

LB/TH/41/2025
TH6006

**DEEP LEARNING BASED U-NET VARIANTS FOR
CARDIAC MRI SEGMENTATION**

C. B. Wijesinghe

239375E

Master of Science in Computer Science & Engineering

Department of Computer Science & Engineering
Faculty of Engineering

University of Moratuwa
Sri Lanka

May 2025

DEEP LEARNING BASED U-NET VARIANTS FOR CARDIAC MRI SEGMENTATION

C. B. Wijesinghe

239375E

Thesis submitted in partial fulfillment of the requirements for the degree
Master of Science in Computer Science & Engineering

Department of Computer Science & Engineering
Faculty of Engineering

University of Moratuwa
Sri Lanka

May 2025

DECLARATION

I declare that this is my own work and this Thesis does not incorporate without acknowledgement any material previously submitted for a Degree or Diploma in any other University or Institute of higher learning and to the best of my knowledge and belief it does not contain any material previously published or written by another person except where the acknowledgement is made in the text. I retain the right to use this content in whole or part in future works (such as articles or books).

Signature:

Date: 2025/05/21

The supervisor should certify the Thesis with the following declaration.

The above candidate has carried out research for the Master of Science in Computer Science & Engineering Thesis under my supervision. I confirm that the declaration made above by the student is true and correct.

Name of Supervisor: Prof. Dulani Meedeniya

Signature of the Supervisor:

Date: 22/ 05/ 2025

DEDICATION

I dedicate this thesis report to the unwavering support and boundless love of my parents, whose encouragement and belief in my abilities have been the foundation of my academic journey. Their sacrifices and dedication to my education have shaped me into the person I am today.

To the University of Moratuwa, my academic home since my undergraduate years, I extend my gratitude for providing an environment that fosters intellectual growth and innovation. The knowledge and skills I have gained here have been instrumental in undertaking this research endeavor.

I express my deepest appreciation to my supervisor, Prof. Dulani Meedeniya, for her invaluable guidance, mentorship, and unwavering support throughout the research process. Her expertise and encouragement have been a guiding light, propelling me forward in my academic endeavors.

I extend heartfelt thanks to the hardworking Cardiologists and Radiologists whose dedication to patient care inspired this research. Their expertise and collaborative spirit have not only elevated the quality of this study, but also emphasized the importance of bridging the gap between medical practitioners and researchers in the pursuit of improved diagnostic tools and methodologies.

I am also grateful to the contributors of publicly available datasets, whose commitment to advancing research by sharing valuable resources has significantly contributed to the success of this research. Their generosity has broadened the scope of this study, enabling a more comprehensive analysis and understanding of cardiac MRI segmentation of ventricular structures and myocardium.

This work is dedicated to all those who have played a role, big or small, in shaping this academic endeavor. Your support and contributions have been instrumental in bringing this thesis to fruition.

ACKNOWLEDGEMENT

I would like to express my sincere gratitude to my supervisor, Prof. Dulani Meedeniya, whose unwavering support and guidance have been instrumental in the successful completion of this research. Her expertise, encouragement, and valuable insights have significantly contributed to the development and refinement of my research in the field of Cardiac MRI segmentation of ventricular structures and Myocardium.

I would like to express my sincere gratitude to Mr. Dharshana Muthtettugoda for his invaluable support in facilitating access to the High Performance Computing (HPC) server and for his continuous assistance in troubleshooting connectivity issues throughout the course of this research. I also extend my sincere gratitude to the Department of Computer Science and Engineering for granting me access to the HPC resources, which were crucial for the successful completion of my work.

I extend my heartfelt appreciation to the contributors of related studies, whose groundbreaking work has laid the foundation for my research. Their pioneering efforts have enriched my understanding and provided a robust framework for the exploration of cardiac MRI segmentation.

I am also grateful to those who generously made publicly available datasets, without which the empirical validation of my research would not have been possible. Their commitment to advancing scientific knowledge by sharing resources has been a vital aspect of my research journey.

I would like to acknowledge the invaluable feedback and support from my colleagues and peers. Their constructive criticisms, discussions, and encouragement have played a pivotal role in shaping the direction of my research and refining its methodologies.

Last but not least, I want to express my deepest gratitude to my parents for their unwavering support, understanding, and encouragement throughout this academic journey. Their love and encouragement have been a constant source of inspiration, motivating me to strive for excellence.

This thesis represents the culmination of the collective efforts and support from these individuals, and I am truly grateful for their contributions to the successful completion of this research project.

