UNIVERSITI TEKNOLOGI MARA

A FULLY AUTOMATED MYOCARDIAL INFARCTION SEGMENTATION AND DETECTION IN LATE GADOLINIUM ENHANCEMENT CARDIAC MAGNETIC RESONANCE IMAGING USING DEEP LEARNING

DAYANG SUHAIDA BINTI AWANG DAMIT

PhD

February 2026

UNIVERSITI TEKNOLOGI MARA

A FULLY AUTOMATED MYOCARDIAL INFARCTION SEGMENTATION AND DETECTION IN LATE GADOLINIUM ENHANCEMENT CARDIAC MAGNETIC RESONANCE IMAGING USING DEEP LEARNING

DAYANG SUHAIDA BINTI AWANG DAMIT

Thesis submitted in fulfilment
of the requirements for the degree of
Doctor of Philosophy
(Electrical Engineering)

Faculty of Electrical Engineering

February 2026

CONFIRMATION BY PANEL OF EXAMINERS

I certify that a Panel of Examiners has met on 16th December 2025 to conduct the final examination of Dayang Suhaida Binti Awang Damit on her Doctor of Philosophy thesis entitled "A Fully Automated Myocardial Infarction Segmentation and Detection in Late Gadolinium Enhancement Cardiac Magnetic Resonance Imaging Using Deep Learning" in accordance with Universiti Teknologi MARA Act 1976 (Akta 173). The Panel of Examiners recommends that the student be awarded the relevant degree. The Panel of Examiners was as follows:

Zainal Hisham Che Soh, PhD
Associate Professor
Faculty of Electrical Engineering
Universiti Teknologi MARA
(Chairman)

Iza Sazanita Isa, PhD
Associate Professor
Faculty of Electrical Engineering
Universiti Teknologi MARA
(Internal Examiner)

Fakroul Ridzuan Hashim, PhD
Associate Professor
Department of Electrical and Electronics
Faculty of Engineering
Universiti Pertahanan Nasional Malaysia
(External Examiner)

PROFESSOR DR HJH ZURAEDA IBRAHIM
Dean
Institute of Postgraduates Studies
Universiti Teknologi MARA
Date: 19 February 2026

AUTHOR'S DECLARATION

I declare that the work in this thesis was carried out in accordance with the regulations of Universiti Teknologi MARA. It is original and is the results of my own work, unless otherwise indicated or acknowledged as referenced work. This thesis has not been submitted to any other academic institution or non-academic institution for any degree or qualification.

I, hereby, acknowledge that I have been supplied with the Academic Rules and Regulations for Post Graduate, Universiti Teknologi MARA, regulating the conduct of my study and research.

Name of Student	Dayang Suhaida Binti Awang Damit
Student ID. No.	2021803278
Programme	Doctor of Philosophy (Electrical Engineering) - CEEE950
Faculty	Electrical Engineering
Thesis Title	A Fully Automated Myocardial Infarction Segmentation and Detection in Late Gadolinium Enhancement Cardiac Magnetic Resonance Imaging Using Deep Learning
Signature of Student	
Date	February 2026

ABSTRACT

Late gadolinium enhancement cardiac magnetic resonance imaging (LGE-CMR) has emerged as the gold standard for non-invasive myocardial tissue characterization, particularly for identifying scarred or infarcted myocardial regions. The current diagnostic workflow remains reliant on manual segmentation and expert interpretation, which are time-consuming and subject to inter-observer variability. Although deep convolutional neural networks (DCNNs) have demonstrated strong potential in automated medical image segmentation, existing models suffer from low segmentation accuracy due to the intrinsic complexity of the modality. This study proposes a fully automated, end-to-end framework for MI detection, comprising two main components: a segmentation of myocardial scars and a subsequent detection module that employs sequential analysis to identify MI cases based on the segmented regions. As a foundational step, this study proposes a new automated classification method to address a gap in previous studies that relied on manual slice preselection. A shallow deep convolutional neural network (S-DCNN) model is introduced to automatically classify LGE-CMR slices into those containing the left ventricle (LV) and non-LV categories. This classification step significantly streamlines the automated segmentation pipeline by ensuring that only relevant slices are processed further. Central to the proposed approach is a novel dual-stage segmentation model designed for fully automated myocardial infarction detection from LGE-CMR images. This model comprises two task-specific stages, each built upon a customized DeepLabV3+ segmentation network. The first stage, referred to as DeepLabV3+ Tailor LV (DLT-LV), is tailored for accurate segmentation of the LV region. The second stage, DeepLabV3+ Tailor Scar (DLT-Scar), focuses on segmenting myocardial scar tissue within the segmented LV area. To improve the reliability of this pipeline, a morphological post-processing procedure is introduced between the two stages to refine the LV segmentation output and address common issues such as small gaps and isolated mis-segmentations. To further enhance the performance of the second stage, a hybrid optimization-segmentation approach is proposed. This method integrates Particle Swarm Optimization (PSO) into the DLT-Scar model by automatically tuning the hyperparameters of the Tversky loss function used in its classification layer, specifically addressing the class imbalance inherent in scar segmentation. Finally, in the MI detection phase, a Scar Sequential Slice Reconstruction (3SR) module is integrated with a S-DCNN classification model to exploit the spatial continuity of myocardial scars across consecutive slices. This module constructs a localized sequence of consecutive image slices with the largest segmented scar identified during the segmentation stage. During the performance evaluation stage, the results were compared with patient radiology reports to assess their clinical relevance. Experimental results demonstrate that the proposed method achieves high segmentation accuracy and robust MI detection, with a mean Dice Score of 74.31% for scar segmentation and 96.7% accuracy, sensitivity of 100% and specificity of 95% for MI classification. These findings offer a scalable and fully automated solution for improving diagnostic consistency and efficiency in clinical cardiac imaging workflows.

ACKNOWLEDGEMENT

Alhamdulillah, Subhanallah, Allahuakbar.

Firstly, I wish to thank Allah the Merciful and the All-Knowing for granting me the strength and perseverance to successfully complete this long and challenging PhD journey.

My deepest gratitude goes to my dear supervisor, Prof. Ir. Ts. Dr. Hjh Siti Noraini Sulaiman, and co-supervisors, Assoc. Prof. Dr Muhammad Khusairi Osman, Assoc Prof. Dr Noor Khairiah A. Karim and Assoc. Prof. Ir. Ts. Dr. Samsul Setumin, who provided valuable insights, feedback, and assistance throughout the study duration. Their expertise and guidance have significantly contributed to the quality and depth of this research.

I would also like to express my sincere appreciation to the staff of the Advanced Medical and Dental Institute (AMDI) for providing the necessary facilities, knowledge, and assistance that were essential to this work. A special thanks goes to my colleagues and friends from the Advanced Control System and Computing Research Group (ACSCRG) for their encouragement, insightful ideas, and support, which greatly advanced my understanding during this study.

I am genuinely grateful for the wonderful friends I have been blessed with throughout this PhD journey, especially Firdaus, Atiqah and Azura, who have been by my side from start to finish, always spreading positivity and uplifting my spirits. To those whom I may not have mentioned by name, please know that I am forever indebted to your invaluable support. Only Allah the Almighty can truly repay your kindness.

This thesis is dedicated to my beloved parents, Abah (Awang Damit) and Mama , my husband, Azim Izzuddin, and our precious son, . Their tremendous love, endless compassion and continuous emotional and financial support have been the backbone of my journey.

Last but not least, I am grateful to myself for enduring, persevering, and continuing on, no matter how hard the journey became. Alhamdulillah, by His grace, I am able to turn this page and embrace another chapter in my life.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

Abbreviations

3SR	Scar Sequential Slice Reconstruction
Adam	Adaptive Moment Estimation
AMDI	Advanced Medical and Dental Institute
ASPP	Atrous Spatial Pyramid Pooling
DCNN	Deep Convolutional Neural Networks
DLT-LV	DeepLabV3+Tailor LV
DLT-Scar	DeepLabV3+ Tailor Scar
GDL	Generalized Dice loss
GT	Ground Truth
LGE-CMRI	Magnetic Resonance Imaging
LV	Left Ventricle
MI	Myocardial Infarction
MRI	Magnetic Resonance Imaging
PSO	Particle Swarm Optimization
S-DCNN	Shallow DCNN
SGDM	Stochastic Gradient Descent with Momentum
USM	Universiti Sains Malaysia
WCE	Weighted Cross-Entropy

CHAPTER 1
INTRODUCTION

1.1 Research Background

The heart is a vital organ responsible for circulating blood throughout the body. When the coronary arteries become blocked, the blood supply carrying oxygen is restricted, preventing oxygen from reaching the heart muscle. This deprivation causes irreversible tissue damage, leading to the formation of myocardial scars. The condition caused by these arterial blockages is known as ischemic heart disease (IHD) [1], [2]. The severity of the blockage determines the extent of tissue damage, and a more significant obstruction may result in myocardial infarction (MI), commonly known as a heart attack.

According to the World Health Organization (WHO), cardiovascular diseases (CVDs) cause about 32% of all global deaths with 85% of these fatalities due to ischemic heart diseases and strokes [3], [4]. In Malaysia, ischemic heart disease has been the leading cause of death for over a decade, and the number of cases continues to grow annually [5], [6]. Among these conditions, MI stands out as a major contributor and are often referred to as a "silent killer" due to their sudden and life-threatening nature. Recently, large-scale studies have reported a long-term increase in the risk of cardiovascular events, including MI, among individuals recovering from COVID-19 and among vaccinated individuals who experienced heart-related side effects [7], [8], [9], [10]. These trends point to a growing population at risk and a rising demand for timely and accurate cardiac assessments. Therefore, the adoption of automated techniques offering high-speed and precise assessment should be carried out in order to improve the growing interest in medical image analysis and to enhance early detection and ultimately improve patient outcomes.

MI detection relies on multiple diagnostic tools, each offering specific advantages and limitations. Among these, late gadolinium enhancement-cardiac magnetic resonance (LGE-CMR) has emerged as the reference standard for non-invasive myocardial tissue characterization, particularly in the left ventricle (LV) [11], [12], [13]. LGE-CMR provides high-resolution visualization of fibrotic and scarred myocardial regions, which are critical indicators of prior or ongoing infarction. The

early identification and accurate delineation of these scarred regions are essential for guiding clinical decision-making, enabling timely therapeutic interventions, and potentially reducing the risk of adverse cardiac events [14], [15],[16]. However, the current standard for myocardial scar analysis involves manual segmentation, wherein radiologists delineate scar boundaries slice by slice. This approach is not only time-consuming and labor-intensive but also subject to inter-observer variability, as it relies heavily on expert judgment [17], [18]. As the demand for cardiac imaging continues to grow, these limitations underscore the growing need for automated segmentation techniques that can improve efficiency, minimize observer dependence, and facilitate large-scale deployment in healthcare settings.

In the context of MI detection, segmentation serves as a critical prerequisite, as identifying and quantifying myocardial scars are fundamental steps in diagnosing infarction from LGE-CMR images. Recent studies have shown increasing interest in automated segmenting of left ventricular scar tissue due to its significant role in clinical applications. However, in CMR imaging, automatic segmentation of myocardial infarction in LGE-CMR remains highly challenging due to the intrinsic complexity of the imaging modality. These scars often exhibit structural variability, including differences in size and shape, as well as heterogeneity in signal intensity within the LGE-CMR images [19], [20]. This heterogeneity arises from both intrinsic tissue characteristics and extrinsic imaging limitations, which frequently hinder the clear visualization of infarcted regions and their boundaries. As a result, infarcted areas may appear poorly defined or visually similar to surrounding healthy tissue, thereby reducing the effectiveness of conventional threshold-based segmentation methods.

In recent years, deep convolutional neural networks (DCNNs) have shown great promise in medical image segmentation [18], [21], [22] . Unlike traditional threshold-based or handcrafted feature methods, DCNNs can automatically learn hierarchical spatial and intensity-based features directly from imaging data. However, only a few studies [18], [23], [24] aimed primarily at fully automated myocardial scar segmentation with LGE-CMR modality, with reported relatively low accuracy in scar segmentation. Many of these models are primarily designed for general-purpose segmentation with limited adaptation to the unique characteristics of LGE-CMR images.

Hence, as motivated by the fact that the size of the infarct can be significantly reduced when early treatment is administered, and that timely diagnosis plays a critical role in improving patient outcomes, new research must consider a task-specific DCNN

framework designed to address the anatomical and imaging complexities of cardiac structures, with a focus on fully automating myocardial infarction detection from LGE-CMR images. This includes recognizing that each segmentation task may require different tuning strategies depending on the nature of the target region and the LGE-CMR modality characteristics. In addition, a fully automated framework that operates from the point at which radiologists receive the LGE-CMR images is essential to support fast and clinically practical myocardial infarction segmentation and detection.

1.2 Problem Statement

Early detection of myocardial infarction is essential for minimizing severe cardiac complications and reducing the risk of sudden cardiac death. The presence and extent of infarcted myocardial tissue can be accurately assessed through tissue characterization using LGE-CMR imaging. While expert visual interpretation of cardiac images remains a fundamental aspect of clinical practice, the increasing demand for cardiac imaging, coupled with the time-intensive and subjective nature of manual analysis, presents a significant challenge. Therefore, there is an urgent need for an automated diagnostic system that can deliver rapid, accurate, and reproducible results to support early detection and facilitate improved clinical decision-making in myocardial infarction management.

In typical CMR protocols, the acquired image stack includes both LV and non-LV cardiac segments [25], where non-LV refers to any image that does not contain the LV region. However, only the slices containing the LV region are clinically relevant for myocardial scar assessment, while the outermost non-LV slices are usually excluded from further analysis. This selection process is currently performed manually by radiologists, adding unnecessary burden to their workflow and reducing overall efficiency. Several researchers have claimed to have developed an automated MI detection system as in [26], [27], [28], [29], [30], [31]. However, these methods apply their algorithms directly to manually selected LV images, requiring radiologists to predetermine the LV region. Without an initial classification mechanism to distinguish LV from non-LV slices, segmentation pipelines remain semi-automated and dependent on expert input. Therefore, there is a need for a fully automated approach that includes an initial classification stage to correctly identify LV cardiac segments. This step is vital

to streamline the segmentation process and reduce the computational burden associated with the subsequent process.

Despite advancements in deep learning techniques for cardiac image analysis, only a limited number of studies have dedicated efforts to achieving fully automated myocardial scar segmentation directly from LGE-CMR images. The process of extracting myocardial scar from LGE-CMR images using a single segmentation model often results in suboptimal performance [32], [33]. This is primarily due to class imbalance and the difficulty in learning distinct anatomical features, each of which introduces its own challenges. Current methods often overlook the need for task-specific modeling strategies that can effectively adapt to the anatomical variations of the LV and the diverse sizes of myocardial scars. These challenges are further complicated by heterogeneity in signal intensity within LGE-CMR images, intensity inconsistencies within the LV cavity, and the influence of structures such as papillary muscles and trabeculations [23], [28], [33], [34]. Additionally, background artifacts and surrounding tissues that resemble the LV, along with low scar-to-background contrast, further hinder accurate segmentation [18], [19]. These factors have contributed to low Dice scores in previous studies, underscoring the need for more robust and adaptive segmentation approaches.

Loss functions play a critical role in the performance of deep learning-based medical image segmentation models, particularly in addressing the issue of class imbalance when the target region occupies a small portion of the image, especially in medical imaging. This challenge is particularly evident in myocardial scar segmentation from LGE-CMR images, where scar tissue often constitutes a minimal fraction compared to the background myocardium. Conventional loss functions such as binary cross-entropy and Dice loss tend to bias predictions toward the majority class [35], resulting in under-segmentation of clinically important scar regions and increased false negatives. To address class imbalance in medical image segmentation, several studies have adopted adaptive loss functions including the Tversky loss. The Tversky loss introduces tuneable parameters, a and P, which adjust the weighting between false positives and false negatives,depending on the specific characteristics of the segmentation task or clinical application. In the context of myocardial scar segmentation, where reducing false negatives is crucial, this flexibility allows the model to be more sensitive to small or subtle scar regions. However, in most existing studies, the values of a and P are manually set or empirically defined without systematic

optimization [36], [37], [38]. Such rigid parameter settings limit the adaptability and generalizability of the loss function across different medical imaging contexts. Therefore, there is a need to propose an optimized method to automatically search accurate a and P values.

