropriate, since the U-Net algorithm performs optimally in images with superior contrast between the myocardium and surrounding structures, which is the case in CINE images. The reported DSC metrics for contrast enhanced T1-maps demonstrated in Table 3-3 support this idea; wherein DSC increased from $65.54 \pm 18.24\%$ using the direct U-Net technique to $73.24 \pm 7.29\%$ using the CINE-registration technique for patients in the ICM group.

Both the U-Net-based and CINE-registration-based approaches produced high agreement in Bland-Altman plots comparing global myocardial T1-values to those computed from manual segmentation, with 95% of measurements from both methods lying within the limits of agreement for both native and contrast enhanced T1-maps. In addition, high correlation was demonstrated when comparing global T1 values computing using both U-Net-based and CINE-registration-based techniques to manual delineation in both native and contrast enhanced T1-maps, wherein Pearson's correlation coefficient r exceeded 0.87 with a P-value < 0.0001 in all cases.

In this work, separate models were used for native and contrast-enhanced myocardial segmentation using the direct U-Net based approach. Similarly, a U-Net-based model was used for automated CINE segmentation, followed by MIND-registration using a single set of parameters for both native and contrast-enhanced T1-maps. The MIND registration parameters were optimized based on experimental observation to yield native and contrast-enhanced myocardial segmentations with the highest accuracy. Nevertheless, the U-Net-based CNN model allows for accurate, fast and direct delineation of myocardial contours in both native and contrast-enhanced T1-maps, without the risk of carrying forward prior segmentation errors via the registration of CINE contours.

There are several limitations of our study. Firstly, our test dataset in this study is somewhat small. However, due to the exclusivity of this dataset and resource constraints, we were precluded from recruiting further patients to participate. However, we aimed to address this limitation by using five-fold cross validation in our methodology. With recent advancements geared toward the use of machine learning-based algorithms in clinical cardiac imaging software, our methodology is highly translatable to clinical application. There is an overlap of 20 ICM patients in this dataset with our prior study.[88] However, we have also included an additional 40 patients in this study who have different forms of cardiomyopathy relating to interstitial and diffuse fibrosis. Further, our methodology was not implemented on ECV maps generated following registration of pre- and post-contrast images. This was due to a lack of ubiquitous hematocrit data for all subjects. However, by achieving myocardial segmentation of both native and post-contrast T1 maps, we establish generalizability to myocardial segmentation of ECV maps, such that the automated contours of native and post-contrast maps can be co-registered and motion corrected to quickly and efficiently translate prior myocardial contours to ECV maps. We intend to evaluate this methodology in ECV maps in future work.

3.8 Conclusion

Native and post-contrast T1-mapping techniques produce images reflective of diffuse and interstitial myocardial fibrosis. Automated myocardial segmentation of both native and contrast-enhanced T1-maps is achievable using convolutional neural networks based on the U-Net architecture. We incrementally demonstrated that myocardial borders can alternatively be segmented from spatially matched CINE images and applied to T1-maps via deformation and propagation through a modality-independent neighborhood

descriptor (MIND). Further, we demonstrate average global T1 values in native and contrast enhanced maps at 3T for patients with three common forms of cardiomyopathy. For automated myocardial segmentation of native T1-maps, the direct U-Net approach is more efficient and eliminates cross-technique dependence. However, the CINE-registration technique may be more accurate and appropriate, particularly for contrast enhanced T1-maps and/or cases exhibiting dense regions of replacement fibrosis, such as is the case for patients with ICM.

3.9 Chapter 3 References

- [1] J. R. Burt, S. L. Zimmerman, I. R. Kamel, M. Halushka, and D. A. Bluemke, “Myocardial T1 Mapping: Techniques and Potential Applications,” *RadioGraphics*, vol. 34, no. 2, pp. 377–395, Mar. 2014, doi: 10.1148/rg.342125121.
- [2] L. Iles *et al.*, “Myocardial Fibrosis Predicts Appropriate Device Therapy in Patients with Implantable Cardioverter-Defibrillators for Primary Prevention of Sudden Cardiac Death,” *JAC*, vol. 57, no. 7, pp. 821–828, 2011, doi: 10.1016/j.jacc.2010.06.062.
- [3] N. Morita, W. J. Mandel, Y. Kobayashi, and H. S. Karagueuzian, “Cardiac fibrosis as a determinant of ventricular tachyarrhythmias,” vol. 30, pp. 389–394, 2014, doi: 10.1016/j.joa.2013.12.008.
- [4] A. Gulati *et al.*, “Association of fibrosis with mortality and sudden cardiac death in patients with nonischemic dilated cardiomyopathy,” *Jama*, vol. 309, no. 9, pp. 896–908, 2013, doi: 10.1001/jama.2013.1363.
- [5] N. Mewton, C. Y. Liu, P. Croisille, D. Bluemke, and J. A. C. Lima, “Assessment of myocardial fibrosis with cardiovascular magnetic resonance,” *J Am Coll Cardiol*, vol. 57, no. 8, pp. 891–903, 2011, doi: 10.1016/j.jacc.2010.11.013.
- [6] N. A. Farrag and G. A. Wright, “Effect of T1-mapping technique and diminished image resolution on quantification of infarct mass and its ability in predicting appropriate ICD therapy”.
- [7] F. M. Caballeros *et al.*, “T1 mapping of the myocardium: what a resident need to know,” *European Cardiology Review*, pp. 1–17, 2016, doi: 10.1594/ecr2016/C-0385.
- [8] P. R. Scully, G. Bastarrika, J. C. Moon, and T. A. Treibel, “Myocardial Extracellular Volume Quantification by Cardiovascular Magnetic Resonance and Computed Tomography,” 2018.
- [9] A. Gonzalez, E. B. Schelbert, J. Diez, and J. Butler, “Myocardial Interstitial Fibrosis in Heart Failure,” *Journal of the American College of Cardiology*, vol. 71, no. 15, 2018, doi: 10.1016/j.jacc.2018.02.021.
- [10] S. Donekal and J. A. C. Lima, “Translational Medicine Diffuse Interstitial Myocardial Fibrosis by T1 Myocardial Mapping: Review,” *Translational Medicine*, vol. 3, no. 1, pp. 1–5, 2013, doi: 10.4172/2161-1025.1000111.

- [11] D. R. Messroghli *et al.*, “Myocardial T1 mapping: Application to patients with acute and chronic myocardial infarction,” *Magnetic Resonance in Medicine*, vol. 58, no. 1, pp. 34–40, 2007, doi: 10.1002/mrm.21272.
- [12] E. B. Schelbert and T. C. Wong, “Clinical Applications of T1 Mapping,” *American College of Cardiology*, 2015. https://www.acc.org/latest-in-cardiology/articles/2015/02/13/09/24/clinical-applications-of-t1-mapping?w_nav=TI
- [13] P. Haaf, P. Garg, D. R. Messroghli, D. A. Broadbent, J. P. Greenwood, and S. Plein, “Cardiac T1 Mapping and Extracellular Volume (ECV) in clinical practice: A comprehensive review,” *Journal of Cardiovascular Magnetic Resonance*, vol. 18, no. 1, pp. 1–12, 2016, doi: 10.1186/s12968-016-0308-4.
- [14] J. C. Moon *et al.*, “Myocardial T1 mapping and extracellular volume quantification: a Society for Cardiovascular Magnetic Resonance (SCMR) and CMR Working Group of the European Society of Cardiology consensus statement,” *Journal of Cardiovascular Magnetic Resonance*, vol. 15, no. 1, p. 92, 2013, doi: 10.1186/1532-429X-15-92.
- [15] C. R. Hamilton-craig, “T 1 Mapping for Myocardial Fibrosis by Cardiac Magnetic Resonance Relaxometry — A Comprehensive,” vol. 3, no. March, pp. 1–11, 2017, doi: 10.3389/fcvm.2016.00049.
- [16] A. J. Taylor, M. Salerno, R. Dharmakumar, and M. Jerosch-Herold, “T1 Mapping Basic Techniques and Clinical Applications,” *JACC: Cardiovascular Imaging*, vol. 9, no. 1, pp. 67–81, 2016, doi: 10.1016/j.jcmg.2015.11.005.
- [17] V. O. Puntmann *et al.*, “Native T1 mapping in differentiation of normal myocardium from diffuse disease in hypertrophic and dilated cardiomyopathy,” *JACC: Cardiovascular Imaging*, vol. 6, no. 4, pp. 475–484, 2013, doi: 10.1016/j.jcmg.2012.08.019.
- [18] E. B. Schelbert and D. R. Messroghli, “State of the Art: Clinical Applications of Cardiac T1 Mapping 1,” vol. 278, no. 3, 2016.
- [19] D. R. Messroghli *et al.*, “Clinical recommendations for cardiovascular magnetic resonance mapping of T1, T2, T2 * and extracellular volume: A consensus statement by the Society for Cardiovascular Magnetic Resonance (SCMR) endorsed by the European Association for Cardiovascular Imaging (EACVI),” pp. 1–24, 2017, doi: 10.1186/s12968-017-0389-8.
- [20] R. J. Perea, J. T. Ortiz-perez, and M. Sole, “T1 mapping: characterisation of myocardial interstitial space,” pp. 189–202, 2015, doi: 10.1007/s13244-014-0366-9.
- [21] S. K. Piechnik *et al.*, “Shortened Modified Look-Locker Inversion recovery (ShMOLLI) for clinical myocardial T1- mapping at 1.5 and 3 T within a 9 heartbeat breathhold,” pp. 1–11, 2010.
- [22] C. M. Kramer, J. Barkhausen, S. D. Flamm, R. J. Kim, and E. Nagel, “Standardized cardiovascular magnetic resonance (CMR) protocols 2013 update,” pp. 1–10, 2013.
- [23] C. Petitjean and J. Dacher, “A review of segmentation methods in short axis cardiac MR images,” *Medical Image Analysis*, vol. 15, no. 2, pp. 169–184, 2011, doi: 10.1016/j.media.2010.12.004.
- [24] Q. Tao, S. R. D. Piers, H. J. Lamb, and R. J. van der Geest, “Automated left ventricle segmentation in Late Gadolinium-Enhanced MRI for Objective Myocardial Scar Assessment,” *Journal of Magnetic Resonance Imaging*, vol. 42, no. 2, pp. 390–399, 2014, doi: 10.1002/jmri.24804.
- [25] R. S. Xu, P. Athavale, Y. Lu, P. Radau, G. A. Wright, and M. G. Canada, “MYOCARDIAL SEGMENTATION IN LATE-ENHANCEMENT MR IMAGES VIA REGISTRATION

- AND PROPAGATION OF CINE CONTOURS University of Toronto, Department of Medical Biophysics,” pp. 856–859, 2013.
- [26] O. S. H. Wei D., Sun Y., Chai P., Low A., “Myocardial Segmentation of Late Gadolinium Enhanced MR Images by Propagation of Contours from Cine MR Images.,” in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2011*, 2011, pp. 428–435.
- [27] Y. Lu, G. Wright, and P. E. Radau, “Automatic myocardium segmentation of LGE MRI by deformable models with prior shape data,” vol. 15, no. Suppl 1, pp. 14–15, 2013, doi: 10.1186/1532-429X-15-S1-P14.
- [28] M. R. Avendi, A. Kheradvar, and H. Jafarkhani, “A combined deep-learning and deformable-model approach to fully automatic segmentation of the left ventricle in cardiac MRI R,” *Medical Image Analysis*, vol. 30, pp. 108–119, 2016, doi: 10.1016/j.media.2016.01.005.
- [29] L. K. Tan, R. A. Mclaughlin, E. Lim, Y. Faridah, A. Aziz, and Y. M. Liew, “Fully Automated Segmentation of the Left Ventricle in Cine Cardiac MRI Using Neural Network Regression,” 2018, doi: 10.1002/jmri.25932.
- [30] T. Anh, Z. Lu, and G. Carneiro, “Combining deep learning and level set for the automated segmentation of the left ventricle of the heart from cardiac cine magnetic resonance,” vol. 35, pp. 159–171, 2017, doi: 10.1016/j.media.2016.05.009.
- [31] P. V. Tran, “A Fully Convolutional Neural Network for Cardiac Segmentation in Short-Axis MRI,” pp. 1–21, 2016.
- [32] F. Zabihollahy, J. A. White, and E. Ukwatta, “Fully automated segmentation of left ventricular myocardium from 3D late gadolinium enhancement magnetic resonance images using a U-net convolutional neural network-based model,” *Proc. SPIE 10950*, no. Computer-Aided Diagnosis, 2019, doi: 10.1117/12.2512146.
- [33] A. S. Fahmy, H. El-rewaidy, M. Nezafat, S. Nakamori, and R. Nezafat, “Automated analysis of cardiovascular magnetic resonance myocardial native T 1 mapping images using fully convolutional neural networks,” vol. 1, pp. 1–12, 2019.
- [34] O. Ronneberger, P. Fischer, and T. Brox, “U-Net: Convolutional Networks for Biomedical Image Segmentation,” pp. 1–8.
- [35] I. Goodfellow, Y. Bengio and A. Courville. “Deep Learning”. (Book). MIT Press, 2016.
- [36] O. Oktay et al., “Attention U-Net: Learning Where to Look for the Pancreas”, <https://arxiv.org/abs/1804.03999> , 2018.
- [37] G. Huang and K. Q. Weinberger, “Densely Connected Convolutional Networks”.
- [38] M. P. Heinrich *et al.*, “MIND: Modality independent neighbourhood descriptor for multi-modal deformable registration,” *Medical Image Analysis*, vol. 16, no. 7, pp. 1423–1435, 2012, doi: 10.1016/j.media.2012.05.008.
- [39] N. A. Farrag, J. A. White, and E. Ukwatta, “Semi-automated myocardial segmentation in native T1-mapping CMR using deformable non-rigid registration of CINE images,” *Proc. SPIE 10953*, vol. Biomedical, no. March 2019, doi: 10.1117/12.2513054.
- [40] N. A. Farrag, J. A. White, and E. Ukwatta, “Semi-automated myocardial segmentation in native T1-mapping CMR using deformable non-rigid registration of CINE images,” *Proc. SPIE 10953*, vol. Biomedical, no. March 2019, doi: 10.1117/12.2513054.

Chapter 4: Automated myocardial segmentation of extra-cellular volume mapping cardiac magnetic resonance images using fully convolutional neural networks

4.1 Purpose

Extra-cellular volume (ECV) mapping is a quantitative cardiac magnetic resonance (CMR) imaging technique that allows for characterization and measurement of myocardial tissue composition[61], wherein pixel intensities are reflective of an estimation of each tissue's ECV. Changes to the myocardial ECV is a diagnostic marker of myocardial tissue remodeling[50], [54], [60], [61]. In the absence of amyloid or edema, expansion of the myocardial extracellular space is a common feature of myocardial fibrosis (MF), a condition that is characterized by excessive deposition of collagen in the extracellular myocardial space[49], [50], [60]. Thus, ECV mapping is becoming increasingly explored as a diagnostic tool for characterization of cardiac diseases that manifest MF, as this technique offers a number of advantages over late gadolinium enhanced (LGE), native and post-contrast T1-mapping CMR techniques, which are variable to different field strengths, imaging vendors, and metabolic/renal function[48], [52], [61], [71], [72]. Prior to measurement of myocardial ECV, manual delineation of the endocardial and epicardial borders is required. This task is challenging, tedious, and highly prone to inter- and intra-operator error. Erroneous inclusion of the blood pool and/or pericardial space at endocardial and epicardial regions can significantly confound ECV measurements. Thus, automated myocardial segmentation in ECV mapping is highly preferred.

In the last decade, deep learning algorithms have been implemented in a wide variety of CMR segmentation problems. Deep learning-based algorithms offer advantages over traditional segmentation techniques, which are typically sensitive to noise and inhomogeneities in image intensity. Convolutional neural networks (CNNs) are one of the most prevalent deep learning algorithms used for left ventricular (LV) myocardial segmentation and MF segmentation in CINE and LGE imaging, respectively[115], [116], [118], [130]. CNNs have also been implemented in LV myocardial segmentation of T1 mapping imaging[89], [122]. However, no study to date has implemented the use of fully CNNs for myocardial segmentation of ECV mapping CMR images. Therefore, we propose a technique for automated segmentation of the myocardium in ECV mapping CMR images using fully CNNs. In this work, we compare two CNNs, based on the standard U-Net architecture[112] and the U-Net++[131], which incorporates nested and dense skip connections to the original architecture.

4.2 Methods

4.2.1 Study Population and CMR Protocol

Seventy-three (73) patients with a clinical diagnosis and confirmed MRI-based diagnosis of MF associated with dilated cardiomyopathy (DCM = 35) or ischemic cardiomyopathy (ICM=38) underwent CMR imaging using a Siemens 3T magnet (Prisma or Skyra, Siemens Healthineers, Erlangen), at the Stephenson Cardiac Imaging Centre in Calgary, AB. The CMR imaging protocol was inclusive of short-axis, long-axis and axial CINE, short-axis LGE imaging using an inversion-recovery gradient echo pulse sequence, and short-axis native, post-contrast T1-mapping and ECV images using a phase sensitive inversion recovery (PSIR) MOLLI pulse sequence. Synthetic PSIR LGE images were

also derived in short axis from the post-contrast T1 maps. Synthetic PSIR LGE, T1 and ECV maps were acquired at three locations: apical, basal and mid-ventricular region. This resulted in 219 (i.e., 3 x 73 images) available for model training. The study was approved by the Conjoint Health Research Ethics Board at University of Calgary and all subjects provided written informed consent.

4.2.2 Image processing

Images were pre-processed before being split into training and testing subsets. A flowchart of the myocardial segmentation frameworks for U-Net and U-Net++ are demonstrated in Figures 4-1 and 4-2, respectively. Output myocardial masks were validated against contours manually delineated by an expert using DSC, accuracy, precision, recall and relative error index.

ECV maps were preprocessed by resizing to standard dimensions of 192 by 176 pixels. In order to ensure equivalent representation of ICM and DCM patient groups in training the CNN model, we used five-fold cross validation wherein ICM and DCM patient groups were split into five sub-sets each, and four subsets each from the ICM and DCM groups were used toward training, while the remaining subset from ICM and DCM groups was used for testing. This process was repeated five times, each with a different subset from both ICM and DCM groups for testing, such that the entire dataset could be used toward evaluation of the proposed technique.

4.2.2.1 Standard U-Net

A standard U-Net architecture [112] was implemented with an encoder/decoder pipeline. Random grid search was used to optimize hyperparameters. The kernel size was

fixed at 3x3 for all but the last convolutional layer, which was adjusted to 1x1. Filter sizes increased by a factor of two from 64 to 1024 in the encoder path and decreased at the same interval back to 64 in the decoder path using max pooling. Each Convolution layer was followed by a ReLU activation function. Batch Normalization was used in the encoder path for regularization. The final output was obtained using a sigmoid activation, which outputs a probability map from 0 – 1, indicating whether each pixel belonged to the myocardium. Probabilities >0.5 were mapped to 1 and <0.5 were mapped to 0 to generate the output binary mask. The model architecture can be seen in Figure 4-1.

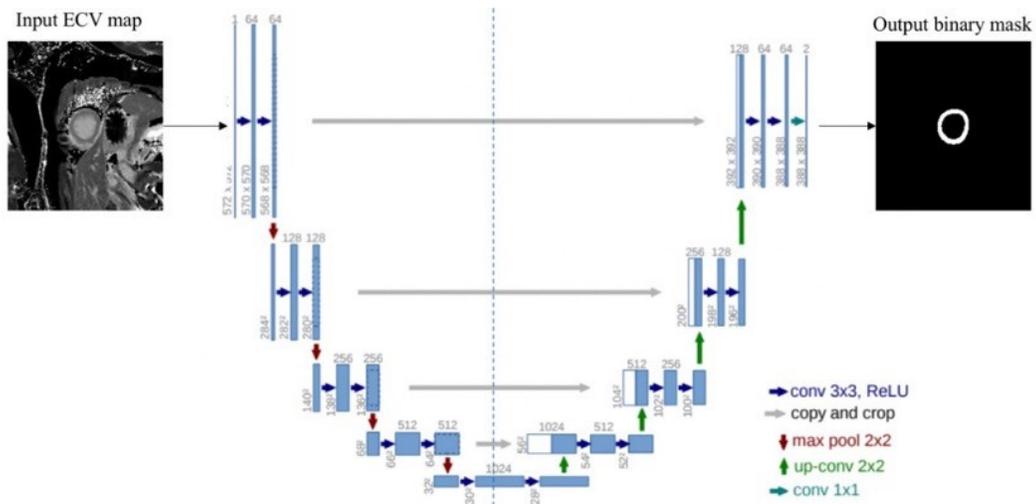


Figure 4-1 Framework used for CNN model based on standard U-Net architecture in ECV map segmentation.

The pre-processed ECV map is used as input and output is a binary mask based on a probability map generated by a sigmoid function. This figure is referenced from Ronneberger et al.’s (2015) manuscript.

4.2.2.2 U-Net++

The U-Net++ model started with an encoder sub-network (i.e., backbone) followed by a decoder sub-network, connected using re-designed skip pathways. Random grid search was used optimize hyperparameters. Deep Supervision was not used, i.e., only the output at the end of the decoder path was used for computing loss and as the final output

of the model. The binary mask was generated using a sigmoid function, as described in section 2.2.2 The model architecture can be seen in Figure 4-2.

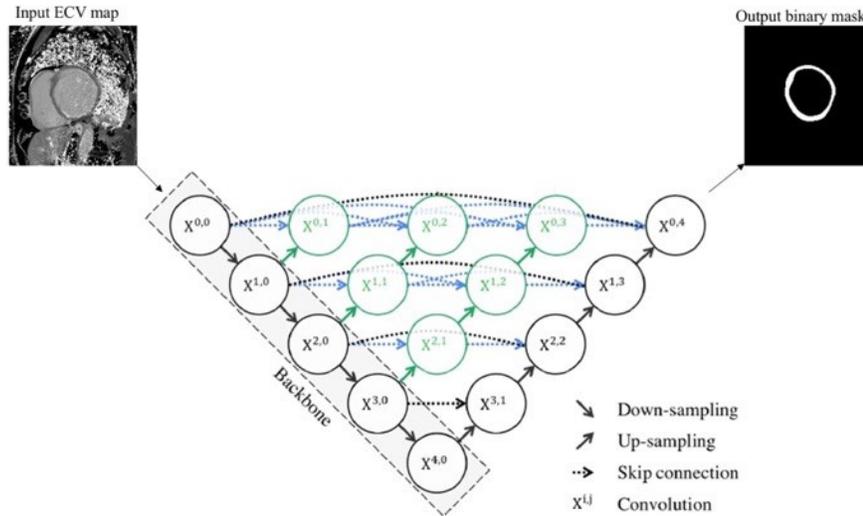


Figure 4-2 Framework used for CNN model based on U-Net++ for ECV map segmentation. The pre-processed ECV map is used as input and output is a binary mask. This figure is referenced from Zhou et al.'s (2018) manuscript.

4.2.2.3 Training procedure

Training code and model implementations were done using PyTorch. For both CNN models, images were augmented at runtime using the Albumentations library. The five folds were trained for 500 epochs using Adam Optimizer with an initial learning rate of 0.0001, decreasing to 0.00005 after 250 epochs for better fine-tuning. The networks were trained on a P100 GPU on Shared Hierarchical Research Computing Network (SHARCNET). The total training time was approximately 7.5 hours for all folds.

4.2.3 Inter-operator variability analysis

In order to further validate the methods proposed in this work, we performed an inter-operator variability analysis on a sub-set of the data across three operators. For this analysis, twelve (12) patients from the DCM group and fourteen (14) from the ICM group were randomly selected, yielding 78 images. Each operator manually segmented the

myocardial region using the polygon tool on ITK-SNAP software[18] in order to create a binary mask. The dice similarity coefficient (DSC) was used to evaluate the degree of spatial overlap between binary masks created by three operators against each other, and three operators against U-Net++ and basic U-Net. Further analysis was performed by assessing the differences between myocardial ECV values among the methods, which were computed by applying each operator's and each CNN-based method's binary myocardial mask onto the original image.

Manually derived myocardial ECV values were compared against those derived by U-Net, U-Net++, and those derived by other operators on an image-wise basis. For each image used in this analysis, we performed a Kruskal-Wallis test across the distribution of myocardial ECV values derived by all operators versus both U-Net and U-Net++, wherein we report the mean P-value across all images. Furthermore, we computed the Spearman correlation coefficient ρ on the global (i.e., mean) myocardial ECV value computed by each operator versus both U-Net++ and U-Net across all images. For this analysis, we compute ρ for each operator versus the other operators to evaluate operators-versus-operators variability and report the mean value. To evaluate operators-versus-CNN (i.e., operators vs. U-Net and operators vs. U-Net++) variability, we compute ρ for each individual operator versus U-Net and U-Net++ and report the mean value across the three operators.

4.3 Results

4.3.1 Segmentation Results

DSC, accuracy, precision and relative error index metrics yielded from the original U-Net and U-Net++ are demonstrated in Table 4-2. DSC metrics are reported >87% for both ICM, DCM and combined groups for both networks. Accuracy was reported over 99%, while precision and recall exceeded 85% for all patient groups in both models. The recall metric was higher among the DCM group for both models. The relative error of the myocardial area was lowest in the ICM group for both models. Output segmentations generated by both models are shown in Figure 4-3.

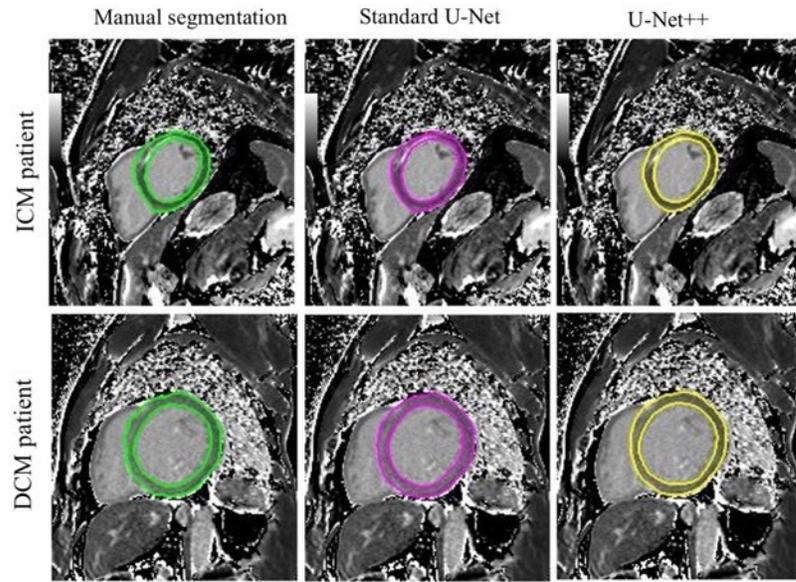


Figure 4-3 Segmentation results on ECV maps using U-Net and U-Net++
Manual delineation (green), U-Net (magenta) and U-Net++ (yellow) masks are demonstrated on a patient with ICM (top row) and DCM (bottom row).

Table 4-1 Segmentation results on ECV maps using U-Net and U-Net++

Table 4-1				
DSC, accuracy, precision and relative error index metrics computed on myocardial segmentations from U-Net and U-Net++ techniques vs. manual delineations among ICM and DCM patient groups.				
Patient group	DSC (%)	Precision (%)	Recall (%)	Relative error of myocardial area (%)
Standard U-Net				
ICM (n=38)	87.12	88.58	85.72	11.17

DCM (n=35)	88.32	85.64	91.20	11.49
All (n=73)	87.61	87.28	87.97	11.31
U-Net++				
ICM (n=38)	87.37	88.90	85.90	9.87
DCM (n=35)	88.61	86.02	91.40	11.15
All (n=73)	87.89	87.63	88.15	10.43

While accuracy metrics demonstrate high agreement across both the U-Net and U-Net++ models, performance was not always adequate in all cases. For example, the myocardial region should resemble a “donut” like structure, wherein the outer border of the donut is the epicardial border and the inner border of the donut is the endocardial border. In some cases, U-Net (n = 6 images, 4 ICM, 2 DCM) and U-Net++ (n = 7 images, 5 ICM, 2 DCM) produced incomplete myocardial segmentations, whereby an open circular donut shape was output by the model. Similarly, extra objects (e.g., more than one segmented region per output mask) were segmented in addition to the donut by the U-Net (n = 5, 2 ICM, 3 DCM) and U-Net++ (n = 6, 3 ICM, 3 DCM) models. Similarly, the models each failed to segment the myocardium (i.e., no output segmentation mask, only background pixels) in 1 image (ICM apical image). While imperfections such as incomplete mask and extra regions will impact segmentation accuracy, they can be mitigated through application of morphological image processing techniques, such as opening and closing. Sub-par myocardial segmentations output by U-Net and U-Net++ are demonstrated in Figure 4-4.

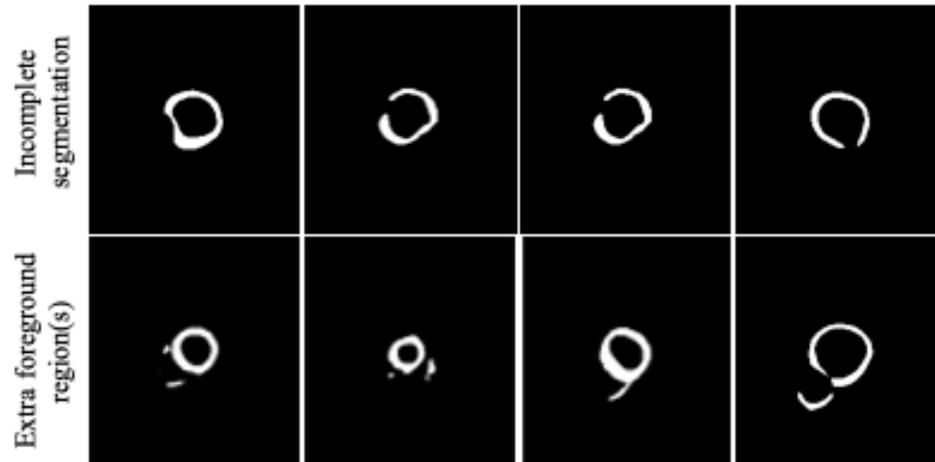


Figure 4-4 Sub-par myocardial segmentations output by U-Net++. Incomplete segmentations are demonstrated in the top row and segmentations output with extra foreground objects are demonstrated in the bottom row. All given examples are of different patients.

4.3.2 Inter-operator variability analysis

Average DSC metrics computed in our operators-versus-operators and operators-versus-U-Net++ and basic U-Net variability analyses are demonstrated in Table 4-2. We computed an average DSC of 78.7% among manual segmentations delineated by three operators across both patient groups, as compared to 83.6% and 83.4% for the operators versus U-Net++ and operators versus basic U-Net, respectively. Among all comparisons, the best DSC agreement was demonstrated in the DCM patient group.

Differences in myocardial ECV values were assessed among three operators and by each operator versus U-Net++ and basic U-Net on an image-wise basis using the Kruskal-Wallis test (across the full distribution of ECV values) and the Spearman correlation coefficient ρ (across the mean myocardial ECV value). Statistical results of the Kruskal-Wallis test and Spearman correlation coefficients are demonstrated in Tables 4-3 and 4-4, respectively. Significant differences in the distributions of myocardial ECV values across operators was demonstrated by Kruskal-Wallis test (P-value=0.04 across all patient groups). The Kruskal-Wallis test failed to show significant differences in the distributions

of ECV values derived by the operators versus U-Net++ and basic U-Net. Figure 4-5 demonstrates the histogram distribution of ECV values and associated myocardial contours derived by all users, U-Net++ and basic U-Net techniques from one mid-ventricular ECV map in a patient with DCM.