ABSTRACT

Accurate segmentation of ventricular structures and the myocardium from Cardiac Magnetic Resonance (CMR) images is essential for the diagnosis and management of cardiovascular diseases. This study presents a comprehensive approach to cardiac MRI segmentation by developing and evaluating six U-Net variants: Original U-Net, Residual U-Net, Attention U-Net, Feature Pyramid U-Net, Feedback Residual U-Net, and Transformer-Based U-Net, each incorporating architectural enhancements tailored to address specific challenges in segmenting complex cardiac anatomy. These architectures incorporate advanced enhancements such as deeper encoder levels, attention mechanisms, residual connections, multi-scale feature fusion, transformer modules, and feedback mechanisms. To improve segmentation robustness, a novel hybrid loss function, combining Dice Loss and Cross-Entropy Loss, was proposed to effectively manage class imbalance and improve segmentation precision. Among the evaluated models, the Feature Pyramid U-Net achieved the highest performance, with Dice coefficients of 0.9388 (Left Ventricle), 0.8759 (Right Ventricle), and 0.8426 (Myocardium), demonstrating its superior ability to capture multi-scale contextual information. To bridge the gap between research and clinical application, an interactive web application was developed and deployed, enabling real-time inference, visual inspection of annotated segmentations, and region-specific descriptions through a user-friendly interface. This work not only advances the design of deep learning architectures for medical image segmentation, but also demonstrates a practical pathway for integrating these models into clinical workflows.

Keywords: Cardiac MRI, Segmentation, U-Net

TABLE OF CONTENTS

Declaration of the Candidate & Supervisor	i
Dedication	ii
Acknowledgement	iii
Abstract	iv
Table of Contents	v
List of Figures	viii
List of Tables	x
List of Abbreviations	x
List of Appendices	xii
1 Introduction	1
1.1 Background	1
1.2 Research Problem	2
1.3 Motivation and Novelty	4
1.4 Research Objectives	5
1.5 Research Statement	5
1.6 Research Questions	5
1.7 Thesis Structure	6
2 Literature Review	8
2.1 Overview of CMRI Image Segmentation	8
2.2 Current Clinical Practices	9
2.3 Publicly Available Datasets	11
2.4 Data Pre-Processing Techniques	11
2.5 Traditional CMRI Segmentation Methods	13
2.5.1 Thresholding	13
2.5.2 Region-Growing	13
2.5.3 Pixel Classification	14
2.5.4 Deformable Methods	15

2.5.5	Atlas-Based Methods	16
2.5.6	Statistical Shape Models (SSM)	16
2.6	Deep Learning Based CMRI Segmentation Methods	17
2.6.1	Convolutional Neural Network (CNN) Based Methods	18
2.6.2	Fully Convolutional Neural Network (FCN) Based Methods	19
2.6.3	Recurrent Neural Network Based Methods	20
2.6.4	Generative Adversarial Network Based Methods	21
2.6.5	U-Net Based Methods	22
2.6.6	Attention Based Methods	26
2.6.7	State-of-the-Art Model	29
2.7	Comparison of Related Studies	31
2.7.1	Traditional Methods vs. Deep Learning Methods	31
2.7.2	Comparison of Techniques Used for CMRI Segmentation	33
2.7.3	Comparison of State-of-the-Art Segmentation Methods Evaluated on ACDC Dataset	33
2.8	Evaluation Metrics	33
2.9	Limitations and Challenges in Existing Methods	37
3	Methodology	38
3.1	Process Flow	38
3.2	Dataset	38
3.3	Data Pre-Processing	41
3.4	U-Net Based Variants	42
3.4.1	Original U-Net (O-UN)	44
3.4.2	Residual U-Net (Res-UN)	46
3.4.3	Attention U-Net (Atn-UN)	48
3.4.4	Feature Pyramid U-Net (FP-UN)	49
3.4.5	Feedback Residual U-Net (Feed-Res-UN)	52
3.4.6	Transformer-Based U-Net (Trans-UN)	54
3.5	Loss Function	55
3.6	Experimental Setup	56
3.7	Web Application Development	57

4	Results	60
4.1	Evaluation Metrics	60
4.2	ACDC Test Set Performance	60
4.3	Segmentation Results Analysis	64
4.4	Dice Score and Loss Graphs	68
5	Discussion	73
5.1	Study Contribution	73
5.1.1	Achieving Research Objectives and Research Questions	76
5.2	Comparison with Existing Studies	77
5.3	Challenges and Limitations	80
5.3.1	Computational Resource Limitations	81
5.3.2	Platform Constraints	81
5.3.3	Changes in External Infrastructure	81
5.3.4	Challenges in Model Complexity and Training	82
5.4	Future Work	82
6	Conclusion	85
	References	86
	Appendix A Performance Evaluation of DL-Based Studies	94
	Appendix B Publications	99