Most existing studies on MI detection from LGE-CMR images focus on analyzing a single segmented image slice to determine the presence of infarcted regions[23]. However, this approach does not align with standard clinical practice, where radiologists assess infarction based on the continuity of enhancement across sequential image slices to confirm pathological findings and reduce the risk of misinterpretation caused by artifacts or isolated noise. A common approach in previous studies involves applying voxel count thresholds to classify each case as either MI or non-MI [39], [40]. Therefore, there is a need for a post-segmentation framework that incorporates information from the entire LV sequence rather than relying on isolated slices.

Thus, this study proposes a fully automated, end-to-end framework for segmentation and detection of MI in short-axis LGE-CMR images, incorporating a CNN-based method to improve diagnostic accuracy. It introduces a task-specific design tailored to the unique challenges of myocardial scar segmentation and clinical interpretation in LGE-CMR, addressing several limitations that have been overlooked in previous research. As for the outcome of this study, the proposed system is expected to provide clinicians with reliable and efficient assistance in early MI detection and treatment.

1.3 Research Objectives

The primary objective of this research is to develop a fully automated, end-to-end framework for segmentation and detection of MI in short-axis LGE-CMR images using a CNN-based model. To achieve this aim, the work is divided into several stages and carried out systematically with the associated objectives that have been identified as follows:

1. To propose an automated method to classify LV and non-LV chambers in short-axis LGE-CMR images using a DCNN model.

2. To model a new Dual-stage DeeplabV3+-based model to improve the segmentation performance of myocardial scar in LGE-CMR images.
3. To enhance the performance of the proposed model in (2) by automatically identifying the optimal Tversky loss hyperparameter values through a particle swarm optimization (PSO) based algorithm.
4. To develop an automated MI detection system from sequences of images obtained in (3) and validate its performance through the radiologist reports.

1.4 Scope of Research

This research is focused on developing an automated detection of MI in LGE-CMR images by incorporating neural network design and image processing. The research scope is outlined as follows:

i) Image acquisition and data collection:

This study uses a private dataset obtained from the Imaging Department, Advanced Medical and Dental Institute (AMDI), Universiti Sains Malaysia (USM), Kepala Batas, Pulau Pinang, with ethical approval. The dataset comprises LGE-CMR images from patients scanned using the Philips Achieva 3.0T Scanner. A total of 60 patient data sets were collected retrospectively from 2018 to 2024, which included 31 patients with confirmed MI diagnoses and 29 subjects from 'normal' cardiac patients. It is important to note that no individual patient assessments or interviews were conducted before, during, or after the imaging process. The data collection for this study was specifically limited to patients with myocardial scars, indicative of prior myocardial infarction events related to IHD only.

ii) Selection of methods:

The study introduced a new framework for the fully automated detection of MI using exclusively the LGE-CMR imaging modality. The overall methodological design is aligned with the research objectives, emphasizing the integration of DCNN and image processing techniques tailored for LGE-CMR data. The classification component of the proposed system employs a hierarchical DCNN-based approach to differentiate images

containing LV and non-LV regions. A comparative analysis was conducted involving both shallow and deeper DCNN architectures, including pretrained models.

In the segmentation stage, the study focuses on myocardial scar segmentation through a dual-stage segmentation framework built upon a modified DeepLabV3+, an encoder-decoder CNN architecture. To enhance the accuracy of LV segmentation, a novel post-processing morphological procedure is introduced between the two stages, utilizing image processing techniques to refine the initial segmentation output before proceeding to the scar segmentation phase.

Furthermore, to address the challenge of class imbalance and improve the robustness of scar detection, the second stage integrates a heuristic optimization technique, specifically PSO. The PSO is employed to automatically optimize the hyperparameters of the loss function, thereby enhancing the model's ability to segment scar tissue accurately and consistently across diverse anatomical variations.

iii) Evaluation and analysis:

The performance evaluation of the proposed method implemented on the MATLAB software is divided into two categories: qualitative and quantitative evaluation. Quantitative assessments are done on each output of the objectives using evaluation metrics from the confusion metrics, such as accuracy, sensitivity/recall, and F-score for classification. Dice Similarity Index and Intersection over Union are used for segmentation results. For MI detection, the MI detection results will also be compared with radiologist reports to ensure the validity of the final outputs.

iv) Platform

The automated system is developed using MathWorks: MATLAB version R2023b software. It is designed to detect MI from LGE-CMR images using customized CNNs including DeepLabV3+ segmentation model that built and trained within the MATLAB environment using built-in functions from the Deep Learning Toolbox. Suitable and well-established morphological post-processing methods were also applied to the segmented images in both the segmentation and detection stages using image processing techniques.

1.5 Thesis Layout

This thesis is divided into five (5) chapters that explain the entire operation of this study. It begins with a brief introduction to MI and its clinical importance. It discusses the role of LGE-CMR imaging in non-invasive MI diagnosis and outlines the challenges in accurately segmenting myocardial scars, particularly due to anatomical variability, image noise, and class imbalance. The chapter also presents the problem statement. The objectives and scopes of the thesis are defined at the end of this chapter.

Chapter 2 continues presenting the descriptions of MI, the differences in MI detection tools, issues, and challenges. The chapter includes a comprehensive literature review of existing studies that address MI segmentation and detection through advanced deep-learning imaging techniques and medical image analysis. It discusses traditional methods as well as recent advancements in DCNN and optimization strategies, including innovative architectures and loss functions designed to manage class imbalance effectively. Additionally, the review highlights significant gaps in the current methodologies, laying the groundwork for the proposed advancements in MI detection and segmentation algorithms.

Chapter 3 describes the details of the proposed methodology for automated MI detection using LGE-CMR images. Starting from data collection and preparation stages, including the image labelling process and augmentation. Next, it describes an automated slice classification method using a Shallow DCNN (S-DCNN) to differentiate between slices with and without the LV region for the next stage. The core of the methodology lies in the development of a novel dual-stage segmentation model, designed specifically for the sequential segmentation of the LV region and myocardial scar tissue. Each stage in this model is built upon a customized DeepLabV3+ segmentation network, optimized for the respective segmentation task. A morphological post-processing procedure is introduced between the two stages to refine the LV segmentation before scar detection. The chapter also presents a hybrid optimization-segmentation approach that integrates PSO into the scar segmentation model, enabling automated tuning of the Tversky loss function parameters for enhanced scar segmentation accuracy. Finally, the Scar Sequential Slice Reconstruction (3SR) module and the S-DCNN classification model are described for robust MI detection based on spatial continuity.

Chapter 4 presents the experimental results of the proposed framework, including both quantitative and qualitative evaluations of classification and

segmentation performance. Comparisons are made between different network architectures, loss functions, and classification inputs (single vs. sequential images). The effectiveness of post-processing, PSO optimization, and the 3 SR module are also evaluated. Performance metrics such as Dice Score, IoU, sensitivity, specificity, and confusion matrices are analyzed in detail, and the results are correlated with expert radiology reports to assess clinical relevance.

Finally, Chapter 5 concludes the thesis by summarizing the main findings of this study, as well as the contributions made from each part of the study. Limitations of the current work are acknowledged, and directions for future research are suggested.

CHAPTER 2
LITERATURE REVIEW

2.1 Introduction

Myocardial infarction (MI), commonly referred to as a heart attack, remains one of the leading causes of morbidity and mortality worldwide. Timely and accurate detection of MI is crucial for initiating effective clinical interventions and improving patient outcomes. The identification of myocardial scar tissue, which is an indicator of infarction, is essential not only for diagnosis but also for treatment planning, risk stratification, and prognosis assessment. Late Gadolinium Enhancement Magnetic Resonance Imaging (LGE-MRI) has emerged as the gold standard for non-invasive myocardial tissue characterization. It provides high-resolution contrast between healthy and infarcted myocardium, enabling precise visualization of scars. However, the manual interpretation and delineation of scar tissue in LGE-MRI is time-consuming, subject to inter-observer variability, and requires specialized expertise. These limitations underscore the growing need for automated, end-to-end deep learning frameworks capable of accurately segmenting myocardial infarction regions from LGE-MRI with minimal human intervention.

This chapter presents a comprehensive literature review of existing methodologies related to MI detection. It begins by outlining the clinical background of MI, followed by a discussion of the challenges faced by radiologists in accurately identifying infarcted tissue from LGE-CMR images. The subsequent section provides an overview of deep learning techniques relevant to medical image optimization, including a review of key model architectures, loss functions, and hyperparameter optimization approaches, with a particular attention to PSO. The chapter then critically reviews prior research on the automated segmentation and detection of MI, covering both conventional and deep learning-based methods. This review highlights the evolution of fully automated pipelines, their limitations, and how recent advances address key clinical and technical challenges. The chapter concludes with a summary of the main findings, identifying research gaps that form the foundation for the proposed methodology in this study.

2.2 Cardiac Anatomy and Myocardial Infarction

The human heart is a complex muscular pumping organ composed of four chambers, including two large ventricles and two atria, which work in tandem to facilitate blood circulation throughout the body. The left ventricle (LV) is the biggest chamber has a cone-shaped chamber enclosed by the myocardium, whereas the right ventricle (RV) has a more complex crescent shape. LV plays a vital role in pumping oxygenated blood to the entire body organs and tissues via the aorta with every heartbeat, and lung by the RV [41]. There are three main coronary arteries of the heart: (i) left anterior descending artery (LAD), (ii) left circumflex artery (LCX) and (iii) right coronary artery (RCA) [42]. These major arteries supply oxygen-rich blood to the heart muscle, which allows the heart to beat and pump blood throughout the body. Ischemic heart disease (MD) refers to a condition where the coronary arteries, which supply blood to the heart muscle, become narrowed or blocked usually caused by a buildup of plaque (atherosclerosis) within the artery walls. When the coronary arteries are narrowed, the blood flow to the heart muscle is reduced, leading to a shortage of oxygen called ischemia. Prolonged ischemia can lead to cell death, reducing the heart's ability to pump effectively, and may result in a myocardial infarction (MI). During a heart attack, the myocardium undergoes cell death and is gradually replaced by scar tissue over several weeks [1]. The specific characteristics of that scar such as the size, location, shape and mechanical properties can significantly influence the patient's long-term survival and risk of future heart problems [1], [43] . MI is often called a "silent killer" because it doesn't always present with obvious symptoms. Instead of classic chest pain, some experience milder symptoms like discomfort in the chest, arm, or jaw, shortness of breath, fatigue, or indigestion [44]. These vague symptoms can be easily dismissed or misattributed to other conditions, leading to dangerous delays in seeking treatment. The infarcted heart due to narrowed artery illustrate in Figure 2.1 and the cross anatomical section of the normal and infarcted heart are illustrated as in Figure 2.2.

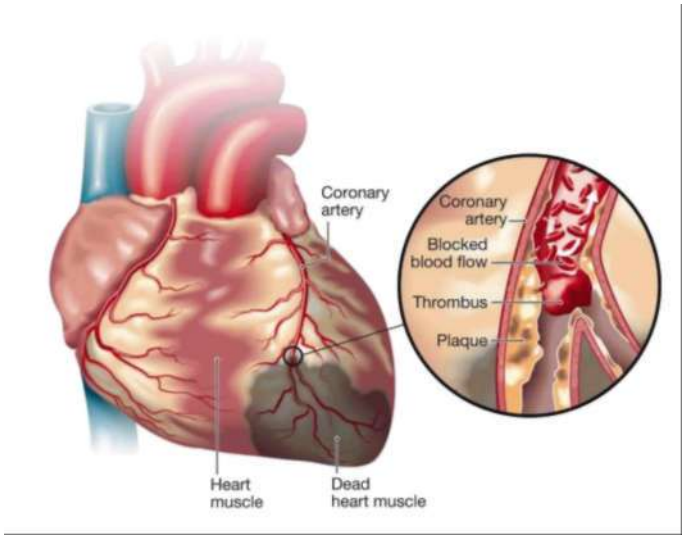
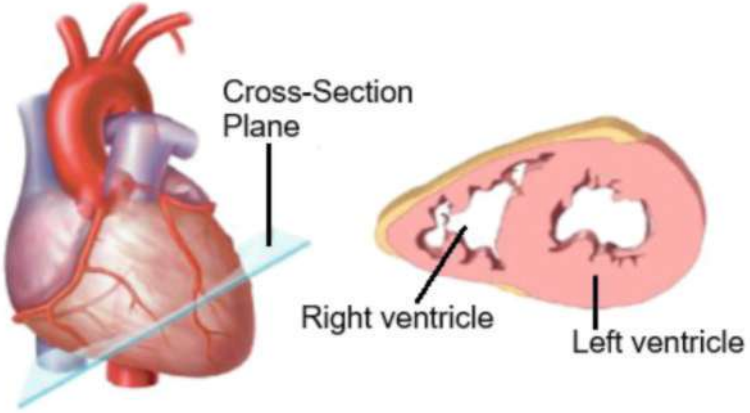
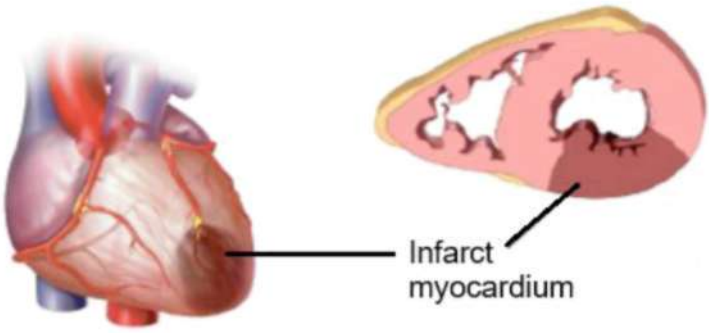


Figure 2.1 The Infarcted Heart Due To Narrowed Artery [45]



(a) Normal Heart



(b) Infarcted heart

Figure 2.2 Cross-Section Illustration of the Human Heart For (a) Normal and (b) Infarcted, reproduced from [46]

The infarcted tissue or myocardial scarring is the accumulation of fibrosis tissue from the formation of excess tissue in replacement of necrotic tissue [43]. It is referred to as the final result and pathological correlate of myocardial infarction. Scar formation following a myocardial infarction typically progresses from the endocardium (the innermost layer of the heart wall) outward toward the epicardium (the outermost layer) in a "wavefront" pattern. Depending on the presence of collateral circulation, persistent or intermittent coronary arterial blockage, sensitivity of myocytes to ischemia, preconditioning, and individual demand for myocardial oxygen and nutrients, complete necrosis can take 2-4 hours or longer. Figure 2.3 illustrates the anatomical classification of MI based on the extent of myocardial involvement[47]. Transmural infarcts involve the full thickness of the myocardium, extending from the epicardium to the endocardium. Subendocardial infarcts affect the inner layer of the myocardium adjacent to the endocardium. Intramural infarcts involve the middle layer of the myocardium. Subepicardial infarcts affect the outer layer of the myocardium adjacent to the epicardium. The ventricular scars can be indicative of an earlier episode of myocardial infarction.