Correlation of the global ECV value derived by each operator demonstrated high agreement in the operators-versus-operators comparison ($\rho = 0.83$) and operators versus both U-Net++ and basic U-Net ($\rho = 0.87$ and $\rho = 0.88$, respectively).

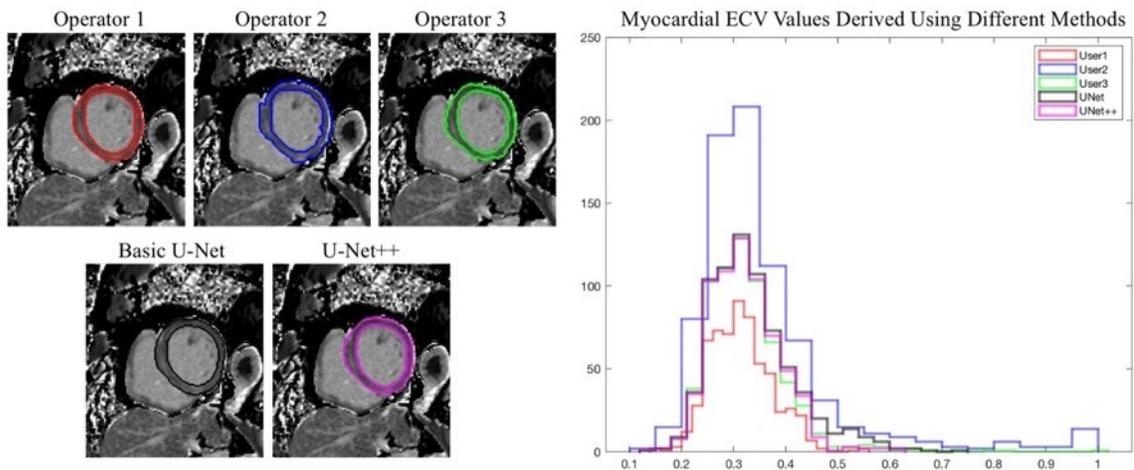


Figure 4-5 Manual myocardial delineation by three operators versus U-Net and U-Net++
Contours delineated by three operators (red, blue and green), U-Net (black) and U-Net++ (magenta) are demonstrated from a mid-ventricular ECV map in a patient with DCM. The histogram of the ECV value distribution associated with each method is demonstrated on the right-hand side, wherein histogram colors coordinate with the respective colors of myocardial contours.

Table 4-2 metrics computed across myocardial masks generated by three operators versus U-Net++ and basic U-Net.

Patient group	DSC metric (%)		
	Operators vs. Operators	Operators vs. U-Net++	Operators vs. basic U-Net
DCM (n=12)	80.2	85.0	84.9
ICM (n=14)	77.5	82.5	82.3
All patients (n=26)	78.7	83.6	83.4

Table 4-3 Results of the Kruskal-Wallis (KW) test computed on the full distribution of myocardial ECV values derived by operators vs. U-Net and U-Net++

P-values computed by the KW test for operators vs. operators, operators vs. U-Net++ and operators vs. basic U-Net. Myocardial ECV values were derived by applying the myocardial masks provided by each of the methods. The KW test was applied on an image-wise basis across three groups for the operators vs. operators metric. For operator vs. U-Net++ and operator vs. basic U-Net metrics, we performed the KW test across two groups (i.e., each operator versus the respective U-Net) and report the three-way average *P*-value.

	Kruskal-Wallis test <i>P</i>-value
--	---

Patient group	Operators vs. Operators	Operators vs. U-Net++	Operators vs. basic U-Net
DCM (n=12)	0.041	0.053	0.072
ICM (n=14)	0.038	0.065	0.054
All patients (n=26)	0.040	0.055	0.060

Table 4-4 Spearman correlation coefficient ρ computed on global (i.e., mean) ECV values derived by operators vs. operators, U-Net++ and basic U-Net.

For the operator vs. operator metric, we computed ρ on global ECV values derived by three users and report the three-way average ρ . For both the operators vs. U-Net++ and operators vs. basic U-Net metrics, we computed ρ on global ECV values derived by each user versus the respective U-net and report the three-way average ρ

Patient group	Spearman correlation coefficient ρ		
	Operators vs. operators	Operators vs. U-Net++	Operators vs. basic U-Net
DCM (n=12)	0.770	0.818	0.818
ICM (n=14)	0.847	0.907	0.906
All patients (n=26)	0.828	0.88	0.877

4.4 Discussion

Segmentation of the myocardium is an important, yet challenging step that is required prior to measurement of ECV values and subsequent characterization of myocardial disease in ECV mapping CMR. In this work, we use fully convolutional neural networks based on the U-Net and U-Net++ architectures to achieve fully automated myocardial segmentation of ECV maps acquired in short-axis view at the base, mid-point and apex of the left ventricle. To our knowledge, our work is the first deep learning-based study to achieve fully automated segmentation of the myocardium in ECV maps.

We demonstrate that deep learning models based on the U-Net architecture may be used as appropriate means toward myocardial segmentation in ECV maps. Our study demonstrated a minor outperformance of the U-Net++ compared to the standard U-Net. Both models performed best in the DCM patient group in the DSC metric, however, both models yielded lower relative error index in the ICM group. Resultant precision metrics were higher in the ICM group, but recall metrics were lower in this patient group as demonstrated by results from both models.

This study has several limitations. Firstly, our dataset of 73 patients is somewhat small. However, we implemented five-fold cross validation to address the dataset size in training of the models. Also, patient characteristics, such as average heart rate are not reported as they were not recorded during image acquisition in this study. Since higher heart rates and/or arrhythmias can cause increased motion artifact, the performance of U-Net and U-Net++ models may vary across patients experiencing these phenomena. We intend to explore the idea of comparing model accuracies across patients with variable heart rates in the future. In addition, we were not able to compare segmentation results and ECV measurements yielded by our automated techniques versus those generated by multiple trained users through variability analysis. However, this will be explored in further work. Additionally, we were limited from doing a direct comparison of segmented MF regions (as opposed to the entire myocardium) using the models used in this study. Manual delineation of fibrosis is a considerably difficult task, especially in cases where interstitial MF is present. Thus, we are instead limited to compute the global ECV measurements across the entire myocardium and compare those measured by the models versus manual delineation. In future work, we intend to perform inter- and intra- observer analysis on individual endocardial and epicardial borders across different disease states. Moreover, we aim to achieve direct segmentation of MF regions within the myocardial boundaries provided, by the U-Net or U-Net++ model, through clustering and/or thresholding based on knowledge of the studied ranges of healthy ECV in 3T CMR images. These regions may then be corroborated against conventional LGE using image registration.

4.5 Conclusion

Fully convolutional neural networks based on the U-Net architecture are suitable for automated segmentation of the myocardium in ECV mapping CMR images in short-axis view. The U-Net and U-Net++ techniques produce myocardial contours with negligible differences in accuracy.

4.6 Chapter 4 References

- [1] P. Haaf, P. Garg, D. R. Messroghli, D. A. Broadbent, J. P. Greenwood, and S. Plein, “Cardiac T1 Mapping and Extracellular Volume (ECV) in clinical practice: A comprehensive review,” *J. Cardiovasc. Magn. Reson.*, vol. 18, no. 1, pp. 1–12, 2016.
- [2] R. J. Perea, J. T. Ortiz-perez, and M. Sole, “T1 mapping: characterisation of myocardial interstitial space,” pp. 189–202, 2015.
- [3] J. C. Moon *et al.*, “Myocardial T1 mapping and extracellular volume quantification: a Society for Cardiovascular Magnetic Resonance (SCMR) and CMR Working Group of the European Society of Cardiology consensus statement,” *J. Cardiovasc. Magn. Reson.*, vol. 15, no. 1, p. 92, 2013.
- [4] D. Radenkovic, S. Weingärtner, L. Ricketts, J. C. Moon, and G. Captur, “T 1 mapping in cardiac MRI,” pp. 415–430, 2017.
- [5] F. M. Caballeros *et al.*, “T1 mapping of the myocardium: what a resident need to know,” *Eur. Cardiol. Rev.*, pp. 1–17, 2016.
- [6] C. L. Jellis and D. H. Kwon, “Myocardial T1 mapping: modalities and clinical applications,” vol. 4, no. 2, pp. 126–137, 2014.
- [7] P. Kellman and M. S. Hansen, “T1-mapping in the heart: accuracy and precision,” *J. Cardiovasc. Magn. Reson.*, vol. 16, no. 2, 2014.
- [8] F. aus dem Siepen *et al.*, “T1 mapping in dilated cardiomyopathy with cardiac magnetic resonance: quantification of diffuse myocardial fibrosis and comparison with endomyocardial biopsy,” *Eur. Heart J. Cardiovasc. Imaging*, vol. 16, no. 2, 2015.
- [9] E. B. Schelbert and D. R. Messroghli, “State of the Art: Clinical Applications of Cardiac T1 Mapping 1,” vol. 278, no. 3, 2016.
- [10] M. R. Avendi, A. Kheradvar, and H. Jafarkhani, “A combined deep-learning and deformable-model approach to fully automatic segmentation of the left ventricle in cardiac MRI R,” *Med. Image Anal.*, vol. 30, pp. 108–119, 2016.
- [11] L. K. Tan, R. A. McLaughlin, E. Lim, Y. Faridah, A. Aziz, and Y. M. Liew, “Fully Automated Segmentation of the Left Ventricle in Cine Cardiac MRI Using Neural Network Regression,” 2018.
- [12] T. Anh, Z. Lu, and G. Carneiro, “Combining deep learning and level set for the automated segmentation of the left ventricle of the heart from cardiac cine magnetic resonance,” vol. 35, pp. 159–171, 2017.

- [13] F. Zabihollahy, J. A. White, and E. Ukwatta, “Fully automated segmentation of left ventricular myocardium from 3D late gadolinium enhancement magnetic resonance images using a U-net convolutional neural network-based model,” *Proc. SPIE 10950*, no. Computer-Aided Diagnosis, 2019.
- [14] A. S. Fahmy, H. El-rewaidy, M. Nezafat, S. Nakamori, and R. Nezafat, “Automated analysis of cardiovascular magnetic resonance myocardial native T1 mapping images using fully convolutional neural networks,” vol. 1, pp. 1–12, 2019.
- [15] N. A. Farrag, A. Lochbihler, J. A. White, and E. Ukwatta, “Evaluation of fully automated myocardial segmentation techniques in native and contrast-enhanced T1-mapping cardiovascular magnetic resonance images using fully convolutional neural networks,” *Med. Phys.*, 2020.
- [16] O. Ronneberger, P. Fischer, and T. Brox, “U-Net: Convolutional Networks for Biomedical Image Segmentation,” pp. 1–8.
- [17] Z. Zhou, M. M. Rahman Siddiquee, N. Tajbakhsh, and J. Liang, “Unet++: A nested u-net architecture for medical image segmentation,” *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 11045 LNCS, pp. 3–11, 2018.
- [18] Y. PA *et al.*, “User-guided 3D active contour segmentation of anatomical structures: significantly improved efficiency and reliability,” *Neuroimage*, vol. 31, no. 3, pp. 1116–1128, Jul. 2006.

Chapter 5: Evaluation of the Progression of Left Atrial Fibrosis in Canines Following Rapid Ventricular Pacing Using 3D Late Gadolinium Enhanced CMR Images

5.1 Introduction

Atrial fibrillation (AF), the most common form of cardiac arrhythmia, is characterized by rapid and irregular atrial rhythm. AF is often a precursor of several cardiac and cerebrovascular complications, including ischemic stroke, transient ischemic attack, and heart failure [1]– [3]. Extracellular matrix (ECM) remodelling is an important contributor to the AF substrate, and thus, AF often co-exists with interstitial myocardial fibrosis (MF) [4],[5]. The underlying pathophysiological factors that contribute to AF are yet to be fully understood. ECM remodelling is known to contribute to the onset and progression of AF [4]– [6]. Likewise, atrial ECM remodelling is known to create re-entry pathways for abnormal electrophysiological activity, which can perpetuate AF [7]– [10]. However, ECM remodelling may also occur as a consequence of a prior fibrillation process [4],[5],[7]. While MF is frequently observed in AF, a definitive causal relationship has yet to be established.

Detection left atrial (LA) MF via late gadolinium enhanced (LGE) cardiac magnetic resonance (MR) has been of increasing interest for pre- and post-ablation evaluation [11], indicating the need for objective analysis methods for quantification of LA MF. The image intensity ratio (IIR) analysis method [11]– [14] defines LA fibrosis based on IIR thresholds after normalizing LA myocardial intensity using the mean intensity of the blood pool volume as reference. This method has been validated at 1.5T and 3T mitigates signal

intensity variability caused by intrinsic physiological factors, therefore producing more consistency in atrial fibrosis quantification. The method has been validated at 1.5T and 3T [11],[13],[15], but has not yet been applied in a serial imaging study with histological validation. For LA fibrosis quantification to be translated into clinical practice, further validation of quantification methodologies and thresholds are required. Therefore, the goal of this study was to validate the IIR method for atrial MF quantification after rapid ventricular pacing using a canine model and histopathology.

5.2 Methods

5.2.1 Image Acquisition

Twelve (12) canines (9 experimental, 3 control) underwent serial 3D LGE-CMR imaging at 3T using a Siemens Biograph Positron Emission Tomography (PET)-MR system at the Lawson Health Research Centre in London, Ontario, Canada. All canines were female, bred-for-research adult hounds, weighing approximately 20 kg on average. Animal Use Protocol (AUP) and ethics approval was provided by the Animal Care Committee (ACC) on behalf of Western University's Council on Animal Care. Animals used in this study were cared for in alignment with the policies and procedures approved by Western University's Senate and the ACC and Western University's Manual of Policies on institutional use of animals in research in training, testing and research. This study was prepared according to the Animal Research: Reporting of In vivo Experiments (ARRIVE) essential 10 guidelines [17].

The 3D LGE cardiac MR imaging was performed at sequential time points, including baseline (i.e., before pacing), in-vivo (after pacing) and in-situ (post-mortem). After establishing a reference level of MF during baseline imaging, the canines underwent

surgical insertion of a pacing device as per the canine model developed by Dr. Stanley Nattel at the Montreal Heart Institute [18]. This model involves 5 weeks of rapid ventricular pacing (220 to 240 bpm) to achieve progressive atrial fibrosis and has shown favourable findings compared to alternate canine models using direct LA pacing. Although prior canine models used to assess LA fibrosis and sympathetic innervation have primarily used atrial pacing-based protocols [19], the degree of LA fibrosis using these protocols may be only modest. In contrast, the expected LA fibrosis burden using the ventricular pacing model is 10-15% at 5 weeks [18]. Therefore, the rapid ventricular pacing model was used in this study. A ventricular pacemaker was implanted in all dogs under anesthesia and attached to a pacing lead in the right ventricular (RV) apex [18]. Control canines underwent surgical insertion of the pacing device but no rapid ventricular pacing, in order to assess the effect of surgical insertion on LA enhancement and MF, as determined by histopathological imaging.

Following the 5-week period, the pacing device was removed in order to perform 3D LGE imaging in-vivo. Following in-vivo imaging, canines were euthanized and 3D LGE imaging was repeated in-situ. Finally, the hearts were excised and ex-vivo 3D LGE-CMR and histopathological imaging was performed. The structure, geometry and position of the heart relative to other organs is lost during ex-vivo imaging, thus, comparison to in-vivo and baseline imaging is not possible. Therefore, we unfortunately were not able to use ex-vivo 3D LGE-CMR images for any analysis in this study due to inherent alterations to the metabolic function and anatomical structure of the heart from in-vivo to ex-vivo. The 3D LGE images were acquired using fast low angle shot (FLASH) MR imaging, i.e., a spoiled 3D inversion recovery gradient echo pulse sequence (“3D IR-FLASH”) [128],

[129] approximately 7-15 min post injection of 0.2 mmol/kg Gadovist® (Bayer Inc.) at baseline and in-vivo time points. A second dose of 0.2 mmol/kg gadolinium was administered 5 minutes prior to euthanization for in-situ (post-mortem) imaging. A flow-chart demonstrating the study protocol is demonstrated in Figure 5-1. Average sequence parameters for the 3D IR-FLASH sequence are given in Table 5-1.



Figure 5-1 Flow-chart of study protocol and 3D LGE-CMR imaging time points.

Six-step study protocol underwent by experimental canines. Control canines underwent the same sequence of events excluding the 5 weeks of rapid ventricular pacing (i.e., the third step in the flow-chart).

Table 5-1 Average pulse sequence parameters for 3D LGE-CMR sequence

Slice thickness (mm)	0.975
Resolution (mm ²)	0.625
TR (ms)	1.68
TE (ms)	1.34
TI (ms)	341.25
Flip angle (°)	19.5
Trigger time (ms)	484.06
Percent sampling (%)	100

5.2.2 Image Processing

5.2.2.1 Segmentation of LA myocardium and enhanced regions

Several image processing steps were performed on the 3D IR-FLASH images at baseline, in-vivo, and in-situ time points. The image processing pipeline used to quantify MF in canine hearts at each time point is shown in Figure 5-2.

Firstly, the LA chamber (i.e., blood pool) was segmented manually by a trained operator using the Cardiac MRI Toolkit (Scientific Computing and Imaging Institute,

University of Utah, 2019) in 3D Slicer software [22]. To delineate the borders of the LA myocardial wall, axial dilation and Boolean remove algorithms were applied to the manually-generated LA chamber segmentations. The axial dilate algorithm performs a 4-pixel dilation in the axial plane to expand the manual chamber segmentation into the myocardial wall boundaries. The Boolean remove algorithm subtracts the blood pool mask from the dilated mask, generating a segmentation of the LA myocardial wall. The operator is required to make manual adjustments, as necessary, to ensure non-myocardial tissue regions are erased from the Boolean remove output (i.e., the region between the LA and RA septum). An example of the execution of this workflow is depicted in Figure 5-3. This process was performed at each time point. In situ images allowed for increased accuracy in manual blood pool chamber segmentation, since the images were not corrupted by cardiac and respiratory motion. These images served as an anatomical guideline for baseline and post-pacing time points of the same animal.

To quantify MF, the image intensity ratio (IIR) thresholding method [13] was used in order to compare computed fibrosis percentages across different animals. Voxel intensities of the LA myocardium were first normalized by dividing each voxel in the LA by the mean intensity of the LA blood pool, as in Equation 5-1. An alternative, commonly used normalization technique [23] instead normalizes the LA wall intensities by subtracting the mean intensity of the LA blood pool from raw LA wall intensities, and dividing the result by the LA blood pool standard deviation. However, this technique was not appropriate for the canine CMR images, as average blood pool intensities exceeded most LA wall intensities among most canines due to dosing and slow washout of gadolinium, leading to stagnancy of gadolinium, and therefore, excessive brightness in the blood pool.

Normalized left atrial wall intensity is given by the IIR normalization technique,

$$NI(x) = \frac{I(x)}{\mu_{BP}} \quad \text{Equation 5-1}$$

where $NI(x)$ is normalized left atrial wall intensity, $I(x)$ is the original intensity and μ_{BP} is the average signal intensity of the left atrial blood pool.

Atrial MF was then quantified using an IIR threshold $>$ mean IIR + 2 IIR SDs (Equation 5-2). This technique is notably different from the conventional signal threshold versus reference mean (STRM) [24] technique, wherein the fibrosis threshold is based on the mean intensity plus 2 – 4 standard deviations of a manually-selected remote (normal) LA myocardial region of interest (ROI). The thin nature of the LA myocardium constitutes unavoidable partial volume effect from inclusion of MF and/or adjacent blood pool in the remote LA myocardium ROI, which may significantly confound MF measurements. On the contrary, the IIR thresholding technique does not rely on manually selected healthy tissue ROI's. Instead, the fibrosis threshold is based solely on the normalized signal intensity of the LA myocardium. The IIR+2SDs technique was selected after establishing that this threshold yielded the best quantification results for LA enhancement versus other thresholding techniques, including IIR+ 3 or 4 SDs, which exhibited too few fibrotic signal intensities to quantify voxel values, and STRM using a mean reference intensity derived from remote left ventricle (i.e., not the left atrium, in order to mitigate partial volume effect) and 2 – 4 SDs, which significantly over-segmented enhanced LA regions.

The threshold for LA fibrosis quantification is given by,

$$T_{fib} = \mu_{NI} + 2\sigma_{NI} \quad \text{Equation 5-2}$$

where μ_{NI} is the mean IIR normalized intensity of the left atrial wall and σ_{NI} is the standard deviation of IIR normalized intensities.

MF volumes were computed by first finding the sum of all enhanced voxels within the LA myocardial wall and multiplying that sum by the voxel dimensions. We computed the percentage of MF (%MF) in the entire LA myocardium by dividing MF volumes by the overall myocardial wall volume, which was computed using the aforementioned approach. Normalization, IIR thresholding and MF volume quantification were computed using Matric Laboratory (MATLAB) 2020a. This process was performed at all time points. However, we report LA enhancement volumes only in baseline and post-pacing time points. This is because additional dosing of a gadolinium-based contrast agent prior to euthanization and lack of physiological blood flow caused a pooling and stagnation of gadolinium within the blood pool chamber in the in-situ images, which served chiefly to assist in delineating reference myocardial contours, uncorrupted by physiological artifacts encountered on in-vivo imaging. This phenomenon impaired fibrosis quantification, as the bright gadolinium caused a substantial amplification of the blood pool signal, thereby illegitimizing normalization using the mean blood pool intensity. Since the IIR thresholding technique relies on blood pool-based normalization, MF quantification in these images using this technique is not appropriate, as quantification results would be unreliable. The baseline and post-pacing images do not present these complications, and therefore LA enhancement volumes are more reliable at these time points.

The %MF measured across the entire LA myocardium is measured on a 3D volumetric basis, wherein every LA myocardial slice is used toward the computation, inclusive of image slices where no fibrosis is present. However, validation using histopathological imaging is measured on a 2D slice-wise basis, wherein the %MF is computed across the area of only one imaging slice. Therefore, in order to better

corroborate our fibrosis quantification method with histopathology, we also computed the area-wise %MF using the IIR thresholding technique described above in a single slice of the 3D LGE imaging stack. We chose the slice demonstrating the greatest extent of LA enhancement for both baseline and post-pacing time points. Area-wise %MF was computed by taking the sum of all enhanced MF pixels in the slice and dividing by the overall myocardial wall area.

The evolution of LA enhancement was assessed quantitatively and qualitatively across baseline and post-pacing time points. Quantitative assessment was performed by comparing %MF sequentially between baseline and post-pacing images for individual dogs. We also computed the mean and median %MF across experimental and control groups at baseline and post-pacing. In addition, we compute the absolute %MF difference (D_{abs}) between post-pacing and baseline time points for all dogs using the following approach,

$$D_{abs} = V_p - V_b \quad \text{Equation 5-3}$$

where V_b and V_p are the %MF values measured at baseline and post-pacing, respectively. Notably, the absolute difference D_{abs} is not expressed as an absolute value. The mean and median values of D_{abs} were computed and compared between experimental and control dog groups. Qualitative assessment was performed by assessing the spatial distribution of MF between images at all consecutive time points following landmark-based registration, described below.

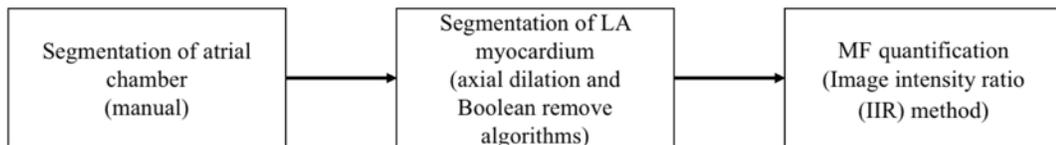


Figure 5-2 Image processing pipeline for MF segmentation of the LA.

Image processing pipeline used to quantify MF in 3D IR-FLASH LGE-CMR images at baseline, in-vivo and in-situ time points.

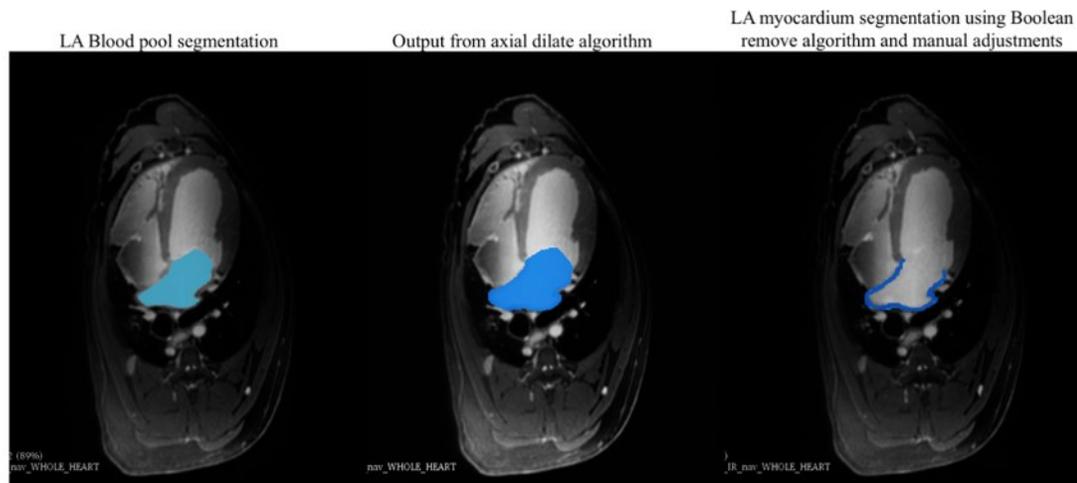


Figure 5-3 Example workflow of canine LA myocardial segmentation.

Example workflow of LA blood pool and myocardial segmentation using axial dilate and Boolean remove algorithms in in-situ IR-FLASH CMR images. Manual segmentation of the LA blood pool is shown on the left. The axial dilate algorithm expands the blood pool segmentation, as demonstrated in the middle image. The Boolean remove algorithm subtracts the blood pool chamber (left) from the axially dilated segmentation (middle) to reveal the LA wall boundaries. Corrections are done manually to remove erroneous areas as needed, before revealing the final LA myocardial wall segmentation (right).

5.2.2.2 Landmark-based 3D image registration

To localize the progression of fibrosis across time points, image registration must be performed. Towards this aim, we performed a landmark-based registration to mitigate motion differences and any disparities in anatomical positioning/shape between the 3D IR-FLASH at baseline and post-device insertion. Thirteen landmarks were selected within the 3D IR-FLASH imaging stacks at each time point using the Fiducial module in 3D Slicer. These landmarks included the ascending and descending aorta, aortic arch, left and right pulmonary arteries, inferior/superior left and right pulmonary veins, left atrial/ventricular septum, LA appendage, left anterior descending artery, and aortic valve.

Following selection of landmarks, rigid registration was performed using the Fiducial Registration module in 3D Slicer. Fixed landmarks were designated in the post-pacing images, while baseline and in-situ fiducials were chosen as moving landmarks (i.e., the resultant transform maps fiducials from the baseline/in-situ space to the fixed landmarks

in post-pacing space). In rigid registration, the resultant transform minimizes the sum of squared differences between fixed and moving landmarks.

The landmark-based registration technique allowed myocardial volumes to be rendered in the same space for qualitative assessment across time points. However, inherently, myocardial positioning and/or shape will change among images acquired post-operatively at different time points. Thus, quantitative overlap metrics were not computed on MF volumes across different time points.

5.2.2.3 Histopathological imaging

Histopathological imaging of the LA was performed on a slice-by-slice basis in six of the nine experimental canines to permit fibrosis quantification at Sunnybrook Hospital Pathology Department (Toronto, ON). Immediately after sacrifice, canine hearts were excised and immersed in formalin fixative for a minimum of 7 days, after which, the LA was isolated for tissue analysis. The LA was embedded in paraffin to allow for slice-by-slice analysis. Paraffin blocks were cut every 5 mm through the LA, and each block had sequential slices cut 4 μm thick using a large-scale microtome. Slices were put on positively charged slides and deparaffinized prior to staining with Masson's Trichrome. After staining, slides were imaged on brightfield mode with a Tissuescope LE (Huron Digital Pathology) on 20x magnification. Images were created as non-proprietary 24-bit RGB Pyramidal Big TIFF with JPEG 2000 compression. Between 2 and 7 slices were acquired per canine in the LA. Fibrosis was quantified using FIJI v. 1.49v software and analyzed using a script that quantifies the percentage of fibrotic tissue in each sample by setting thresholds for fibrosis (blue) and nonfibrotic tissue (red). These thresholds are optimized across each sample in the study in order to normalize fibrosis quantification.

5.2.3 Method Evaluation and Statistical Analysis

5.2.3.1 Observer variability

In order to validate our MF quantification technique, we performed an inter- and intra- observer variability analysis using two operators. Due to resource and time restraints, we randomly selected three animals (2 experimental, 1 control) for this analysis. Each operator performed the MF segmentation as per the pipeline in Figure 5-2 three times at each time point per animal. To assess the agreement of overlap of segmentation volumes across operators, we computed the dice similarity coefficient (DSC) for the blood pool chamber, LA myocardial wall and MF regions across each operator's three attempts. We also computed the DSC metric on the aforementioned regions among all attempts between both operators. The equation for DSC is given in section 2.4.1 of this document. In addition to computation of the DSC, we also used the blood pool chamber, LA myocardial wall and MF volumes per each attempt/operator to compute boundary F1 score, precision, recall, and relative volume error (see Appendix B).

For each of the operators' three attempts per time point, we computed the volumetric %MF across the entire LA myocardium. Intra-observer variability of %MF measurements was evaluated using the intraclass correlation coefficient on %MF calculated from the operators' three segmentation attempts. The calculation for ICC is given as follows:

$$ICC = (1 + \frac{\sigma_w^2}{\sigma_b^2})^{-1} \quad \text{Equation 5-4}$$

where σ_w^2 is the variance within measurements reported by one operator and σ_b^2 is the variance between measurements reported by both operators [25].

Inter-observer variability of %MF measurements was evaluated using the inter-rater reliability metric, given by:

$$R = \left(1 + \frac{\sigma_w^2}{k\sigma_b^2}\right)^{-1} \quad \text{Equation 5-5}$$

where σ_w^2 and σ_b^2 are defined as denoted in Equation 4-4 and k indicates the number of operators [25]. The *ICC* and *R* metrics range between 0 and 1, such values closer to 0 denote poor agreement (high variation) between observations and values closer to 1 denote high agreement (low variation) between observations.

5.2.3.2 Statistical Tests

In addition to the aforementioned evaluation and reliability metrics, we validated our segmentation technique using various statistical tests. Non-parametric statistics were computed in this study due to the small dataset size. To assess the significance between baseline and post-pacing MF measurements within experimental and control groups, we used the Wilcoxon signed rank test on dogs paired %MF measurements within respective groups. To compare differences between post-pacing %MF measurements computed among experimental versus control groups, we used the Mann Whitney U test. The Mann Whitney U Test compares two nonpaired groups with which data is not normally distributed.