LIST OF FIGURES

Figure	Description	Page
Figure 1.1	3D anatomical structure of the cardiac	2
Figure 1.2	Applications of imaging modalities. Adapted from [1]	3
Figure 2.1	Number of DL-based papers published from January 1, 2019 to December 31, 2024, related to Cardiac MR, CT and US segmentation.	9
Figure 2.2	An application of thresholding for LV endocardium segmentation. Adapted from [2]	14
Figure 2.3	Pixel classification using GMM: (a) the input image; (b) GMM with 3 components; (c) the output image with pixel classification. Adapted from [3]	15
Figure 2.4	Architecture of a CNN. Adapted from [1]	18
Figure 2.5	CNN for patch-based segmentation. Adapted from [1]	19
Figure 2.6	Architecture diagram of a Fully Convolutional Neural Network (FCN). Adapted from [1]	20
Figure 2.7	Proposed FCN architecture for initial segmentation. Adapted from [4]	20
Figure 2.8	The architecture of TCN. Adapted from [5]	21
Figure 2.9	The architecture of GAN. Adapted from [1]	22
Figure 2.10	The U-Net architecture. Adapted from [6]	23
Figure 2.11	The Feedback U-Net architecture. Adapted from [7]	24
Figure 2.12	The proposed workflow for CNN + U-Net based composite model. Adapted from [8]	24
Figure 2.13	The proposed initial segmentation network: Res U-Net. Adapted from [5]	24
Figure 2.14	The architecture of the proposed MTL-UNet. Adapted from [9]	25
Figure 2.15	The end-to-end workflow of the AI-based framework. Adapted from [10]	26
Figure 2.16	The Attention U-Net model. Adapted from [11]	27
Figure 2.17	The Shape Attentive U-Net (SAUNet) model. Adapted from [12]	28
Figure 2.18	The Attention U-Net model with input image pyramid and deep supervised output layers. Adapted from [13]	28
Figure 2.19	The architecture of the 2D ARW-Net model. Adapted from [14]	29
Figure 2.20	The architecture of the FPN. Adapted from [15]	30
Figure 3.1	Overall Process of Cardiac MRI Segmentation.	39
Figure 3.2	2D CMR image and its ground truth from the ACDC dataset.	40
Figure 3.3	The structure of the ACDC dataset.	41
Figure 3.4	The structure of the processed ACDC dataset.	43
Figure 3.5	Image and Mask Augmentation.	44

Figure 3.6	The Original U-Net Architecture.	45
Figure 3.7	The Residual U-Net Architecture. Adapted from [16].	47
Figure 3.8	The Attention Block. Adapted from [11].	49
Figure 3.9	The Attention U-Net Architecture	49
Figure 3.10	The Feature Pyramid U-Net Architecture	50
Figure 3.11	The Feature Pyramid Block	50
Figure 3.12	The Feedback Res U-Net Architecture	53
Figure 3.13	The Transformer-Based U-Net Architecture	54
Figure 3.14	The Web Application	58
Figure 3.15	Sample Image Segmented Using the Web Application	58
Figure 4.1	Model Predictions for Sample Image - I	65
Figure 4.2	Model Predictions for Sample Image - II	66
Figure 4.3	Model Predictions for Sample Image - III	66
Figure 4.4	Model Predictions for Sample Image - IV	67
Figure 4.5	Dice Score and Loss of U-Net Variants	68
Figure 4.6	Training Hybrid Loss of U-Net Variants	71

LIST OF TABLES

Table	Description	Page
Table 2.1	Public Datasets for Cardiac Magnetic Resonance Imaging (CMRI) Segmentation.	11
Table 2.2	Comparison of traditional methods used for CMRI segmentation.	34
Table 2.3	Segmentation accuracy (Dice) of SOTA methods evaluated on ACDC dataset.	35
Table 2.4	Feature comparison of the SOTA methods using ACDC dataset	36
Table 3.1	Summary of the study population in ACDC dataset	39
Table 3.2	Dataset Distributions	42
Table 4.1	Evaluation of U-Net Variants Using Test Set Dice Scores	61
Table 4.2	Evaluation of U-Net Variants Using Test Set Jaccard Coefficients	63
Table 5.1	Dice Score Comparison with Existing Studies Evaluated on ACDC Dataset	79
Table 5.2	Comparison with Existing Studies Evaluated on ACDC Dataset	80

LIST OF ABBREVIATIONS

Abbreviation	Description
ACDC	Automated Cardiac Diagnosis Challenge
Atn-UN	Attention U-Net
CMR	Cardiac Magnetic Resonance
CMRI	Cardiac Magnetic Resonance Imaging
CNN	Convolutional Neural Network
DC	Dice Coefficient
DL	Deep Learning
ED	End Diastolic
ES	End Systolic
FCN	Fully Convolutional Neural Network
Feed-Res-UN	Feedback Residual U-Net
FP-UN	Feature Pyramid U-Net
FPB	Feature Pyramid Block
FPN	Feature Pyramid Network
GANs	Generative Adversarial Networks
GRU	Gated Recurrent Unit
JC	Jaccard Coefficient
LSTM	Long Short-Term Memory
LV	Left Ventricle
ML	Machine Learning
MLP	Multi-Layer Perceptron
MRI	Magnetic Resonance Imaging
MYO	Myocardium
O-UN	Original U-Net
Res-UN	Residual U-Net
RNN	Recurrent Neural Network
ROI	Region of Interest
RV	Right Ventricle
SSMs	Statistical Shape Models
w.r.t	with respect to

LIST OF APPENDICES

Appendix	Description	Page
Appendix -A	Performance Evaluation of DL-Based Studies	94
Appendix -B	Publications	99