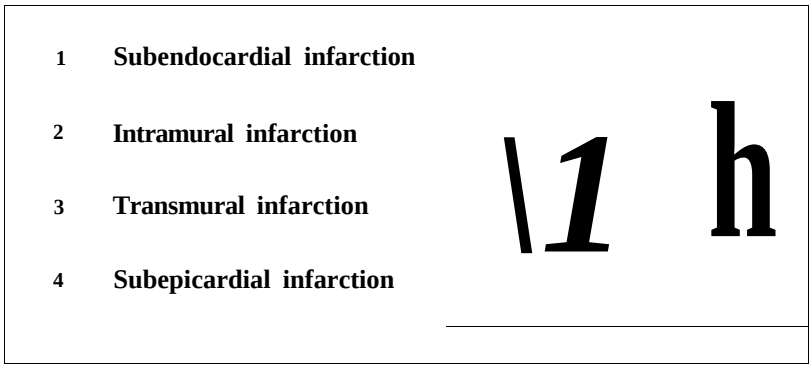


Figure 2.3 Lateral Cross-Section Anatomical Classification of MI [47]

2.2.1 Diagnosis of Myocardial Infarction

MI may represent a major catastrophic event leading to severe hemodynamic failure or sudden death, or it may occur repeatedly in patients with established heart disease. It requires timely and accurate screening methods to ensure effective treatment and reduce mortality. Several methods and technologies are used for MI screening, each with its own advantages and limitations. An electrocardiogram (ECG) is the most common non-invasive and rapid technique for MI diagnosis [47]. It detects ischemic

changes such as ST-segment elevation (STIEMI), where there's a complete blockage of a coronary artery or depression by identifying characteristic electrical patterns. However, its sensitivity in non-ST-segment elevation MI (NSTEMI) and prior MI detection is limited[44]. In this context, the role of imaging techniques may become useful for the understanding of the determinants in a preclinical setting before acute coronary events, and for an accurate and correct diagnosis of MI. Echocardiography uses ultrasound waves to create images of the heart [48]. It provides valuable insights into cardiac function and wall motion abnormalities, useful for assessing MI-related damage. Yet, its accuracy can be affected by factors like body habitus and operator expertise [49]. Three-dimensional imaging techniques, such as Cardiac Computed Tomography (Cardiac CT), specifically Coronary computed tomography angiography (CCTA) and Cardiac magnetic resonance imaging (CMR) are both valuable imaging techniques used in MI screening, but they have distinct focuses and characteristics. CCTA excels at visualizing coronary artery anatomy with high spatial and temporal resolution, making it ideal for detecting blockages or narrowing. Its fast acquisition time and wide availability make it a practical choice, especially in acute settings. However, CCTA involves radiation exposure and may be limited by calcifications or rapid heart rates, with a limited capacity to directly assess myocardial tissue [50], [51]. In contrast, CMR provides comprehensive information about both coronary anatomy and myocardial tissue characteristics, including infarct size and viability without radiation exposure. It also assesses cardiac function and wall motion. Therefore, CMR has been established as a method to detect myocardial infarction, particularly useful in cases where the diagnosis is uncertain or when further information is needed to guide treatment decisions. Figure 2.4 shows an example of a cardiac MRI scan.

Figure 2.4 CMR Scanner

2.2.2 Manual Diagnosis of Myocardial Infarction Using LGE-CMR

CMR plays a crucial role in assessing cardiac mass. The segmentation of heart chambers and myocardium presents significant challenges primarily due to variations in heart anatomy among individuals and CMR imaging acquisition parameters. This complexity is compounded by the need for precise delineation of anatomical structures and pathological lesions, which is critical for accurate clinical diagnosis and treatment planning in cardiovascular medicine [52], [53], [54]. MI mainly affects the LV, which is the heart's primary pumping chamber, but the damage may extend into the right ventricle (RV) or the atria. While MRI can be applied to both ventricles, the LV has historically been the focus of more research due to its greater workload and clinical significance. Its thicker wall, compared to the thinner-walled RV, also allows for better resolution on CMR imaging. The LV is cone-shaped and positioned posterior to the right ventricle, sloping from its base in the plane of the atrioventricular groove to the cardiac apex [55]. Left ventricular segmentation in cardiac images typically involves dividing the chamber into three main segments: apical, middle, and basal. These segments are defined using the standard imaging plane for cardiac imaging, known as the short-axis plane, which is oriented perpendicular to the long (apex-base) axis of the heart [56] and shown in Figure 2.5. In the field of cardiac magnetic resonance imaging (MRI), a stack of short-axis cardiac MR images frequently includes above or out of outmost slices of the left ventricle (LV) coverage [57], [58].

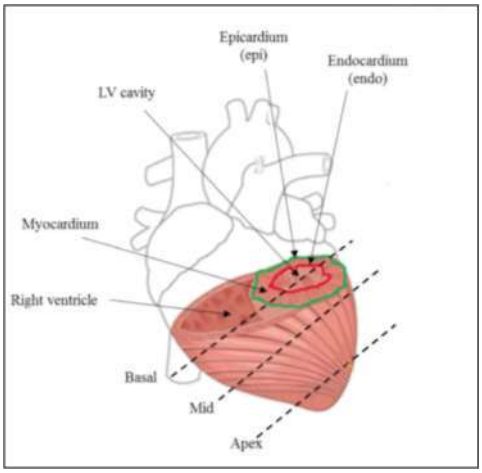


Figure 2.5 Short-axis CMR Imaging Planes Illustrating the Division of the Left Ventricle Into Basal (Base), Mid (Middle), and Apex (Apical) Segments reproduced from [31]

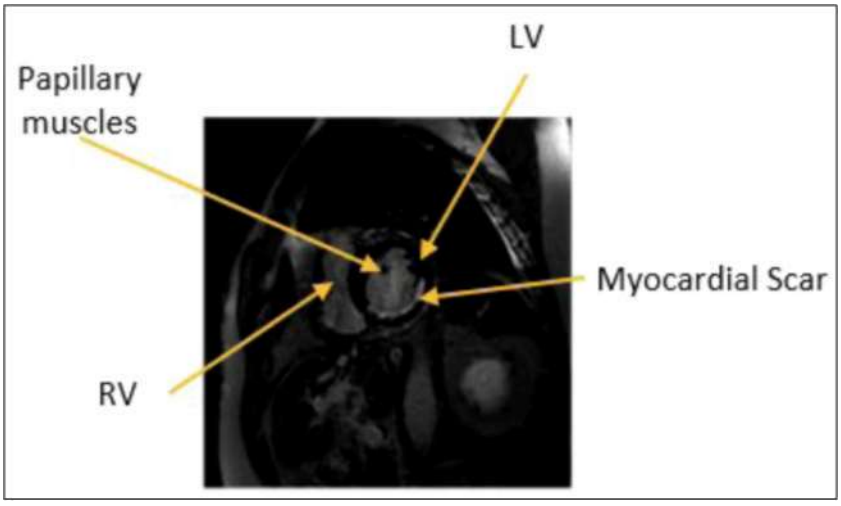


Figure 2.6 An Example of Short-Axis LGE-CMR Image, LV Region and the Myocardium Scar

The choice of cardiac imaging modality CMR imaging depends on the disease being investigated. Late Gadolinium Enhancement Cardiac Magnetic Resonance imaging (LGE-CMR) has emerged as a gold-standard modality for identifying and quantifying regions of scarring in [47], [59]. The imaging is performed 10 to 20 minutes after of the injection of a gadolinium-enhanced contrast agent, which accumulates in damaged tissues and causes over-enhancement of the scar region [60]. In contrast, normal myocardium tissue exhibits rapid gadolinium wash in and wash out, with no abnormal accumulation of the contrast agent. This contrasting feature on the LGE-CMR images clearly delineate the regions of scar tissue, as presented in Figure 2.6. By visually assessing the contrast between enhanced and non-enhanced areas on LGE

images, experienced cardiac radiologists or cardiologists can accurately pinpoint the location and extent of myocardial infarction.

The location of myocardial scar on LGE is visually assessed using the American Heart Association (AHA) 17-segment model [56] (Figure 2.7), a widely adopted standard in clinical practice. This model facilitates the correlation of abnormalities observed in short-axis cardiac segments with the corresponding territories of the three major coronary arteries, LAD, RCA and LCX. In current clinical practice, the radiologist usually presents slices of images through the entire cardiac. Medical experts visually estimate and manually draw contours on pathological tissues to quantify the disease's severity according to the amount and transmural extent of myocardial scar [57], [61]. For example, a large, transmural scar might indicate a more severe heart attack, and a higher risk of complications compared to a small, subendocardial scar. It can appear in many different ways, and these variations are important for diagnosis, prognosis, and treatment planning. However, this inspection procedure is highly time-consuming, and the reliability of the diagnostics largely depends on the expert's experience, availability, and skills.

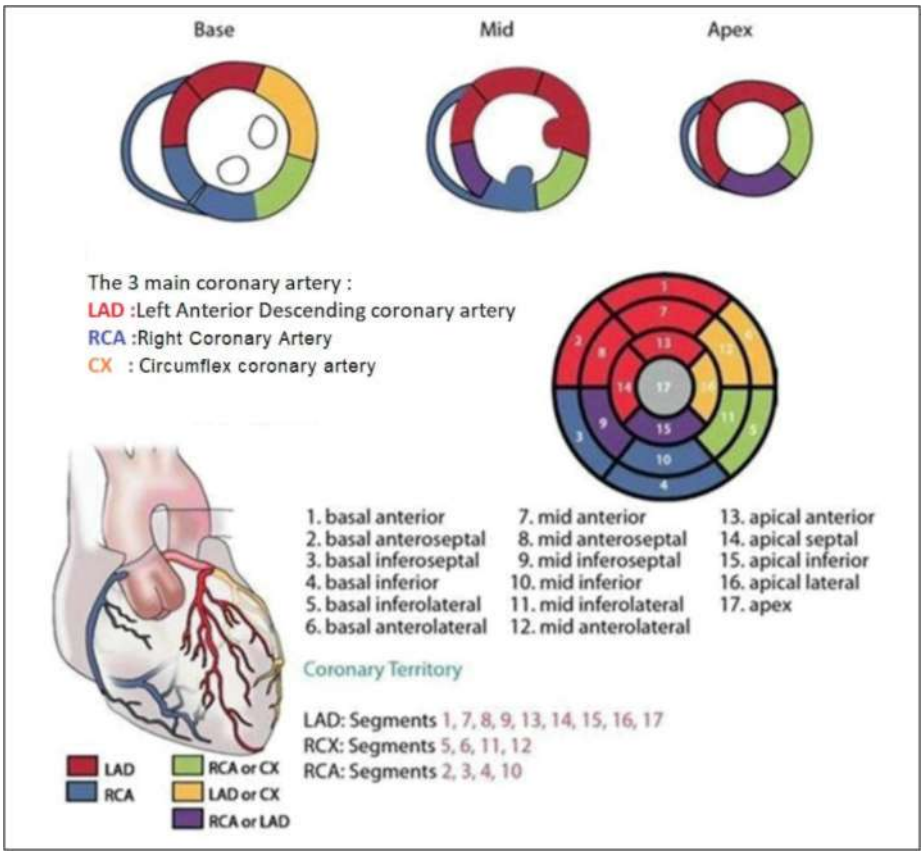


Figure 2.7 The American Heart Association (AHA) 17 Segment Model [40]

2.2.3 Issue and Challenges in Manual Delineation of Myocardial Scar

Current issues in LGE-CMR imaging reflect both the potential of this diagnostic tool and the challenges that hinder its widespread application in clinical practice. Despite being recognized as the gold standard for assessing myocardial infarction tissue characteristic, CMR faces significant barriers, including resource limitations, interpretation difficulties, and variability in diagnostic performance across different clinical contexts. These challenges include:

i) LGE-CMR Image Quality Variations [62], [63], [64]: These variations arise from several sources. Cardiac and respiratory motion during image acquisition can induce blurring artifacts, distorting the true morphology of the scar. Furthermore, inconsistencies in MRI hardware, acquisition protocols, and patient-specific factors can lead to variations in image noise and artifacts. Finally, the degree of gadolinium contrast enhancement can vary both between patients and within different regions of the same myocardium. This variability poses a challenge for consistent scar tissue identification. Furthermore, the intensity distribution of pathological myocardium is inherently heterogeneous.

ii) Anatomical Variability [64], [65]: Hearts exhibit significant variations in size, shape, and orientation across individuals. Moreover, scar tissue itself can have irregular shapes, varying sizes, and different transmural extents (from subendocardial to transmural), making them challenging to segment accurately.

iii) Subjectivity in Manual Segmentation:
Manual segmentation of cardiac structures in current clinical practice is labour-intensive and time-consuming process, prone to observer variability [66], [67]. Including LGE in MRI significantly extends the scanning time. The complexity of cardiac anatomy and physiology necessitates a high level of expertise among radiologists and cardiologists, which is often lacking due to insufficient training opportunities [68]. This shortage of qualified personnel can lead to delays in diagnosis and treatment, particularly in urgent cases.

iv) High Workload: The increasing demand for cardiac imaging has resulted in a substantial workload for radiologists, who face challenges managing the volume of cases within limited timeframes using traditional manual methods [69], [70]. Therefore, automated image analysis tools that assist in scar identification can

improve radiologist efficiency and potentially mitigate burnout in this high-stress environment

Given the inherent challenges in manual scar delineation from LGE-CMR images, automated segmentation and detection of myocardial scar tissue has become a critical area of research. This Computer-Aided Detection, particularly those employing deep learning techniques, offers the potential to address these challenges. These technological tools can play a significant role in early diagnosis, improved screening time, and optimized treatment planning, ultimately enhancing patient care and potentially mitigating the burden on clinicians.

2.3 Deep Convolutional Neural Network for Medical Image Optimization

Deep Convolutional Neural Networks (DCNN) have emerged as a powerful and widely adopted class of deep learning models, demonstrating state-of-the-art performance in various computer vision tasks, including image recognition, object detection, and medical image analysis[18], [71]. Characterized by their hierarchical and multi-layered architecture, DCNN is designed to extract and learn relevant features from input images automatically. While these networks have demonstrated remarkable learning capacity, especially in tasks involving large-scale datasets, early deep neural network research highlighted challenges in optimizing architectures and struggled to generalize effectively when exposed to complex object categories [72] . Given their potential, particularly in the medical domain, this section focuses on the optimization of DCNNs for medical image analysis. It covers key components including convolutional neural networks, architectural advancements, loss function selection, and modern hyperparameter tuning strategies to enhance model robustness, accuracy, and clinical applicability.

2.3.1 Convolutional Neural Networks

A fundamental building block of DCNN is the Convolutional Neural Network (CNN). processing structured grid-like data, such as images, by exploiting spatial correlations between pixels. Their design is inspired by the biological visual cortex, enabling them to automatically learn hierarchical feature representations from raw input

data without manual feature engineering. The core components of CNN are convolutional layers, pooling layers, activation functions, and fully connected layers [18], [71]. The convolutional layer forms the fundamental building block, employing small learnable matrices known as kernels (filters) that slide across the spatial dimensions of the input to produce feature maps. These kernels detect local patterns, such as edges or textures, in early layers and progressively capture more abstract features in deeper layers. By applying weight sharing and sparse connectivity, convolutional layers significantly reduce learnable parameters compared to fully connected networks, improving generalization and computational efficiency. Following convolution, pooling layers perform downsampling to reduce spatial dimensions while preserving the most salient information, thereby enhancing translation invariance, lowering computational demands, and mitigating overfitting. Pooling operations, such as max pooling or average pooling, must be carefully designed to avoid the excessive loss of small features like lesions. Non-linearity is introduced through activation functions, which enable CNNs to model complex input-output relationships. Functions such as the Rectified Linear Unit (ReLU) and its variants are widely adopted in shallow models for their computational efficiency, mitigation of the vanishing gradient problem, and faster convergence compared to traditional sigmoid or tanh functions. Finally, fully connected (FC) layers integrate features extracted by previous layers, enabling high-level reasoning and classification; the output is typically normalized via a Softmax function to produce probabilistic predictions. Although FC layers are highly discriminative, they are also parameter-intensive, and in shallow networks their size and number must be carefully minimized to maintain high accuracy while controlling overfitting and computational cost.

The strength of CNNs lies in their localized receptive fields, weight sharing, and hierarchical feature maps, which enable them to detect low-level patterns such as edges, corners in earlier layers and progressively more complex structures like shapes, textures in deeper layers [72]. These properties reduce the need for handcrafted feature extraction, allowing end-to-end learning optimized for the specific dataset and task. However, the relationship between depth and performance is not strictly linear. While deeper models can improve accuracy in challenging datasets, they also come with drawbacks, including higher computational costs, longer training times, increased memory requirements, and a higher risk of overfitting when data is limited [73]. Moreover, very deep networks can suffer from vanishing or exploding gradients, which

impair training stability unless mitigated by advanced architectural strategies such as batch normalization and skip connections.