5.3 Results

5.3.1 Segmentation of MF volumes using 3D-FLASH CMR

Example CMR images at each time point for experimental dog 7 are demonstrated in Figure 5-4. MF volumes computed using the IIR technique were quantified in millimetres and expressed as a percentage of the overall LA myocardial wall for each subject (%MF), as shown in Table 5-2. The threshold used to define fibrosis was the mean

normalized value + 2 SDs, ranging between 0.90 – 1.3 across all time points for each dog. The average baseline IIR threshold, expressed as average threshold \pm standard deviation, was 1.10 ± 0.12 at baseline and 1.0 ± 0.05 post-pacing among experimental canines. Among the three control dogs, average IIR thresholds of 1.10 ± 0.06 and 1.15 ± 0.07 were used at baseline and post device-insertion time points, respectively.

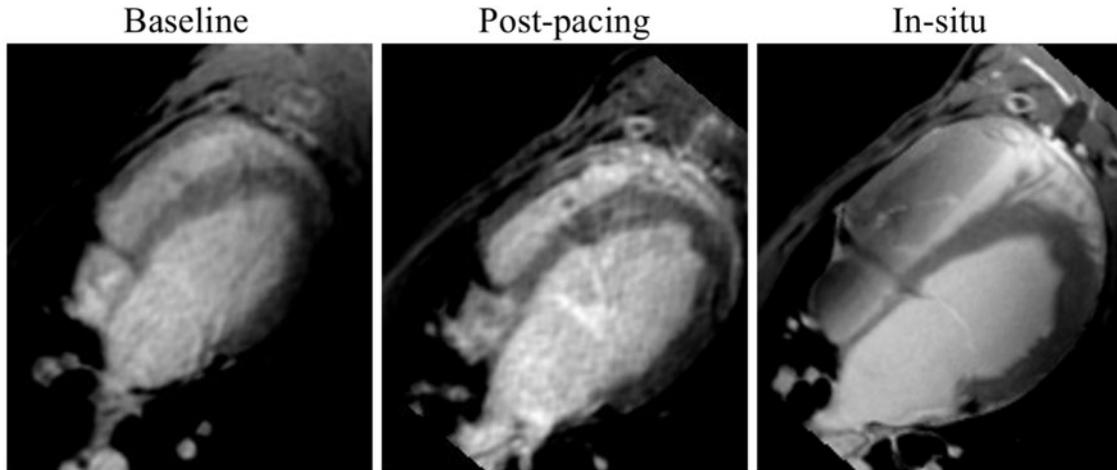


Figure 5-4 Images of the left atrium in canines at different time points.

A single slice from 3D IR-FLASH CMR acquisitions are demonstrated at baseline (left column), post-pacing (middle) and in-situ (right column) in experimental dog 7 at approximately the same slice location. LA enhancement is demonstrated in the anterior aspect of the LA in post-pacing and in-situ images. Post-pacing images are significantly corrupted by motion artifact due to irregular, elevated heart rate that occurred as a result of abnormal pacing.

Table 5-2 LA MF volumes computed using IIR technique

LA fibrosis volumes computed using the IIR technique are expressed as a percentage of MF across the overall LA myocardial wall (%MF) and are shown at baseline and post-pacing (i.e., post device-insertion for the control group) for each canine in experimental and control groups. In the third column, we express the absolute difference of volumetric %MF D_{abs} between baseline and post device-insertion time points. Average and median values of %MF and D_{abs} are demonstrated in the final two rows of experimental and control groups, as mean volume \pm standard deviation and median volume \pm interquartile range (IQR), respectively.

Subject	Baseline %MF over total LA myocardial volume	Post-pacing %MF over total LA myocardial volume	Absolute %MF difference (D_{abs})
Paced dogs			
Dog 1	2.71%	3.59%	0.87%
Dog 2	0.99%	2.00%	1.10%
Dog 3	0.11%	1.93%	1.82%
Dog 4	2.25%	4.33%	2.08%
Dog 5	2.05%	1.40%	-0.65%
Dog 6	2.56%	2.20%	-0.36%
Dog 7	0.12%	2.96%	2.84%
Dog 8	0.38%	3.76%	3.38%

Dog 9	0.29%	3.08%	2.79%
All paced dogs (mean \pm SD)	1.27 \pm 1.10%	2.80 \pm 0.98%	1.53 \pm 1.42%
All paced dogs (median \pm IQR)	0.99 \pm 2.08%	2.96 \pm 1.64%	1.82 \pm 2.24%
Control dogs			
Control 1	0.05%	0.72%	0.67%
Control 2	0.39%	0.95%	0.56%
Control 3	1.64%	1.12%	-0.52%
All control dogs (mean \pm SD)	0.69 \pm 0.83%	0.93 \pm 0.20%	0.23 \pm 0.65%
All control dogs (median \pm IQR)	0.39 \pm 1.2%	0.95 \pm 0.30%	0.56 \pm 0.89%

Figure 5-5 demonstrates a dot and line plot showing the change in volumetric %MF for each animal in experimental and control groups across baseline and post device-insertion time points. The mean %MF across all experimental and control animals was computed by taking the average of all reported %MF values at baseline and post-pacing time points. We computed a mean %MF of 1.27 \pm 1.10 % at baseline and 2.80 \pm 0.98% post-pacing in experimental dogs (mean \pm SD). In control dogs, the mean %MF was computed at 0.69 \pm 0.83% at baseline versus 0.93 \pm 0.20% post-pacing (mean \pm SD). We also reported median %MF values of 0.99 \pm 2.08% versus 2.96 \pm 1.64% (median \pm IQR) at baseline and post-pacing, respectively, in experimental dogs. Median %MF values of 0.39 \pm 1.2% versus 0.95 \pm 0.30% (median \pm IQR) are reported for control dogs at baseline and post device-insertion time points, respectively. We performed a Wilcoxon signed rank test on experimental dogs' paired MF measurements, which demonstrated statistically significant differences at baseline versus post-pacing time points (i.e., P=0.019).

We computed the mean absolute difference of %MF (D_{abs}) between time points by subtracting the %MF reported at post-pacing from the %MF reported at baseline across each animal, then taking the mean of those absolute differences. An increase in the reported %MF was demonstrated in seven of nine experimental dogs (i.e., paced dogs 1, 2, 3, 4, 7,

8 and 9) across baseline to post-pacing time points. The mean D_{abs} computed across these seven animals was measured at $2.11 \pm 0.95\%$ (mean \pm SD). Dogs 5 and 6 demonstrate a decrease in %MF of 0.65% and 0.36%, respectively, from baseline to post-pacing. Taking into consideration dogs 5 and 6, the mean D_{abs} from baseline to post-pacing was computed at $1.53 \pm 1.42\%$. Similarly, we also computed the median D_{abs} by finding the median value of all absolute %MF differences between post-pacing and baseline time points. Across the seven dogs with which an increase in LA enhancement was reported, the median D_{abs} demonstrates an increase of $2.08 \pm 1.62\%$ (median \pm IQR). Taking into consideration dogs 5 and 6 wherein a decrease in LA enhancement is reported from baseline to post-pacing, we computed a median D_{abs} of $1.82 \pm 2.24\%$ (median \pm IQR).

We report a mean D_{abs} of $0.23\% \pm 0.65\%$ (mean \pm SD) and median D_{abs} of $0.56 \pm 0.89\%$ (median \pm IQR) between baseline and post-pacing time points in control dogs. Control dog 3 demonstrates a decrease in %MF from baseline to post device-insertion, while control dogs 1 and 2 demonstrate an increase in %MF. However, the increase in %MF between baseline and post device-insertion that is exhibited in control dogs 1 and 2 is less than 0.67%, compared to an increase of 2.11% (mean) in experimental dogs. A Wilcoxon signed rank test performed on control dogs paired volumetric fibrosis measurements did not demonstrate a statistically significant difference in measurements from baseline to post device-insertion time points (i.e., $P=0.5$). Thus, no significant response was demonstrated in the control dogs between baseline and post-pacing time points.

Post-pacing volumetric fibrosis measurements were compared between the experimental and control dogs. A Mann-Whitney U-Test demonstrated a statistically

significant difference in fibrosis measurements for paced vs. non-paced dogs at the 5-week post-insertion time point (one-tailed P-value = 0.0045, two-tailed P-value = 0.009).

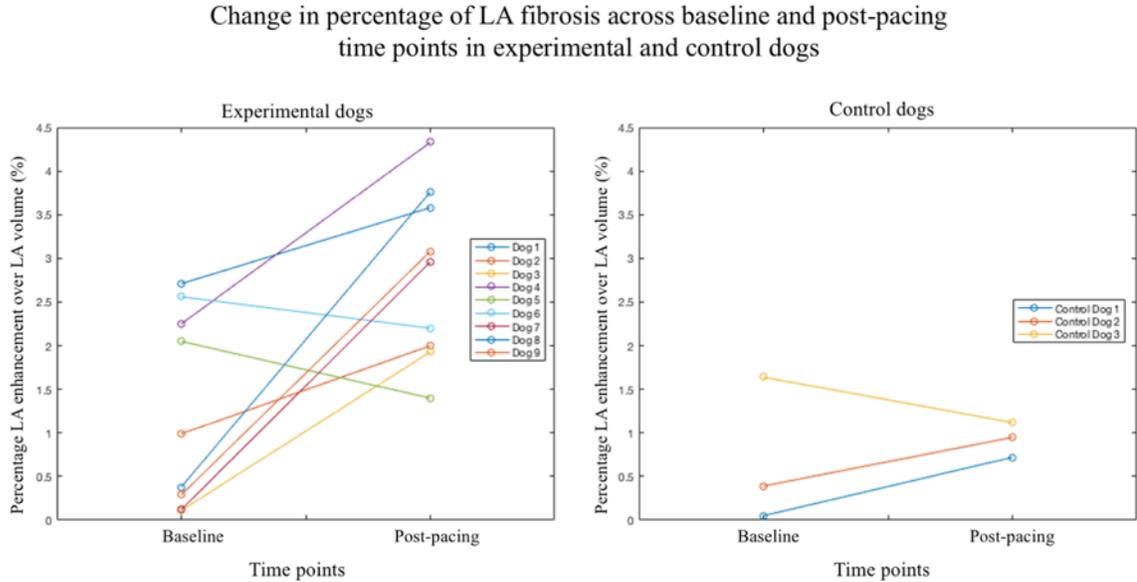


Figure 5-5 Dot and line plot of MF volumes in canines. Dot and line plot demonstrating the change in percent fibrosis over the LA volume for experimental (paced) dogs (left) and control dogs (right). Baseline and post-pacing time points are indicated on the x-axis while LA fibrosis percentage is demonstrated on the y-axis.

5.3.2 Segmentation of slice-wise MF areas using 3D-FLASH CMR

The %MF reported at baseline and post-pacing time points in Table 5-2 incorporates every LA myocardial slice into the computation, inclusive of slices where no myocardial scar is present. However, histopathological imaging is performed on a 2D basis, such that %MF is computed across the area of a single slice. Therefore, we measured the 2D slice-wise %MF in the slice demonstrating the greatest extent of LA enhancement across each time point, demonstrated in Table 5-3.

Table 5-3 Canine slice-wise LA MF measurements at different time points.

Slice-wise (i.e., area-wise) LA fibrosis measurements are expressed as a percentage of MF (%MF) over the slice with which the greatest extent of fibrosis was shown. Area-wise %MF measurements are shown at baseline and post-pacing (i.e., post device-insertion for the control group) for each canine in experimental and control groups. In the third column, we express the absolute difference of area-wise %MF D_{abs} between baseline and post-pacing. Average and median values of %MF and D_{abs} are demonstrated in the final two rows of experimental and control groups, as mean volume \pm standard deviation and median volume \pm interquartile range (IQR), respectively.

Subject	Baseline %MF over single LA myocardial slice area	Post-pacing %MF over single LA myocardial slice area	Absolute %MF difference (D_{abs})
Paced dogs			
Dog 1	11.30%	34.46%	23.16%
Dog 2	7.16%	11.22%	4.06%
Dog 3	3.22%	16.61%	13.39 %
Dog 4	10.63%	22.51%	11.88 %
Dog 5	6.04%	5.13%	-0.91%
Dog 6	6.58%	3.34%	-3.24%
Dog 7	1.63%	8.19%	6.56%
Dog 8	1.64%	19.66%	18.02%
Dog 9	1.16%	8.38%	7.22%
All paced dogs (mean \pm SD)	5.48 \pm 3.84%	14.38 \pm 9.96%	8.90 \pm 8.59%
All paced dogs (median \pm IQR)	6.04 \pm 6.39%	11.22 \pm 12.95%	7.22 \pm 11.73%
Control dogs			
Control 1	9.24%	9.98%	0.74%
Control 2	9.98%	9.28%	-0.70%
Control 3	10.46%	11.52%	1.06%
All control dogs (mean \pm SD)	9.89 \pm 0.61%	10.26 \pm 1.15%	0.37 \pm 0.94%
All control dogs (median \pm IQR)	9.98 \pm 0.92%	9.98 \pm 1.68%	0.74 \pm 1.32%

Similar to the volumetric %MF measurements demonstrated in Table 5-2, an increase in the reported area-wise %MF was demonstrated in seven of nine experimental dogs (i.e., paced dogs 1, 2, 3, 4, 7, 8 and 9) across baseline to post-pacing time points. Dogs 5 and 6 demonstrate a decrease in the absolute area-wise %MF difference (D_{abs}), at 0.91% and 3.24%, respectively, from baseline to post-pacing. The mean area-wise %MF across all experimental animals was computed at 5.48 \pm 3.84% at baseline and 14.38 \pm 9.96% post-pacing, demonstrating an absolute %MF difference D_{abs} of 8.90 \pm 8.59% (mean \pm SD). Median %MF values of 6.04 \pm 6.39% versus 11.22 \pm 12.95% (median \pm IQR) are reported in experimental dogs at baseline and post-pacing time points, respectively, yielding a median absolute difference D_{abs} of 7.22 \pm 11.73%. A Wilcoxon signed rank test demonstrated a statistically significant difference in experimental dogs' slice-wise paired MF measurements at baseline versus post-pacing time points (i.e., $P=0.0195$).

In control dogs, the mean area-wise %MF was computed as $9.89 \pm 0.61\%$ and $10.26 \pm 1.15\%$ at baseline and post-pacing time points, respectively, demonstrating an absolute %MF difference D_{abs} of $0.37 \pm 0.94\%$ (mean \pm SD). Median %MF values of $9.98 \pm 0.92\%$ versus $9.98 \pm 1.68\%$ (median \pm IQR) are reported for control dogs at baseline and post device-insertion time points, respectively. A Wilcoxon signed rank test did not demonstrate a statistically significant difference in control dogs' area-wise %MF measurements from baseline to post device-insertion time points (i.e., $P=0.5$). A Mann-Whitney U test did not demonstrate statistically significant differences between area-wise %MF measurements in experimental dogs versus control dogs ($P=1$).

5.3.3 Registration of Myocardial Volumes

Landmark-based image registration was used to qualitatively assess LA enhancement distribution among canines across all time points. Although the comparison of MF volumes was performed among the same animals, the baseline, post-pacing and in-situ images are acquired at different time points post-operatively, thereby posing disparities in myocardial positioning/shape. Thus, overlap metrics were not computed on fibrosis volumes across different time points. Figure 5-6 demonstrates the evolution of LA enhancement across time points after landmark-based registration in three example dogs. Evolution of LA enhancement in control dogs is demonstrated in Figure 5-7.

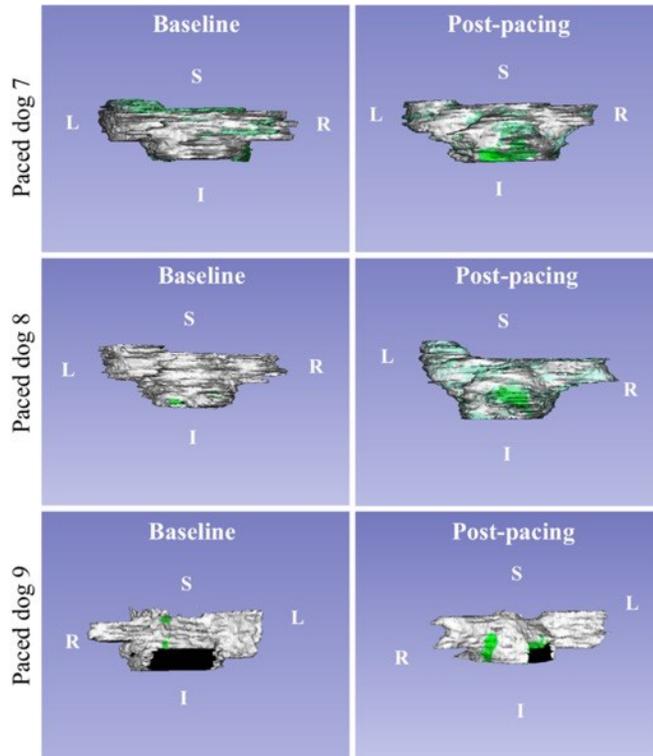


Figure 5-6 Evolution of LA MF volumes across time points (experimental dogs)
 LA enhancement volumes (green color) in three example dogs at baseline (left), post-pacing (middle) and in-situ (right). The LA myocardial volume is rendered in white color, while enhanced regions are depicted in green. The black regions demonstrate regions outside of the myocardium, such as the septum space between the left atrium to left ventricle (such as in Paced dog 9).

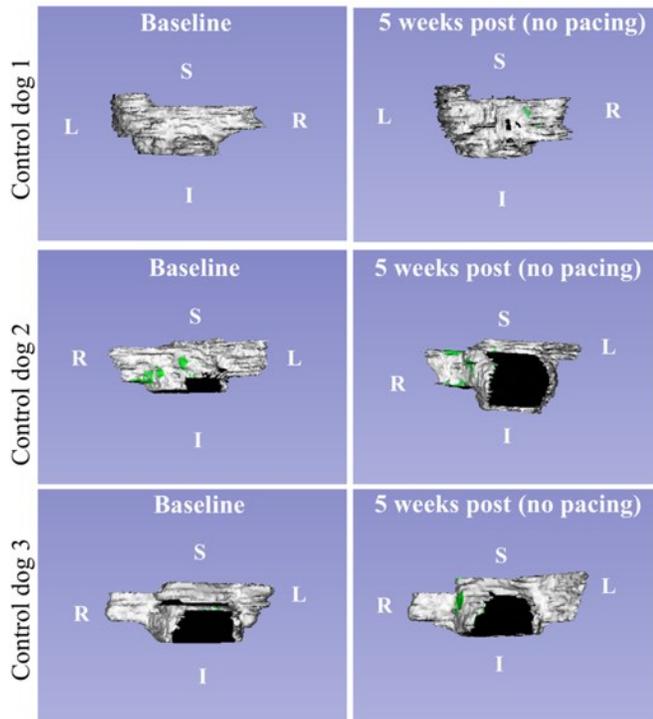


Figure 5-7 Evolution of LA MF volumes across time points (control dogs).

LA enhancement volumes are shown in green in control dogs at baseline (left) and 5 weeks post insertion of pacing device (right) with no rapid ventricular pacing. In-situ 3D-IR-GRE MR images were not acquired in these canines.

5.3.4 Histopathological Validation

Fibrosis regions were quantified in a voxel-by-voxel basis in histopathological images and are expressed as a percentage of the myocardium per slice. Left atrial fibrosis volumes are summarized in Table 5-4. Measured LA fibrosis volumes range between 8.74 – 33.18% of myocardium per histopathological slice in the paced dogs. Among the five paced dogs, we report a mean of $19.42 \pm 4.80\%$ (mean \pm SD) and median of $20.77 \pm 3.92\%$ (median \pm IQR) across the reported average percentages of slice-wise MF per animal. Taking into account the control dog, we report an average %MF of $16.48 \pm 8.36\%$ (mean \pm SD) and median %MF of $20.47 \pm 10.06\%$ (median \pm IQR) across all six dogs with which histopathological imaging was acquired.

Six histopathological slices were acquired from one control dog, with computed MF percentages ranging between 0.34 – 3.06%. The average MF volume computed across histopathological slices acquired from the control dog was 1.85%. Figure 5-8 demonstrates example histopathological images of two experimental and one control dog after staining using Masson’s Trichome.

Table 5-4 MF measurements in corresponding histopathological images

Left atrial fibrosis volumes computed from histopathological images in 5 experimental and 1 control canine. Regions of fibrosis are expressed as a percentage with respect to healthy myocardium for each slice.

<i>Paced dogs</i>								
Slice #	1	2	3	4	5	6	7	Average
Dog 1	N/A	32.43%	23.43%	N/A	13.09%	12.77%	19.17%	20.18%
Dog 2	29.51%	23.38%	25.63%	19.35%	18.53%	17.36%	14.83%	21.23%
Dog 3	N/A	N/A	N/A	22.17%	19.37%	N/A	N/A	20.77%
Dog 5	N/A	N/A	N/A	N/A	N/A	14.28%	33.18%	23.73%
Dog 6	N/A	8.99%	12.87%	8.74%	11.69%	13.57%	N/A	11.17%
<i>Control dogs</i>								
Control 3	3.06%	2.55%	2.24%	0.34%	1.08%	N/A	N/A	1.85%

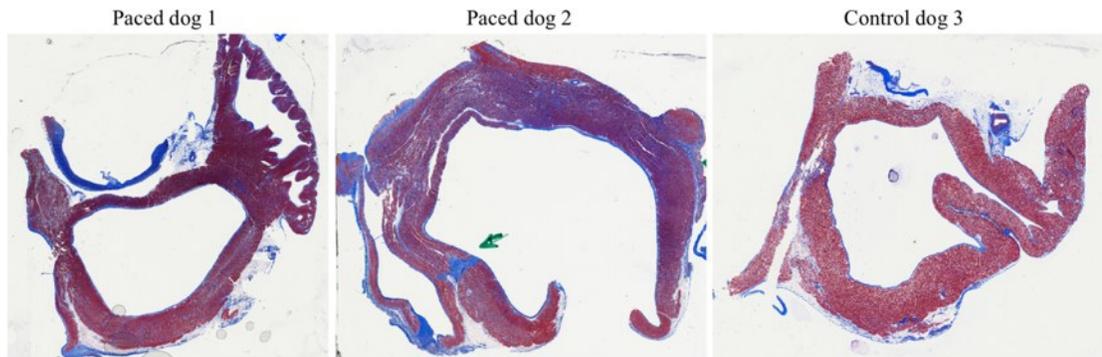


Figure 5-8 Histopathological images of LA samples demonstrating MF regions. Example histopathological images of left atrial samples acquired after Masson's Trichrome staining. Fibrotic areas are shown in blue while normal tissue is shown in red.

5.3.5 Observer Variability

Two operators (redacted) were selected to perform an inter/intra- operator variability analysis on our LA MF segmentation pipeline among three animals (2 experimental, 1 control). The dice similarity coefficient (DSC) was computed on segmentations of the myocardial chamber, myocardial LA wall and MF volumes across both operators, as demonstrated in Table 5-5. We report the %MF computed per the three attempts and the mean %MF across the attempts for each operator in Table 5-6. Boundary F1 score (a measure of the boundary overlap relative to a distance metric), precision, recall and relative volumetric error are reported across each user in Appendix B.

Mean DSC metrics computed across the blood pool chamber demonstrate high agreement in both intra-observer and inter-observer comparison, i.e., >81% reported across both time points for all dogs among both operators. However, accuracy declines as the segmentation becomes refined to smaller regions. For the LA myocardial wall, DSC is reported in the range of 45.03 – 69.49% for Operator 1 and is reported in the range of 51.81 – 71.45% for Operator 2 across all dogs/time points. Among both operators, we report LA myocardial wall DSC agreement in the range of 55.80 – 65.81% across all dogs/time points.

The DSC agreement is lowest among MF volume regions in both intra- and inter-observer comparison. We report MF volume DSC metrics in the range of 8.88 – 20.23% for Operator 1, 7.88 – 38.80% for Operator 2, and 0.28 – 34.26% across both operators.

Both operators also computed the volumetric %MF at baseline and post-pacing time points across for all attempts across all dogs, as shown in Table 5-6. We also computed the intraclass correlation coefficient (ICC) and inter-rater reliability (R) metrics using the %MF (i.e., not the MF volume), also reported among the three attempts per operator in Table 5-6.

Table 5-5 Observer variability DSC metrics

Average dice similarity coefficient (DSC) computed on three segmentation attempts performed by two operators across each time point in a subsample of three animals. The DSC is given as a percentage and is expressed as mean ± SD.

		Paced dog 1		Paced dog 3		Control dog 3	
Operator 1	Region	Baseline	Post-pacing	Baseline	Post-pacing	Baseline	Post-pacing
	DSC Blood pool (%)	88.37 ± 2.09	85.51 ± 1.87	92.13 ± 1.01	89.61 ± 0.82	91.72 ± 1.00	90.49 ± 2.24
	DSC LA wall (%)	49.34 ± 13.57	45.03 ± 7.15	66.01 ± 7.65	64.23 ± 3.02	67.05 ± 3.43	69.49 ± 3.28
	DSC MF volume (%)	14.64 ± 19.52	15.16±20.37	9.80 ± 16.98	20.23 ±30.55	6.84 ± 9.89	8.88 ± 14.26
Operator 2	DSC Blood pool (%)	90.40 ± 1.63	91.08 ± 0.08	94.51 ± 1.08	91.96 ± 1.71	94.57 ± 0.10	94.75 ± 0.15
	DSC LA wall (%)	51.81± 11.21	54.94 ± 3.97	65.58 ± 7.96	65.55 ± 5.58	69.75 ± 2.61	71.45 ± 1.52
	DSC MF volume (%)	17.04± 12.63	24.48±5.24	15.80±12.15	7.88 ± 3.43	25.53±10.81	38.80±11.58
Operator 1 vs. Operator 2	DSC Blood pool (%)	88.56 ± 1.21	87.43 ± 4.43	90.84 ± 1.77	88.06 ± 1.56	81.64 ± 0.80	90.79 ± 0.90
	DSC LA wall (%)	65.81 ± 23.90	64.91 ± 21.05	58.23 ± 5.77	55.80 ± 3.19	59.89 ± 1.88	63.62 ± 1.73
	DSC MF volume (%)	34.26 ± 19.91	25.28 ± 21.58	0.48 ± 0.67	0.28 ± 0.46	2.52 ± 2.71	14.58 ± 12.65

Table 5-6 LA MF volumes measured by different operators

Computed %MF across a subset of three animals recorded by two operators. The %MF is expressed across three attempts and as a mean ± SD value across the attempts per operator at each time point. The intraclass correlation coefficient (ICC) is computed across attempts for each operator. Inter-rater reliability (R) is expressed across all attempts among both operators.

		Paced dog 1			Paced dog 3			Control dog 3		
		Baseline	Post-pacing	Absolute difference (D _{abs})	Baseline	Post-pacing	Absolute difference (D _{abs})	Baseline	Post-pacing	Absolute difference (D _{abs})
Operator 1	Attempt 1 MF%	2.45	4.32	1.87	0.40	1.86	1.46	1.71	0.94	-0.77
	Attempt 2 MF%	2.23	4.84	2.61	1.35	2.75	1.40	1.54	1.31	-0.23
	Attempt 3 MF%	2.70	3.58	0.88	1.07	1.93	0.86	1.63	1.11	-0.52
	Mean %MF±SD	2.46±0.23	4.25±0.63	1.79±0.89	0.94±0.49	2.18±0.49	1.24±0.33	1.62±0.09	1.12±0.19	-0.51±0.27
	Median %MF±IQR	2.45±0.35	4.32±0.94	1.87±1.30	1.07±0.71	1.93±0.67	1.40±0.45	1.63±0.13	1.11±0.28	-0.52±0.41
	ICC	0.90	0.53	0.64	0.58	0.45	0.48	0.84	0.91	0.88
Operator 2	Attempt 1 MF%	2.24	3.11	0.87	2.00	2.78	0.78	1.34	1.28	-0.06
	Attempt 2 MF%	3.80	3.18	-0.62	1.19	2.00	0.81	1.23	2.55	1.32
	Attempt 3 MF%	3.82	3.93	0.11	1.89	2.70	0.81	1.30	1.49	0.19
	Mean %MF±SD	3.29±0.90	3.41±0.45	0.12±0.75	1.69±0.43	2.49±0.43	0.80±0.02	1.29±0.06	1.77±0.68	0.48±0.74
	Median %MF±IQR	3.80±1.18	3.18±0.62	0.11±1.11	1.89±0.61	2.70±0.59	0.81±0.02	1.30±0.08	1.49±0.95	0.19±1.03
	ICC	0.40	0.68	0.71	0.64	0.52	0.99	0.93	0.41	0.49
R (Operator 1 vs. Operator 2)		0.58	0.69	0.81	0.74	0.62	0.82	0.92	0.59	0.66

The %MF measured across baseline to post-pacing time points demonstrates an increasing trend across all segmentation attempts by Operator 1 in both paced animals. Similar to the reported MF volumes in Table 5-2, all segmentation attempts by Operator 1 show a decrease in %MF from baseline to post-pacing in control dog 3.

Likewise, Operator 2 demonstrates an increasing trend across all segmentation attempts from baseline to post-pacing in paced dog 3 and two of three attempts for paced dog 1. However, Operator 2's %MF measurements for control dog 3 show both decreasing and increasing trends from baseline to post-pacing. Notably, a significant increase is shown at control dog 3's post-pacing time point in attempt 2 by Operator 2.

The %MF computed by Operator 1 demonstrates moderate to high agreement across baseline and post-pacing time points, wherein ICC ranges between 0.45 – 0.91

across all dogs and all time points. The highest agreement is reported across post-pacing %MF measurements in control dog 3 and lowest agreement is reported in post-pacing measurements in paced dog 3 for Operator 1.

Operator 2 also demonstrates moderate to high agreement in %MF measurements computed across all dogs and all time points, wherein ICC values range between 0.40 – 0.93. The highest agreement in %MF measurements is denoted at the baseline timepoint for control dog 3, while the lowest agreement is computed at the post-pacing timepoint in paced dog 1 for Operator 2. The ICC of absolute %MF differences (D_{abs}) also demonstrates moderate to high agreement, ranging from 0.48 – 0.88 and 0.49 – 0.99 for Operators 1 and 2, respectively.

Overall, moderate to high agreement is denoted by the inter-rater reliability metric (R) computed using %MF measurements across both operators, ranging between 0.58 – 0.92. The lowest inter-observer agreement is demonstrated at baseline for paced dog 1 and highest agreement is shown at baseline for control dog 3. On absolute %MF differences (D_{abs}), moderate to high agreement is also demonstrated, ranging between 0.66 – 0.82.

5.4 Discussion

In this study, we aimed to validate an objective method for left atrial MF quantification using confirmation from histopathological images and to corroborate causality of AF to MF in a canine model. We used the IIR method with threshold of IIR+2SDs to define fibrosis in 9 experimental and 3 control dogs. Overall, a trend of increasing fibrosis was observed across sequential time points. LA enhancement increased from baseline to post-pacing in 7 out of 9 experimental dogs. In control dogs, a marginal

increase was demonstrated on average across baseline to post device-insertion points but was negligible compared to the trend observed in animals who underwent pacing.

Two animals in the experimental group demonstrated a decrease in LA enhancement from baseline to post-pacing time points. This decrease in enhancement was less than 1% in both cases and may reflect differences in gadolinium dosing or timing of image acquisition in those animals. Alternatively, an increase in the enhancement threshold may reverse this trend. For example, one animal demonstrated a decrease from baseline to post-pacing time points using the IIR+2SDs threshold; however, an increase in enhancement was demonstrated in the same dog when using a threshold of IIR+3SDs.

Histopathological imaging was acquired in several samples for six experimental dogs. Fibrosis on LGE CMR imaging was identified and quantified and compared with histological evidence of fibrosis. The measured LA fibrosis volumes ranged between 8.74 – 33.18% of myocardium per slice, wherein we report an average %MF of $16.48 \pm 8.36\%$ (mean \pm SD) and median %MF of $20.47 \pm 10.06\%$ (median \pm IQR) across all six canines (i.e., inclusive of 5 experimental dogs and 1 control dog). Among the five paced dogs, we report a mean of $19.42 \pm 4.80\%$ (mean \pm SD) and median of $20.77 \pm 3.92\%$ (median \pm IQR) %MF. Dogs 5 and 6, which demonstrated a decrease in LA enhancement measurements across baseline to post-pacing CMR images, exhibit confirmed fibrosis in Masson Trichome-stained slides. Thus, the discrepancy in the increasing trend from baseline to post-pacing demonstrated in these dogs through LGE-CMR is likely artifactual, possibly attributable to partial volume effect. However, this is an important consideration demonstrating the difficulty of measuring the extent of LA fibrosis using LGE-CMR alone.

Notably, there is a discrepancy in volumetric %MF measured in post-pacing time points as compared to histopathology. Certain limitations contribute to this discrepancy, namely, coarse spatial resolution of LGE images, partial volume effect, and the fact that volumetric %MF quantifies scar across the entire LA myocardial volume as opposed to a single slice area as with histopathology. Thus, we computed a slice-wise area-wise %MF for better comparison against histopathology. For example, in paced dog 1, we report a volumetric extent of 3.59%MF post-pacing across the entire myocardium (i.e., all image slices). However, in the same dog at the post-pacing time point, the slice containing the highest number of MF pixels yields a computed area-wise extent of %MF of 34.46%, which is in better keeping with the %MF reported in histopathological imaging.

There are several limitations in this study. Firstly, our sample size of nine experimental and three control dogs is small. However, we have expressed the results on a subject-by-subject basis so as to demonstrate an increasing trend across time points in experimental dogs, and an absence of this trend across control dogs. In this study, we consider LGE of the LA to be indicative of the presence of MF. However, late enhancement may also be used to identify other pathophysiological conditions that manifest in expansion of the myocardial space, such as amyloid or myocardial edema. Thus, enhanced regions of the LA are assumed to be MF while in theory, these are areas of expanded myocardial space that could be a product of another disease process. However, through histopathological analysis we confirmed the presence of LA myocardial MF. Therefore, classifying enhanced areas as MF was a reasonable assumption.

We also acknowledge the limitations involved in quantification of MF using LGE-CMR technique caused by partial volume effect. Although 3D LGE-CMR slice thickness

and spatial resolution were $<1\text{mm}$ as demonstrated in Table 4-1, the typical thickness of the left atrial wall generally ranges between 2 – 4 mm; thus, the LGE-CMR method bears the risk of wrongly characterizing adjacent tissue as myocardium which would further confound measurements of MF. Moreover, the reliance of operator precision in correctly segmenting the myocardial wall further elicits this risk. While respiratory and cardiac navigated pulse sequences improve spatial resolution to mitigate partial volume, these sequences are additionally limited by image artifacts and motion-induced respiratory and/or cardiac blurring. Therefore, MF volumes computed in this study may be confounded by image quality. Pulse sequences that mitigate the effects of partial volume, motion and blurring due to arrhythmic heart beats and suboptimal spatial resolution should be developed in further studies.

This study was further limited by the tachycardia exhibited by paced animals during CMR image acquisition. Following 5 weeks of rapid ventricular pacing, the experimental dogs developed persistent atrial fibrillation, characterized by abnormally rapid heart rates. It is particularly challenging to select an inversion time that is long enough to result in satisfactory nulling of the myocardium when R-R intervals are short due to fast heart rates. Thus, enhanced LA regions in post-paced LGE CMR images demonstrated suboptimal contrast with healthy myocardium, potentially confounding the computed MF volumes (due to partial volume effect). Moreover, other confounding factors including contrast dosage, circulation kinetics, metabolic function, and surface coil proximity affect image quality and contrast between normal myocardium and fibrosis. Consequently, a threshold of mean IIR + 2 standard deviations was chosen in order to mitigate the effects of these factors and adequately account for regions consisting of intermediate degrees of MF. The

faster heart rates of canines also impacted the ability to acquire LGE CMR images in ventricular diastole, which introduced substantial motion artifacts in post-pacing images. Although these artifacts led to challenges in the manual segmentation of the atrial chamber, these challenges were mitigated by registering the post-pacing images with uncorrupted post-mortem in situ images. Nevertheless, manual segmentation of the canine LA region is a complex task and requires anatomical expertise, which can result in significant variation in measurements computed using the IIR thresholding technique between different operators and among multiple attempts by a single operator. This is demonstrated in our intra-/inter- observer variability analysis of %MF. Paced animals show variation in %MF measurements computed at post-pacing time points by both observers, wherein the ICC of absolute %MF differences (D_{abs}) ranges between 0.48 to 0.99 across both operators. However, the mean \pm SD of %MF measurements computed by both operators nevertheless demonstrates an increasing trend from baseline to post-pacing among both paced animals. In the control dog, Operator 1 computes a mean decrease while Operator 2 demonstrates a mean increase in %MF from baseline to post-pacing (mean \pm SD). Operator 2 computes a significant increase at the post-pacing time point for attempt 2 in control dog 3, however, this is likely an outlier as compared to metrics computed by the same operator in attempts 1 and 2. Nevertheless, this phenomenon demonstrates the impact that errors can have on %MF measurements.

We demonstrate high and moderate spatial agreement of blood pool and LA myocardial wall regions, respectively, among both operators. However, poor spatial agreement is reported in MF volume segmentations provided by both operators, as demonstrated by the low DSC metrics reported in this region both operators in Table 5-5.

This is likely due to the increasing necessity of operator interaction across the MF segmentation pipeline. The Boolean remove and Axial Dilate algorithms depend on the initial chamber segmentation but are imperfect in nature and require further alteration by the operator (i.e., removal of valves, septal regions, inclusion or exclusion of LA appendage, manual modifications to errors, etc.). Therefore, the decreasing trend of the DSC metric across blood pool, LA wall and MF regions is plausible, and in keeping with other studies [26]. Nevertheless, better spatial agreement is demonstrated by the boundary F1 score of the MF region (see Appendix B, Section B1, Table B1), which indicates more consistency in the boundary-wise detection of MF (i.e., across the edges, not the entire region). The BF score ranges between 29.77 – 87.09% across the MF volumes computed by both operators in our variability analysis, which is vastly superior to the DSC (entire region) metrics computed across both operators, which ranges between 0.28 – 34.26%. Moreover, moderate to high agreement was denoted by ICC and R metrics computed across both operators, which indicates that the percentages (not spatial volumes) of MF computed by both operators demonstrate some consistency. However, the ICC and R metrics are computed using a single value (i.e., the percentage of MF) as opposed to a volume. Thus, spatial agreement of %MF is not taken into consideration in the computation of these metrics. Improved spatial agreement of MF volumes may be facilitated through the development of additional segmentation pipelines wherein minimal observer interaction is required.

We also faced limitations in the correspondence of LA enhancement results derived from CMR images and histopathological quantification results. Inherent differences in image acquisition prevented us from doing more rigorous analysis of the relationship

between the CMR results to the histopathological results. Because the %MF reported in histopathological imaging is reflective of a single slice area measurement, we computed slice-wise %MF from the slice containing the largest amount of MF in LGE-CMR in order to better compare the two modalities. However, we faced limitations in this comparison. Firstly, we were unable to compute slice-wise %MF measurements from the same slice(s) as measured in histopathology. This was because of inherent differences between image acquisition capabilities of each modality. Moreover, it was not feasible to match the slice orientation and location of ex-vivo histopathological images to in-vivo and in-situ LGE-CMR. Ideally, a registration of post-pacing and/or in-situ images to histopathological and/or ex-vivo images would allow direct spatial and geometric comparison of CMR and histological quantification. However, differences among these images are too numerous for this type of analysis to be attempted. For example, slice thickness of CMR images is 0.975 mm compared to 4 μm in histopathological images. That is, the slice thickness of CMR images is close to 1000 times larger than the slice thickness of histopathological images. Likewise, pixel-wise resolution of CMR images is on the order of millimetres as compared to micrometers in histopathological imaging, making direct quantitative comparison of fibrosis volumes problematic. Similarly, volumetric comparison of LA fibrosis measurements in histopathological imaging vs. CMR imaging was not possible in this study, as histopathological sampling of the entire LA volume was not feasible. Instead, a maximum of 7 slices were sampled for histopathological analysis, which may not be reflective of the measurement of fibrosis across the entire LA. Nevertheless, histopathological imaging confirmed the presence of fibrosis in the LA myocardium and therefore served as validation for MF measurements derived by CMR imaging using the

IIR thresholding method. Lastly, histopathological imaging of each dog was not performed due to resource constraints, which limited our analysis of fibrosis trends across each sample group.

5.5 Conclusion

Atrial fibrillation (AF) typically co-exists with myocardial fibrosis (MF); but underlying pathophysiological factors that contribute to AF are yet to be fully understood, making the definitive causal relationship controversial. In this study, we demonstrate that persistent AF leads to an increase in MF using quantification metrics based on a normalized image intensity ratio (IIR) in 3D IR-FLASH CMR imaging in a canine model and histopathological validation. Quantification of MF using LGE-CMR technique remains a challenging task limited by numerous factors, including inherent image acquisition parameters and observer variability. Improved MF computation may be facilitated by development of arrhythmia insensitive CMR imaging and/or techniques for automated segmentation of LA myocardial volumes. Additional studies are needed to further corroborate the causality of AF and MF.

5.6 Chapter 5 References

- [1] M. S. Dzeshka, G. Y. H. Lip, V. Snezhitskiy, and E. Shantsila, “Cardiac Fibrosis in Patients with Atrial Fibrillation: Mechanisms and Clinical Implications,” *J Am Coll Cardiol*, vol. 66, no. 8, 2015, doi: 10.1016/j.jacc.2015.06.1313.
- [2] A. Ruigomez, L. A. Garcia, S. Johanasson, M.-A. Wallander, and N. Edvardsson, “Risk of Cerebrovascular Accident After a First Diagnosis of Atrial Fibrillation,” *Clinical Cardiology*, vol. 30, no. 12, pp. 624–628, 2007, doi: 10.1002/clc.20178.
- [3] H. Kamel, P. M. Okin, S. V. M. Elkind, and C. Iadecola, “Atrial Fibrillation and Mechanisms of Stroke: Time for a New Model,” *Stroke*, vol. 47, no. 3, pp. 895–900, 2016, doi: 10.1161/STROKEAHA.115.012004.
- [4] H. Cochet, R. Dubois, S. Yamashita, A. Jefairi, and B. Berte, “Relationship Between Fibrosis Detected on Late Gadolinium-Enhanced Cardiac Magnetic Resonance and Re-Entrant Activity Assessed with Electrocardiographic Imaging in Human Persistent Atrial Fibrillation Relationship Between Fibrosis Detected on Late Gadolin,” *JACC: Clinical Electrophysiology*, vol. 4, no. 1, 2018, doi: 10.1016/j.jacep.2017.07.019.

- [5] S. Nattel, B. Burstein, and D. Dobrev, "Atrial Remodelling and Atrial Fibrillation: Mechanisms and Implications," *Circulation: Arrhythmia and Electrophysiology*, vol. 1, no. 1, pp. 62–73, 2008, doi: 10.1161/CIRCEP.107.754564.
- [6] B. Burstein and S. Nattel, "Atrial Fibrosis: Mechanisms and Clinical Relevance in Atrial Fibrillation," vol. 51, no. 8, 2008, doi: 10.1016/j.jacc.2007.09.064.
- [7] P. G. Platonov, "Atrial fibrosis: an obligatory component of arrhythmia mechanisms in atrial fibrillation?" pp. 233–237, 2017, doi: 10.11909/j.issn.1671-5411.2017.04.008.
- [8] I. V Kazbanov, K. H. W. J. Tusscher, and A. V Panfilov, "Effects of Heterogeneous Diffuse Fibrosis on Arrhythmia Dynamics and Mechanism," *Nature Publishing Group*, no. December 2015, pp. 1–14, 2016, doi: 10.1038/srep20835.
- [9] M. Pop *et al.*, "Quantification of fibrosis in infarcted swine hearts by *ex vivo* late gadolinium-enhancement and diffusion-weighted MRI methods," *Physics in Medicine and Biology*, vol. 58, no. 15, pp. 5009–5028, 2013, doi: 10.1088/0031-9155/58/15/5009.
- [10] N. A. Farrag, V. Ramanan, G. A. Wright, and E. Ukwatta, "Effect of T1-mapping technique and diminished image resolution on quantification of infarct mass and its ability in predicting appropriate ICD therapy," *Medical Physics*, vol. 45, no. 4, pp. 1577–1585, 2018, doi: 10.1002/mp.12840.
- [11] L. Bertelsen *et al.*, "Verification of threshold for image intensity ratio analyses of late gadolinium enhancement magnetic resonance imaging of left atrial fibrosis in 1.5T scans," *The International Journal of Cardiovascular Imaging*, vol. 36, no. 3, pp. 513–520, 2019, doi: 10.1007/s10554-019-01728-0.
- [12] I. M. Khurram *et al.*, "Magnetic resonance image intensity ratio, a normalized measure to enable interpatient comparability of left atrial fibrosis," *Heart Rhythm*, vol. 11, no. 1, pp. 85–92, 2014, doi: 10.1016/j.hrthm.2013.10.007.
- [13] E. M. Benito *et al.*, "Left atrial fibrosis quantification by late gadolinium-enhanced magnetic resonance: a new method to standardize the thresholds for reproducibility," pp. 1272–1279, 2017, doi: 10.1093/europace/euw219.
- [14] J. Siebermair, E. G. Kholmovski, and N. Marrouche, "Assessment of Left Atrial Fibrosis by Late Gadolinium Enhancement Magnetic Resonance Imaging: Methodology and Clinical Implications," *JACC: Clinical Electrophysiology*, vol. 3, no. 8, pp. 791–802, 2017, doi: 10.1016/j.jacep.2017.07.004.
- [15] G. Pontecorboli, R. M. Figueras, A. Carlosena, E. Benito, S. Prat-gonzales, and L. Padeletti, "Use of delayed-enhancement magnetic resonance imaging for fibrosis detection in the atria: a review," *European Society of Cardiology*, vol. 19, pp. 180–189, 2017, doi: 10.1093/europace/euw053.
- [16] N. P. du Sert *et al.*, "The arrive guidelines 2.0: Updated guidelines for reporting animal research," *PLoS Biology*, vol. 18, no. 7, pp. 1–12, 2020, doi: 10.1371/journal.pbio.3000410.
- [17] J. A. White *et al.*, "Fused Whole-Heart Coronary and Myocardial Scar Imaging Using 3-T CMR," vol. 3, no. 9, 2010, doi: 10.1016/j.jcmg.2010.05.014.
- [18] F. S. Prato *et al.*, "Can the Inflammatory Response Be Evaluated Using 18 F-FDG Within Zones of Microvascular Obstruction After Myocardial Infarction?" pp. 299–305, 2015, doi: 10.2967/jnumed.114.147835.
- [19] F. L. Goerner and G. D. Clarke, "Measuring signal-to-noise ratio in partially parallel imaging MRI," *Medical Physics*, vol. 38, no. 9, 2011, doi: 10.1118/1.3618730.

- [20] National Electrical Manufacturers Association, “Determination of Signal-to-Noise Ratio (SNR) in Diagnostic Magnetic Resonance Imaging,” *Standards Publication MS1-2008*, 2008.
- [21] D. Li, S. Fareh, T. K. Leung, and S. Nattel, “Promotion of atrial fibrillation by heart failure in dogs: Atrial remodeling of a different sort,” *Circulation*, vol. 100, no. 1, pp. 87–95, 1999, doi: 10.1161/01.CIR.100.1.87.
- [22] A. Fedorov *et al.*, “3D Slicer as an image computing platform for the Quantitative Imaging Network,” *Magnetic Resonance Imaging*, vol. 30, no. 9, pp. 1323–1341, 2012, doi: 10.1016/j.mri.2012.05.001.
- [23] A. Giannakidis *et al.*, “Rapid automatic segmentation of abnormal tissue in late gadolinium enhancement cardiovascular magnetic resonance images for improved management of long-standing persistent atrial fibrillation,” *BioMedical Engineering Online*, vol. 14, no. 1, pp. 1–20, 2015, doi: 10.1186/s12938-015-0083-8.
- [24] O. R. Coelho-Filho *et al.*, “The Role of Transcytolemmal Water Exchange in Magnetic Resonance Measurements of Diffuse Myocardial Fibrosis in Hypertensive Heart Disease,” *Circ Cardiovasc Imaging*, vol. 6, no. 1, p. 134, Jan. 2013, doi: 10.1161/CIRCIMAGING.112.979815.
- [25] G. Wu *et al.*, “N-terminal pro-brain natriuretic peptide and sudden cardiac death in hypertrophic cardiomyopathy,” *Heart*, vol. 107, no. 19, pp. 1576–1583, Oct. 2021, doi: 10.1136/HEARTJNL-2020-317701.
- [26] J. Ouyang, M. Guzman, F. Desoto-Lapaix, M. R. Pincus, and R. Wieczorek, “Utility of desmin and a Masson’s trichrome method to detect early acute myocardial infarction in autopsy tissues,” *International Journal of Clinical and Experimental Pathology*, vol. 3, no. 1, p. 98, 2010.
- [27] A. Weymann *et al.*, “Total Aortic Arch Replacement: Superior Ventriculo-Arterial Coupling with Decellularized Allografts Compared with Conventional Prostheses,” *PLOS ONE*, vol. 9, no. 7, p. e103588, Jul. 2014, doi: 10.1371/JOURNAL.PONE.0103588.
- [28] C. A. Bobak, P. J. Barr, and A. J. O’Malley, “Estimation of an inter-rater intra-class correlation coefficient that overcomes common assumption violations in the assessment of health measurement scales,” *BMC Medical Research Methodology*, vol. 18, no. 1, pp. 1–11, 2018, doi: 10.1186/s12874-018-0550-6.
- [29] C. Hu *et al.*, “T1-refBlochi: high resolution 3D post-contrast T1 myocardial mapping based on a single 3D late gadolinium enhancement volume, Bloch equations, and a reference T1,” pp. 1–17, 2017, doi: 10.1186/s12968-017-0375-1.
- [30] O. Razeghi *et al.*, “Fully Automatic Atrial Fibrosis Assessment Using a Multilabel Convolutional Neural Network,” *Circulation: Cardiovascular Imaging*, no. December, pp. 1–11, 2020, doi: 10.1161/CIRCIMAGING.120.011512.

Chapter 6: Effect of T1-mapping Technique and Diminished Image Resolution on Quantification of Infarct Mass and its Ability in Predicting Appropriate ICD Therapy

6.1 Introduction

Myocardial infarction (MI) is a condition whereby the tissue of the heart is damaged due to a prolonged lack of oxygen, typically resulting from a blocked coronary artery or coronary artery disease (CAD). Myocardial tissue in the region of the blocked coronary artery endures focal (i.e., replacement) fibrosis which may consist of two zones, infarct core (IC), which is the area of non-viable replacement fibrosis, and the gray zone or border zone (BZ), which is a heterogeneous, semi-viable region.

Patients with chronic MI in the left ventricular (LV) myocardium are at risk of developing arrhythmias, which have been found to increase the risk of sudden cardiac death (SCD) [1],[2]. This can be prevented via implantable cardioverter defibrillator (ICD) therapy [1], whereby arrhythmic events are detected and corrected by an ICD device. A decreased LV ejection fraction (LVEF) is widely used as a risk stratification technique for SCD in patients with chronic MI, and hence a primary deciding factor for ICD therapy [3]. However, several studies have shown that patients who experience SCD do not necessarily have significantly low LVEF [3],[4], indicating the need for development of new risk stratification strategies.

Indices based on IC and BZ masses have been recognized as strong indicators of adverse cardiac outcomes [5],[6]; and recently, studies have focused on developing alternative risk stratification techniques for SCD based on direct measurement of MI masses [5]– [8]. Particularly, BZ regions may cause propagation of electrical signals across

the heart to slow down; potentially creating re-entry pathways of abnormal electrical activity that can lead to ventricular tachycardia. Thus, BZ masses may be stronger predictors of adverse ICD events [7],[8].

LGE CMR techniques based on inversion-recovery fast gradient echo (IR-FGRE) are typically used for quantification of MI mass [9]. However, quantification of MI masses using IR-FGRE relies on accurate segmentation of the endocardial and epicardial boundaries, which is challenging due to the similar intensities of the blood pool and the sub-endocardial infarcts [8]. Additionally, segmentation of the myocardial boundaries is manually delineated by clinicians, which is time consuming, tedious, and subject to operator variability. These circumstances lead to inaccuracies in MI mass quantification.

Quantitative imaging techniques, such as T1-mapping, have been increasingly investigated for MI quantification and may be more reproducible than IR-FGRE in quantifying MI tissue regions [7],[10]. The MOLLI method is widely used for T1 mapping, which requires 11 – 17 heartbeats and several breath holds for full sequence acquisition [12]. Multi-contrast late enhancement (MCLE), an alternative T1-mapping technique, enables infarct and cardiac wall assessment in a single acquisition with reduced scan times [13],[14]. The 2D-MCLE technique typically employs a balanced steady-state free precession (b-SSFP) pulse sequence over about 13 heartbeats, allowing acquisition of higher quality images compared to both IR-FGRE and MOLLI [8],[15]. This method has been shown to be more reproducible than the conventional IR-FGRE technique in estimating infarct regions [7],[16].

Accurate reconstruction of IC and BZ in T1-mapping requires adequate image quality and spatial resolution. However, improving spatial resolution typically leads to

longer image acquisition times, and therefore longer breath hold requirements for the patient. In addition, intrinsic differences in T1-mapping pulse sequences will inherently yield images varying in contrast, spatial resolution and quality, leading to variation in MI mass quantification. Several studies have compared the efficacy of the MOLLI technique or the MCLE technique vs. IR-FGRE for infarct mass quantification [7],[17], but multiple T1-mapping techniques against conventional IR-FGRE has not been investigated. Moreover, previous studies have not evaluated the implication of reduced spatial resolution on MI mass quantification for predicting ICD-corrected arrhythmic events. Therefore, the objectives of this study were to compare MI masses determined by different T1-mapping CMR techniques to conventional IR-FGRE, and to investigate the effect of diminishing spatial resolution of T1-mapping images on the sensitivity of MI mass quantification for predicting appropriate ICD therapy.

6.2 Methods

6.2.1 Image Acquisition

6.2.1.1 Study Populations and CMR Protocol

The first objective of our study was to investigate the differences in MI masses determined by T1-mapping methods. We chose MCLE and MOLLI as T1-mapping techniques in our experiment, as the methods differentially trade off motion compensation and spatial resolution for acquisition durations fitting within a breath hold [8],[13]–[15]. The study consisted of 38 patients with chronic MI. Each patient was referred for CMR imaging at the Sunnybrook Research Institute (SRI) in Toronto, ON, and provided written and informed consent to undergo the imaging procedure. Patients underwent MCLE, MOLLI, and IR-FGRE acquisitions approximately 20 minutes after a double dose injection

of Gadolinium (Magnevist®, Bayer Inc. Canada; equivalent to 0.2 mmol/kg) on a 1.5 T GE Signa MRI system. The approximate largest cross-section of the infarct in short-axis view was evaluated by each method. Eighteen patients were excluded from further analysis for varying reasons, including no visible infarct (13), missing or corrupt data (2), and severe motion artifact (3).

The second objective of our study was to investigate the effect of diminished image resolution in the sensitivity of predicting ICD therapy using masses determined by T1-mapping. Toward this objective, we used a separate group of 27 patients with chronic MI, mean age 62.3 ± 11.2 years, 89% men, and average LVEF $25.7 \pm 8.65\%$, who had already been referred for ICD implantation. Our analysis of this study population is retrospective and overlaps largely with previously published results from our group [7]. All patients in this group were referred for ICD implantation based on LV functional parameters, where a LVEF $<35\%$ was largely used as a deciding factor for ICD implantation. Patient characteristics for this group are demonstrated in Table 1. Participants underwent CMR scanning on a 1.5 Tesla General Electric (GE) Signa system and were sent for ICD implantation following imaging. The CMR imaging protocol included LV functional assessment and implemented pre- and post-contrast imaging sequences. A steady-state free precession (SSFP) sequence was implemented prior to contrast injection for baseline reference. The IR-FGRE and MCLE images were acquired 10 – 20 minutes after double-dose of Gadolinium injection. The MOLLI pulse sequence was not implemented on this patient group. The IC and BZ masses of the LV were computed for the entire stack of slices.

All patients received an ICD device according to clinical guidelines and were followed up for 6 – 46 months after imaging in intervals of 3 months, or more frequently if shocks were delivered, with a median follow-up of 30 months. The ICD data was reviewed by two experienced electrophysiologists to determine all relevant ICD events. The measured outcome of ICD data was appropriate ICD therapy, defined as one or more device shocks for relevant ventricular arrhythmic events, including sustained and non-sustained ventricular tachycardia (VT) or ventricular fibrillation (VF) [7].

6.2.1.2 CMR Pulse Sequence Parameters

Pre-contrast SSFP images consisted of a stack of several short-axis oblique slices covering the entirety of the LV from base to apex. Twenty phase-resolved images were acquired in one breath-hold over the whole cardiac cycle. These images were used to evaluate several LV functional parameters, such as LV ejection fraction (LVEF), LV stroke volume (LVSV) etc., as shown in Table 6-1.

Table 6-1 Patient characteristics and LV functional parameters in 27 patients with ICD devices
Categorical data (patient characteristics) are expressed as percentages. Continuous variables are expressed as mean \pm SD.

Patient characteristics	Total (n = 27)	With ICD therapy (n = 12)	Without ICD therapy (n = 15)	P-value
Age (years)	62.30 \pm 11.2	63.7 \pm 9.3	61.2 \pm 12.8	0.58
Male	24 (89%)	10 (83.3%)	13 (86.7%)	0.8
Primary prevention	16 (59%)	6 (50%)	10 (66.7%)	0.49
NYHA functional class	1.44 \pm 0.97	1.8 \pm 1.0	1.2 \pm 0.9	0.15
Anti-arrhythmic	5 (18%)	2 (16.7%)	3 (20.0%)	0.83
Smoking	15 (56%)	6 (50%)	8 (61.5%)	0.62
Hypertension	20 (74%)	10 (83%)	9 (60.0%)	0.34
Diabetes	5 (18%)	2 (16.7%)	3 (20.0%)	0.83
Hyperlipidemia	22 (81%)	11 (91.7%)	11 (73.3%)	0.24
QRS duration (ms)	113.7 \pm 28.57	120.7 \pm 36.4	108.1 \pm 20.1	0.27

Left bundle-branch block	5 (18%)	3 (27.3%)	2 (13.3%)	0.51
CMR LV function				
LV EF (%)	25.7 ± 8.65	22.1 ± 8.5	28.6 ± 9.2	0.053
LV ESV (ml)	174.9 ± 79.8	203.5 ± 82.7	157.6 ± 76.2	0.11
LV EDV (ml)	231.8 ± 81.1	256.2 ± 82.2	212.2 ± 77.4	0.17
LV SV (ml)	56.5 ± 18.2	52.7 ± 18.9	59.6 ± 17.6	0.34
LVM (g)	106.6 ± 32.2	110.9 ± 34.8	103.2 ± 30.8	0.55

Table 6-2 demonstrates image acquisition parameters for MOLLI, MCLE, and IR-FGRE. The IR-FGRE sequence employed an inversion time (TI) varying between 200 to 300 ms and a delay time of 500 ms. Images were acquired in the short-axis oblique view. A single IR-FGRE acquisition required approximately 20 heartbeats and an average breath-hold time of 18 s. IR-FGRE and SSFP sequences were acquired with in-plane resolution of 1.48 x 1.48 mm².

Table 6-2 CMR imaging parameters of IR-FGRE, MOLLI and MCLE imaging techniques

	MOLLI	IR-FGRE	MCLE
Receive bandwidth (rBW)	±83 kHz	±31.5 kHz	±125 kHz
flip angle	35°	20°	30°
views per segment (VPS)	8	20	16
TR/TE (ms)	4.2/1.2	6.0/3.0	3.3/1.4
field of view (FOV) (cm)	32	32	32
image matrix	160×160	192×192	192×192
in-plane resolution (mm ²)	2.47×2.47	1.5×1.5	1.37×1.37
slice thickness (mm)	8	8	8
number of slices per acquisition	1	8 – 10	8 – 10
number of excitations (NEX)	1	2	1

The MOLLI technique allowed for selective image acquisition during end diastole. This technique acquired multiple images with each IR experiment at increasing TI values. The 3,5 MOLLI sequence was employed, using two IR pulses with 3 and 5 acquisitions,

respectively. A delay time of three heartbeats was required before the between IR pulses. The total sequence acquisition required 11 heartbeats. Resultant signal intensities were constructed to a parametric curve to determine T1 values. Acquired in-plane resolution of MOLLI images was $2.47 \times 2.47 \text{ mm}^2$.

The MCLE technique implemented in both groups uses a b-SSFP pulse sequence following inversion recovery acquired after IR-FGRE, requiring about 13 heartbeats for each acquisition. The first heartbeat established a steady-state and the 12 following heartbeats enabled acquisition of twenty separate images during different phases of the cardiac cycle at different effective TIs. Data acquisition was continuously implemented in a segmented fashion between each TI. The MCLE sequence implemented a delay time of 500 ms and total sequence acquisition required a breath-hold of about 11 s. Analysis of signal recovery was then performed using nonlinear regression to fit the signal intensities to a parametric curve [13]. The acquired in-plane image resolution of MCLE was $1.37 \times 1.37 \text{ mm}^2$.

6.2.2 Image Processing

We post-processed MCLE, IR-FGRE, and MOLLI images to compute IC and BZ masses and determined the differences in the MI mass reconstructed by each technique. Due to the inherent technical differences of the methods, different previously validated methods were used to segment infarct regions in each case. While MCLE and IR-FGRE images consisted of several short-axis slices of the heart from apex to base, MOLLI was limited to a single slice imaged across the mid cross-section of the infarct. We therefore segmented the slices of MCLE and IR-FGRE that corresponded to the same mid cross-section of the MOLLI sequence for analysis in the first patient group.

6.2.2.1 IR-FGRE Images

We first manually segmented LV epicardial and endocardial contours to define the epicardial and endocardial borders of the myocardium prior to MI segmentation, in order to avoid the segmentation algorithm leaking into the background, due to similar appearances in intensity. Epicardial and endocardial contours were also used to quantify the left ventricular mass (LVM). A region of interest (ROI) was chosen from remote myocardium to calculate reference values for signal intensity (SI) mean, peak, and standard deviation of healthy myocardium. Signal intensities of IC and BZ regions (SI_{IC} and SI_{BZ} , respectively) were determined using a full-width half-maximum (FWHM) approach determined by peak intensities of remote myocardium and infarct as follows:

$$SI_{IC} > 0.5 * Peak_{infarct} \quad \text{Equation 6-1}$$

$$Peak_{remote} < SI_{BZ} < 0.5 * Peak_{infarct} \quad \text{Equation 6-2}$$

where $Peak_{infarct}$ refers to the peak intensity infarct and $Peak_{remote}$ refers to peak intensity of remote myocardium [7].