Recent studies [72], [74] emphasize that optimal depth is task-dependent, as adding layers beyond a certain point may yield diminishing returns or even degrade performance. This is particularly relevant for medical imaging tasks, which also depend on task-based figures of merit, indicating that what is optimal depends not only on the network but also on the diagnostic or clinical task at hand. In such scenarios, shallow yet well-optimized CNN architectures can match or outperform deeper models by focusing on efficient feature extraction while avoiding overfitting.

2.3.2 Model Architecture

Segmentation is one of the most common and essential tasks in medical image analysis, and CNN-based encoder-decoder architectures have become the foundation of many modern segmentation networks. In these structures, an encoder is often used to extract image features while a decoder is often used to restore extracted features to the original image size and output the final segmentation results. Common encoder-decoder frameworks include Fully Convolutional Neural Networks (FCNNs), U-Net, SegNet and DeepLab variants. Among these, segmenting myocardial scars, especially within the complex and variable multiscale shape of the LV region, poses unique challenges that require more advanced multiscale feature aggregation strategies. Motivated by the strength of the ASPP module [75], [76], [77], [78] in capturing contextual features at multiple scales, this study explores the use of DeepLabV3+[79] for myocardial scar segmentation.

Originally developed as an extension of the DeepLab family, DeepLabV3+ builds upon the strengths of its predecessors, DeepLabV2 [80] and DeepLabV3 [81] by integrating a decoder module into the atrous convolution-based encoder architecture. DeepLabV3+ introduces a lightweight yet effective decoder module that refines the segmentation output, particularly along object boundaries. In parallel with architectural refinements, the DeepLab series has also evolved through successive improvements in backbone network integration. Early versions of DeepLab, such as DeepLabV1 and DeepLabV2, employed VGG-16 and ResNet-101 respectively. In DeepLabV3, ResNet-101 was retained as the default backbone, leveraging atrous convolution to extract dense, multi-scale features. However, the computational cost of applying atrous

convolutions over deeper residual blocks posed scalability limitations, especially when targeting finer resolution outputs. To address this, DeepLabV3+ introduced a more efficient backbone choice by adopting a Xception architecture to improve segmentation performance while significantly reducing computational complexity. Beyond the DeepLab series, numerous studies have explored the integration of other backbones in encoder. For U-Net, numerous variants have emerged that replace the original encoder with more powerful pre-trained backbones such as ResNet, MobileNetV2, VGG16 [82], [83], [84]. FCNN-based and SegNet architectures have also adopted advanced backbones to improve segmentation outcomes. For instance, EfficientFCN[85] utilizes a standard ImageNet pre-trained network (such as ResNet) as its encoder backbone, omitting dilated convolutions to reduce computational complexity and memory usage. In the SegNet architecture proposed by Badrinarayanan et al.[86], the encoder network leverages pre-trained weights from VGG16, which facilitates better generalization and accelerates convergence during training. These backbones are often adopted to leverage their superior feature representation capabilities, which improve segmentation accuracy.

Motivated by these insights, this study investigates task-specific backbone configurations within a dual-stage modified DeepLabV3+ framework for myocardial segmentation. The backbone networks were empirically evaluated alongside the original DeepLabV3+ backbone, Xception, to determine the optimal trade-off between segmentation accuracy and computational efficiency for each task. To the best of the author's knowledge, this is the first study to implement the DeepLabV3+ architecture specifically for myocardial scar segmentation using LGE-CMR images.

2.3.3 Loss Function

Loss functions play a crucial role in guiding the learning process, particularly in tasks that involve significant class imbalance, such as segmentation of small medical structures like scars, tumors, or lesions. In these scenarios, the foreground regions represent only a minor fraction of the entire image, resulting in the model biasing toward the dominant background class and causing false-negative predictions (high precision but low recall) [87], [88], [89], [90]. During the training phase of deep learning algorithms, backpropagation updates model parameters in accordance with the optimization goal defined by the loss function. While standard approaches such as

Cross-entropy [91] hold its own values, the quest for better performance has opened new doors for various modified loss functions. These modified solutions tackle specific limitations of traditional losses method in medical image segmentation tasks, focusing on enhanced robustness towards class imbalance.

Notably, generalized Dice loss addresses the sensitivity to foreground volume by incorporating weight penalty terms. Generalized Dice loss function (GDL) as formulated in Equation (2.1), is a class-specific weighting approach proposed by Sudre et al. [92] to control each class's contribution to the overall loss in segmentation tasks. Built upon the Dice similarity coefficient, which quantifies the overlap between predicted and ground truth segmentation, GDL accounts for the loss in similarity between the predicted probability, $P;j_k$ of class k at pixel (i,j), and the corresponding ground truth g_{ik} (binary) of class *k* at pixel (i,j). To address class imbalance, the weighting factor w_k for each class is computed as the inverse of the expected region area for class k, as defined in Equation (2.2), across a total number of classes (n). This formulation reduces the well-known dependency between region size and the DSC, thus ensuring that the model does not become biased toward larger regions

The GDL used in this study is given by:

$$GDL(P,g)=l-2 \frac{Zk\text{-}iW_kZijP_{ijk}g_{ijk}}{Zk\text{-}iw_kZijPi_{jk}+Zijg_{ijk}} \quad (2.1)$$

$$w_k=\frac{1}{\wedge i\ j\ \S ijk)} \quad (2.2)$$

Similarly, weighted cross-entropy [93] combats class imbalance by assigning higher weights to the underrepresented class during error calculation. The weighted cross-entropy (WCE) has been notably used in [93], [94] . It is a modification of the standard cross-entropy loss that addresses the issue of class imbalance in classification tasks. The WCE loss function is calculated by multiplying the standard cross-entropy loss for each class by its corresponding weight *(a_t).* This weight is typically determined based on the inverse frequency of the class, meaning that underrepresented classes are assigned higher weights. By penalizing misclassifications of underrepresented classes more severely, WCE encourages the model to focus on these classes and improve its

performance on imbalanced datasets. The WCE loss function used in this study can be expressed as Equation (2.3):

$$WCEL(P,g)= WCEL(p_t)=-o_t\log(p_t) \quad (2.3)$$

where- $= \begin{cases} P & \text{if } \S = 1 \\ (1\text{-}p & \text{otherwise} \end{cases}$ (1)

Focal loss [95] takes a different approach, dynamically reduces the weight of easily classified examples to focus on harder ones, leading to improved convergence. Focal loss function (FL) introduced by Lin et al. [95] is another loss function designed to address class imbalance in object detection tasks. It builds upon the cross-entropy loss by introducing a modulating factor that focuses learning on difficult examples. This factor downweights easy samples, allowing the model to concentrate on the more challenging ones, which are often from the minority class. The modulating factor in Focal Loss as in Equation (2.4) is controlled by a hyperparameter y, which determines the degree of downweighting for easy samples. In our comparative study, we adopted the proposed hyperparameter value of $y = 2$ for the Focal Loss function, as recommended by Lin et al.

$$FL(p_t)=-\log(1\text{-}p_t)^T\log(p_t) \quad (2\text{-}4)$$

Introduced by Salehi et al. [36], the Tversky loss function is an extension of the Dice loss function, which has become well-known in the field of medical image segmentation [38] because of its adaptability in managing class imbalance and imposing distinct penalties on false positives and false negatives via its a and P parameters based on specific requirements of the segmentation task. Tversky loss function is defined as Equation (2.5). This study proposes to use this loss function as in scar segmentation task, it is highly desirable to have a higher rate of false positives than false negatives in clinical settings, as missing a scar can lead to incorrect treatment decisions. By increasing p parameter , the penalty for false negatives is reduced, making the segmentation model more sensitive to scar tissue and minimizing the risk of overlooking these critical areas. This feature is particularly beneficial in scenarios where different tissues or structures possess varying degrees of importance.

$$Z_{n-l}^{t} \quad \text{TP} \quad \backslash \qquad (2\ 5)$$
$$\text{V} \sim \text{TP+aFP+BFN/}$$

Traditionally, choosing hyperparameters like a and P has been a manual, trial-and-error process. Researchers would train multiple models, each with a different hyperparameter setting, and compare their validation results. This manual tuning is common but often requires extensive trial-and-error to achieve optimal results [36], [37]. Most studies utilized the proposed hyperparameter values of a=0.3 and P=0.7 as suggested in [36]. However, some studies have demonstrated that Tversky hyperparameters strongly affect performance. For instance, in their research on stroke lesion segmentation, Hui et al. [96] systematically varied the parameter P from 0.5 to 0.95. They observed significant changes in the Dice score, with the optimal segmentation performance achieved at P = 0.8 and a =0.2 . In another study on segmentation of carotid artery calcification from panoramic radiographs in [97], the authors applied the Tversky loss function to mitigate the effects of class imbalance in segmentation tasks. The two weighting parameters, a and P, were adjusted manually through an ablation experiment in which their values were varied asymmetrically in increments of 0.1. In their study, the optimal Dice score for the dataset was at a = 0.4 and P = 0.6. While this manual tuning approach was systematic and provided insight into the trade-off between false positives and false negatives, it was constrained by coarse search resolution, high computational cost, and potential dataset-specific bias. The authors' findings underscore the influence of a and P on segmentation performance, and the study exemplifies the common practice of manual hyperparameter selection in medical image segmentation, highlighting the need for improvement, such as adopting automated optimization methods for more efficient and generalizable results.

2.3.4 PSO for Automated Optimization Techniques

In recent years, automated methods for parameter tuning in CNN, particularly using metaheuristic algorithms, have attracted significant attention. Among these, Particle Swarm Optimization (PSO) and Genetic Algorithms (GA) are widely adopted due to their ability to navigate complex, nonlinear, and multimodal search spaces efficiently. Inspired by natural processes, these methods iteratively improve candidate

solutions based on defined objective functions and have been successfully applied to optimize diverse CNN parameters, including the number of Network architectures [98], Learning Parameters [99] and specific application parameters [100], [101], [102]. The choice between these metaheuristic algorithms typically depends on factors such as the complexity of the hyperparameter search space, the availability of computational resources, and the desired model performance.

Empirical evidence suggests that PSO often demonstrates competitive or superior solution quality compared to GA and other metaheuristics in CNN hyperparameter optimization tasks. For instance, Ru et al. [99] investigated the prediction of rheological behavior in multi-walled carbon nanotubes and ethylene glycol nanofluids using machine learning models optimized with metaheuristic algorithms. Their results showed that PSO-tuned models achieved higher predictive accuracy and faster convergence than those optimized with other metaheuristics, including GA. The authors attributed PSO's advantage to its collaborative information, a common limitation in GA when population diversity is lost early. In a different domain, Sahu et al. [100] developed a hybrid Cellular Potts-PSO model for optimizing energy consumption and latency in edge computing environments. Their hybrid approach leveraged the spatial modeling strengths of Cellular Potts and the continuous-space optimization capabilities of PSO, resulting in superior performance over GA. The study demonstrated PSO's flexibility to integrate with domain-specific models, a concept transferable to deep learning contexts where PSO could be embedded within model training pipelines. In the medical imaging domain, Gangwar et al. [102]applied PSO to optimize enhancement parameters for an exposure region-based modified adaptive histogram equalization method in chest X-ray preprocessing. The PSO-optimized parameters yielded higher image contrast and improved diagnostic clarity compared to GA-optimized configurations.

From a theoretical perspective, PSO is a population-based metaheuristic optimization algorithm inspired by the social behavior of birds flocking or fish schooling [100], [103]. It operates by having a swarm of particles, where each represents a candidate solution. The algorithm iteratively updates each particle's position in the search space based on both its own historical best position and the best position identified by the swarm. At iteration t, the velocity of the i-th particle in the *j*-th dimension, denoted as $v_{;j}(t)$, is updated according to Equation (2.6) [98]

$$v;j(t+1)= w \times v_{lJ}(t) + c_p \times r_p \times (pBest_{1;}j - x_{1;J}(t)) \quad (^{2}\text{-}^{6})$$

$$+ c_g x r_g x (gBest_j - x_{1J}(t)),$$

where x is the current particle position, w is the inertia weight controlling the influence of the previous velocity, c_p and c_g are cognitive and social acceleration coefficients, respectively, r_p and r_g are uniformly distributed random values in (0,1), $pBest_{i}j$ represents the best position previously achieved by the t-th particle, and *gBestj* denotes the best position discovered by the entire swarm in dimension *j*.

Subsequently, the position of the t-th particle in they-th dimension is updated as in Equation (2.7) [98] . This update mechanism ensures that particles progressively converge toward promising regions of the search space.

$$x_{ij}(t+1) = x_{ij}(t) + v_{ij}(t+1) \quad (2.7)$$

The general workflow of PSO is illustrated in Figure 2.8 [104]. The PSO update mechanism balances exploration and exploitation in a simple yet effective manner. The process begins with random initialization of particle positions and velocities, followed by iterative evaluation of fitness, updating of personal and global best solutions, velocity and position updates, and termination once convergence criteria are met. It requires fewer parameters to configure than GA and avoids the use of complex genetic operators such as crossover and mutation [98] . This simplicity, combined with PSO's collaborative search strategy, often results in faster convergence and competitive optimization performance. While most existing works have applied PSO to optimize network architecture parameters, learning rate schedules, or image preprocessing settings, they have rarely addressed loss function hyperparameters within the classification layer of a CNN. To the best of current knowledge, no prior study has employed PSO to automate the hyperparameter tuning of the loss function in the classification layer specifically for the Tversky loss function (a and P) in medical image segmentation or classification tasks, which motivates this study to adopt PSO as the automated method for hyperparameter tuning.

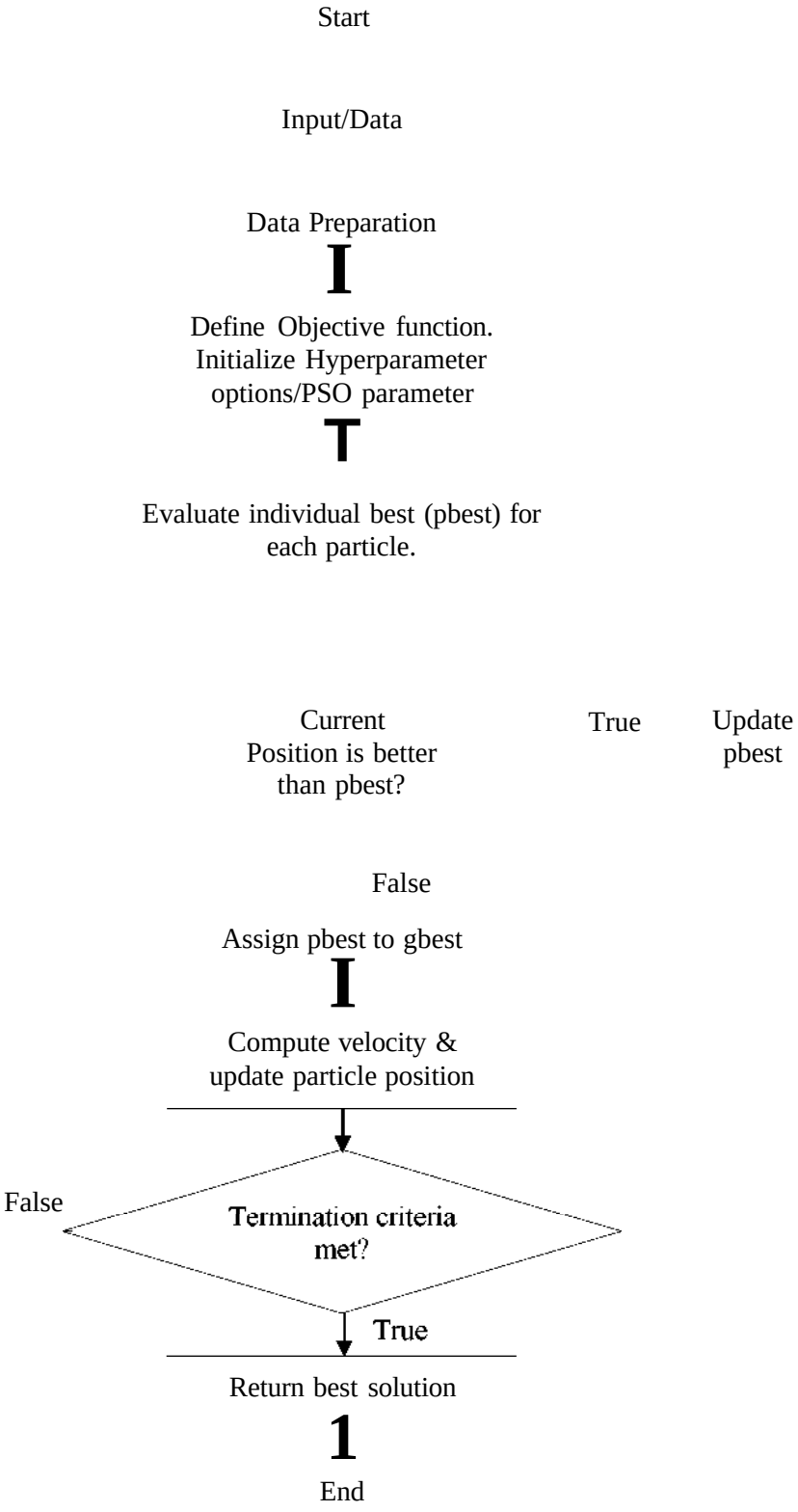


Figure 2.8 The Flowchart of the PSO Algorithm [104]

2.4 Review of Past Research in Automated Segmentation and Detection of MI

The limitations of manual MI assessment from LGE-CMR imaging, as discussed in the Section 2.2, highlight the need for automated solutions. Numerous automated segmentation and detection method and analysis have been successfully applied for MRI medical image applications. This Section is structured to first address traditional and conventional methods in Subsection 2.4.1, followed by a comprehensive overview of the deep learning-based methods for Myocardial Scar Segmentation approaches in Subsection 2.4.2. Thus, the limitations of existing methods and the research gaps motivating the present work, particularly concerning the use of single LGE modality input for deep learning models, will be summarized. Lastly the detection method will be discussed in 2.4.3.