6.2.2.2 MOLLI and MCLE Images

For MCLE and MOLLI images, segmentation of myocardial boundaries is usually not required for accurate segmentation of MI regions. However, we also manually segmented the LV boundaries in MCLE and MOLLI images prior to segmentation of BZ and IC in order to maintain consistency with the approach used for IR-FGRE images. The BZ and IC regions were segmented using a fuzzy c-means clustering approach, which used both $T1^*$ and steady-state maps [13]. We first used an exponential curve-fitting approach to generate the $T1^*$ and steady-state maps from each image, which are demonstrated in Figure 6-1. A scatter plot which showed the $T1^*$ vs. steady-state value for each voxel

allowed the fuzzy c-means algorithm to characterize IC, BZ, healthy myocardium and blood. The fuzzy c-means algorithm determines the probability of each voxel belonging to one of these three clusters based on a distance metric derived from the scatterplot. A voxel probability of 75% or higher was required to classify voxels as belonging to a certain cluster. Voxels with a probability of less than 75% of belonging to the IC cluster but greater than 25% probability belonging to the remote myocardium were classified as border-zone pixels. The total size of each mass was determined by the summation of all voxels given to each class.

For an acquisition of fixed number of heartbeats, the resolution of MOLLI images was coarser than that of MCLE images. Therefore, to evaluate whether differences in image resolution affects estimation of BZ mass and consequently impacts sensitivity of predicting ICD therapy, we also downsampled the MCLE images to the lower resolution of MOLLI images and recomputed the infarct masses for secondary analysis. T1* and steady-state maps were separately generated for both the original and downsampled MCLE images (see Fig. 6-1), which were then segmented to quantify BZ and IC masses using the fuzzy c-means clustering algorithm mentioned above. The T1* vs. steady-state scatterplots showing fuzzy c-means clustering in original and downsampled MCLE images are demonstrated in Figure 6-3.

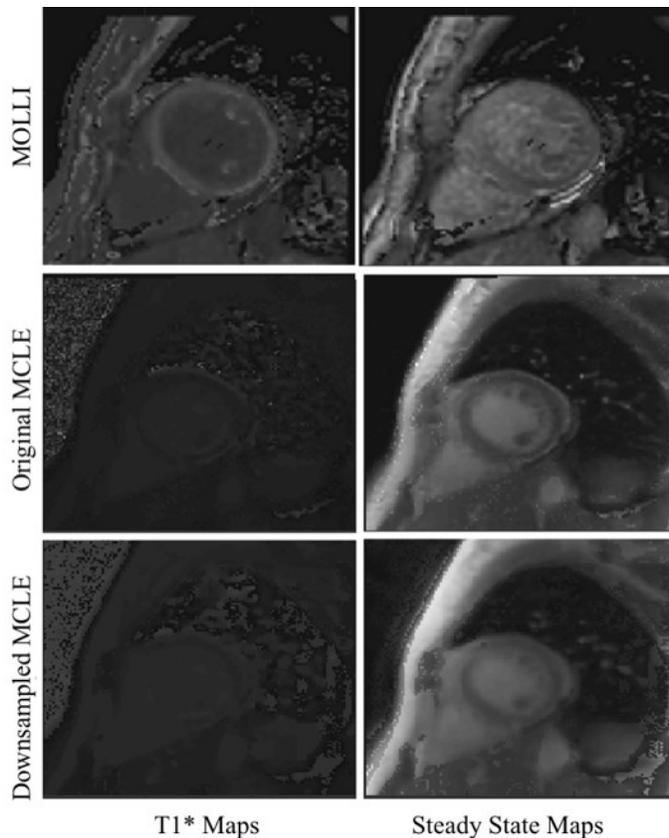


Figure 6-1 T1* and steady-state maps for MOLLI, MCLE and downsampled MCLE

6.2.3 Statistical Analyses

Statistical analyses of infarct comparisons were performed using International Business Machines Statistical Package for Social Sciences (IBM SPSS) Statistical version 19 (IBM Corporation, 2010). Categorical data are expressed as percentages and compared using the Fisher exact test. Continuous data are expressed as mean \pm SD and compared using student's t-test, in which results were considered significant when the probability of making a type I error was less than 5% ($p < 0.05$). For analysis of multiple comparisons, we used the paired t-test and applied a Bonferroni correction to determine significance. For these comparisons, results were considered significant when the probability of making a Type I error was less than 0.83% ($p < 0.0083$).

6.3 Results

Figure 6-2 shows segmentations of the IC and BZ masses for two patient cases overlaid on images acquired by the different CMR techniques. Table 6-3 shows a quantitative summary of the IC and BZ masses determined by different CMR techniques, which were reported as a percentage with respect to the LVM. Manual LVM measurements varied between MOLLI, MCLE and IR-FGRE; therefore, the BZ and IC masses are expressed as continuous values normalized by the LVM found by the respective imaging technique. Between MOLLI, MCLE and IR-FGRE, the estimated differences in BZ and IC masses with respect to the total LVM were smaller than 1%.

Table 6-3 Normalized IC and BZ masses quantified by different CMR techniques

The values are reported as a percentage of LV mass are for a single slice

Summary of infarct mass quantification for 20 patients undergoing MCLE, MOLLI and IR-FGRE. The values reported as a percentage of LV mass are for a single slice.				
Region	MCLE	Downsampled MCLE	MOLLI	IR-FGRE
IC MI/LVM (%)	7.34±3.46	8.05±3.68	7.49±3.54	7.91±3.61
BZ MI/LVM (%)	3.47±1.57	4.20±1.67	4.15±1.65	3.34±1.24

The statistical significance for the pair-wise comparisons of estimated infarct masses by different methods are shown in Table 6-4. A Bonferroni correction was applied to determine the significance for these comparisons, where results were considered significant when the probability of making a Type I error was less than 0.83% ($p < 0.0083$). Bonferroni corrected p values are reported in Table 6-4. The IC masses determined by any pair of methods were not significantly different, and BZ masses determined by MCLE and IR-FGRE were not significantly different. However, BZ masses determined by MOLLI were significantly higher compared to those determined by MCLE and IR-FGRE (p value=0.0022 and 0.0003, respectively). The BZ masses determined by the downsampled

MCLE were greater in comparison to those determined by the original MCLE (p value=0.0003). However, BZ masses determined by the downsampled MCLE were not significantly different than those determined by MOLLI.

Table 6-4 Comparison of IC and BZ regions quantified by different CMR techniques
 Statistical significance for comparing MCLE, MOLLI, IR-FGRE and downsampled MCLE. The Bonferroni corrected p-values that are significant are shown in bold.

Comparison	P-value	
	BZ	IC
MCLE – MOLLI	0.0022	0.5923
MCLE – IR-FGRE	0.5458	0.1563
MOLLI – IR-FGRE	0.0003	0.2951
MOLLI – downsampled MCLE	0.8394	0.197
MCLE – downsampled MCLE	0.0003	0.0537
IR-FGRE – downsampled MCLE	0.0033	0.7439

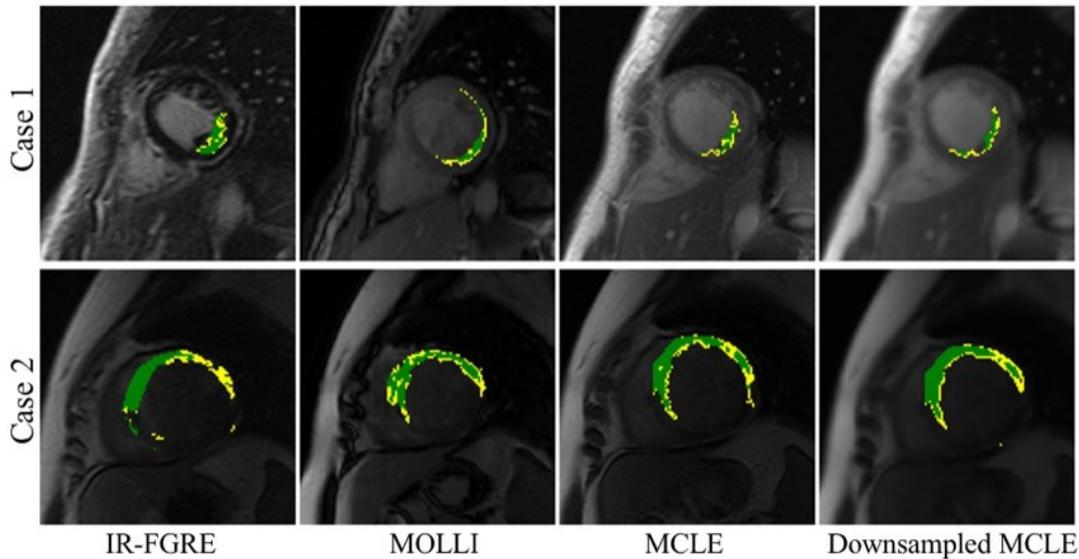


Figure 6-2 Segmentation of IC and BZ masses in different CMR techniques for two example patients. The IC is represented as a green region while the BZ is represented as a yellow region

Among the patient group with ICD implants and clinical outcomes, 44.4% (12/27) received ICD therapy (i.e., one or more delivered ICD shock) during the follow-up stage. Patient characteristics and LV functional parameters did not differ significantly between patients with and without appropriate ICD therapy.

Figure 6-3 shows the T1* and steady-state scatterplots used in the fuzzy c-means clustering algorithm for a series of MCLE slices in the first row, followed by the corresponding scatterplots for downsampled MCLE slices in the second row. The segmentation results of the IC and BZ masses determined by original and downsampled MCLE are shown in Figure 6-4, overlaid on the respective images. The BZ regions appear to be overestimated in the downsampled images as compared to the original MCLE images. The summary of quantitative evaluation of the IC and BZ masses are shown in Table 6-5.

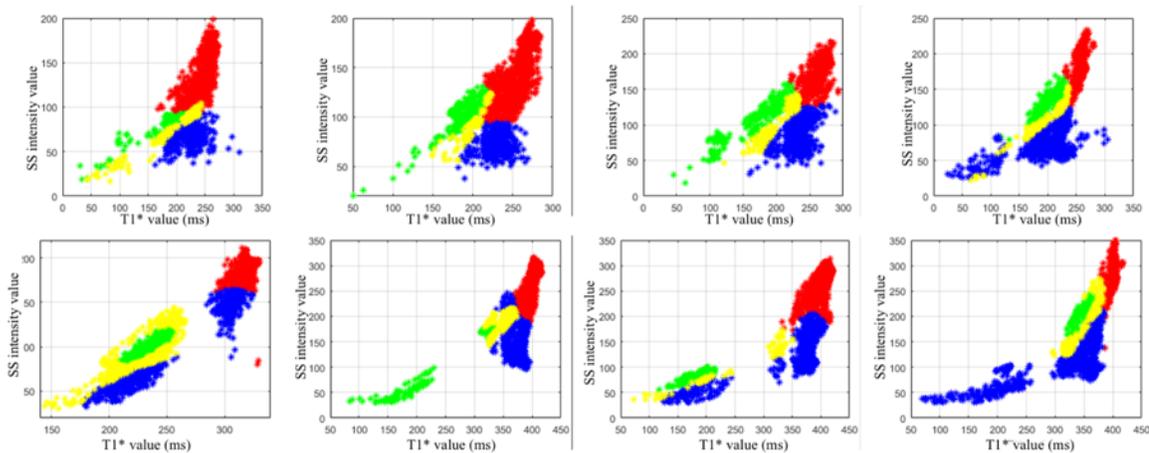


Figure 6-3 T1* and steady-state scatterplots for a series of original and downsampled MCLE slices. T1* (x-axis) and steady-state (y-axis) scatterplots for original MCLE images from a single patient in the first row. The second row demonstrates scatterplots for the same series of MCLE slices after downsampling. The fuzzy c-means algorithm classified pixels as blood pool (red), healthy myocardium (blue), IC (green) or border zone (yellow).

For the entire group of patients, the BZ and the total infarct mass determined by downsampled MCLE was significantly higher than that of original MCLE. However, IC mass determined by original and downsampled MCLE were not significantly different. Between the patients with and without ICD therapy, only the BZ mass determined by original MCLE was statistically significant (p value = 0.044). However, when the MCLE images were downsampled and reanalyzed, the difference in BZ masses between patients with and without ICD therapy were no longer significant.

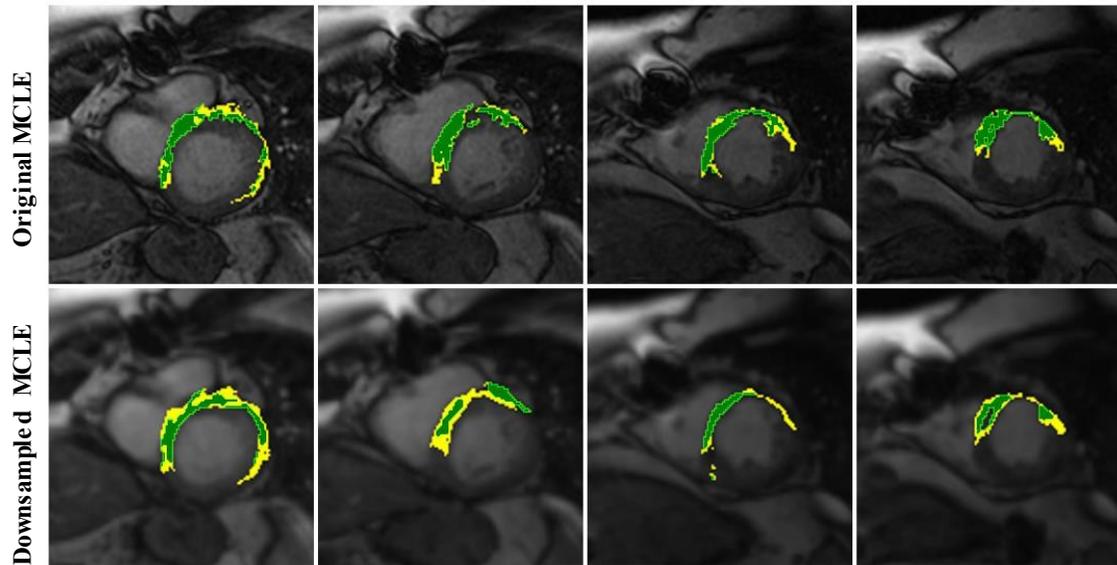


Figure 6-4 Example IC and BZ regions in MCLE and downsampled MCLE images. A series of MCLE slices for a single patient are demonstrated in the first row. The second row demonstrates the same series of MCLE slices after downsampling. Segmented infarct regions, determined by fuzzy c-means clustering, are demonstrated within each slice of original and downsampled MCLE images. The IC is represented as a green region while the BZ is represented as a yellow region.

Table 6-5 Comparison of infarct masses found by original vs. downsampled MCLE in patients with and without ICD therapy

Patients undergoing MCLE imaging prior to ICD implantation									
	Entire sample (n = 27)			With ICD therapy (n = 12)		Without ICD therapy (n = 15)		P value	
	Original MCLE	Down- sampled MCLE	p-value	Original MCLE	Down- sampled MCLE	Original MCLE	Down- sampled MCLE	Original MCLE	Down- sampled MCLE
Core MI/LVM (%)	26.1 ± 11.1	32.4 ± 12.3	0.08	29.8 ± 8.6	36.4 ± 8.0	23.1 ± 12.9	29.3 ± 14.8	0.094	0.161
BZ MI/LVM (%)	14.4 ± 6.45	19.9 ± 11.0	0.032	17.4 ± 6.4	22.8 ± 10.1	12.0 ± 5.6	17.6 ± 12.1	0.044	0.32
Total MI/LVM (%)	40.5 ± 16.5	52.3 ± 21.3	0.041	47.2 ± 13.1	59.2 ± 15.0	35.1 ± 18.3	46.9 ± 25.5	0.081	0.27

6.4 Discussion

ICD therapy is a reliable treatment method in patients at risk of ventricular arrhythmias following myocardial infarction [18]. The most common deciding factor influencing ICD implantation is a low LVEF, yet several studies have shown that patients

with low LVEF may not benefit from ICD therapy [2]– [4]. Recent studies have found infarct mass as a sensitive predictor of adverse arrhythmic events, and therefore risk of SCD, in patients with MI [9],[19]. As multiple CMR techniques are available for quantifying infarct mass, it is important to estimate the differences in infarct mass quantification between techniques. The objective of this study was to compare the sensitivity of infarct mass determined by two different T1-mapping techniques to those of LGE-CMR using IR-FGRE and evaluate the effect of spatial resolution on sensitivity of infarct mass for predicting ICD therapy. For this comparison, we chose MOLLI and MCLE as T1-mapping techniques, as the MRI pulse sequences in both techniques are fundamentally different.

The results of this study demonstrate that T1-mapping techniques produce comparable IC masses to those determined by IR-FGRE, however, the BZ masses as determined by MOLLI were significantly higher compared to those determined by MCLE and IR-FGRE. Furthermore, the BZ masses determined by the downsampled MCLE were significantly larger than the masses determined by original MCLE and IR-FGRE, while they were not statistically different from MOLLI.

Several ex-vivo animal studies have shown that diminished image resolution leads to overestimation of BZ mass, but estimation of the IC mass is not substantially affected [8],[20]. Pop et al.'s 2013 ex-vivo study on swine hearts [21] found that diminished image resolution led to misclassification of BZ pixels and accentuated the difference between BZ vs. IC pixels in diffusion weighted imaging (DWI)-CMR. Schuleri et al. 2012 produced similar findings in mini-pigs [22], demonstrating inaccurate discrimination of BZ vs. dense scar as a result of partial volume effect with slice thicknesses of 8–10 mm in LGE-CMR

images and DWI-CMR. In our study, we demonstrated similar effects in-vivo using T1-mapping images for patients. The MOLLI technique, as compared to other CMR techniques, yielded greater estimations of the BZ mass, likely due to the coarser image resolution that is inherent to MOLLI images. This suggests that diminished image resolution leads to an increase in estimation of the BZ mass, which is likely due to partial volume averaging effect. The BZ masses were not significantly different when they were determined by MOLLI and MCLE images at the same image resolution, indicating that the differences in BZ masses observed at the original acquired resolution may have been mostly caused by the differences in spatial resolution between the imaging techniques.

In our analysis of the follow-up patient group with ICD implants, we evaluated several patient characteristics as well as LV functional characteristics between the patients with and without ICD therapy. None of the patient or CMR LV characteristics measured in this patient group were significantly different between the patients with and without ICD therapy. However, the average BZ measurements in patients with ICD therapy were significantly higher than those found in patients without ICD therapy; suggesting that indices based on BZ measurement may lead to more effective risk stratification strategies in patients with chronic MI. When the BZ was determined using the downsampled MCLE images, the difference in BZ mass between the patients with and without ICD therapy was no longer significant. Thus, spatial resolution of the CMR technique is an important determining factor in BZ mass quantification, and low spatial resolution leads to low sensitivity in predicting ICD therapy using BZ mass.

This study has several limitations. Firstly, the sample sizes in both groups were small. This affects our statistical testing, especially in analysis of the group of 38 patients who

had undergone MOLLI, MCLE, and IR-FGRE imaging. The analysis of the MOLLI pulse sequence was limited to a single slice per acquisition, and thus a reduced volume of tissue was assessed in this group compared to the other pulse sequences. Therefore, in order to maintain consistency among the three techniques, we only analyzed the single corresponding slice of the MCLE and IR-FGRE pulse sequences in the comparisons with MOLLI. This led to an underrepresentation of tissue volume to characterize scar in the group analysis. A multi-slice analysis would be most ideal to address this issue; however, this was not feasible given the single-slice limitation of the MOLLI pulse sequence.

Another limitation in our study is the exclusion of the MOLLI imaging technique in the group of 27 patients with ICD follow-up. The study on ICD follow-up was performed retrospectively, and unfortunately, the MOLLI pulse sequence was not available at SRI at the time. Due to the delay in access to the MOLLI pulse sequence, we were unable to include it in the CMR procedure for our patient group with ICD follow-up. Nonetheless, we were able to implement this imaging technique in the other patient group involved in the study. However, since the MOLLI technique was limited to single slice, a much larger sample size would likely be needed to determine the significance of BZ measurements between patients with and without appropriate ICD therapy. Infarct area measurements, as compared to infarct volume measurements, may be less sensitive in quantifying infarct masses and therefore may not be as sensitive in predicting appropriate ICD therapy.

Our study involved comparison of the MCLE T1-mapping technique to two other conventional techniques. Due to inherent differences in acquisition parameters between these sequences, the acquired image resolutions differed among the methods. To explore the impact of resolution, we artificially downsampled the MCLE images to the resolution

of MOLLI images. We recognize, however, that artificial downsampling may not ideally mimic changes in the signal-to-noise ratio and/or motion blur affecting the distribution of the underlying biological material when the images are acquired directly at a lower resolution. However, due to resource and time constraints, we were unable to acquire a second, lower resolution MCLC acquisition on the same patient set in the current study.

6.5 Conclusions

T1-mapping techniques produce adequate CMR images and comparable IC masses to those of IR-FGRE. The BZ mass index is an appropriate measure in risk stratification for adverse ICD events and risk of SCD. Spatial resolution is an important determining factor in quantification of BZ infarct regions, and diminished image resolution causes an increase in the estimation of the BZ due to partial volume effects. Poor image resolution is associated with a reduced sensitivity in the prediction of appropriate ICD therapy from BZ mass.

6.6 Chapter 6 References

- [1] R. J. Myerburg, “Implantable Cardioverter–Defibrillators after Myocardial Infarction,” pp. 2245–2253, 2008.
- [2] B. A. Koplman, W. G. Stevenson, and O. N. Cardiovascular, “Ventricular Tachycardia and Sudden Cardiac Death,” *Mayo Clin. Proc.*, vol. 84, no. 3, pp. 289–297, 2009.
- [3] A. Gulati *et al.*, “Association of fibrosis with mortality and sudden cardiac death in patients with nonischemic dilated cardiomyopathy.,” *Jama*, vol. 309, no. 9, pp. 896–908, 2013.
- [4] G. I. Fishman *et al.*, “Sudden Cardiac Death Prediction and Prevention: Report From a National Heart, Lung, and Blood Institute and Heart Rhythm Society Workshop,” *Circulation*, vol. 122, no. 22, pp. 2335–2348, Nov. 2010.
- [5] S. Heidary *et al.*, “Quantitative Tissue Characterization of Infarct Core and Border Zone in Patients With Ischemic Cardiomyopathy by Magnetic Resonance Is Associated With Future Cardiovascular Events,” *J. Am. Coll. Cardiol.*, vol. 55, no. 24, pp. 2762–2768, 2010.
- [6] A. Schmidt *et al.*, “Infarct Tissue Heterogeneity by Magnetic Resonance Imaging Identifies Enhanced Cardiac Arrhythmia Susceptibility in Patients With Left Ventricular Dysfunction,” *Circulation*, vol. 115, no. 15, pp. 2006–2014, Apr. 2007.
- [7] Y. Yang *et al.*, “Multi-contrast late enhancement CMR determined gray zone and

- papillary muscle involvement predict appropriate ICD therapy in patients with ischemic heart disease.,” *J. Cardiovasc. Magn. Reson.*, vol. 15, no. 1, p. 57, 2013.
- [8] M. Pop *et al.*, “High-resolution 3-D T1*-mapping and quantitative image analysis of gray zone in chronic fibrosis,” *IEEE Trans. Biomed. Eng.*, vol. 61, no. 12, pp. 2930–2938, 2014.
- [9] K. C. Wu, “Sudden Cardiac Death Substrate Imaged by Magnetic Resonance Imaging: From Investigational Tool to Clinical Applications.,” *Circ. Cardiovasc. Imaging*, vol. 10, no. 7, p. e005461, Jul. 2017.
- [10] J. R. Burt, S. L. Zimmerman, I. R. Kamel, M. Halushka, and D. A. Bluemke, “Myocardial T1 Mapping: Techniques and Potential Applications,” *RadioGraphics*, vol. 34, no. 2, pp. 377–395, Mar. 2014.
- [11] E. B. Schelbert and T. C. Wong, “Clinical Applications of T1 Mapping,” *American College of Cardiology*, 2015. [Online]. Available: https://www.acc.org/latest-in-cardiology/articles/2015/02/13/09/24/clinical-applications-of-t1-mapping?w_nav=TI.
- [12] A. J. Taylor, M. Salerno, R. Dharmakumar, and M. Jerosch-Herold, “T1 Mapping Basic Techniques and Clinical Applications,” *JACC Cardiovasc. Imaging*, vol. 9, no. 1, pp. 67–81, 2016.
- [13] J. S. Detsky, G. Paul, A. J. Dick, and G. A. Wright, “Reproducible classification of infarct heterogeneity using fuzzy clustering on multicontrast delayed enhancement magnetic resonance images,” *IEEE Trans. Med. Imaging*, vol. 28, no. 10, pp. 1606–1614, 2009.
- [14] K. A. Connelly *et al.*, “Multicontrast late gadolinium enhancement imaging enables viability and wall motion assessment in a single acquisition with reduced scan times,” *J. Magn. Reson. Imaging*, vol. 30, no. 4, pp. 771–777, Oct. 2009.
- [15] M. Pop, S. Oduneye, L. Zhang, S. Newbigging, and G. Wright, “In Vivo Parametric T1 Maps Correlate with Structural and Molecular Characteristics of Focal Fibrosis BT - Functional Imaging and Modelling of the Heart: 9th International Conference, FIMH 2017, Toronto, ON, Canada, June 11-13, 2017, Proceedings,” M. Pop and G. A. Wright, Eds. Cham: Springer International Publishing, 2017, pp. 13–22.
- [16] M. Fontana *et al.*, “Comparison of T1 mapping techniques for ECV quantification. Histological validation and reproducibility of ShMOLLI versus multibreath-hold T1 quantification equilibrium contrast CMR.,” *J. Cardiovasc. Magn. Reson.*, vol. 14, no. 1, p. 88, 2012.
- [17] D. R. Messroghli, A. Radjenovic, S. Kozerke, D. M. Higgins, M. U. Sivananthan, and J. P. Ridgway, “Modified Look-Locker inversion recovery (MOLLI) for high-resolution T1 mapping of the heart,” *Magn. Reson. Med.*, vol. 52, no. 1, pp. 141–146, Jul. 2004.
- [18] A. J. Moss *et al.*, “Prophylactic Implantation of a Defibrillator in Patients with Myocardial Infarction and Reduced Ejection Fraction,” *N. Engl. J. Med.*, vol. 346, no. 12, pp. 877–883, Mar. 2002.
- [19] D. Bello *et al.*, “Infarct morphology identifies patients with substrate for sustained ventricular tachycardia,” *J. Am. Coll. Cardiol.*, vol. 45, no. 7, pp. 1104–1108, Apr. 2005.
- [20] E. B. Schelbert *et al.*, “Late gadolinium-enhancement cardiac magnetic resonance identifies postinfarction myocardial fibrosis and the border zone at the near cellular

- level in ex vivo rat heart,” *Circ. Cardiovasc. Imaging*, vol. 3, no. 6, pp. 743–752, 2010.
- [21] M. Pop *et al.*, “Quantification of fibrosis in infarcted swine hearts by *ex vivo* late gadolinium-enhancement and diffusion-weighted MRI methods,” *Phys. Med. Biol.*, vol. 58, no. 15, pp. 5009–5028, 2013.
- [22] K. H. Schuleri *et al.*, “Cardiovascular magnetic resonance characterization of peri-infarct zone remodeling following myocardial infarction,” *J. Cardiovasc. Magn. Reson.*, vol. 14, no. 1, p. 24, 2012.

Chapter 7: Radiomic analysis-based classification of cardiomyopathy states manifesting with interstitial myocardial fibrosis using native T1 and ECV mapping CMR

7.1 Introduction

Cardiac magnetic resonance (CMR) imaging is considered the gold-standard diagnostic imaging approach for assessing myocardial disease. Recently, quantitative imaging approaches, such as native T1-mapping and ECV mapping, have been proposed for evaluation of various cardiovascular diseases that present with either interstitial or replacement fibrosis (MF), as these methods allow for improved characterization of the myocardial tissue space [1]– [4].

Cardiomyopathy phenotypes may be distinguished based on differences in myocardial T1 and/or ECV values, and recent studies have applied machine learning approaches toward this aim. Such approaches can broadly be categorized into two kinds: (1) radiomic analysis (RA), wherein hand-selected image features are extracted using various image filters and transforms, and (2) convolutional neural networks (CNN), wherein a computer automatically recognizes image features on its own. In both cases, data-characterization algorithms are used to extract quantitative features from segmented regions, including texture, shape, and statistical features, which are specific to the segmented region of interest (ROI) [5]– [9]. These image features which may be reflective of the underlying pathophysiology of the tissue but may be imperceptible to the human eye [7]– [10]. Following feature extraction, various machine learning methods can be implemented for characterization and classification of cardiomyopathy disease states. Both RA and CNN

methods have been successfully applied to classify cardiomyopathy phenotypes, however, CNNs typically require a larger amount of training data in order to perform suitably [11].

Recent studies have used RA to differentiate hypertrophic cardiomyopathy (HCM) phenotypes and ischemic disease states in CINE and LGE images with semi-manually derived myocardial contours [12]– [15]. HCM phenotypes have also been differentiated using RA of native T1-mapping CMR and manually derived myocardial contours [6],[12],[16]– [18]. Similarly, RA has also been applied to extracellular volume (ECV) mapping CMR for differentiation between reversible versus irreversible myocardial damage following myocardial infarction [19], prediction of heart failure [20], as well as HCM versus athlete’s heart [21] using manually traced myocardial ROIs. Thus, recent studies support the clinical value for RA to differentiate between various cardiovascular diseases; however, current methods rely on manual or semi-manual delineation of the myocardial ROI, which introduces the risk of inter- and intra-operator variability affecting the derived features. Inadvertent inclusion of pericardial fat and/or blood pool regions to the myocardial ROI can significantly confound T1 and ECV measurements, potentially leading to erroneous radiomic features and/or deficits in reproducibility of features [22].

Deep learning-based methods, such as convolutional neural networks (CNN), have also been implemented for CMR image analysis tasks, including image segmentation and disease classification. However, the latter objective has been achieved by only few studies, such as classification of normal, HCM and dilated cardiomyopathy (DCM) [23] and HCM genotypes [24] using CINE imaging.

Therefore, we propose fully automated cardiomyopathy disease classification using T1 and ECV mapping CMR and compare two approaches: (1) RA-based disease classification using ensembled machine learning methods and (2) deep learning-based disease classification using CNNs, both of which implemented using myocardial contours derived automatically by a convolutional neural network (CNN) based on the U-Net architecture [25] and transfer-learning.

7.2 Methods

7.2.1 Patient population and CMR imaging protocol

Two hundred and seventy-two (272) patients were prospectively recruited to the Stephenson Cardiac Imaging Centre, Libin Cardiovascular Institute, in Calgary, AB, Canada. The study was approved by the Conjoint Health Research Ethics Board at University of Calgary and all subjects provided written informed consent. Patients with both a clinical diagnosis and confirmed CMR imaging-based diagnosis of a cardiomyopathy state commonly associated with interstitial MF were selected for this study. The patient population included the following disease states: cardiac amyloidosis (n = 50), Fabry's disease (n = 23), HCM (n = 75), hypertensive heart disease (HTN) (n = 39), and healthy volunteers (n = 85). All patients underwent a standardized imaging protocol using a Siemens 3T magnet (Prisma or Skyra, Siemens Healthineers, Erlangen), inclusive of short and long-axis CINE and LGE imaging using SSFP and inversion-recovery gradient echo (IR-GRE) pulse sequences, respectively. LGE imaging was performed 10 minutes following the intravenous administration of 0.1 mmol/kg Gadovist (Bayer Inc., Canada). Native (i.e., pre-contrast), post-contrast and ECV mapping was performed in short-axis view at the apex, base and/or mid-point of the LV region using a PSIR MOLLI pulse

sequence. Thus, two or three mapping images at the aforementioned LV locations were available per patient. All patients underwent native T1-mapping imaging; however, post-contrast T1-mapping images were not acquired in all patients in this study. Therefore, ECV mapping images were not available for all study participants. Thus, in total, 813 native T1-mapping and 497 ECV mapping images were available for analysis. The number of images acquired per cardiomyopathy disease state is demonstrated in Table 7-1.

Table 7-1 Image availability per patient population

Cardiomyopathy state	CMR imaging sequence	
	Native T1-mapping	ECV mapping
Cardiac amyloidosis	150	90
Fabry's disease	68	56
HCM	225	204
HTN	116	79
Healthy volunteers	254	68
Total	813	497

7.2.2 Image processing

The image processing and classification pipeline used in this study is demonstrated in Figure 1, consisting of three main steps: (1) preprocessing, (2) myocardial segmentation, and (3) classification. Pre-processing steps (1) included image resizing and scaling the bit depth of pixels. The myocardial segmentation step (2) was achieved with a transfer learning-based approach using a U-Net [25] CNN architecture that was pre-trained on ECV [26] or native T1-mapping images (respectively) of patients with ischemic and dilated cardiomyopathies (ICM and DCM, respectively). Binary myocardial segmentation masks were required as input along with mapping images for both the classification step (3), which was implemented using two approaches: RA-based and deep-learning based classification algorithms.

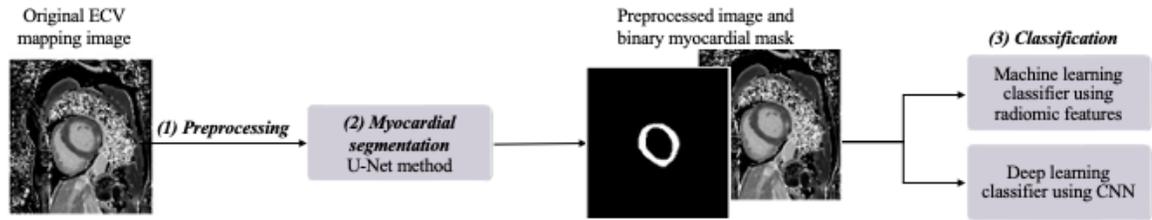


Figure 7-1 Image processing pipeline used for cardiomyopathy disease classification. All images were preprocessed before being passed through a U-Net-based algorithm for myocardial segmentation. The output binary myocardial mask was used with the pre-processed image as input for two classification algorithms which were implemented to compare performance accuracy, i.e., (1) random forest algorithm using radiomic features and (2) deep learning algorithm using CNNs.

7.2.2.1 Preprocessing

All images were originally saved in Digital Imaging and Communications in Medicine (DICOM) format with a pixel depth of 12 bits per pixel. Images were preprocessed by converting to 16-bit double precision Nifti format and resizing to standardized dimensions of 192 by 176 pixels in length and width, respectively, as in our prior study (Chapter 4) [26].

7.2.2.2 Myocardial segmentation

Preprocessed images were passed through a pre-trained CNN based on the U-Net architecture [25],[26] to generate myocardial segmentation masks. By using this transfer-learning based approach, the U-Net model did not require training on the current dataset, thereby allowing all 813 native and 497 ECV mapping images to be segmented for the classification task. A standard encoder and decoder pipeline was used with a fixed kernel size of 3x3 for all but the last convolutional layer, which was adjusted to 1x1 in order to generate the output segmentation mask. Filter sizes increased by a factor of two from 64 to 1024 across convolution layers in the encoder path and decreased at the same interval back to 64 in the decoder path using max pooling. Each Convolution layer was followed by a rectified linear unit (ReLU) activation function. Batch normalization was used in the

encoder path for regularization. The myocardial segmentation mask was obtained using a sigmoid activation function which output a probability map from 0 – 1, wherein probabilities >0.5 were mapped to 1 and <0.5 were mapped to 0 to generate a binary image [26]. Figure 7-2 demonstrates the U-Net model used for myocardial segmentation.

Following myocardial segmentation, images and binary masks were used as input for disease classification using two classification algorithms for comparison, based on (1) RA and (2) deep learning.

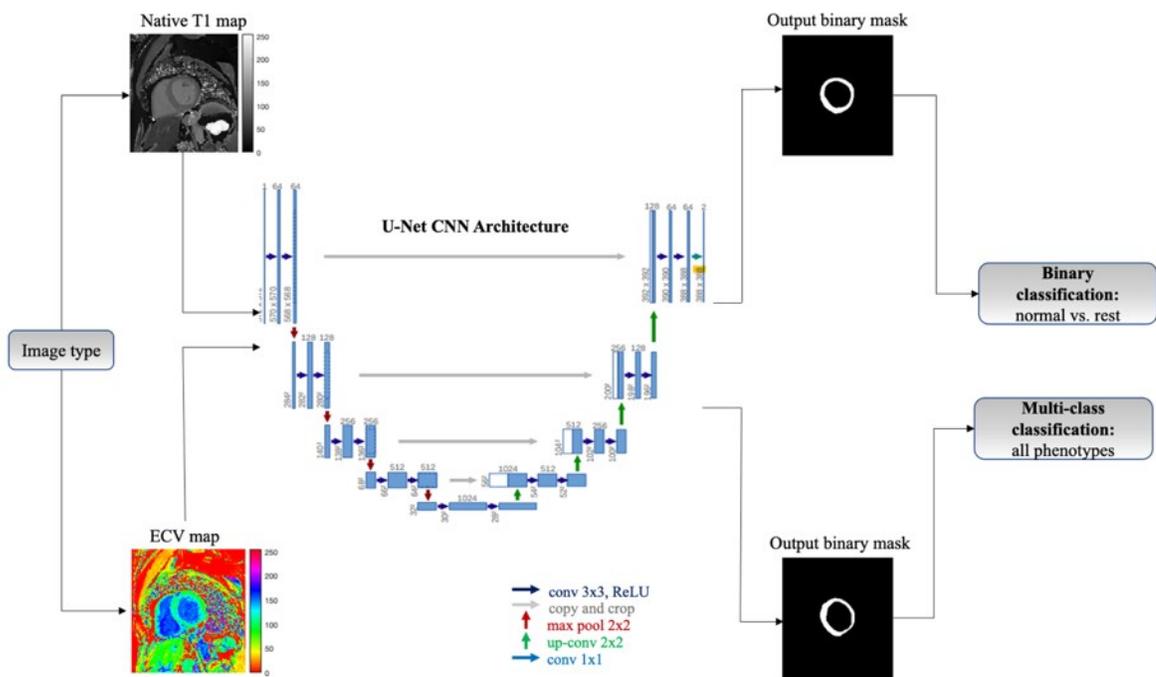


Figure 7-2 Classification approach used for native T1-mapping and ECV mapping images. Native T1 images were used for binary classification wherein the normal phenotype was differentiated against all other cardiomyopathy disease states, i.e., normal versus cardiac amyloidosis, Fabry’s disease, HCM or HTN. ECV mapping images were used for multiclass classification wherein all 5 phenotypes were differentiated.

7.2.3 Cardiomyopathy phenotype classification

As demonstrated in the image processing pipeline in Figure 7-1, preprocessed images and binary masks were used as input into two algorithms, which were employed for performance comparison, to achieve classification of cardiomyopathy phenotypes. To this

aim, we experimented with binary (normal vs. abnormal) and multiclass classification of the five disease states. Native T1 mapping is sensitive to myocardial processes that affect the myocardium globally and best demonstrates areas of replacement MF [27]– [29], whereas ECV mapping is predominantly reflective of the myocardial ECV space and best demonstrates ECV expansion, and may consequently be more beneficial in the assessment of cardiomyopathies associated with interstitial fibrosis [29]–[31]. Since the disease states studied in this work are all associated with interstitial fibrosis, ECV mapping images were selected for multiclass classification, while native T1 mapping images were selected for binary classification, as demonstrated in Figure 7-2.

7.2.3.1 RA-based classification method

Two machine-learning algorithms were compared for RA-based classification of cardiomyopathies, namely, a gradient boosting regression trees (GBRT) algorithm and a k-nearest neighbour (KNN) algorithm. An ensemble learning approach was implemented, wherein both GBRT and KNN algorithms were stacked with a linear support vector classifier (SVC) to boost classification performance.

The GBRT algorithm is a model that is similar to the random forest algorithm, but instead uses an ensemble of regression trees. Regression trees output continuous values rather than discrete values (as in decision trees). Gradient boosting is implemented, whereby difference between the current prediction and the target value is calculated (called the residual difference). The output of each regression tree is summed, and the residual difference is again computed. In this manner, each subsequent regression tree corrects the error of the previous tree. The process is repeated along the model, allowing the residual difference to be minimized which improves prediction accuracy.

The KNN method makes predictions based on the similarity of an instance to k nearest neighbours in feature space based on a weighting factor and distance metric. The parameter k (i.e., number of nearest neighbours) is chosen by the user, typically using a grid search to find the optimized parameter.

Both RA-based classification algorithms consisted of three steps: (1) feature extraction, (2) feature pre-processing, and (3) model training. These steps were implemented using Python v. 3.9 using PyRadiomics, Scikit-learn, and Imbalanced-learn packages[32]– [34].

7.2.3.1.1 Feature extraction

Radiomic feature extraction was achieved using PyRadiomics software [34], an open-source platform capable of extracting a large variety of radiomic features from a region of interest across various user-defined parameters. We applied PyRadiomics' built-in preprocessing methods prior to feature extraction, namely, image normalization (i.e., intensity scaling), histogram bin discretization, and application of several filters and mathematical transforms. For normalization, pixel intensities were scaled between a range of 0 and 1. Bin discretization is performed on pixels within the given binary ROI, wherein intensities are grouped according to specific range intervals. Bin widths were chosen based on a heuristic approach. For binary classification with native T1-mapping images, a bin width of 16 was used, while a bin width of 25 was used for ECV mapping images. The filters and mathematical transforms applied to images prior to feature extraction included the Laplacian of Gaussian (LOG) filter with sigma (σ) = 1.0, 3.0 and 5.0, wavelet transform, logarithm, square, and exponential transforms, and gradient.

PyRadiomics extracted five feature categories from the original image and all aforementioned filtered/transformed images [34], including (1) first-order statistical, (2) shape descriptors, (3) gray level cooccurrence matrix[10], (4) gray level run length matrix [35], [36], and (5) gray level size zone matrix [37]. This resulted in a total of 1125 unique features to be used for classification.

7.2.3.1.2 Pre-processing of radiomic features

Radiomic features were normalized such that all feature values ranged between minimum and maximum values of 0 and 1, respectively, before being randomized and subdivided into training (90%) and testing (10%) subsets. A stratified approach was used to subdivide the data such that proportions of the classes were maintained in the train and test subsets.

As demonstrated in Table 7-1, Fabry's disease patients were fewest among all cardiomyopathy groups, representing only 9% and 11% of all native T1-mapping and ECV mapping images, respectively. To address the class imbalance, synthetic minority oversampling technique (SMOTE)[38] was applied to the training subset in order to up-sample the minority classes to match the representation of the majority class (i.e., the HCM group for ECV maps). The SMOTE technique creates artificial minority samples based on the interpolation between a given minority sample and its nearest minority class neighbors in feature space, thereby increasing the representation of minority samples while preserving the structure of the original data. Synthetic feature data of Fabry's, cardiac amyloidosis, HTN and healthy groups were created to increase the number of each subtype in the training set 184 cases for ECV mapping image multiclass classification. For binary classification of native T1-mapping images, all disease phenotypes were combined and

considered a single class to be differentiated from the healthy (normal) group. Therefore, SMOTE resampling was used to up-sample the normal group to match the representation of the abnormal group, bringing up the total of normal cases to 498 in the training set.

Following SMOTE, we implemented feature selection to reduce noise, redundancy, and computation time prior to training the algorithms. We compared three common methods: principal component analysis (PCA), least absolute shrinkage and selection operator (LASSO) and mutual information. The mutual information method was ultimately chosen for feature selection as it generated the best reproducibility across repeated experiments as compared to PCA and LASSO methods. Mutual information is a measure of the relative entropy between the joint probability and the product probability of the predictor variable (i.e., the features) and target variable (i.e., the classes), thereby providing a relative estimate of the dependence between each feature and the classes and is given in Equation 7-1 [39]. A criterion of mutual information > 0.22 was required for any given feature to be selected.

$$I(X, Y) = \int_x \int_y p(x, y) \log \frac{p(x, y)}{p(x)p(y)} dx dy \quad \text{Equation 7-1}$$

where $p(x, y)$ is the joint probability density function and $p(x)$ and $p(y)$ are the marginal distribution functions of variable X (feature) and variable Y (target).

7.2.3.1.3 Model training

A randomized search [40] was used to find the best combination of hyperparameters to tune the GBRT and KNN classification algorithms. The models were then trained with the optimized hyperparameters and validated using 5-fold cross validation. Since feature data was randomized prior to being split into train and test sets,

the possibility exists of images from the same patient (e.g., apical and base LV images) existing within a single fold during the 5-fold cross validation procedure. However, this is unlikely to impact classification results since inherent anatomical differences exist across base, mid-point and apical regions of the heart (e.g., level of hypertrophy, degree of interstitial fibrosis, sub-endocardial involvement, etc.). The hyperparameters used for the GBRT and KNN algorithms are demonstrated in Tables 7-2 and 7-3, respectively.

Table 7-2 Hyperparameters used for GBRT binary and multi-class classification algorithm

Hyperparameter	Binary classification (Native T1 mapping)	Multi-class classification (ECV mapping)
Number of trees in forest	1008	1924
Maximum depth of each tree	4	9
Minimum samples to split internal node	2	2
Minimum samples at each leaf node	1	15
Maximum # features to consider when looking for best split	43	14

Table 7-3 Hyperparameters used for KNN binary and multi-class classification algorithm

Hyperparameter	Binary classification (Native T1 mapping)	Multi-class classification (ECV mapping)
Number of nearest neighbours	1	3
Weight function used in prediction	Distance (neighbour points weighted inversely proportional to their distance)	Uniform (all points in each neighborhood are weighted equally)
Distance metric	Minkowski	Minkowski
Approach used to compute the nearest neighbours	Naïve	Ball tree

7.2.3.1.4 Ensemble approach using stacking with linear support vector classifier

Both the GBRT and KNN classifiers were separately stacked with a linear SVC to build the final classifier, as shown in Figure 7-3. The linear SVC used a “one versus rest” approach to classify each disease state in multiclass classification of ECV images while a “one versus one” approach was used for binary classification. L2 regularization was used

to avoid overfitting. The linear SVC was optimized using a squared hinge loss function. The final meta-model used for classification was a logistic regression algorithm which was built by stacking estimators from the GBRT and KNN with the linear SVC, respectively.

Typically, machine learning classifiers based on decision trees, such as GBRT, perform comparably to KNN. Since stacking is a machine learning method based on boosting model performance using weak learners, stacking models whose performances are comparable may not result in improved accuracy. Therefore, we decided not to implement a three-way stacking model inclusive of all GBRT, KNN and linear SVC.

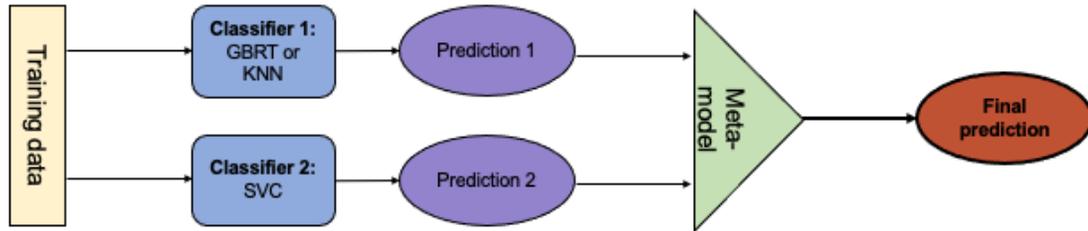


Figure 7-3 Stacking ensemble approach for RA-based classification

7.2.3.2 Deep learning-based classification method

7.2.3.2.1 CNN training and testing procedure

The CNN-based classification method used images and binary masks as input to achieve binary (native T1-mapping) and multiclass (ECV mapping) disease classification. Native T1 and ECV mapping datasets were subdivided into training (90%) and testing (10%) subsets (respectively), using a stratified approach such that the proportions of the classes were maintained in the train and test subsets. Images were normalized such that pixel intensities ranged between 0 – 1 prior to training and testing. Due to the class imbalance noted in Table 1, data augmentation was used for multiclass classification of ECV mapping images, wherein images were rotated, flipped and shifted in height and

weight to up-sample the minority classes and avoid over-fitting. Training and testing were performed on GPU in Kaggle using Scikit-learn, Keras and Tensorflow libraries.

7.2.3.2.2 CNN Architecture

Two separate CNN models were trained for the multiclass and binary classification tasks. The CNN models were built with four convolution layers, each followed by a max pooling layer. Two dense layers were placed after the final max pooling layer, wherein the final dense layer applied the activation function to generate the class labels. Softmax and sigmoid activation functions were used for multiclass and binary classification, respectively. Hyperparameters were tuned manually and are shown in Table 4. Both CNN models were trained for 15 epochs. The architecture of CNN classification models is shown in Figure 7-4.

Table 7-4 Hyperparameters for binary and multiclass classification of ECV mapping images

Hyperparameter	Value	
	Binary classification of native T1 mapping	Multi-class classification of ECV mapping
Optimizer	Adam	RMS prop
Learning rate	0.0001	0.00001
Loss function	Binary cross entropy	Categorical cross entropy
Batch size	16	16
Data augmentation	N/A	Yes

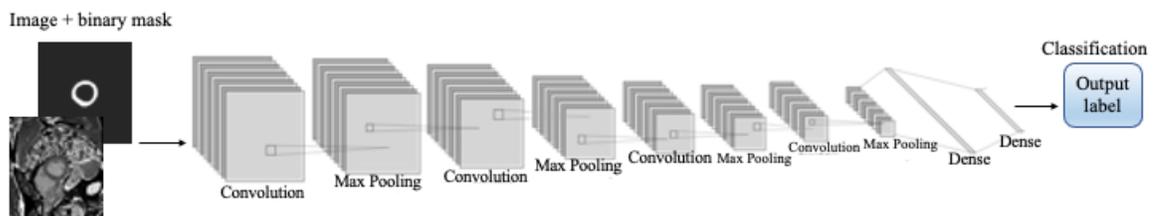


Figure 7-4 CNN architecture used for cardiomyopathy classification.

An example normal phenotype ECV mapping image (shown in grayscale) and binary mask is shown. The activation function at the final dense layer of the CNN generated the classification label (multiclass for ECV mapping images and binary for native T1-mapping images).

7.2.4 Performance Evaluation and Statistical Analysis

The performance of RA-based and CNN-based classification algorithms were assessed using conventional metrics including accuracy, precision, recall (sensitivity), and F1 score. All aforementioned metrics were computed per class (i.e., per each disease group) and across all classes. For computation of per-class metrics, the class-specific number of true positive, true negative, false positive, and false negative (i.e., TP, TN, FP, and FN, respectively) were used, while overall metrics were computed using the total number of TP, TN, FP, and FN instances across all classes.

Classification performance was also assessed using receiver operator curves (ROC) and area under the curve (AUC). The ROC and AUC metrics were computed on a per-class basis for multiclass classification on ECV mapping images and across both classes for binary classification on native T1-mapping images.

To assess the correlation between the multiclass and binary predicted labels vs. true labels, the Spearman Correlation Coefficient ρ was computed. To assess the significance between RA-based classification labels (i.e., GBRT vs. KNN ensemble), a Wilcoxon signed rank test was used. The Kruskal-Wallis test was used to determine significance between the predicted RA-based and predicted CNN-based classification labels against the true class labels, respectively, for both multiclass and binary classification. Lastly, McNemar's chi square test was used to assess the difference between the proportions of correctly classified observations (i.e., accuracy) among all three models. The McNemar test is a non-parametric method that compares the discordance of proportions of paired observations and is applied to a 2x2 contingency table to assess row and column marginal homogeneity with one degree of freedom [41]. The contingency table used for this analysis

is shown in Table 7-5. The McNemar test statistic is given by Equation 7-2, by which a value of 0 denotes no difference between paired proportions of observations and larger values denote larger differences.

Table 7-5 Contingency table for the McNemar test

	Method 2 correct	Method 2 misclassified
Method 1 correct	<i>A</i>	<i>B</i>
Method 1 misclassified	<i>C</i>	<i>D</i>

A represents the number of observations correctly classified by both methods, *B* is the number of observations correctly classified by Method 1 but misclassified by Method 2, *C* is the number of observations misclassified by Method 1 but correctly classified by Method 2, and *D* is the number of observations incorrectly classified by both methods.

$$\chi^2 = \frac{(B-C)^2}{B+C} \quad \text{Equation 7-2}$$

7.3 Results

7.3.1 Myocardial Segmentation

Myocardial segmentations automatically derived using the U-Net technique were compared to manual contours delineated by an expert. Dice similarity metric (DSC), precision, recall and boundary F1 scores are demonstrated in Table 6 for native T1-mapping and ECV images. The U-Net method achieves results greater than 86% across all performance metrics in myocardial segmentation of both native T1 maps and ECV maps. Example segmentations derived by the U-Net architecture versus manual ground truth delineations are demonstrated across all cardiomyopathy phenotypes in Figure 7-5.

Table 7-6 Accuracy metrics of myocardial segmentations generated using U-Net and transfer learning

Metric	Image type	
	Native T1 maps	ECV maps
DSC	86 ± 7.3	88 ± 6.1
Precision	87 ± 13	89 ± 10
Recall	88 ± 12	90 ± 10
Boundary F1	87 ± 13	89 ± 10

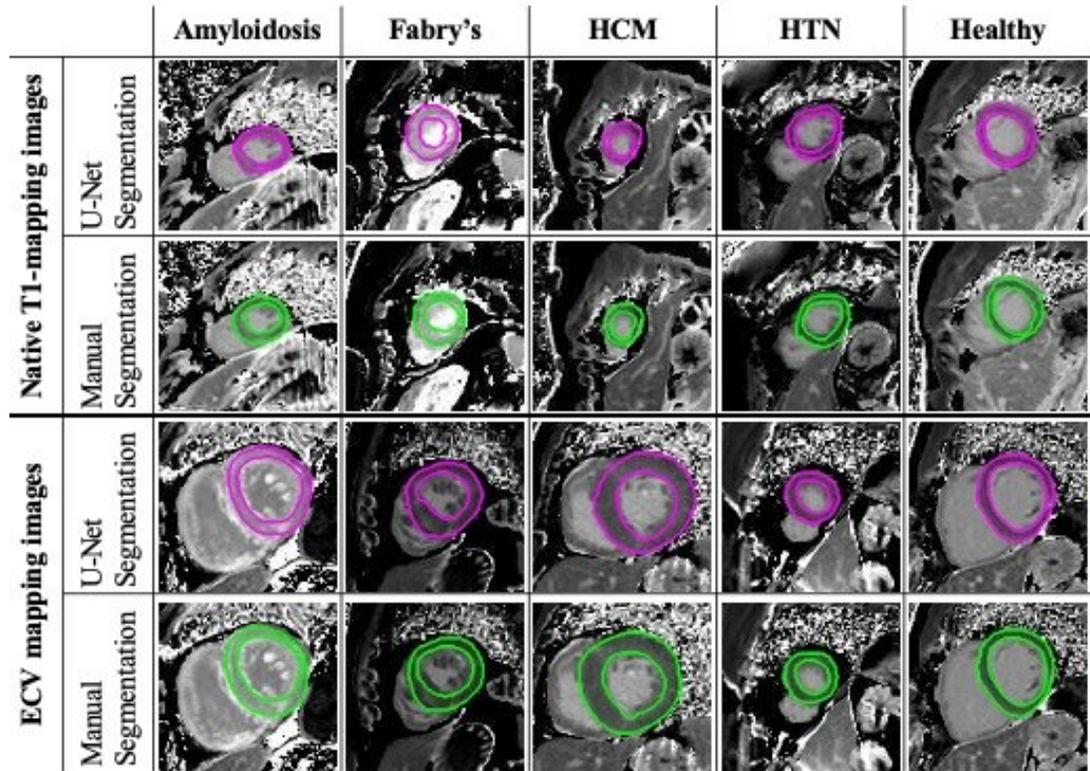


Figure 7-5 Myocardial segmentations derived using U-Net versus manually.

U-Net contours are shown in magenta and manual contours are shown in green in native T1-maps (first two rows) and ECV maps (second two rows). The ECV mapping images are demonstrated in grayscale rather than colour, in order to better demonstrate the coloured myocardial contours. An example image is demonstrated for each cardiomyopathy phenotype.

7.3.2 Binary classification using native T1-mapping images

For the RA-based classifiers, feature selection was performed using a criterion of mutual information >0.22 , thereby reducing the number of features from 1125 to 98. A pie chart demonstrating the distribution of feature importance across various image filters and transforms is shown in Figure 7-6. The logarithm transform provided the most significant features by mutual information, followed by the gradient transform. Table 7-7 demonstrates the MI distribution of feature categories across all filters and transforms. The most significant feature category was the GLRLM, followed by first order statistics, across all filters and transforms.

Pie chart distribution of filters and transforms with the highest MI used for binary classification

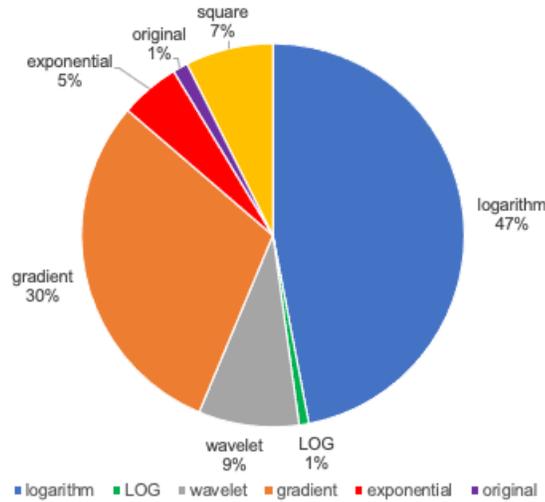


Figure 7-6 Pie chart of filters and transforms with highest MI for binary classification

Table 7-7 Distribution of binary feature importance across categories

	Feature category				
	First order	Shape	GLCM	GLRLM	GLSZM
% used for classification	9.51	0	12.84	66	3.08

Both GBRT and KNN ensemble models were trained using a stratified 5-fold cross validated approach, wherein the average training accuracies are reported at $93 \pm 2.3\%$ and $93 \pm 1.6\%$, respectively. Table 7-8 demonstrates per-class testing accuracy, precision, recall, and F1 score metrics given by the two RA-based and the deep learning (CNN)-based binary classification methods. Although both RA-based classification methods outperformed the deep learning CNN classification algorithm, in particular, the KNN and SVC ensemble method performed superiorly across all metrics, excluding precision. All models performed best in classification of the abnormal subgroup versus normal (healthy). Normalized confusion matrices for GBRT and SVC ensemble, KNN and SVC ensemble, and CNN methods are demonstrated in Figure 7-7.

Table 7-8 Accuracy metrics of binary abnormal vs. normal classification using native T1-mapping images

Metric	Patient group	Classification Method		
		GBRT and SVC ens.	KNN and SVC ens.	CNN
Accuracy (%)	All patients	88	89	85
Precision (%)	Normal (healthy)	76	80	85
	Abnormal	93	93	85
Recall (%)	Normal (healthy)	83	83	74
	Abnormal	90	91	92
F1 Score (%)	Normal (healthy)	79	82	79
	Abnormal	91	92	89

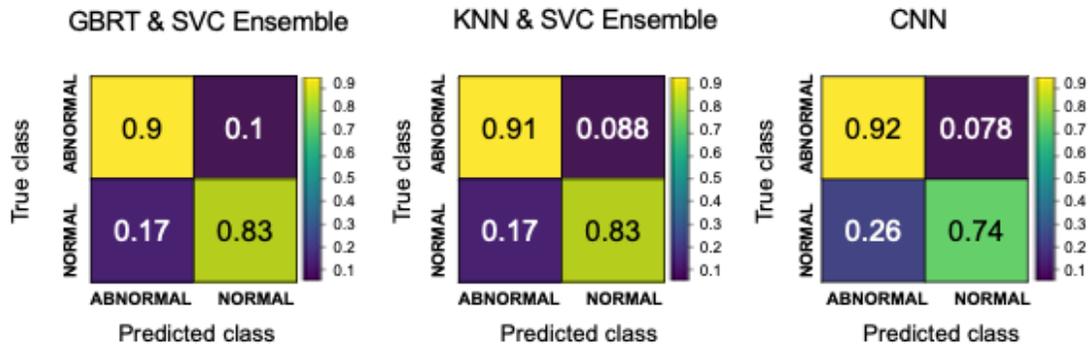


Figure 7-7 Normalized confusion matrices for three binary classification models. Predicted labels (columns) are normalized by the true number of classes (rows).

Figure 7-8 demonstrates the binary ROC, wherein the AUC metric is reported in the legend. The AUC of the GBRT and SVC ensemble and the KNN and SVC ensemble methods were computed at 0.939 and 0.934, respectively, which exceeded that of the CNN-based method, computed at 0.88.

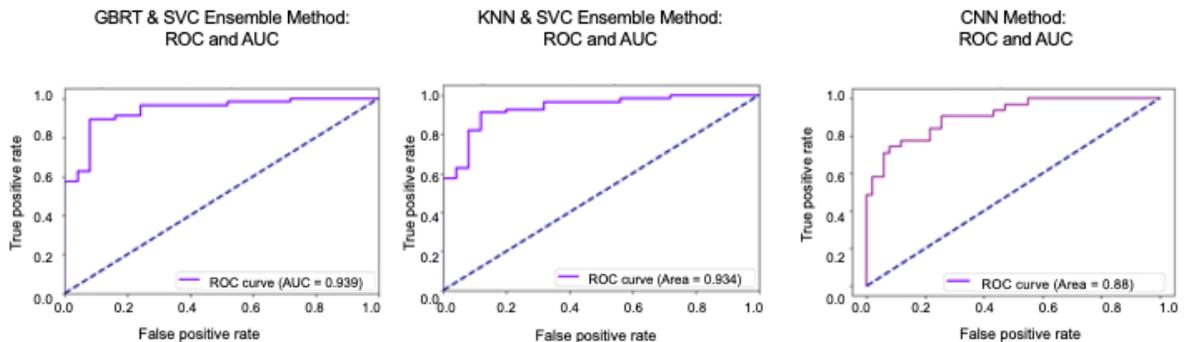


Figure 7-8 Binary ROC and AUC for three classification models.

7.3.3 Multiclass classification using ECV mapping images

RA-based classification of five cardiomyopathy states, i.e., amyloidosis, Fabry’s, HCM, HTN and normal (healthy), was performed after feature selection a criterion of mutual information > 0.22 , reducing the number of features from 1125 to 125. A pie chart demonstrating the distribution of feature importance across various image filters and transforms is shown in Figure 7-9. The exponential transform yielded the most significant features by mutual information, followed by features derived from the original image. Table 7-9 demonstrates the MI distribution of feature categories across all filters and transforms. The most significant feature category was the GLCM, followed by first order statistics, across all filters and transforms.