2.4.1 Conventional Techniques for Myocardial Scar Segmentation

The segmentation of left ventricular (LV) myocardial scars from LGE-CMR remains a critical area of research, with ongoing advancements in algorithmic approaches underscoring its essential role in future clinical developments. Early efforts to delineate the scar regions accurately predominantly employed rule-based algorithms that leveraged pixel intensity thresholds to identify scarred regions. Notable techniques include the Full-Width at Half-Maximum (FWHM) method [105] and the n-standard deviation approach [106], [107], which defines infarcted tissue based on signal intensities deviating from the mean of healthy myocardium by a fixed number of standard deviations (n = 2 , 3 , 4 , 5 , 6). Clustering algorithms such as k-means and fuzzy c-means were also employed to group pixels based on signal similarity. Yet, these intensity-based segmentation techniques struggle to address the heterogeneity in LGE-CMR images, where non-scar cardiac structures may also appear enhanced. Moreover, many rule-based techniques require manual pre-segmentation of the myocardium to narrow down the region of interest, reducing scalability and increasing the burden on clinicians. This dependence on manual input significantly hinders their integration into fully automated clinical workflows. Thus, these limitations have led to a growing interest in deep learning techniques, which offer the ability to learn hierarchical features directly from data, eliminate the need for handcrafted inputs, and enable end-to-end segmentation frameworks.

2.4.2 Deep Learning-Based Methods for Myocardial Scar Segmentation

Deep learning-based algorithms excel at automatically extracting complex features from data for object detection and segmentation. These features are learned directly from data through an end-to-end training process using general-purpose optimization techniques, bypassing the need for manually crafted feature extraction rules. This makes deep learning-based algorithms easy to apply in various image analysis applications. Recent studies have shown that CNNs have the potential to segment medical images in a variety of ways [22]. However, in a comprehensive review from Chen et al. [18] on deep learning for cardiac image segmentation, noting that a majority of works have focused on segmenting cardiac chambers or regions such as the left ventricle (LV), right ventricle (RV), and left atrium (LA). In contrast, relatively fewer studies have focused on segmenting myocardial scars, which are fibrotic regions formed as a result of myocardial infarction, using deep learning approaches applied to LGE-MRI of the left ventricle. This is likely due to the limited relevant public datasets as well as the challenging task associated with accurately segmenting myocardial scars due to variability in the size, heterogeneity, and intensity of scar in the image itself.

Precise LV segmentation is critical as it provides a reliable foundation for scar segmentation. During the early 21st century, radiologists often relied on cross-referencing LGE-CMR with cine MRI to manually delineate the myocardium. This practice inspired the development of computational methods that integrated other image modalities such as cine CMR, bSSFP CMR and T2 CMR for LV myocardial segmentation [20], [27], [108], [109], [110]. While some of these approaches demonstrated promising accuracy, their reliance on multiple data sources introduced several challenges, including increased model complexity, higher computational demands, and a greater risk of misalignment between modalities. These limitations have prompted researchers to shift their focus toward utilizing single-modality LGE-CMR data, which is widely regarded as the primary imaging modality for myocardial scar assessment. Furthermore, following the breakthroughs in CNN architectures for computer vision, most researchers aimed to develop efficient and practical solutions that ensured high accuracy while simplifying the segmentation process for clinical and research applications.

Several studies have explored whether a single-network approach or task-specific (multi-network) models yield better results for complex segmentation tasks.

Moccia et al., in [28] and [33] were among the first to explore this area, utilizing a Fully Convolutional Neural Network based on the modified Efficient Neural Network (Enet) architecture. Both studies utilized a total of 250 slices short-axis LGE-CMR images from 30 patients with ischemic cardiomyopathy. In the 2018 study, the direct scar segmentation from cropped and resized 64x64 LGE-CMR images achieved a Dice score of 55% using a single model configuration. In their 2019 follow-up, the authors further supported this observation by showing performance improvements when restricting segmentation to predefined anatomical regions. Their results demonstrated that performance improved when segmentation was restricted to the manually predefined anatomical regions, with the Dice score increasing from 54% to 71.25%). However, a key limitation in both studies is their reliance on manually cropped or segmented LV regions as input, hindering the realization of a truly end-to-end solution. Additionally, neither study incorporated an initial LV slice classification mechanism to exclude non-relevant slices, a necessary step in clinical settings where entire MRI volumes are presented.

Sharma et al. [34] reported a low Dice score of 25%> when using a standard U-Net with a geometric transformation's augmentation technique for scar segmentation in LGE-MRI images, highlighting its poor performance in detecting subtle, low-contrast infarct areas. To address these challenges, Sharma et al. proposed a two-stage chained U-Net architecture, where the first network segments the LV myocardium and the second focuses on infarcted regions within the previously segmented myocardium. They reported a low DSC of 32%> suggesting further refinement and optimization. While the concept of chained U-Nets is novel, the resulting scar segmentation performance remains low. The authors also noted that the public dataset's training set lacked sufficient examples with visible infarction and included cases where the annotated ground truth scar regions were not visibly distinguishable in the corresponding images. Similarly, Jani et al. [32] investigated the benefits of a cascaded U-Net approach, where LV myocardium segmentation is followed by scar segmentation. Their results demonstrated an improved Dice score for infarcted regions compared to single-network U-Net, confirming that anatomical pre-localization can enhance segmentation performance.

Chen et al. [23] pioneered early work on end-to-end myocardial scar segmentation from LGE-MRI. Their work proposes two frameworks using modified U-Net architecture, one incorporating pre-segmented left ventricular myocardium and

another without such anatomical guidance. While achieving relatively low Dice scores, their work demonstrated the potential of incorporating pre-segmented myocardial boundaries, which led to improved scar segmentation accuracy compared to the lower performance observed when this anatomical guidance was not included. A limitation of their work is the suboptimal performance, reflected in a low Dice score of 67%. This highlights the need for further improvements in segmentation accuracy. Notably, their results were calculated only on the middle slices, excluding the most basal and apical slices. This exclusion may have contributed to an overestimation of the model's performance, as these more complex LV regions were not fully considered in the evaluation.

Another study from Fahmy et al. [111] presented an initial proof-of-concept study to quantify LV myocardium and scar segmentation using the U-Net architecture on LGE-CMR images. The U-Net architecture is built upon the FCN with skip connections between the downsampling path and the upsampling path, applying a concatenation operator instead of summation using concatenation rather than summation to preserve spatial information. The study reported a Dice score of 82% for LV identification and 57% for scar segmentation.

In summary, these studies reveal that single-stage models often struggle with low performance for myocardial scar segmentation due to large class imbalance and limited contextual awareness. In cardiac MRI (e.g. LGE-CMR), myocardial scar regions are often tiny and have subtle intensity differences from surrounding tissue, making them hard to distinguish. Scar tissue typically occupies less than 10% of the myocardial volume, creating a severe bias toward dominant classes like the LV and myocardium [19]. In contrast, task-specific or multi-stage frameworks have demonstrated improvements in scar segmentation accuracy, although they still face challenges in low accuracy, demonstrating potential for improved performance by leveraging prior anatomical localization. Moreover, robust architectural models are required due to several factors, such as variations in LV size and heterogeneous intensity distribution that further complicate segmentation accuracy [18], [19], [119]. The presence of infarcted scars adjacent to cavities or blood pools leads to low contrast between the LV myocardium and surrounding tissues, making accurate boundary delineation difficult.

In addition to these challenges, current studies have not addressed an essential processing step, which automatically excludes images that do not contain the LV region.

In clinical practice, radiologists receive MRI scans that include a range of slices, some of which do not feature the LV as stated in Section 2.2. Manual preselection of slices containing the left ventricle remains a standard practice in many myocardial scar segmentation pipelines using LGE-CMR. Most existing studies begin by manually selecting slices that contain only the relevant LV slices, requiring radiologists to identify and crop these regions beforehand. This not only increases workload but introduces subjectivity and inefficiency, undermining the goal of end-to-end automation. To achieve a fully automated segmentation approach, this initial exclusion step must be systematically integrated into the workflow to ensure the model processes only relevant images. Table 2.1 summarizes past literature on the deep learning approach for segmentation of myocardial scar on LGE-CMR images.

Table 2.1
Overview of Prior Deep Learning Approaches for Segmenting Myocardial Scar Regions in LGE-CMR

Author (year)	**Dataset /MRI modalities**	**DCNN Model**	**Performance**	**Advantage/Disadvantage/Gap**
A. Fahmy et al. (2021) [27]	LGE CMR Cine CMR	U-Net	Accuracy (successfully segmented % of total slices) = 95%	High performance Scar segmentation. Segmentation on manually cropped LGE and Cine images. Two U-Net were trained for myocardium and scar segmentation for LGE-Cine image fusion.
Y. Liu et al. (2020) [20]	T2-weighted CMR bSSFP LGE CMR	U-net	Average Dice score = 57%	M-shape network based on U-net was trained on bSSFP images to locate LV myocardium. Greater risk of misalignment between modalities.
W.Ding et. al. (2023) [109]	bSSFP LGE CMR	U-MyoPS	Average Dice score : 64.7%	Model complex, higher computational demand
M.Chen et al. (2019) [110]	LGE CMR Cine CMR	DNN	Accuracy = 87.6% precision = 86.2%.	LV myocardium image information from the cine CMRI and motion information described by the optical flow were fused with LGE and fed into the stack denoising autoencoder (SDAE) with a support vector machine (SVM) to assess myocardial scar.
Moccia et al. (2018) [28]	LGE CMR	E-Net	Scar segmentation Average Median Dice score (single E-net) =55%	Automated Scar segmentation on manually cropped LV myocardium region. Low performance Scar segmentation.
Dan et al (2022) [112]	LGE CMR	ACSNet	Average Dice score : 57%	Low performance Scar segmentation.

Author (year)	Dataset /MRI modalities	DCNN Model	Performance	Advantage/Disadvantage/Gap
Moccia et al. (2019) [33]	LGE CMR	FCNN- E-Net	Protocol 1 (single E-net) : Scar segmentation- Average Median Dice score = 54% Protocol 2 (manual LV myocardium segmentation) : Scar segmentation- Average Median Dice score =71.25%	Protocol 1: Low performance Scar segmentation. Input is manually cropped LV images. Protocol 2: Input is manually identifying LV myocardium images.
Sharma et al., 2020 [34]	LGE CMR	Chained two U-Net	Scar Average Dice score: Single Unet =25% Chained Unet = 32%	No LV slice classification step to exclude irrelevant slices before segmentation. Low performance of Scar segmentation.
Jani et al. 2024 [32]	LGE CMR	U-Net and Cascaded U-Net	Scar Average Dice score: Single Unet = 46% Cascaded Unet = 50%	No LV slice classification step to exclude irrelevant slices before segmentation. Low performance of Scar segmentation.
Chen et al.(2022) [23]	LGE CMR	Modified Unet	Scar segmentation Average Dice score = 67%.	No LV slice classification step to exclude irrelevant slices before segmentation. Results were calculated on the LV middle slices only, excluding the most basal and apical slices. Low performance of Scar segmentation.
Fahmy et al. (2018) [111]	LGE CMR	Unet	Scar segmentation Average Dice score = 57%.	Low performance of Scar segmentation.
Felix Lau et al. (2018) [29]	LGE CMR	GAN + U-Net	Scar segmentation improves Median Dice score = 75.9% to 80.5%.	Employing GAN to augment scar data. Manually crop myocardium LV region.

2.4.3 Detection of Myocardial Infarction from Sequences of LGE-MRI

The presence of myocardial scars typically visualized in LGE-CMR are strong indicator of prior MI. Consequently, the segmentation of scar tissue has been widely used as an indirect method for identifying MI in previous research [23], [107], [113]. However, the risk of false segmentation that often due to low contrast, heterogeneous intensity distribution, and imaging artifacts present in LGE-CMR images can lead to the misclassification of healthy tissue as infarcted regions, thereby reducing the reliability of MI detection systems. A common approach in previous research is to use voxel count thresholds to refine segmentation outcomes or the presence and absence of segmented scar tissue as a decision rule to classify each case as either non-MI or MI [39], [40]. For instance, Ma et al. [114] proposed post-processing method that, if the segmented scars are absent or contain fewer than 10 voxels after post-processing, the case is classified as non-MI; otherwise, it is labeled as MI. The classification achieved an accuracy of 92% in distinguishing between normal and pathological cases based on segmentation results. However, the specific voxel thresholds used to define infarct presence vary across studies, influenced by factors such as image resolution and the spatial representation of scar tissue.

De la Rosa et al. [115] aims to develop an automatic method for MI detection and quantification from LGE-MRI using a cascade approach. Their method consists of three main stages. In the first stage, LGE-MRI slices are classified as either healthy or pathological using a fine-tuned VGG19 network as a feature extractor, followed by a Support Vector Machine (SVM) with a linear kernel for classification. In the second stage, scar segmentation is performed using morphological operations to obtain an initial infarct mask after image enhancement method using non-parametric top-hat transform. The third stage involves refining this segmentation through an ensemble of CNNs, which perform boundary-voxel reclassification. A key limitation of their approach lies in the reliance on traditional morphological operations for the initial segmentation, which may be sensitive to image quality, noise, and the heterogeneous appearance of scar tissue. Furthermore, the classification of slices as either healthy or pathological is performed as a separate preliminary step before segmentation, rather than being integrated into an end-to-end framework. This separation of tasks, combined with the need for manual configuration of enhancement and morphological parameters, makes the pipeline semi-automated rather than fully automated. While the method

demonstrated promising results, achieving a classification accuracy of 90.63% and an average Dice coefficient of 77%, further evaluation on larger and more diverse datasets is needed to assess the generalizability and robustness of their method.

Abdulkareem et al. [108] investigated the possibility of predicting post-contrast cardiac MRI (CMR) information, specifically myocardial scar presence using pre-contrast cine CMR images and machine learning techniques. The classification component of the study employed feature engineering to capture relevant information from the cine short-axis images, including optical flow to myocardial motion patterns, rate of change of myocardial area, and radiomics data to extracted texture and shape features in order to distinguish between patients with MI and healthy controls. LGE-MRI was used as the ground truth to establish the presence of myocardial infarction (MI). The patterns of gadolinium enhancement on LGE-MRI were used as the reference standard for confirming the presence of MI, with patterns of gadolinium enhancement indicating necrosis and fibrosis. The study compared deep learning (ResNet50) with traditional machine learning approaches (SVM and Decision Tree). Despite these efforts, the results showed only modest performance, with SVM achieving the highest classification accuracy at 68%.