Pie chart distribution of filters and transforms with the highest MI used for multiclass classification

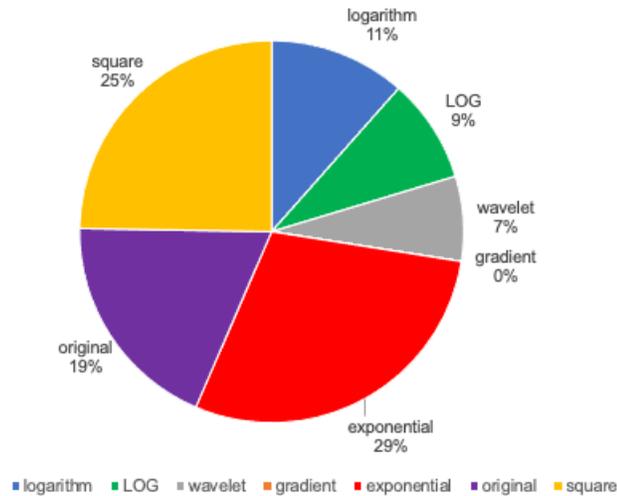


Figure 7-9 Pie chart of filters and transforms with highest MI for multiclass classification

Table 7-9 Distribution of multiclass feature importance across categories

	Feature category				
	First order	Shape	GLCM	GLRLM	GLSZM
% used for classification	31.24	0	36.13	13.44	5.41

Both GBRT and KNN ensemble models were trained using a stratified 5-fold cross validated approach, wherein the average training accuracies are reported at $77 \pm 9\%$ and $86 \pm 2.6\%$, respectively. Table 7-10 demonstrates testing per-class and overall performance metrics, including accuracy, precision, recall and F1 score across two RA-based classification models and the CNN-based model. The RA-based classification methods outperformed the deep learning CNN classification algorithm across the majority of class-specific metrics. Among the two RA-based algorithms, the GBRT ensemble method outperformed the KNN ensemble method across all metrics. The GBRT and KNN ensemble methods yielded the lowest precision, recall, and F1 score metrics among HTN group followed by the Fabry’s disease group, whereas the CNN method yielded the lowest aforementioned scores in the healthy and Fabry’s disease group. The GBRT ensemble method demonstrated the highest metrics among the healthy group, compared to the amyloidosis group and the HCM group for KNN ensemble and CNN methods, respectively. The normalized confusion matrices demonstrating classification accuracies for each subgroup is demonstrated across each method in Figure 7-10.

Table 7-10 Performance metrics computed on algorithms used for multiclass classification of ECV mapping images

Metric	Patient group	Classification Method		
		GBRT & SVC ens.	KNN & SVC ens.	CNN
Accuracy (%)	Amyloidosis	90	90	90
	Fabry’s disease	86	86	94
	HCM	78	74	68
	HTN	84	76	88
	Healthy	94	90	84
	All patients	66	58	62
Precision (%)	Amyloidosis	78	78	83
	Fabry’s disease	50	50	0
	HCM	70	65	61
	HTN	38	25	33
	Healthy	86	57	0
	All patients	64	55	59

Recall (%)	Amyloidosis	70	70	56
	Fabry's disease	43	43	0
	HCM	74	68	100
	HTN	50	25	20
	Healthy	75	67	0
	All patients	62	55	62
F1 Score (%)	Amyloidosis	74	74	67
	Fabry's disease	46	46	0
	HCM	72	67	76
	HTN	43	25	25
	Healthy	80	62	0
	All patients	63	55	55

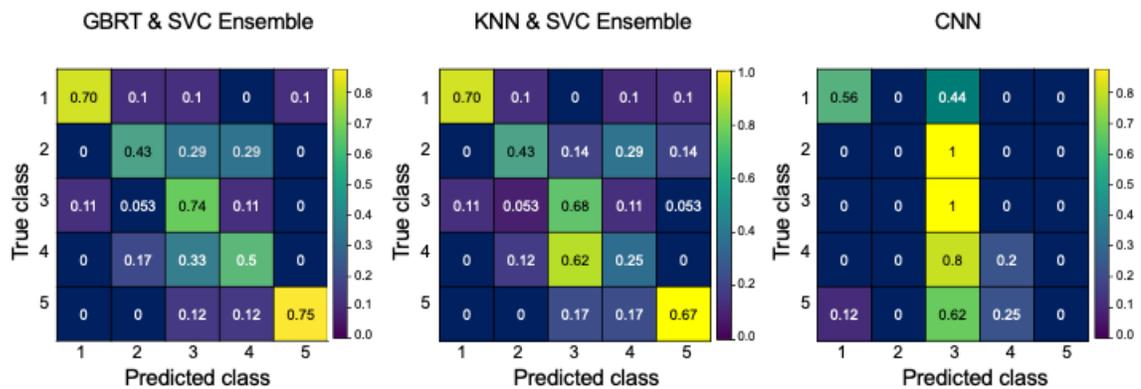


Figure 7-10 Normalized confusion matrices for multiclass classification algorithms.

Confusion matrices for GBRT and SVC ensemble (left), KNN and SVC ensemble (middle), and CNN classification algorithms (right) demonstrate per-class classification accuracies across five cardiomyopathy phenotypes. Predicted labels (columns) are normalized by the true number of samples (rows) per each phenotype. The cardiomyopathy groups are numbered as follows: amyloidosis (class 1), Fabry's disease (class 2), HCM (class 3), HTN (class 4), and healthy (class 5).

Figure 7-11 demonstrates the receiver operating curve (ROC) computed for each of the five cardiomyopathy classes. The area under the curve (AUC) metric is reported in the figure legend. The class demonstrating the highest AUC was amyloidosis for all algorithms. The lowest AUC was demonstrated in the Fabry class for the RA-based algorithms, while Fabry and healthy were tied for the lowest AUC yielded by the CNN-based algorithm.

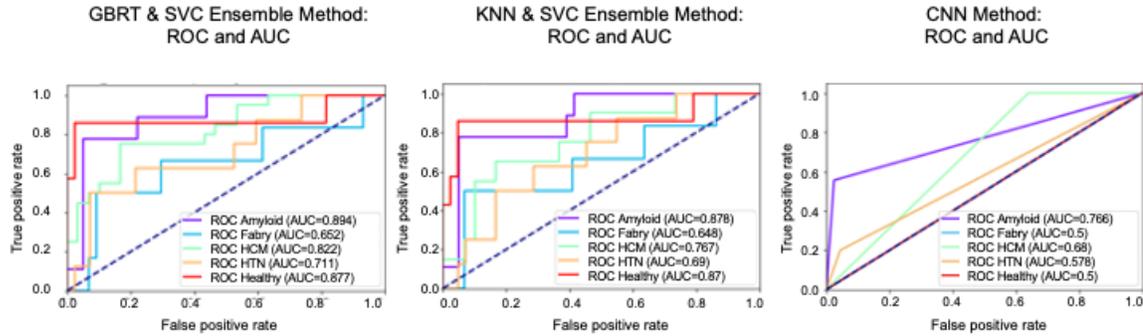


Figure 7-11 Multiclass ROC and AUC for three classification models.

ROC curves and AUC metric per-class for RA-based classification models (GBRT and SVC ensemble method, KNN and SVC ensemble method) and CNN model. The colour/class legend demonstrates the corresponding AUC for each class as follows: Amyloidosis (purple ROC), Fabry's disease (blue ROC), HCM (lime green ROC), HTN (yellow ROC), and healthy (red ROC).

7.3.4 Statistical Analyses

The Spearman correlation coefficient was computed to determine the correlation between predicted classification labels output by each model versus the true class labels, as shown in Table 7-11 for binary and multiclass classification. The KNN ensemble method was the most highly correlated to the true binary class labels, while the GBRT ensemble method was the most highly correlated to the true multiclass class labels.

Table 7-11 Spearman correlation coefficients of GBRT, KNN and CNN classification models versus true labels

Classification method	Spearman correlation coefficient	P-value
Binary classification		
GBRT ensemble method	0.705	<0.0001
KNN ensemble method	0.737	<0.0001
CNN method	0.685	<0.0001
Multiclass classification		
GBRT ensemble method	0.659	<0.0001
KNN ensemble method	0.514	0.00011
CNN method	0.475	0.00042

The Kruskal-Wallis test did not demonstrate significance between the predicted binary class labels output by RA models (P-value=0.964) or CNN model (P-value=0.586) versus the true class labels. To determine the significance between the classification labels predicted by the RA-based methods against each other (i.e., GBRT ensemble method vs. KNN ensemble method), a Wilcoxon signed rank test was applied, which did not

demonstrate statistically significant differences (P-value=1.0). Similarly, McNemar’s chi square test did not demonstrate significance in the marginal homogeneity of paired proportions for the GBRT vs. KNN ensemble methods (P=1.0, McNemar’s test statistic = 0), GBRT ensemble vs. CNN method (P=0.824, McNemar’s test statistic = 9.0), nor KNN ensemble vs. CNN method (P=0.648, McNemar’s test statistic = 8.0); Binary contingency tables for three-way classification model comparisons are demonstrated in Figure 7-12.

For multiclass classification, the Kruskal-Wallis test did not demonstrate significance between predicted class labels output by both RA-based models (P-value=0.926); however, statistical significance was demonstrated on predicted labels from the CNN model versus the true class labels (P-value=0.0042). The Wilcoxon signed rank test was applied to determine the significance between the classification labels predicted by the RA-based classifiers (i.e., GBRT ensemble method vs. KNN ensemble method), which did not demonstrate significance (P-value=0.747). For analysis of the marginal homogeneity of paired proportions, McNemar’s chi square test was used, which did not demonstrate significance for the GBRT vs. KNN ensemble methods method (P=0.344, McNemar’s test statistic = 3.0), the GBRT ensemble vs. CNN method (P=0.851, McNemar’s test statistic = 13.0), nor KNN ensemble vs. CNN (P=0.839, McNemar’s test statistic = 14.0). Multiclass contingency tables for three-way classification model comparisons are demonstrated in Figure 7-13.

Proportion of correct vs. misclassified binary labels for GBRT versus KNN ensemble methods			Proportion of correct vs. misclassified binary labels for GBRT ensemble versus CNN methods			Proportion of correct vs. misclassified binary labels for KNN ensemble versus CNN methods		
	KNN ens. correctly classified	KNN ens. misclassified		CNN correctly classified	CNN misclassified		CNN correctly classified	CNN misclassified
GBRT ens. correctly classified	71	0	GBRT ens. correctly classified	19	4	KNN ens. correctly classified	20	4
GBRT ens. misclassified	1	9	GBRT ens. misclassified	50	8	KNN ens. misclassified	49	8

Figure 7-12 Contingency tables for three-way comparison of binary classification models

Proportion of correct vs. misclassified multiclass labels for GBRT versus KNN ensemble methods			Proportion of correct vs. misclassified multiclass labels for GBRT ensemble versus CNN methods			Proportion of correct vs. misclassified multiclass labels for KNN ensemble versus CNN methods		
	KNN ens. correctly classified	KNN ens. misclassified		CNN correctly classified	CNN misclassified		CNN correctly classified	CNN misclassified
GBRT ens. correctly classified	26	3	GBRT ens. correctly classified	10	23	KNN ens. correctly classified	9	20
GBRT ens. misclassified	7	14	GBRT ens. misclassified	6	11	KNN ens. misclassified	7	14

Figure 7-13 Contingency tables for three-way comparison of multiclass classification models

7.4 Discussion

Quantitative imaging techniques including native T1-mapping and ECV mapping CMR permit characterization of myocardial tissue space, allowing for differentiation between various cardiovascular diseases that present with MF. However, current methods rely on automated or semi-automated delineation of myocardial contours prior to classification. In this work, we implemented a fully automated pipeline inclusive of myocardial segmentation using a transfer-learning approach for binary (normal versus abnormal) and multiclass classification of five cardiomyopathy phenotypes.

The transfer-learning approach used in our pipeline allowed for all images to be used for training of the classification algorithms. We were able to achieve myocardial segmentation with DSC exceeding 86% in both native T1 and ECV mapping images by using a U-Net method that had been pre-trained on native T1 and ECV mapping images in patients with ICM and DCM.

We compared two RA-based and one deep learning-based method for both multiclass and binary classification using ECV mapping and native T1-mapping images, respectively. For binary classification of normal vs. abnormal phenotypes, all methods achieved classification accuracy metrics exceeding 83%; however, both RA-based methods performed superiorly over CNN-based methods. Although the KNN and SVC ensemble

method performed better than the alternative RA-based method, the Wilcoxon signed rank test did not demonstrate statistical significance between the two methods (P-value=1.0). The RA-based methods outperformed the CNN-based method particularly in the classification of the normal subgroup. Nevertheless, output binary classification labels provided by all three methods were highly correlated with the true class labels, wherein Spearman ρ was computed > 0.685 for all methods. Statistical significance was not demonstrated between binary predicted labels vs. true class labels among any method.

For multiclass classification, the GBRT ensemble method performed superiorly compared to the other two classification methods, wherein all accuracy metrics exceeded 62%. The GBRT ensemble method also demonstrated the highest AUC across all cardiomyopathies as compared to the other methods. While both RA-based methods performed well in classification of the healthy and amyloidosis groups, the CNN method performed poorly across these groups, as demonstrated in the ROC curves and confusion matrices in Figures 9 and 10. The CNN method over-predicted the HCM class and under-predicted the Fabry and healthy subgroups, likely due to class imbalance. Thus, low AUC scores are demonstrated across these two groups as well as the HTN group. Thus, significance between true classification labels was not demonstrated against the RA-based methods, but was demonstrated against the CNN method.

Both RA-based and deep learning methodologies typically require large amounts of data to be available in order to produce the best performing classification models. Without adequate data, it can be particularly challenging to train high-performing classification models for multiclass problems, as classification error can increase proportionally to the problem complexity. Moreover, class imbalance can contribute to this challenge. In this

work, we were limited by both the inadequate amount of data and class imbalance for the training of multiclass classification models, leading to moderate accuracies in classification of five cardiomyopathy states. This was particularly problematic for the multiclass classification using the CNN-based technique, which was to be expected, as CNN based methodologies require a larger data pool to in order train the best model [11]. On the contrary, this phenomenon was not demonstrated in the binary classification task across all models, since reducing the number of classes inherently increases the number of samples per class, thereby yielding increased data availability for training. Therefore, in this work, we demonstrate that both RA-based and CNN-based methods are suitable for binary classification of normal vs. abnormal cardiomyopathy phenotypes. However, for multiclass classification, RA-based methods may be more appropriate than CNN-based methods.

In order to improve multiclass classification results using any method, more data is required. Specifically, increasing the data of the minority classes (e.g., Fabry's disease and HTN) would be optimal to improve model performance. However, there are few T1- and native mapping datasets available that are both based on the MOLLI imaging sequence and inclusive of such diverse cardiomyopathy states. While further imaging could be performed to acquire more data, challenges will ultimately still exist in underrepresentation of the minority classes, especially rare diseases, such as Fabry's disease and Amyloidosis. For example, both Fabry's disease and HTN are underrepresented in our dataset. However, Fabry's disease is quite rare, occurring in only 1 in 40,000–60,000 people as compared to HTN which affects approximately 1 in 3–5 adults. Thus, it may be feasible to acquire additional imaging data in the HTN class, but feasibility of additional data acquisition in the Fabry class will inevitably be difficult. Nevertheless, alternative options to increase the

data available for classification may exist. For example, additional imaging data may be available in alternative native T1- and ECV mapping pulse sequences, such as the shMOLLI or SASHA techniques. However, radiomic features may not be directly translatable across different T1-mapping pulse sequences. Nevertheless, this could be explored in further work.

7.5 Conclusion

Native T1-mapping and ECV mapping permits differentiation between various cardiomyopathy states by either binary (normal vs. abnormal) classification via RA or deep learning-based methods. RA-based methods outperform deep learning for multiclass classification; however, multiclass model optimization requires additional data. Further studies and/or the addition of supplemental imaging data could permit improved multiclass classification accuracy.

7.6 Chapter 7 References

- [1] J. R. Burt, S. L. Zimmerman, I. R. Kamel, M. Halushka, and D. A. Bluemke, “Myocardial T1 Mapping: Techniques and Potential Applications,” *RadioGraphics*, vol. 34, no. 2, pp. 377–395, Mar. 2014, doi: 10.1148/rg.342125121.
- [2] P. K. Kim *et al.*, “Myocardial T1 and T2 Mapping: Techniques and Clinical Applications,” *Korean Journal of Radiology*, vol. 18, no. 1, p. 113, 2017, doi: 10.3348/kjr.2017.18.1.113.
- [3] A. J. Taylor, M. Salerno, R. Dharmakumar, and M. Jerosch-Herold, “T1 Mapping Basic Techniques and Clinical Applications,” *JACC: Cardiovascular Imaging*, vol. 9, no. 1, pp. 67–81, 2016, doi: 10.1016/j.jcmg.2015.11.005.
- [4] V. O. Puntmann, E. Peker, Y. Chandrashekar, and E. Nagel, “T1 Mapping in Characterizing Myocardial Disease,” *Circulation Research*, vol. 119, no. 2, 2016.
- [5] E. K. Oikonomou *et al.*, “A novel machine learning-derived radiotranscriptomic signature of perivascular fat improves cardiac risk prediction using coronary CT angiography,” *European Heart Journal*, vol. 40, no. 43, pp. 3529–3543, Nov. 2019, doi: 10.1093/eurheartj/ehz592.
- [6] A. S. Antonopoulos *et al.*, “Machine learning of native T1 mapping radiomics for classification of hypertrophic cardiomyopathy phenotypes,” *Scientific Reports 2021 11:1*, vol. 11, no. 1, pp. 1–11, Dec. 2021, doi: 10.1038/s41598-021-02971-z.
- [7] S. Rizzo *et al.*, “Radiomics: the facts and the challenges of image analysis,” *European Radiology Experimental*, vol. 2, no. 1, Dec. 2018, doi: 10.1186/S41747-018-0068-Z.

- [8] M. E. Mayerhoefer *et al.*, “Introduction to Radiomics,” *Journal of Nuclear Medicine*, vol. 61, no. 4, pp. 488–495, Apr. 2020, doi: 10.2967/JNUMED.118.222893.
- [9] J. E. van Timmeren, D. Cester, S. Tanadini-Lang, H. Alkadhi, and B. Baessler, “Radiomics in medical imaging— ‘how-to’ guide and critical reflection,” *Insights into Imaging*, vol. 11, no. 1, Dec. 2020, doi: 10.1186/S13244-020-00887-2.
- [10] R. M. Haralick, I. Dinstein, and K. Shanmugam, “Textural Features for Image Classification,” *IEEE Transactions on Systems, Man and Cybernetics*, vol. SMC-3, no. 6, pp. 610–621, 1973, doi: 10.1109/TSMC.1973.4309314.
- [11] D. Truhn, S. Schradling, C. Haarburger, H. Schneider, D. Merhof, and C. Kuhl, “Radiomic versus Convolutional Neural Networks Analysis for Classification of Contrast-enhancing Lesions at Multiparametric Breast MRI,” *Radiology*, vol. 290, no. 3, pp. 290–297, Mar. 2019, doi: 10.1148/radiol.2018181352.
- [12] S. Chang, K. Han, Y. J. Suh, and B. W. Choi, “Quality of science and reporting for radiomics in cardiac magnetic resonance imaging studies: a systematic review,” *European Radiology*, pp. 1–13, Mar. 2022, doi: 10.1007/S00330-022-08587-9/TABLES/5.
- [13] C. Izquierdo *et al.*, “Radiomics-Based Classification of Left Ventricular Non-compaction, Hypertrophic Cardiomyopathy, and Dilated Cardiomyopathy in Cardiovascular Magnetic Resonance,” *Frontiers in Cardiovascular Medicine*, vol. 0, p. 1502, Oct. 2021, doi: 10.3389/FCVM.2021.764312.
- [14] E. Avard *et al.*, “Non-contrast Cine Cardiac Magnetic Resonance image radiomics features and machine learning algorithms for myocardial infarction detection,” *Computers in Biology and Medicine*, vol. 141, Feb. 2022, doi: 10.1016/J.COMPBIOMED.2021.105145.
- [15] T. di Noto *et al.*, “Radiomics for distinguishing myocardial infarction from myocarditis at late gadolinium enhancement at mri: Comparison with subjective visual analysis,” *Radiology: Cardiothoracic Imaging*, vol. 1, no. 5, Dec. 2019, doi: 10.1148/RYCT.2019180026/SUPPL_FILE/RYCT180026SUPPA1.PDF.
- [16] J. Wang *et al.*, “Improved Segmental Myocardial Strain Reproducibility Using Deformable Registration Algorithms Compared with Feature Tracking Cardiac MRI and Speckle Tracking Echocardiography,” 2017, doi: 10.1002/jmri.25937.
- [17] J. Wang *et al.*, “Radiomic Analysis of Native T1 Mapping Images Discriminates Between MYH7 and MYBPC3-Related Hypertrophic Cardiomyopathy,” 2020, doi: 10.1002/jmri.27209.
- [18] U. Neisius, H. El-Rewaady, S. Nakamori, J. Rodriguez, W. J. Manning, and R. Nezafat, “Radiomic Analysis of Myocardial Native T1 Imaging Discriminates Between Hypertensive Heart Disease and Hypertrophic Cardiomyopathy,” *JACC: Cardiovascular Imaging*, vol. 12, no. 10, pp. 1946–1954, Oct. 2019, doi: 10.1016/J.JCMG.2018.11.024/SUPPL_FILE/MMC1.DOCX.
- [19] B. H. Chen *et al.*, “Myocardial extracellular volume fraction radiomics analysis for differentiation of reversible versus irreversible myocardial damage and prediction of left ventricular adverse remodeling after ST-elevation myocardial infarction,” *European Radiology*, vol. 31, no. 1, pp. 504–514, Jan. 2021, doi: 10.1007/s00330-020-07117-9.
- [20] L. M. Wu *et al.*, “A Radiomic MRI based Nomogram for Prediction of Heart Failure with Preserved Ejection Fraction in Systemic Lupus Erythematosus Patients: Insights from a Three-Center Prospective Study,” *Journal of Magnetic Resonance Imaging*, 2022, doi: 10.1002/jmri.28070.

- [21] T. Dresselaers *et al.*, “T1 and ECV mapping texture analysis distinguishing hypertrophic cardiomyopathy from athletes heart better than median values,” *European Heart Journal - Cardiovascular Imaging*, vol. 22, no. Supplement_2, p. jeab090.085, Jun. 2021, doi: 10.1093/ehjci/jeab090.085.
- [22] J. Jang *et al.*, “Reproducibility of segmentation-based myocardial radiomic features with cardiac MRI,” *Radiology: Cardiothoracic Imaging*, vol. 2, no. 3, Jun. 2020, doi: 10.1148/ryct.2020190216.
- [23] P. Germain *et al.*, “Classification of Cardiomyopathies from MR Cine Images Using Convolutional Neural Network with Transfer Learning,” *Diagnostics (Basel)*, vol. 11, no. 9, p. 1554, 2021, doi: 10.3390/diagnostics11091554.
- [24] H. Zhou *et al.*, “Deep learning algorithm to improve hypertrophic cardiomyopathy mutation prediction using cardiac cine images,” *European Radiology*, vol. 31, no. 6, pp. 3931–3940, 2021, doi: 10.1007/s00330-020-07454-9.
- [25] O. Ronneberger, P. Fischer, and T. Brox, “U-Net: Convolutional Networks for Biomedical Image Segmentation,” pp. 1–8.
- [26] N. A. Farrag, S. Bhagavan, D. Sebben, P. Ruwanpura, J. A. White, and E. Ukwatta, “Automated myocardial segmentation of extra-cellular volume mapping cardiac magnetic resonance images using fully convolutional neural networks,” in *Proc.SPIE*, Apr. 2022, vol. 12036. doi: 10.1117/12.2626738.
- [27] C. R. Hamilton-craig, “T 1 Mapping for Myocardial Fibrosis by Cardiac Magnetic Resonance Relaxometry — A Comprehensive,” vol. 3, no. March, pp. 1–11, 2017, doi: 10.3389/fcvm.2016.00049.
- [28] P. Haaf, P. Garg, D. R. Messroghli, D. A. Broadbent, J. P. Greenwood, and S. Plein, “Cardiac T1 Mapping and Extracellular Volume (ECV) in clinical practice: A comprehensive review,” *Journal of Cardiovascular Magnetic Resonance*, vol. 18, no. 1, pp. 1–12, 2016, doi: 10.1186/s12968-016-0308-4.
- [29] E. B. Schelbert and T. C. Wong, “Clinical Benefits of T1 and ECV Mapping,” *Magnetom Flash*, pp. 12–17, 2015.
- [30] J. C. Moon *et al.*, “Myocardial T1 mapping and extracellular volume quantification: a Society for Cardiovascular Magnetic Resonance (SCMR) and CMR Working Group of the European Society of Cardiology consensus statement,” *Journal of Cardiovascular Magnetic Resonance*, vol. 15, no. 1, p. 92, 2013, doi: 10.1186/1532-429X-15-92.
- [31] F. M. Caballeros *et al.*, “T1 mapping of the myocardium: what a resident need to know,” *European Cardiology Review*, pp. 1–17, 2016, doi: 10.1594/ecr2016/C-0385.
- [32] G. Lemaître, F. Nogueira, and C. K. Aridas char, “Imbalanced-learn: A Python Toolbox to Tackle the Curse of Imbalanced Datasets in Machine Learning,” 2017. [Online]. Available: <http://jmlr.org/papers/v18/16-365.html>.
- [33] A. Abraham *et al.*, “Machine learning for neuroimaging with scikit-learn,” *Frontiers in Neuroinformatics*, vol. 8, no. FEB, Feb. 2014, doi: 10.3389/fninf.2014.00014.
- [34] J. J. M. van Griethuysen *et al.*, “Computational radiomics system to decode the radiographic phenotype,” *Cancer Research*, vol. 77, no. 21, pp. e104–e107, Nov. 2017, doi: 10.1158/0008-5472.CAN-17-0339.
- [35] A. Chu, C. M. Sehgal, and J. F. Greenleaf, “Use of gray value distribution of run lengths for texture analysis,” *Pattern Recognition Letters*, vol. 11, no. 6, pp. 415–419, 1990, doi: [https://doi.org/10.1016/0167-8655\(90\)90112-F](https://doi.org/10.1016/0167-8655(90)90112-F).

- [36] M. M. Galloway, "Texture analysis using gray level run lengths," *Computer Graphics and Image Processing*, vol. 4, no. 2, pp. 172–179, 1975, doi: [https://doi.org/10.1016/S0146-664X\(75\)80008-6](https://doi.org/10.1016/S0146-664X(75)80008-6).
- [37] G. Thibault *et al.*, "Texture Indexes and Gray Level Size Zone Matrix Application to Cell Nuclei Classification Academic Works View project SimulFoot View project Texture Indexes and Gray Level Size Zone Matrix Application to Cell Nuclei Classification," 2009. [Online]. Available: <https://www.researchgate.net/publication/255609273>
- [38] N. v Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: Synthetic Minority Over-sampling Technique," 2002.
- [39] R. Battiti, "Using mutual information for selecting features in supervised neural net learning," *IEEE Transactions on Neural Networks*, vol. 5, no. 4, pp. 537–550, 1994, doi: 10.1109/72.298224.
- [40] J. Bergstra, J. B. Ca, and Y. B. Ca, "Random Search for Hyper-Parameter Optimization Yoshua Bengio," 2012. [Online]. Available: <http://scikit-learn.sourceforge.net>.

Chapter 8: Conclusions and Future Work

This chapter discusses the conclusions and contributions made in this thesis and presents limitations and future directions for this research.

8.1 Summary of Findings

This thesis investigated various image processing techniques for the assessment of cardiomyopathies manifesting with MF. The aim for this thesis was to contribute to the development of image processing pipelines, inclusive of segmentation and classification techniques, to improve the objectivity of MF analysis using CMR imaging. We focused on using a combination of objective techniques, including machine learning, quantitative imaging methods (i.e., T1-mapping and ECV mapping pulse sequences) and normalized MF quantification approaches in order to reduce the subjectivity of MF analysis. Using these strategies, we were able to achieve fully automated myocardial segmentation of three quantitative CMR imaging types, including (1) native T1, (2) post-contrast T1, and (3) ECV mapping. We also validated objective methods for MF quantification using various techniques, such as IIR and fuzzy c-means clustering. Lastly, we created a fully automated pipeline for both myocardial segmentation and disease classification of several cardiomyopathy phenotypes. A breakdown of each chapter's contributions is given below.

8.1.1 Fully automated myocardial segmentation methods for native and post-contrast T1-mapping CMR (Chapter 3)

In Chapter 3, we present two CNN methods based on the U-Net architecture for fully automated myocardial segmentation of native and contrast-enhanced T1-mapping CMR images in patients with three distinct cardiomyopathy states: DCM, HCM and ICM. The first presented method applies the U-Net segmentation model directly to the image of

interest (i.e., the “direct U-Net” method) and performs best on native T1-mapping images. The alternative method (the “CINE-registration” method) uses MIND to register and propagate spatially matched CINE myocardial contours (generated by U-Net) to the image of interest and performs best on contrast-enhanced T1-mapping images, particularly in images exhibiting dense regions of replacement fibrosis. The direct U-Net and CINE-registration methods were compared against manually delineated myocardial contours, wherein we report average DSC metrics of $82.7 \pm 12\%$ and $81.4 \pm 6.9\%$, in native T1-mapping images (P-value<0.0001 by Wilcoxon signed rank test) versus $74.2 \pm 18\%$ and $77 \pm 9.6\%$ in contrast-enhanced T1-mapping images (P-value<0.0014 by Wilcoxon signed rank test) across all patient groups, respectively. Specifically, the average DSC metric increases by 7.7% in the ICM patient group when myocardial contours are segmented using the CINE-registration method. We also present the global native and post-contrast T1 values measured across DCM, HCM and ICM by both the direct U-Net and CINE-registration techniques. High linear correlation of global T1 values was demonstrated by Pearson analysis of the direct U-Net and CINE-registration techniques in native T1 maps ($r = 0.93$, P<0.0001 and $r = 0.87$, P<0.0001, respectively) and contrast enhanced T1 maps ($r = 0.93$, P<0.0001 and $r = 0.98$, P<0.0001, respectively).

8.1.2 Fully automated myocardial segmentation and observer variability analysis of myocardial segmentation in ECV mapping CMR (Chapter 4)

Chapter 4 presents and compares two automated methods for segmentation of the myocardium in ECV mapping CMR, based on (1) the standard U-Net model and (2) the U-Net++ model. Negligible differences in accuracy are demonstrated between U-Net and U-Net++ (87.61% vs. 87.89%, respectively) in patients with DCM and ICM phenotypes.

This study draws attention to the importance of automated myocardial segmentation techniques by demonstrating the impact of user variation in manually delineated myocardial borders on computed ECV values, wherein ECV values differed significantly between three operators (average pairwise $P = 0.004$ by Kruskal-Wallis H test) but not between those operators versus either U-Net or U-Net++ (average pairwise $P = 0.060$ and 0.055 , respectively). In addition, correlation of global ECV values improved for operators versus U-Net and U-Net++ (Spearman $\rho = 0.877$ and 0.88 , respectively) compared to operators versus operators (Spearman $\rho = 0.828$).

8.1.3 Validation of the IIR quantification technique for analysis of the MF progression in the LA following persistent AF in a canine model (Chapter 5)

Chapter 5 explores an objective method based on the IIR for quantification of MF in the LA and aims to corroborate MF progression in canines experiencing sustained AF using histopathological imaging as validation. The IIR + 2 SDs method is found to produce the best MF quantification among other methods, such as STRM and FWHM, and demonstrated that 5 weeks of persistent AF leads to an increase in MF (volumetric %MF increase of $2.11 \pm 0.88\%$ from baseline to post-pacing among experimental animals, $P = 0.019$ by Wilcoxon signed rank test), while control animals demonstrated no significant change in measured %MF from baseline to the 5-weeks post time point ($P = 0.5$ by Wilcoxon signed rank test). The median %MF measured at the post-pacing time point was significantly greater among experimental animals compared to control animals ($2.96 \pm 0.98\%$ vs. $0.95 \pm 0.30\%$, $P = 0.009$ by Mann Whitney U test). We validated the %MF measurements computed by CMR imaging using histopathological imaging, wherein the average slice-wise histopathological %MF was computed at $19.42 \pm 4.8\%$ in experimental

animals versus 1.85% in one control animal. We hypothesize that development of arrhythmia-insensitive CMR imaging and/or automated techniques for LA segmentation may lead to improved MF quantification, and propose additional studies to further corroborate the causality of AF to MF.

8.1.4 Analysis of the effects of spatial resolution on quantification of replacement MF volumes in MOLLI and MCLE T1-mapping CMR (Chapter 6)

In Chapter 6, we explore the effects of diminished spatial resolution on quantification of infarct core and border zone (IC and BZ, respectively) tissue regions using post-contrast T1-mapping CMR in patients with chronic myocardial infarction. We use a fuzzy c-means clustering method to characterize tissue regions, wherein T1-mapping methods MOLLI and MCLE are shown to produce comparable IC masses to those of LGE CMR imaging. This study demonstrated that the BZ tissue index is an appropriate measure for risk stratification of adverse ICD events and risk of SCD, wherein spatial resolution is an important determining factor on the quantification of such tissue zones. The BZ masses computed using the MOLLI pulse sequence (i.e., the sequence with the coarsest spatial resolution) were significantly larger than those computed using MCLE and LGE imaging ($P = 0.0022$ and 0.0003 by student's t-test, respectively). Downsampled MCLE images estimated BZ tissue indices significantly larger than LGE and original MCLE imaging ($P = 0.0033$ and 0.0003 by student's t-test, respectively). For patients that received appropriate ICD therapy (i.e., one or more delivered ICD shock for relevant ventricular arrhythmic events), the BZ masses estimated by original MCLE imaging were significantly larger than the BZ masses estimated in patients who did not receive ICD therapy ($P = 0.044$ by student's t-test). however, when spatial resolution of MCLE images were diminished to

that of MOLLI, the BZ indices between patients receiving vs. not receiving ICD therapy were no longer significant ($P = 0.32$ by student's t-test). Thus, poor spatial resolution leads to overestimation of the BZ tissue index and is shown to decrease the sensitivity in prediction of appropriate ICD therapy.

8.1.5 Classification of five cardiomyopathy phenotypes associated with interstitial MF using native T1 and ECV mapping CMR (Chapter 7)

In Chapter 7, we present two RA-based methods and one CNN-based method for binary and multiclass disease classification in native T1 and ECV mapping CMR imaging, respectively, for patients with amyloidosis, Fabry's disease, HCM, HTN and healthy volunteers. The RA-based ensemble methods (i.e., gradient boosted regression trees (GBRT) and KNN, both stacked with a linear SVC) outperformed the CNN-based method for both binary (healthy vs. rest) and multiclass classification tasks. The GBRT and KNN ensemble models yield classification accuracies of 88% and 89% (respectively) compared to 85% as per the CNN-based technique. Predicted binary classification labels were not significant for RA-based or CNN-based methods against the true class labels ($P = 0.926$ and 0.576 by Kruskal-Wallis test, respectively). For multiclass classification, the GBRT and KNN ensemble models yield classification accuracies of 66% and 58% (respectively) versus 62% by the CNN-based method. However, the AUC metrics computed across all patient groups by RA-based methods are much higher than those computed using the CNN-based classification method. The minority class (Fabry's disease) AUC is computed at 0.652 by the GBRT ensemble method versus 0.648 and 0.5 for the KNN ensemble and CNN methods, respectively. Moreover, predicted multiclass labels output by the RA-based methods were not significantly different than the true class labels ($P = 0.926$, Kruskal-

Wallis test), but statistical significance was demonstrated on predicted labels output by the CNN model versus true class labels ($P = 0.0042$, Kruskal-Wallis test). We hypothesize that RA methods are more appropriate for multiclass classification using ECV mapping images, however, model optimization requires additional training data.

8.2 Limitations and Future Work

The projects presented in this thesis can be expanded through continued research, and several avenues may be approached.

8.2.1 Fully automated myocardial segmentation methods for native and post-contrast T1-mapping CMR (Chapter 3)

The myocardial segmentation approaches presented in Chapter 3 for native and post-contrast T1-mapping were applied to the shMOLLI pulse sequence, which is limited by confounding factors such as noise, transverse relaxation time, off-resonance effect, and magnetic transfer effects [1]– [4]. Therefore, these myocardial segmentation methods should be applied to alternative T1-mapping pulse sequences, such as MOLLI (the most widely used technique) and/or saturation recovery-based mapping techniques such as SASHA or SAPPHIRE (which yield more accurate measures of T1 but lower precision) [2]. Moreover, we recommend that the segmentation methods presented in this chapter be applied to a wider patient population inclusive of additional disease states in order to further elaborate on their feasibility to interstitial vs. replacement MF.

8.2.2 Validation of the IIR quantification technique for analysis of the evolution of MF volumes in the LA following progression AF in a canine model (Chapter 5)

The MF quantification technique presented in Chapter 5 was applied on a small animal dataset inclusive of just twelve canines. Financial constraints as well as the nature

of the research problem put limitations on our ability to apply the methods presented in this chapter to a larger dataset. However, the unique nature of this study (wherein CMR imaging was acquired at serial time points pre- and post- AF, and histopathological imaging after sacrifice of the animals and excision of the heart) makes application of these methods to a larger dataset more challenging. However, further corroboration to the causality of AF to MF may be established by exploring MF segmentation on the existing canine dataset using pseudo-quantitative imaging techniques, such as the T1-refBlochi method, which transforms a 3D LGE imaging volume to 3D T1 map using Bloch equation modeling of the LGE signal and a single calibration point [5].

8.2.3 Analysis of the effects of spatial resolution on quantification of replacement MF volumes in MOLLI and MCLE T1-mapping CMR (Chapter 6)

In chapter 6, we compare the IC and BZ tissue zones and their sensitivity in predicting ICD therapy using fuzzy c-means (for post-contrast MOLLI and MCLE T1-mapping techniques) and FWHM (for LGE imaging). However, quantification approaches were not explored among native T1-mapping nor ECV-mapping across various pulse sequences. This could be particularly advantageous in ECV-mapping CMR, which is demonstrated to show improved measurement reproducibility compared to T1-mapping [4],[6]. This could be approached in future work. Moreover, there is value in comparing the effects of manual myocardial border delineation on computation of IC and BZ volumes. This could be approached using inter- and intra- observer variability by comparing the IC and BZ volumes computed using myocardial borders delineated by multiple users, as researched in Chapter 4. Furthermore, automated myocardial segmentation techniques

(e.g., U-Net or U-Net++ methods) could be explored for feasibility in this patient group with chronic MI, and variability metrics could be further analysed.

8.2.4 Fully automated pipeline for myocardial segmentation and classification of cardiomyopathy phenotypes associated with interstitial MF using native T1 and ECV mapping CMR (Chapters 4 and 7)

A U-Net method was used to achieve myocardial segmentation in CMR images for patients with DCM and ICM cardiopathies (Chapter 4, ECV mapping) and amyloidosis, Fabry's disease, HCM, HTN and healthy volunteers (Chapter 7, native T1-mapping and ECV mapping). Chapter 4 was focussed solely on the myocardial segmentation step, while classification was approached in Chapter 7. However, transfer-learning allowed for the U-Net model developed in Chapter 4 to be used towards the segmentation step in Chapter 7, which was a required component in order to achieve RA-based classification. As with any deep learning problem, a sufficient amount of data is needed to optimize the segmentation and/or classification model for the best performance. Thus, the main limitation in both of these chapters is a limited number of images. This was addressed by using data augmentation, including image augmentation and SMOTE; however, segmentation and classification accuracy could be further improved through acquisition of additional imaging data. In addition, improving the class imbalance could further facilitate improved accuracy. Thus, the works presented in Chapters 4 and 7 could be expanded through training of segmentation and classification models with larger datasets, thereby demonstrating the practicality of applying such techniques to larger patient populations across multiple cardiomyopathy states.

8.3 Conclusion

This thesis presents tools to evaluate cardiovascular diseases that present with MF using CMR imaging. The works presented in this thesis demonstrate the applicability of machine/deep learning segmentation techniques and normalized MF quantification methodologies that are independent on user-defined reference regions toward appropriate segmentation of the myocardial region and MF tissue zones. With recent advancements geared towards the use of user-independent and machine learning-based algorithms for segmentation and classification in clinical cardiac imaging software, the methods presented in this thesis are highly translatable to clinical practice.

8.4 Contributions

8.5 Chapter 8 References

- [1] J. F. Heidenreich *et al.*, “T1- and ECV-mapping in clinical routine at 3 T: Differences between MOLLI, ShMOLLI and SASHA,” *BMC Medical Imaging*, vol. 19, no. 1, pp. 1–9, Aug. 2019, doi: 10.1186/S12880-019-0362-0/FIGURES/5.
- [2] A. J. Taylor, M. Salerno, R. Dharmakumar, and M. Jerosch-Herold, “T1 Mapping Basic Techniques and Clinical Applications,” *JACC: Cardiovascular Imaging*, vol. 9, no. 1, pp. 67–81, 2016, doi: 10.1016/j.jcmg.2015.11.005.
- [3] S. K. Piechnik *et al.*, “Shortened Modified Look-Locker Inversion recovery (ShMOLLI) for clinical myocardial T1- mapping at 1.5 and 3 T within a 9 heartbeat breathhold,” pp. 1–11, 2010.
- [4] S. Roujol *et al.*, “Accuracy, precision, and reproducibility of four T1 mapping sequences: a head-to-head comparison of MOLLI, ShMOLLI, SASHA, and SAPPHIRE.,” *Radiology*, vol. 272, no. 3, pp. 683–9, 2014, doi: 10.1148/radiol.14140296.
- [5] C. Hu *et al.*, “T1-refBlochi: high resolution 3D post- contrast T1 myocardial mapping based on a single 3D late gadolinium enhancement volume, Bloch equations, and a reference T1,” pp. 1–17, 2017, doi: 10.1186/s12968-017-0375-1.
- [6] N. Child *et al.*, “Comparison of MOLLI, shMOLLI, and SASHA in discrimination between health and disease and relationship with histologically derived collagen volume fraction,” *European heart journal. Cardiovascular Imaging*, vol. 19, no. 7, pp. 768–776, Jul. 2018, doi: 10.1093/EHJCI/JEX309.

Appendices

Appendix A Semi-automated Segmentation of the Left Ventricular Myocardium in Native T1-mapping CMR Using Deformable Non-Rigid Registration of CINE Images

In this approach, we achieve semi-automated segmentation of the left ventricle in native T1-maps using an indirect approach based on registration of CINE contours. The same image processing pipeline discussed in Chapter 3 is followed, but the CINE myocardial segmentation step is achieved using a deformable model as opposed to the U-Net (see Figure 3-2).

The contents of this Appendix are reproduced in part from our publication in the 2019 Society of Optics and Photonics Conference on Medical Imaging, i.e.,

N. A. Farrag, J. A. White, and E. Ukwatta, “Semi-automated myocardial segmentation in native T1-mapping CMR using deformable non-rigid registration of CINE images,” *Proc. SPIE 10953*, vol. Biomedical, no. March, 2019.

A.1 Method

Preprocessing steps:

This approach follows the same pre-processing steps as described in Chapter 3 (i.e., interpolation of CINE images to the spatial resolution of T1-maps). Twenty (20) patients with ICM were used toward this analysis. Image resizing of CINE images and T1-maps (acquired using the shMOLLI sequence) were achieved using a semi-automated technique based on selection of the LV chamber centroid.

Semi-automated myocardial segmentation of CINE images:

In this approach, we implemented a semi-automated technique to find the myocardial borders of CINE images using Medviso Segment version 2.2; an open-source CMR image analysis software [1],[2]. We implemented the alternative automatic LV segmentation algorithm [2] using a modified approach for single-slice segmentation, as the algorithm is typically used for multi-slice pulse acquisitions. The semi-automated algorithm uses an 8-step framework based on a deformable model for semi-automatic delineation of the endocardial and epicardial boundaries [2].

The myocardial segmentation found using Segment v 2.2's algorithm is created based on endocardial and epicardial balloon forces. The balloon forces are derived using an expectation maximization (EM) algorithm, which estimates the assumed Gaussian distribution of intensities for blood, myocardium and surrounding tissues. The endocardial balloon image is calculated as the Gaussian distribution of blood divided by the sum of Gaussian distributions of blood and myocardium. The epicardial balloon image is calculated as the Gaussian distribution of myocardium divided by the sum of the Gaussian distributions of blood, myocardium, and surrounding tissues.

Endocardial and epicardial initializations are required for the deformable model to deform to the myocardial contours. The endocardial initialization is originally estimated at the mid-mural center line. The epicardium is initialized by thresholding the endocardial balloon image at zero to define the blood-pool region, which is then used along with the estimated LV wall thickness to create the final epicardial and endocardial initializations. Following initialization, the endocardial boundary is then defined using a deformable model with endocardial initialization, the endocardial balloon force and the weighting of

forces optimized for endocardial segmentation. Likewise, the epicardial boundary is defined using the deformable model with epicardial initialization, the epicardial balloon force, and weighting of forces optimized for epicardial segmentation. The myocardial segmentation may be refined to exclude any papillary muscles that are included or closely attached to the LV wall.

The CINE and shMOLLI image sequences were acquired at different points in time. Therefore, in order to take into account inadvertent patient motion from one imaging sequence to the next, the CINE myocardial contours were not directly applied to shMOLLI images. We implemented the MIND [3] registration method (described in Chapter 3) to deform CINE contours to native T1-maps in order to account for inadvertent patient motion and/or potential differences in image acquisition location from one imaging sequence to the next. In our implementation of MIND registration, we chose to use a single phase of the raw shMOLLI acquisition that provided the best overall image contrast (typically the 6th or 7th phase) for registration, as opposed to registering CINE with the T1-map itself. We define the shMOLLI phase image as the target and CINE as the moving image.

A.2 Results

We evaluated our semi-automatic myocardial T1-mapping image segmentation technique by comparing its results to manual (ground-truth) segmentations performed directly on the shMOLLI maps. We also evaluated the performance of two other registration-based segmentation techniques for comparison, mutual information (MI) and sum of square differences (SSD), where contours from CINE images were transformed and applied to shMOLLI maps and again compared to manual ground truth segmentations of the shMOLLI maps. Table Appendix A-1 shows a summary of the Sorensen-Dice Coefficient

(DSC), precision, recall and relative error between shMOLLI ground truth contours and those created by propagation of CINE contours via MIND-based registration (our technique) in the first column. Additionally, we demonstrate the results achieved by comparing shMOLLI ground truth contours against the original (non-registered) CINE contours, and those created by propagating CINE contours to shMOLLI maps using SSD-based and MI-based registration in rows 2-4, respectively. In native shMOLLI images, our non-rigid image registration technique improved the Dice index by 11.74% compared to the DSC computed without registration. Our technique yielded significantly better results compared to the MI-based and SSD-based registration techniques, with which results of any metric did not exceed 60.24%.

Table Appendix A-1 Results of myocardial segmentation using semi-automated CINE-registration technique

	CINE to shMOLLI: after propagation based on MIND registration	CINE to shMOLLI: without registration	CINE to shMOLLI: after SSD based registration	CINE to shMOLLI: after MI based registration
DSC (%)	84.36% ± 4.03%	72.62% ± 10.62%	47.74% ± 18.16%	60.24% ± 11.62%
Precision (%)	91.68% ± 7.89%	71.37% ± 11.79%	47.77% ± 21.60%	54.71% ± 14.10%
Recall (%)	91.33% ± 8.41%	71.65% ± 11.53%	54.75% ± 20.75%	58.54% ± 14.09%
Rel. error (%)	16.29% ± 8.58%	15.90% ± 8.653%	35.70% ± 23.90%	23.97% ± 11.78%

Figure Appendix A-1 demonstrates a series of CINE and shMOLLI image pairs along with their corresponding manual contours, motion vector corrections generated by MIND registration, and the resultant contours on native shMOLLI maps

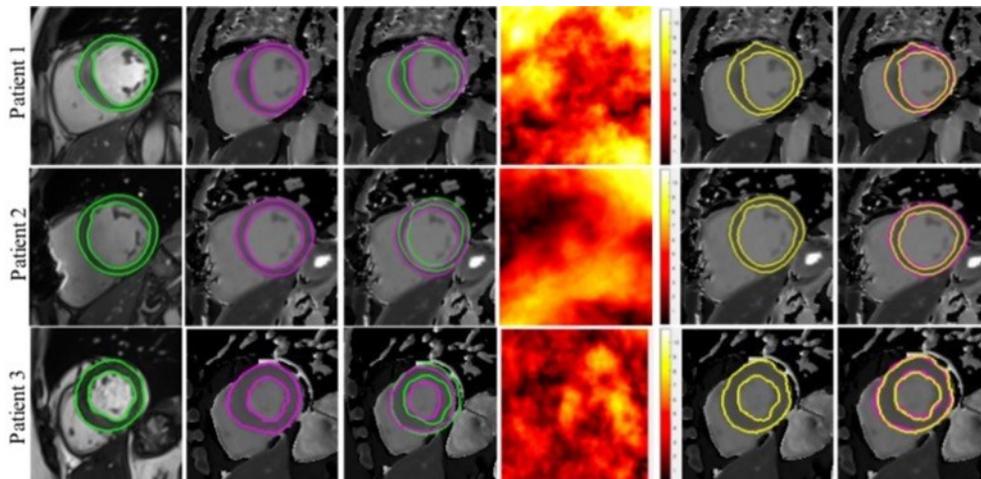


Figure Appendix A-1 Myocardial segmentation results from CINE-registration method using contours derived by deformable model

MIND registration results are demonstrated for three example patients. In each row, the first image demonstrates the contour generated by the LV Automatic Segmentation algorithm in Segment v. 2.2 in the color green. The second image demonstrates the manually drawn ground truth myocardial segmentation of the T1 map in magenta, while the third image overlays the two, demonstrating the extent of segmental mismatch. The fourth image demonstrates the magnitude of motion vectors in mm resulting from MIND registration of the respective CINE to shMOLLI image. The fifth image demonstrates the resultant segmentation generated by vector propagation of the CINE contour (magenta), shown in yellow color. Lastly, the final image demonstrates the extent of segmental mismatch of the manual T1 map contour (magenta) vs. the contour generated by MIND propagation of CINE contours (yellow).

A.3 Discussion & Conclusions

In this study, we developed an image processing pipeline for semi-automated segmentation of the myocardium in native T1-mapping CMR images using epicardial and endocardial contours obtained from CINE images. Our technique takes advantage of the superior image contrast of CINE images, allowing for the semi-automatic detection of myocardial contours that can then be propagated to T1-mapping images based on localized structural differences, as derived from the MIND registration algorithm. Our technique yielded superior accuracy (DSC 84.36%) in segmentation of the myocardium in native T1-maps compared to intensity-based image registration techniques, such as SSD and MI, which produced poor segmentation results wherein the same accuracy metrics fell at or below 60.24%.

The main limitation of this technique is that it is not fully automated, as it relies on some pre-contouring of the CINE images (herein described using a semi-automated approach). However, the approach used here may be beneficial in cases where automated myocardial segmentation from CINE images fails or is suboptimal. Accordingly, our described exploitation of these contours to provide automation for T1-mapping analysis has strong potential for clinical translation.

A.4 Appendix A References

- [1] Heiberg, E., Sjögren, J., Ugander, M., Carlsson, M., Engblom, H. and Arheden, H. “Design and validation of Segment - freely available software for cardiovascular image analysis,” *BMC Medical Imaging*, 10(1) pp. 1–13, (2010).
- [2] Tufvesson, J., Hedström, E., Steding-ehrenborg, K., Carlsson, M., Arheden, H. and Heiberg, E. “Validation and Development of a New Automatic Algorithm for Time-Resolved Segmentation of the Left Ventricle in Magnetic Resonance Imaging,” *BioMed Research International*, vol. 2015, (2015).
- [3] Heinrich M. P., “MIND: Modality independent neighbourhood descriptor for multi-modal deformable registration,” *Med. Image Anal.*, vol. 16, no. 7, pp. 1423–1435 (2012).

Appendix B Inter/Intra Observer Variability of Left Atrial Myocardial Fibrosis Volumes

The contents of this Appendix relate to Chapter 4 of this document, section 4.2.3. In our work on the assessment of MF volumes in the LA myocardium following atrial fibrillation, an inter- and intra-observer analysis was performed wherein two users computed several segmentation evaluation metrics among MF volumes segmented three separate times per observer. In Chapter 4, section 4.3.5, we report observer variability results for the DSC metric. Results for boundary F1 (BF) score, precision, recall and relative error index are herein reported below.

B.1 Boundary F1 Score Results

MF volumes were segmented as per our pipeline (see Chapter 5, Figure 5-2) by two operators for the inter/intra- operator analysis in a sub-set of three animals (2 experimental, 1 control). The boundary F1 (BF) score was computed on segmentations of the myocardial blood pool chamber, myocardial LA wall and MF volumes across both operators, as demonstrated in Table B1. Given two segmentations A and B , the boundary F1 score BF is given as follows:

$$BF = 2 \frac{Precision \times Recall}{Precision + Recall} \quad \text{Equation Appendix B-1}$$

The *Precision* and *Recall* metrics are computed based on object boundaries (i.e., edges) as opposed to the entire foreground regions. The BF score shows how much the boundary of segmentation B overlaps with the boundary of segmentation A . A distance tolerance of 2 voxels is used in the calculation of this measure, i.e., that the boundary of region B is considered to be a true positive if it is within two pixels of a boundary of region A . The BF score is given as a percentage, wherein 0% BF indicates no overlap, and 100%

BF indicates perfect overlap, respectively, between the boundaries of segmentations *A* and *B*. The BF score is not equivalent to the Dice Similarity Coefficient (DSC), which computes similarity based on the entire region, not just the edges.

Table Appendix B-1 Observer variability results of boundary F1 (BF) score

Table B1							
Average boundary F1 (BF) score computed on three segmentation attempts performed by two operators across each time point in a subsample of three animals. The BF score is given as a percentage and is expressed as mean \pm SD.							
		Paced dog 1		Paced dog 3		Control dog 3	
	Region	Baseline	Post-pacing	Baseline	Post-pacing	Baseline	Post-pacing
Operator 1	BF Blood pool (%)	97.37 \pm 2.82	95.95 \pm 2.46	99.88 \pm 0.11	99.14 \pm 0.35	99.65 \pm 0.08	99.88 \pm 0.15
	BF LA wall (%)	97.46 \pm 2.11	92.06 \pm 3.03	98.82 \pm 0.78	98.27 \pm 0.85	98.92 \pm 0.88	99.44 \pm 0.42
	BF MF volume (%)	53.57 \pm 31.19	66.72 \pm 23.70	22.97 \pm 21.15	57.11 \pm 34.17	59.31 \pm 25.13	48.77 \pm 34.89
Operator 2	BF Blood pool (%)	98.91 \pm 0.73	99.52 \pm 0.51	99.37 \pm 1.00	98.46 \pm 1.21	99.96 \pm 0.05	99.94 \pm 0.06
	BF LA wall (%)	98.78 \pm 1.02	99.35 \pm 0.18	99.67 \pm 0.27	98.84 \pm 1.01	99.86 \pm 0.14	99.89 \pm 0.08
	BF MF volume (%)	79.81 \pm 15.42	87.31 \pm 5.08	80.61 \pm 3.34	59.41 \pm 27.35	84.64 \pm 2.02	94.22 \pm 0.91
Operator 1 vs. Operator 2	BF Blood pool (%)	97.95 \pm 1.22	96.74 \pm 2.89	99.26 \pm 0.65	96.77 \pm 1.26	98.08 \pm 0.41	99.33 \pm 0.24
	BF LA wall (%)	98.05 \pm 0.92	95.30 \pm 3.47	98.83 \pm 0.59	97.06 \pm 1.45	97.86 \pm 1.09	99.29 \pm 0.28
	BF MF volume (%)	87.09 \pm 15.15	79.36 \pm 20.57	29.77 \pm 12.22	30.59 \pm 24.45	49.43 \pm 16.84	59.39 \pm 28.77

B.2 Precision and Recall Results

Precision and recall metrics computed on the blood pool chamber, LA myocardial wall and MF volumes are given in Tables B2 and B3, respectively.

Table Appendix B-2 Observer variability results of precision

Table B2							
Average precision computed across three segmentation attempts performed by two operators at each time point in a subsample of three animals. Precision is given as a percentage and is expressed as mean \pm SD.							
		Paced dog 1		Paced dog 3		Control dog 3	
	Region	Baseline	Post-pacing	Baseline	Post-pacing	Baseline	Post-pacing
Operator 1	Precision Blood pool (%)	99.44 \pm 0.38	94.82 \pm 3.40	99.93 \pm 0.01	99.07 \pm 0.43	100 \pm 0.43	99.98 \pm 0.03
	Precision LA wall (%)	99.07 \pm 0.82	90.22 \pm 7.68	98.20 \pm 1.59	97.89 \pm 2.35	97.95 \pm 1.72	99.83 \pm 0.28

	Precision MF volume (%)	70.40 ± 28.10	60.81 ± 32.76	61.07 ± 27.43	97.70 ± 2.85	95.08 ± 5.82	94.51 ± 9.50
Operator 2	Precision Blood pool (%)	99.78 ± 0.06	99.39 ± 0.70	100 ± 0	98.11 ± 1.69	99.92 ± 0.09	99.97 ± 0.04
	Precision LA wall (%)	97.87 ± 1.87	99.17 ± 0.18	99.97 ± 0.03	98.69 ± 1.13	99.92 ± 0.07	99.97 ± 0.03
	Precision MF volume (%)	72.91 ± 22.87	82.10 ± 8.27	74.96 ± 0.82	53.88 ± 27.67	88.49 ± 4.38	92.88 ± 3.78
Operator 1 vs. Operator 2	Precision Blood pool (%)	99.26 ± 0.37	95.78 ± 3.55	99.72 ± 0.40	97.83 ± 2.16	95.99 ± 0.76	99.64 ± 0.47
	Precision LA wall (%)	99.00 ± 0.91	93.51 ± 5.21	98.76 ± 1.13	95.92 ± 2.35	97.16 ± 1.67	99.63 ± 0.48
	Precision MF volume (%)	92.33 ± 10.50	73.36 ± 24.00	60.63 ± 28.43	59.47 ± 34.99	69.73 ± 27.69	81.40 ± 13.86

Table Appendix B-3 Observer variability results of precision

Table B3							
Average recall computed on three segmentation attempts performed by two operators across each time point in a subsample of three animals. Recall is given as a percentage and is expressed as mean ± SD.							
		Paced dog 1		Paced dog 3		Control dog 3	
Region		Baseline	Post-pacing	Baseline	Post-pacing	Baseline	Post-pacing
Operator 1	Recall Blood pool (%)	95.40 ± 5.29	97.13 ± 1.59	99.83 ± 0.20	99.21 ± 0.34	99.74 ± 0.27	99.79 ± 0.32
	Recall LA wall (%)	95.92 ± 3.34	94.32 ± 2.14	99.45 ± 0.37	98.70 ± 1.52	99.93 ± 0.01	99.06 ± 0.58
	Recall MF volume (%)	47.60 ± 41.09	79.19 ± 9.76	18.94 ± 22.57	48.02 ± 42.63	48.66 ± 33.74	42.26 ± 42.99
Operator 2	Recall Blood pool (%)	98.06 ± 1.45	99.65 ± 0.33	98.75 ± 1.97	98.83 ± 1.17	100 ± 0.01	99.92 ± 0.08
	Recall LA wall (%)	99.73 ± 0.28	99.53 ± 0.18	99.37 ± 0.52	98.99 ± 1.14	99.80 ± 0.20	99.81 ± 0.15
	Recall MF volume (%)	90.42 ± 7.11	93.59 ± 4.12	87.49 ± 8.45	67.15 ± 25.23	81.62 ± 7.60	95.83 ± 4.47
Operator 1 vs. Operator 2	Recall Blood pool (%)	96.69 ± 2.20	97.73 ± 2.22	98.81 ± 1.32	95.84 ± 3.25	96.69 ± 2.20	97.73 ± 2.22
	Recall LA wall (%)	97.12 ± 1.51	93.51 ± 1.94	98.92 ± 0.65	98.26 ± 1.30	97.12 ± 1.51	97.23 ± 1.94
	Recall MF volume (%)	83.49 ± 19.47	88.70 ± 12.82	21.72 ± 10.62	21.79 ± 19.22	83.49 ± 19.47	88.70 ± 12.82

B.3 Relative Error Index

metrics computed on the blood pool chamber, LA myocardial wall and MF volumes are expressed in Table B4.

Table Appendix B-4 Observer variability results of relative error index

Table B4							
Average relative error (RE) index computed on blood pool, LA wall and MF volumes across three segmentation attempts performed by two operators at each time point in a subsample of three animals. The RE is expressed as a percentage (mean \pm SD).							
		Paced dog 1		Paced dog 3		Control dog 3	
	Region	Baseline	Post-pacing	Baseline	Post-pacing	Baseline	Post-pacing
Operator 1	RE Blood pool (%)	29.06 \pm 22.82	19.05 \pm 12.36	6.00 \pm 2.59	4.41 \pm 2.46	14.84 \pm 11.86	3.13 \pm 1.58
	RE LA wall (%)	23.02 \pm 17.33	19.68 \pm 12.89	10.99 \pm 4.56	12.78 \pm 5.50	18.98 \pm 9.73	3.15 \pm 2.19
	RE MF volume (%)	18.87 \pm 7.82	33.57 \pm 16.64	81.40 \pm 31.80	72.69 \pm 42.73	47.25 \pm 48.37	32.12 \pm 54.57
Operator 2	RE Blood pool (%)	5.03 \pm 3.32	16.14 \pm 7.11	4.67 \pm 1.97	7.20 \pm 3.35	3.89 \pm 2.63	4.60 \pm 3.66
	RE LA wall (%)	9.28 \pm 3.98	3.27 \pm 1.78	16.00 \pm 6.91	3.62 \pm 1.82	4.06 \pm 1.75	8.11 \pm 5.57
	RE MF volume (%)	21.94 \pm 8.80	13.62 \pm 7.15	31.52 \pm 12.17	42.82 \pm 13.82	5.73 \pm 2.92	31.03 \pm 14.29
Operator 1 vs. Operator 2	RE Blood pool (%)	20.86 \pm 21.70	15.75 \pm 9.72	10.84 \pm 4.56	5.60 \pm 3.41	12.89 \pm 4.93	20.74 \pm 3.25
	RE LA wall (%)	17.82 \pm 14.32	14.00 \pm 10.78	28.54 \pm 9.79	29.68 \pm 6.51	20.37 \pm 12.04	18.61 \pm 5.17
	RE MF volume (%)	23.43 \pm 13.22	25.40 \pm 16.38	89.49 \pm 8.26	72.80 \pm 23.76	63.57 \pm 29.13	47.57 \pm 43.24