Several recent studies have explored 3D deep learning approaches for myocardial scar segmentation from LGE-CMR images. Brahim et al. [30] introduced a 3D U-Net combined with a 3D Autoencoder to enforce anatomical shape constraints for improved scar segmentation, achieving a Dice score of 76.56%, though excluding basal and apical slices due to common artifacts and lower scar prevalence in these regions. Zabihollahy et al. conducted a progressive series of works beginning with semi-automated 2D and 3D CNNs trained on patch-based input, reaching Dice scores above 93%, but requiring manual myocardium delineation and facing high computational demands [26],[116]. They later proposed [117] a fully automated cascaded multi-planar U-Net (CMPU-Net), which processes axial, sagittal, and coronal views, achieving Dice scores of 85.14% for myocardium and 88.61% for scar segmentation. Fahmy et al. [118] applied a deep U-Net with 150 layers for 3D scar quantification but reported a lower Dice score of 54%. Overall, although some 3D segmentation methods demonstrate high performance in detecting myocardial scars, they often incur high computational costs, depend on labour-intensive manual annotations, and primarily focus on segmentation accuracy.

From the reviewed studies, most existing methods detect MI using a single segmented scar slice and the post-processing method using voxel-based thresholds, which does not accurately reflect clinical practice, where radiologists assess the continuity of the scar across multiple slices to confirm infarction. Although some studies have attempted to incorporate multi-slice information, research in this area remains limited. A fully automated, robust, and generalizable MI detection systems still require further development and validation. Table 2.2 summarizes past literature on the MI detection approach on LGE-CMR images and highlights the need for sequential slice analysis.

Table 2.2
Overview of Prior Myocardial Infarction Detection Approaches and the Need for Sequential Slice Analysis

Author (year)	**Method Type**	**MI Detection Approach**	Performance	**Limitations / Gap**
Maetal. [114]	**2D**	Post-processed scar segmentation. Voxel count threshold (<10 voxels = non-Mi)	Accuracy: 92%	Voxel count threshold varies across studies. Don't assess the continuity of scar.
De la Rosa et al. [115]	**2D**	MI Slice classification & segmentation	classification accuracy of 90.63%	Slice classification done separately, semi-automated: Handcrafted segmentation.
Abdulkareem et al. [108]	**2D**	Feature-engineered ML (SVM) from pre-contrast cine	Accuracy: 68%	Indirect MI detection from cine, lacks anatomical segmentation.
K.Brahim et al. (2021)	3D	Scar segmentation only	Average Dice score: Myocardial scai= 76.56%	Input segmentation on manually cropped LGE images. SegU-Net model based on the fusion of two deep learning segmentation techniques (U-Net and Seg-Net) and with exclusion of basal and apical slices from the evaluations.
Zabihollahy et al. (2018) [26]	3D	Scar segmentation only- 2D CNN (4 conv + 2 FC layers, patch-based)	Average Dice score: scar= 94.5%	High accuracy, limited to manually segmented myocardium, semi-automated. lacks integrated MI detection
Zabihollahy et al. (2018) [116]	3D	Scar segmentation only- 3D CNN (patch-based,AdaDelta optimizer)	Average Dice score: scar= 93.63%,	Improved 3D context, still semi-automated; high computation cost. lacks integrated MI detection

Author (year)	Method Type	MI Detection Approach	Performance	Limitations / Gap
Zabihollahy et al.(2020) [117]	3D	Scar segmentation only - CMPU-Net (multi-planar U-Nets: axial, sagittal, coronal)	Average Dice score Scar=88.61%	lacks integrated MI detection no detection/classification with healthy images and complex pipeline/computationally intensive.
Fahmy et al. [118]	3D	Scar segmentation only	Average Dice score Scar : 54%	Focused on segmentation only, no sequential diagnostic logic. Low performance Scar segmentation

2.5 Summary

MI remains a leading cause of morbidity and mortality worldwide, with early and accurate detection being critical for improving patient outcomes. The identification of myocardial scar tissue serves as a key biomarker for diagnosis, treatment planning, and prognosis assessment. While several diagnostic tools, such as ECG and echocardiography, play important roles in MI detection, LGE-CMR has established itself as the gold standard for non-invasive myocardial tissue characterization, offering high spatial resolution and excellent contrast between healthy and infarcted myocardium. However, manual interpretation and delineation of scar tissue from LGE-CMR images are time-consuming, highly dependent on expert experience, and subject to inter-observer variability, which limits efficiency and reproducibility in clinical practice. Furthermore, variability in image quality, anatomical differences, and the subtle or heterogeneous appearance of scars make manual segmentation challenging, while growing clinical workloads exacerbate the problem. These limitations underscore the urgent need for automated, robust, and efficient computational methods

DCNNs have become central to medical image analysis due to their ability to automatically extract hierarchical features from raw data, reducing the need for manual feature engineering. Their architecture, which comprises convolutional layers for localized feature extraction, pooling layers for spatial reduction, activation functions for non-linearity, and fully connected layers for classification, has delivered strong performance in segmentation, detection, and classification tasks [72], [73]. However, evidence from prior studies shows that simply increasing network depth does not necessarily lead to higher accuracy [72], [74], suggesting that simpler and more efficient architectures are needed. Additionally, the challenge of medical imaging segmentation, including for myocardial scar segmentation, is often accompanied by severe class imbalance [87], [88], [89], further complicating these efforts. This imbalance suggests careful selection of loss functions to mitigate bias toward dominant classes and improve sensitivity to small pathological regions [91], [92], [93], [94], [95]. Furthermore, conventional manual tuning of network hyperparameters, including those in loss functions, is labor-intensive and prone to dataset-specific bias [35], [36]. Metaheuristic algorithms such as PSO have shown great promise in overcoming these limitations, offering efficient, automated, and robust parameter optimization for improved performance [97]-[102].

Recent deep learning-based algorithms have demonstrated substantial potential for cardiac imaging tasks, yet relatively few studies have focused specifically on myocardial scar segmentation from LGE-CMR [18],[19]. Existing research shows that single-stage models often underperform in this domain due to severe class imbalance and limited contextual awareness, as scar regions in LGE-CMR are typically small and exhibit subtle intensity differences from surrounding tissues [19],[26],[30],[32]. This imbalance biases models toward dominant classes like the LV and myocardium, while additional factors such as anatomical variability in LV size, heterogeneous intensity distribution, and the close proximity of infarcted scars to blood pools further complicate accurate boundary delineation [18], [19], [118]. Although multi-stage frameworks have shown improved accuracy by incorporating prior anatomical localization, significant opportunities remain to optimize performance further in a fully automated manner. Notably, the review reveals that current pipelines often rely on manual preselection and cropping of LV-containing slices, which increases workload, introduces subjectivity, and hinders the goal of full automation.

Furthermore, most existing MI detection methods base their decision on a single segmented scar slice or apply post-processing voxel-based thresholds. This approach does not reflect clinical practice, where radiologists evaluate scar continuity across multiple slices to confirm infarction. While some studies have attempted to integrate multi-slice information, research in this area remains limited. Addressing these challenges forms the central motivation of this study, which proposes a novel framework to deliver more reliable, scalable, and clinically aligned MI detection from LGE-CMR.

CHAPTER 3
RESEARCH METHODOLOGY

3.1 Introduction

The increasing incidence of myocardial infarction (MI) has driven recent studies to focus on automating MI detection, with the aim of reducing radiologists' workload and improving the efficiency of clinical diagnostics. However, relatively few researchers have focused on achieving fully automated MI detection using short-axis Late gadolinium enhanced cardiac magnetic resonance (LGE-CMR) images [18], [19]. This limited focus can be attributed to several significant challenges, including the scarcity of high-quality annotated datasets and the complexities in accurately segmenting infarcted tissue. To fill the research gaps outlined in Chapter 2, this chapter introduces a novel method for fully automated segmentation and detection of infarcted tissue, commonly referred to as myocardial scar, associated with ischemic heart disease. The proposed approach is based on a deep convolutional neural network (DCNN) framework and is applied exclusively to sequences of LGE-CMR images. This chapter begins with a description of the image acquisition and data collection, pre-processing and expert labeling procedures that are essential for training the proposed DCNN models. It then introduces a novel DCNN-based method for classifying images into LV and non-LV regions, facilitating automatic detection. Following this, a dual-stage DeepLabV3+-based segmentation model is introduced, structured into two key stages: the DeepLabV3+ Tailor LV (DLT-LV) model for segmenting the LV region in the initial stage, followed by the DeepLabV3+ Tailor Scar (DLT-Scar) model focused on segmenting scar tissue within the LV region. Finally, an approach for detecting cases indicative of MI is introduced, utilizing Scar Sequential Slice Reconstruction (3SR) in conjunction with a DCNN-based method. Each of these methods is systematically detailed in its separate subsection.

3.2 Overview of Research Methodology

The primary goal of this research is to provide an automated end-to-end framework for MI segmentation and detection using only LGE-MRI images. This

section provides an overview of the proposed methodology, as illustrated in Figure 3.1, with the main contributions highlighted in red blocks to align with the study's research objectives. The overall process is divided into four (4) main phases. Phase 1 involves data collection and preparation, starting with obtaining ethical approval and gathering the data. This phase involves removing annotations, labeling ground truth (GT) by experts, and augmenting images before proceeding to the next phase. The details of this phase are described in Section 3.3.

The second phase involves the classification of images containing the left ventricle (LV) and those without the LV. The LV, which functions as the heart's primary pumping chamber, is examined from its basal to apical regions to identify potential areas of infarction. LGE-CMR imaging acquires a series of short-axis slices of the heart, covering the outermost parts and the entire cardiac segment. However, when assessing the patient's myocardial scar, only the slices that include the LV region are relevant. To streamline this process, this study proposes an automated approach based on DCNN. The goal is to accurately classify LGE-CMR images to isolate those containing the LV region from non-LV images, thereby preparing them for a subsequent segmentation process. The proposed classification method utilizes both a shallow DCNN model architecture (S-DCNN) and pre-trained DCNN networks. Phase 2 corresponds to Objective 1, which is detailed in Section 3.4.

Once the images containing the LV region have been identified, the third phase introduces a proposed automatic dual-stage myocardial scar segmentation model, where each stage is optimized for a specific segmentation task using a customized DeepLabV3+ segmentation network. In the first stage, the DLT-LV model is used to segment the LV region. In the second stage, the DLT-Scar model focuses on segmenting scar tissue within the previously segmented LV region. On top of that, the study proposed a morphological procedure after first-stage segmentation that links the DLT-LV and DLT-Scar to further refine the LV region segmentation, addressing potential issues such as small gaps within the LV chamber and isolated miss-segmentations after the DLT-LV model. This dual-stage segmentation model corresponds to objective 2, and the development processes for both stages of the segmentation model are discussed in detail in Section 3.5. Following the second-stage scar segmentation, a novel optimization approach is introduced, integrating Particle Swarm Optimization (PSO) with Tversky loss function hyperparameter tuning within the DLT-Scar model to

enhance myocardial scar segmentation performance. This phase relates to objective 3, with details of the proposed optimization method provided in Subsection 3.5.2.2

Finally, the fourth phase continues with the detection of truly segmented scar tissue identification, followed by an overall evaluation process. While the presence of myocardial scars in LGE-CMRI is strongly indicative of prior MI [1], using a single segmented LGE-CMRI image alone is insufficient for reliable scar detection due to limitations in spatial context and variability in scar appearance across slices. To address this, this phase introduces a methodological approach that reconstructs an image from a sequence of LGE-CMR slices, thereby enhancing the visualization of scar features across multiple views. The S-DCNN model, as detailed in Section 3.4.1, is then employed to classify whether the segmented scar image is indicative of MI or represents non-MI myocardial tissue. This phase concludes with an overall evaluation assessing the proposed frameworks. This stage corresponds to objective 4 and detailed methodology is provided in Section 3.6.

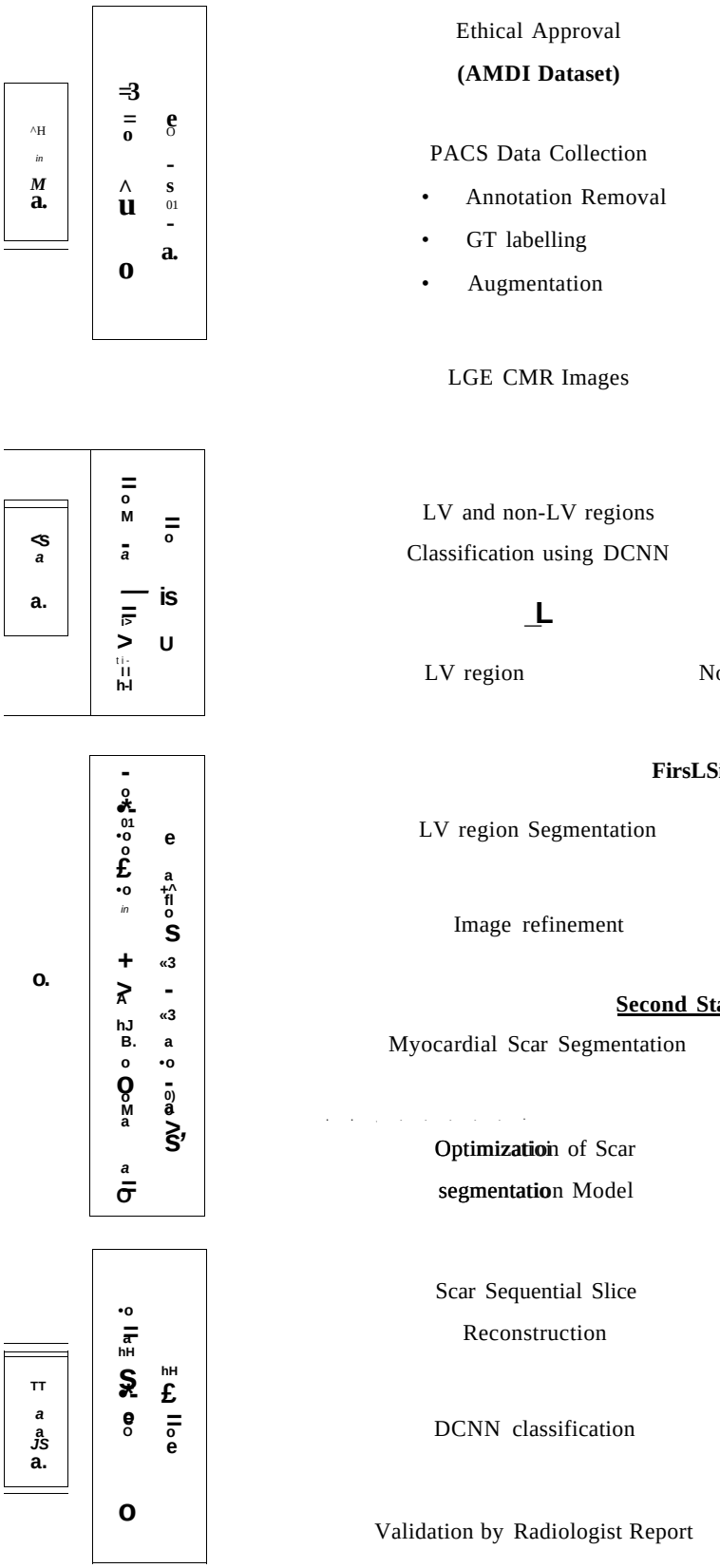


Figure 3.1 Overall Proposed Methodology for MI Segmentation and Detection Model. Red Blocks Indicate the Main Contributions in This Study

3.3 Phase 1: Data Collection and Preparation

This section details the initial data preparation steps carried out in Phase 1 of the study. It includes the data collection procedure, expert labeling, and augmentation processes, as illustrated in Figure 3.1. The general flow of the dataset acquisition and data labelling is illustrated in Figure 3.2. The process begins with obtaining ethical approval, followed by image acquisition and preprocessing steps, before proceeding to ground truth (GT) annotation, which was guided and validated by an expert radiologist. Detailed descriptions of each step are provided in the subsequent subsections. Detailed explanations of each step are provided in the following subsections.

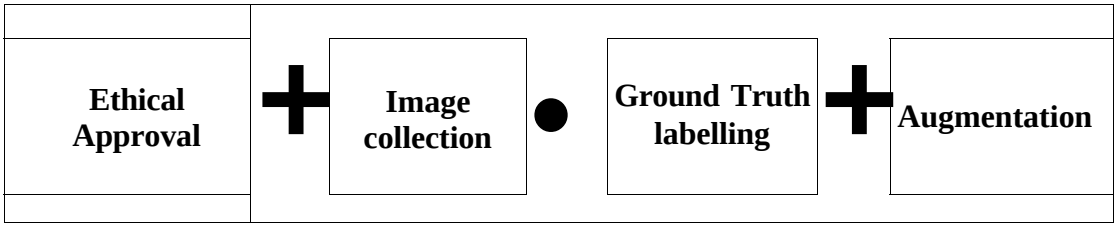


Figure 3.2 The General Flow for Data Collection and Preparation

3.3.1 Data Collection

In order to utilize LGE-CMRI data from patients, this research obtained ethical approval on 22 November 2021 from the Human Research Ethics Committee of USM (JEPeM) under the School of Medical Sciences, USM, IPPT, Bertam, Pulau Pinang. The approval was extended until February 2025 under JEPeM Code: USM/JEPeM/21090623 (See Appendix 1). The flow of the ethical approval process applied to JEPeM-USM is detailed in Appendix 2. All LGE-CMR images used in this study were acquired using a Philips Achieva 3.0T CMRI scanner. The imaging data were retrieved in Digital Imaging and Communications in Medicine (DICOM) format from Picture Archiving and Communication System (PACS) using OsiriX® software (ver. 12.5.2) on an Apple Mac workstation at Imaging Department, AMDI, USM Bertam. LGE-CMR images were retrospectively collected from patients diagnosed with MI with scar-related ischemic heart disease between 2018 and 2024. Short-axis LGE-CMR images with a resolution of 320 × 320 pixels were obtained from a total of 60 patients. For training and testing purposes, the study included a total of 30 patients, comprising 21 individuals with MI and 9 healthy cardiac subjects. An additional 30

patients were involved in the final validation phase to assess the performance of the proposed end-to-end automated MI detection framework, which included 10 subjects with MI and 20 healthy subjects without MI. Each image was accompanied by its corresponding MRI radiology report, providing clinical context for each case. An example of a radiology report is included in Appendix 3. DICOM images often contain patient information, dates, and other metadata unnecessary for the proposed method. To comply with ethical regulations that have been used, these images were anonymised to remove confidential patient information before being transferred to the study domain for storage. Many MATLAB image-processing functions operate on standard numeric image arrays rather than directly on DICOM files, so this study utilizes the *dicomread* function to read DICOM images and then convert them to Bitmap (BMP) format to facilitate the processing workflow in the MATLAB environment. BMP is an uncompressed, lossless image format that preserves pixel intensities accurately, enabling reliable image processing and visualization.

3.3.2 LV and Non-LV Region Labelling

The number of images obtained per patient varied depending on the slice settings and the anatomical size of the patient's heart. This study used MRI scans with an 8 mm slice thickness, which typically produced between 15 and 20 short-axis images for each individual. These images consist of both those containing the LV region, spanning from the basal (base) to the apical of the heart, and those that do not contain the LV region which located below the apical namely non-LV Apical (NLVA) and images that do not contain the LV region located at upper basal namely non-LV Basal (NLVB), as illustrated in Figure 3.3. In this study, both types of non-LV images were grouped under a single category referred to as the non-LV region. The LV region, on the other hand, refers to image slices that contain anatomical structures within the boundaries of the left ventricle, including the myocardial wall, papillary muscles, and the LV cavity, as illustrated in Figure 2.6 of Chapter 2. The labeling task for LV and non-LV region classification was conducted manually by the cardiac radiologist through visual assessment. Figure 3.5 displays LGE-CMR images of LV and non-LV regions based on soft tissue density. Since myocardial scar assessment is primarily focused on the LV region, a classification method is required to distinguish LV from

non-LV regions at the beginning. The proposed method for automated LV and non-LV classification is detailed in Section 3.4.

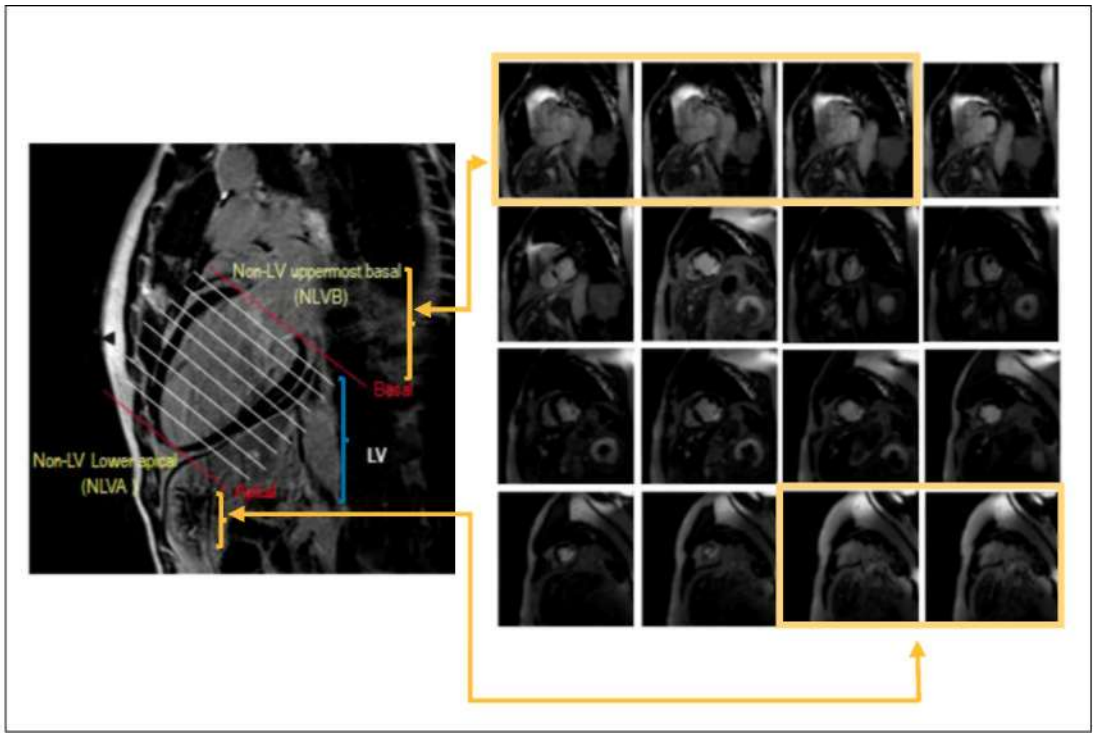


Figure 3.3 Example of LGE-CMR Image Sequences Per Patient. Uppermost Basal and Lower Apical Images Comprise the Non-LV Chamber Class

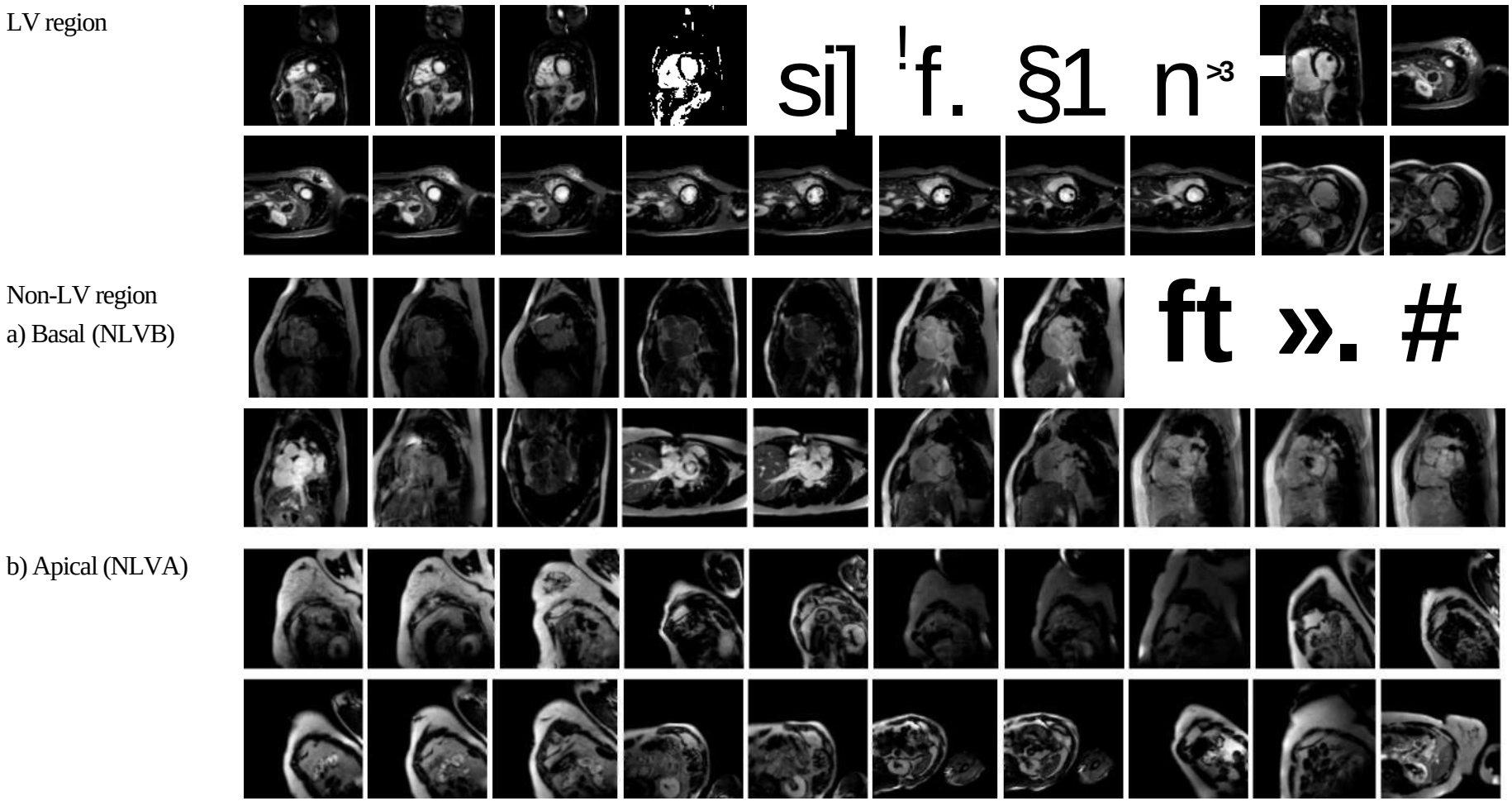


Figure 3.4 Example of Short-Axis LGE-CMR image of LV and Non-LV Region

3.3.3 Manual Labelling of LV Region

Next, all LGE-CMR images containing the LV region were subjected to a manual labeling process for both the LV region and the infarcted scar area. The GT labelling process was manually performed using the Image Labeller tool available in MATLAB R202la under the Image Processing and Computer Vision apps, as shown in Figure 3.5. The Pixel Label feature was selected to obtain the GT for each class. This GT data is utilized in the segmentation and detection processes in Phases 3 and 4, as outlined in Figure 3.1. Figure 3.6 presents examples of GT masks for the LV region and myocardial scar obtained through manual expert annotation. Subfigure (a) displays the original short-axis LGE-CMR image, (b) shows the corresponding GT for the LV region, and (c) illustrates the GT mask for infarcted scar tissue.

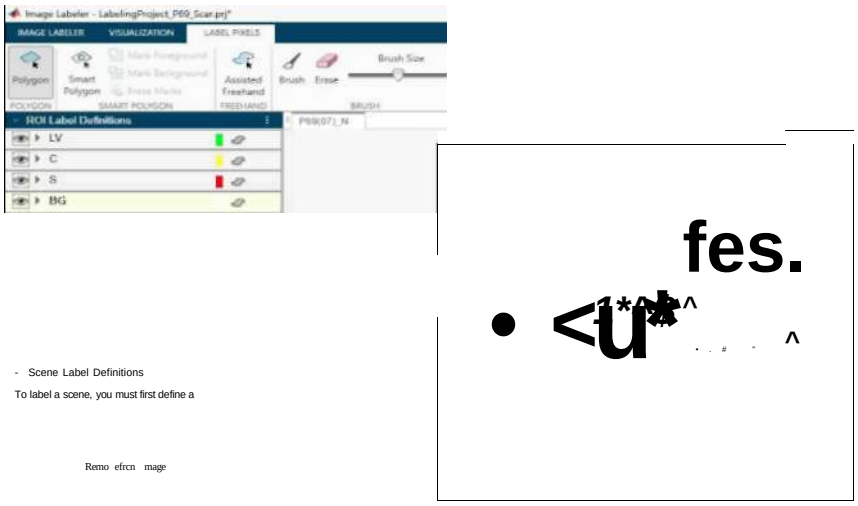
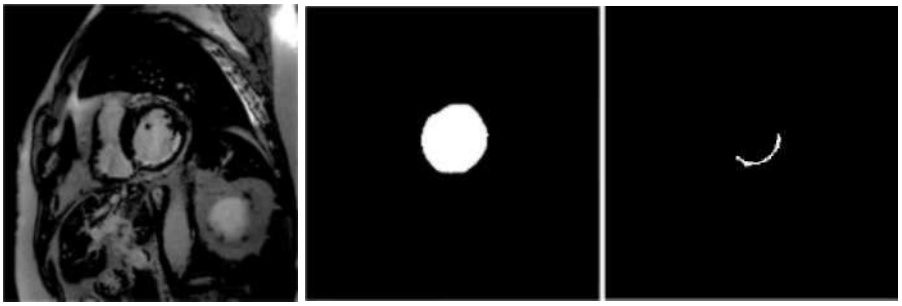


Figure 3.5 The MATLAB Image Labeler Tool for LV Region Labelling



I_____(a) (b) (c)

Figure 3.6 The Examples of GT for the LV region and Myocardial Scar for Validation Purposes, (a) Short Axis of LGE-CMR Image (b) GT for LV Region and (c) GT for infarcted Scar Tissue

3.3.4 Image Augmentation

A well-trained CNN model depends on a substantial amount of data [2], [5]. In the model training phase, we utilized 512 short-axis LGE-CMR images obtained from 30 patients, which were divided into training and testing sets. Specifically, 70% of the images were allocated for training, while the remaining 30% were reserved for testing. This allocation ratio is widely regarded as optimal for achieving strong training and testing performance [124]. Prior to the training process, we employed data augmentation techniques involving geometric transformations [123] to enhance the generalization capability of the DCNN model. These techniques included image rotations (+90, +180, +270 degrees) and image flipping (reversed along the center), which were applied more extensively to the non-LV class than the LV class to address the class imbalance by increasing the representation of underrepresented samples. This geometric transformation approach was chosen not only to expand the dataset but also to preserve key anatomical structures. Table 3.1 presents the total number of images in each class both before and after data augmentation.

Table 3.1
The Total Number of Images in Each Class Before and After Data Augmentation

Dataset	Original Dataset			After Data Augmentation		
	LV region	Non-LV region (NLVA + NLVB)	TOTAL	LV region	Non-LV region (NLVA + NLVB)	TOTAL
Training	252	**105**	357	756	756	1512
Testing	109	**46**	155	327	327	654
Total Images			512			2166

3.4 Phase 2: New Automated Classification of LV Images Using DCNN

This section introduces a new method for classifying images with LV and non-LV regions using DCNNs to automate the manual identification of the short-axis slice in LGE-MRI images prior to segmenting the LV region for MI detection. To the best of the authors' knowledge, this study is the first to introduce an automated method for classifying LV and non-LV regions from a sequence of LGE-CMR images. A new Shallow DCNN (S-DCNN) model was built and trained from scratch with random weight initialization to assess the effectiveness of narrower architecture in handling this task and compared to more complex pre-trained models. The S-DCNN was constructed by first defining the model architecture by experimenting with the number of convolutional layers and incorporating batch normalization layers. Subsequently, optimization techniques including Stochastic Gradient Descent with Momentum (SGDM) and Adaptive Moment Estimation (Adam) used in adjusting the model's weights during training [120] were evaluated, and the learning rate hyperparameter was fine-tuned to achieve the optimal configuration tailored to the classification task. All proposed DCNN models are trained to classify LGE-MRI images into two categories: LV and non-LV regions as shown in Figure 3.4. Consistent with clinical assessment, slices labelled as LV are selected as slices of interest, while both non-LV categories are excluded from further assessment. The flowchart illustrating the proposed image classification method for LV and non-LV regions using a DCNN is shown in Figure 3.7. The following subsections discuss the implementation and architecture details for both the from-scratch S-DCNN model and the pre-trained models.

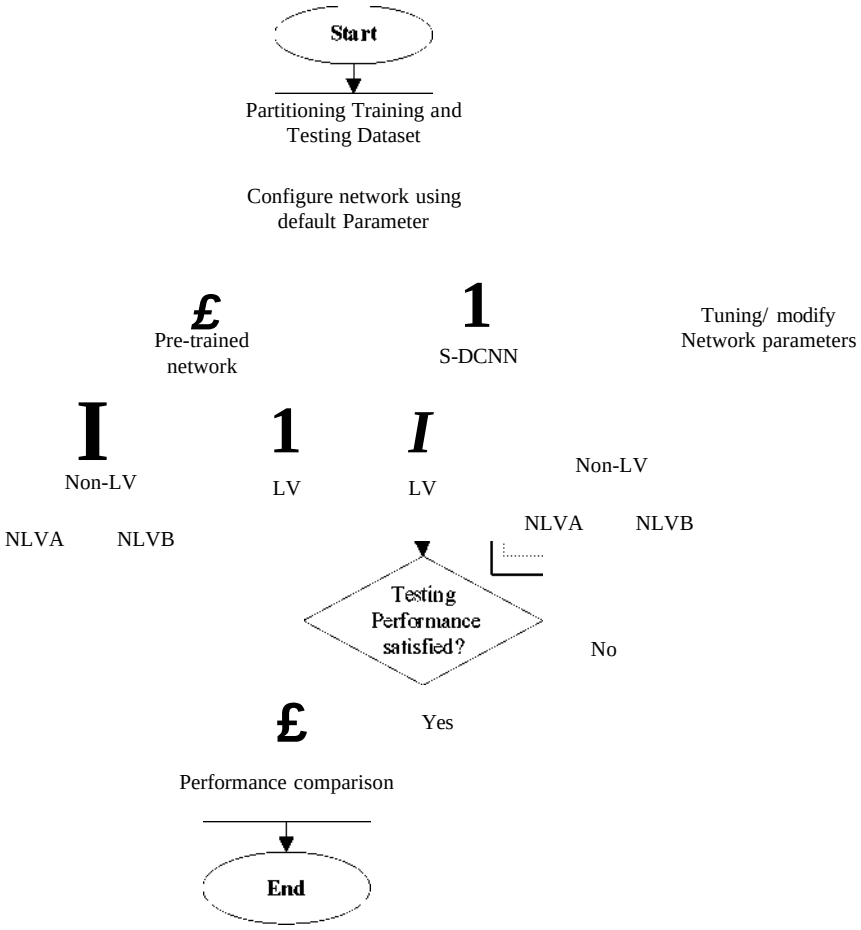


Figure 3.7 Flowchart of the Proposed Classification of LV and Non-LV Regions

3.4.1 Shallow DCNN (S-DCNN) Model

DCNN architecture typically includes multiple convolutional layers that work together to extract hierarchical features from the input image. The initial convolutional layer captures low-level features, such as edges and textures, while subsequent layers learn more complex patterns. Pooling layer is then used to reduce dimensionality while preserving essential features, followed by fully connected layers for classification. In this study, a train-from-scratch approach was adopted to minimize the complexity of DCNN, making it suitable for deployment on lower-specification systems. Furthermore, the network was kept simple and reliable to fulfil the next research objectives: to hybrid with another network model to automatically segment the infarcted tissue in the LV region. A series of experiments were conducted to identify the optimal architecture for classifying LV and non-LV regions, starting by incrementally increasing the number of convolutional layers from one to four. This approach allowed for a thorough assessment

of how additional depth impacts classification accuracy. While deeper networks can potentially capture more complex features, limiting the architecture to four layers helps balance performance with computational efficiency and mitigate issues like vanishing gradients and overfitting [121]. Each DCNN model was trained using LGE-CMR images with a resolution of 320x320x3 pixels as input. A conservative learning rate of 0.0001, a batch size of 16, and the Adam optimizer were initially chosen for training this initial experiment. The training parameters were selected based on recommended practices to establish a stable baseline [122] , and the configuration was set using MATLAB's training parameters. Overall comparison, the concept of increasing number of convolution layers was investigated and the best model will be chosen for further improvement. The detailed structure of the four DCNN models, each with a varying number of convolutional layers, is outlined as follows:

One convolutional layer DCNN

The structure of DCNN that utilising one convolution layer block shown in Figure 3.8. The convolutional layer has 8 kernels (filters) with each of size 3x3. In the context of convolutional operations, features are extracted by applying kernels to the input data. The size of the kernel determines the spatial extent of the receptive field, which influences the scale of features extracted. Smaller kernels typically capture fine-grained details, while larger kernels are more suited for extracting coarse-grained features [123]. In this study, the initial layers of the DCNN employ small kernel sizes to focus on fine-grained details, ensuring that the output feature maps from these layers provide detailed information about the local structure of the input. The stride, which controls the movement of the convolutional kernel, is set to 1, and padding is set to 'same' to preserve the spatial dimensions of the input image, ensuring that the output feature map has the same width and height as the input. This convolution layer followed by a ReLU activation function for non-linear transformation and a max-pooling layer with a 2x2 filter and a stride of 2 is then applied to reduce spatial dimensions to 160 X 160 X 24, preserving essential features while minimizing computational complexity. The max-pooling layer was chosen for its ability to select the maximum value from each pooling region, effectively preserving the most significant features of the input data, such as edges and textures. By focusing on the maximum values, max pooling helps to ignore less relevant background information and noise, which might

otherwise dilute the quality of the feature representation. The output from this process is 'flattened' by the fully connected layer into a single vector, which is then processed by the activation function to compute the probability distribution across all classes. A softmax activation function [124] is applied to determine the probability for each class, represented by two neurons corresponding to the class labels (LV and non-LV). Table 3.2 provides a summary of the detailed parameters for the single-layer DCNN model.

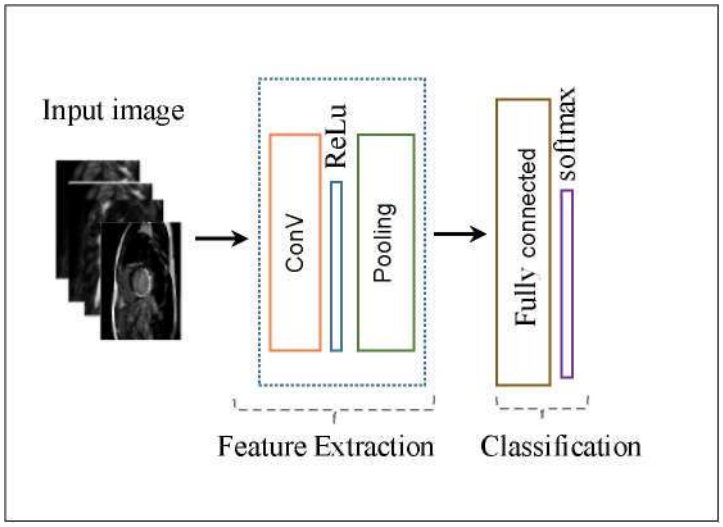


Figure 3.8 The One Convolutional Layer DCNN Structure

Table 3.2
DCNN Layer Configuration for the One Convolutional Layer Model

Layer	Type	Kernal Size	No. of Kernel	Stride	Output size
0	Input		-	-	320x320
1	Conv1	3x3	8	1	320x320
	ReLu	-	-	-	320x320
	Pool1	2	-	2	160x160
2	FC	-	-	-	2
	Softmax	-	-	-	2

Two convolutional layer DCNN

The two-convolution layer DCNN is almost similar to the previous one-convolution layer DCNN, but with an additional convolutional layer block, as illustrated

in Figure 3.9. The first convolutional layer mirrors that of the one-layer DCNN. Following the second convolutional block, which consists of a 3x3 kernel with 16 filters and a stride of 1, a ReLU activation function is applied, and a 2x2 max-pooling layer with a stride of 2 reduces the output size to 80x80. This additional convolutional layer enhances the model's ability to capture more complex features by increasing the number of filters , allowing for richer feature extraction at the cost of minimal added computational complexity. The network retains the same fully connected layer configuration, with the final layers consisting of a fully connected layer followed by a softmax layer, similar to the previous one-convolution layer DCNN model. Table 3.3 provides a summary of the detailed parameters for the two convolution-layer DCNN model.

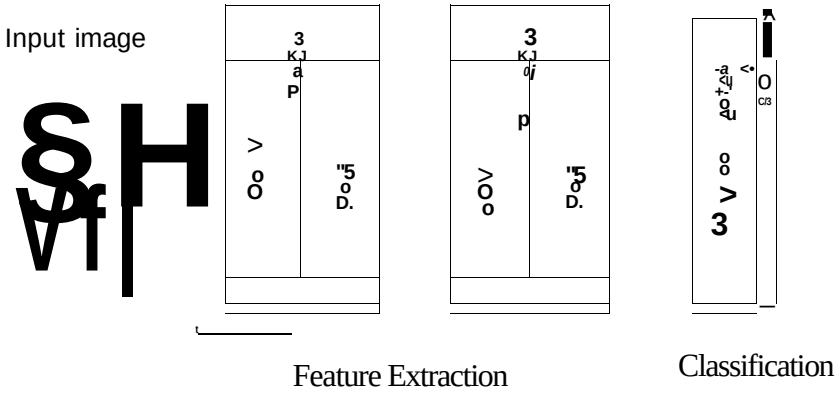


Figure 3.9 The Two Convolutional Layer DCNN Structure

Table 3.3
DCNN Layer Configuration for the Two Convolutional Layer Model

Layer	Type	Kernal Size	No. of Kernel	Stride	Output size
0	Input				320x320
	Convl	3x3			320x320
	ReLu				320x320
	Pooll	2			160x160
	Conv2	3x3	16		160x160
	ReLu				160x160
	Pool2		16		80x80

Layer	Type	Kernal Size	No. of Kernel	Stride	**Output** size
3	FC				
	Softmax				

Three convolutional layers DCNN

The three-convolution layer DCNN model builds on the previous architectures by adding an additional convolutional block, further enhancing its capacity to capture complex features, as illustrated in Figure 3.10. The first two convolutional layers are similar to those in the two-layer model. The additional third convolutional layer (Conv3) applies 32 kernels of size 3x3 with a stride of 1, followed by ReLU activation. This is succeeded by a 2x2 max-pooling layer, which reduces the feature map size to 40x40. This third layer allows the model to capture even more detailed features, enhancing its ability to distinguish finer details within the images. The final layers are also structured similarly to the previous models, with a fully connected layer that flattens the features into a vector and a softmax layer that outputs the probability distribution across the three classes. The detail parameters of this three convolutional - layer DCNN is summarized in Table 3.4.

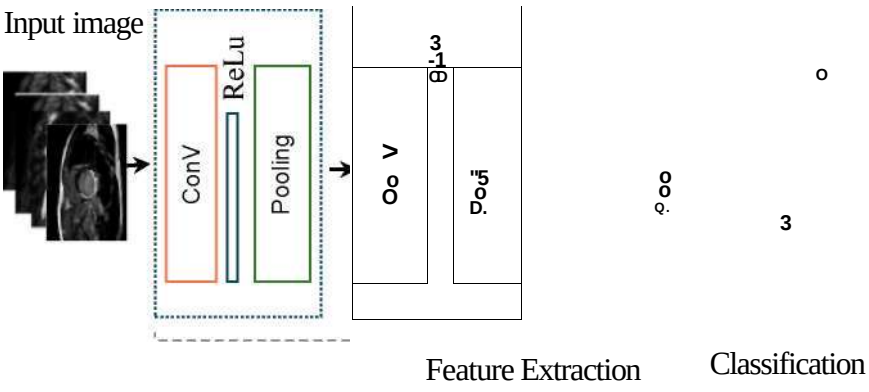


Figure 3.10 The Three Convolutional Layer DCNN Structure

Table 3.4
DCNN Layer Configuration for the Three Convolutional Layer Model

Layer	Type	Kernal Size	No. of Kernel	Stride	Output size
0	Input		-	-	320x320
1	Conv1	3x3	8	1	320x320
	ReLu	-	-	-	320x320
	Pool1	2		2	160x160
2	Conv2	3x3	16	1	160x160
	ReLu	-	-	-	160x160
	Pool2	2		2	80x80
3	Conv3	3x3	32	1	80x80
	ReLu	-	-	-	80x80
	Pool3	2	32	2	40x40
4	FC	-	-	-	2
	Softmax	-	-	-	2

Four convolutional layers DCNN

The four-convolutional layer DCNN model further builds on the previous architectures by incorporating an additional convolutional block, increasing its ability to capture more detailed and complex features, as illustrated in Figure 3.11. This architecture begins similarly to the three-layer model, with the additional third convolutional layer (Conv3) applying 32 kernels of size 3x3 with a stride of 1, followed by a ReLU activation and another 2x2 max-pooling layer, reducing the feature map to 40x40. The additional fourth convolutional layer (Conv4) further extends the feature extraction capacity, using 64 kernels of size 3x3 with a stride of 1, followed by a ReLU activation function and a 2x2 max-pooling layer, reducing the output to 20x20. As the network depth increases, the number of convolutional kernels is increased to 64 to detect more specific and detailed patterns. As with the previous models, the final layers consist of a fully connected layer that flattens the extracted features into a vector and a softmax layer that outputs a probability distribution across the two classes. The detail parameters of this four convolutional -layer DCNN is summarized in Table 3.5.

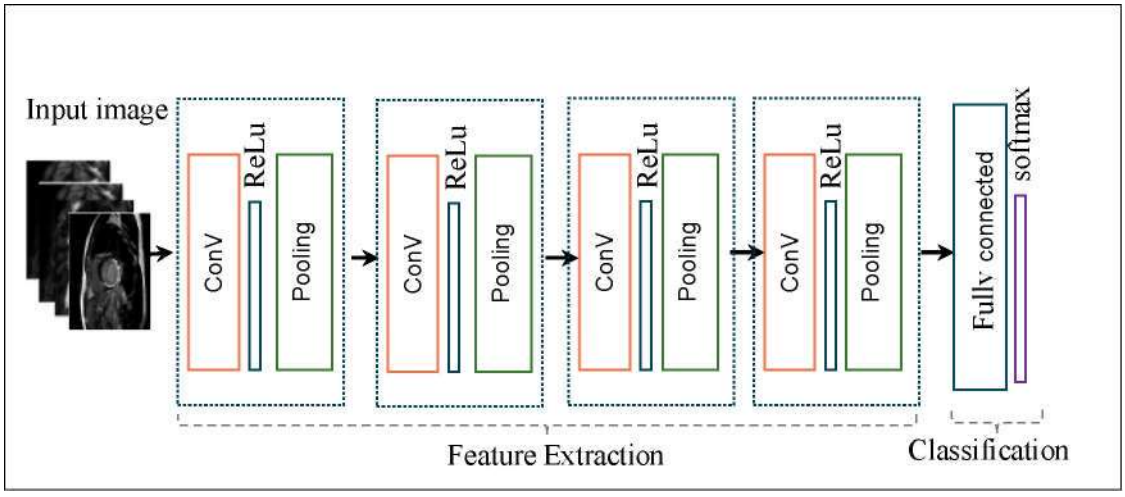


Figure 3.11 The Four Convolutional Layer DCNN Structure

Table 3.5
DCNN Layer Configuration for the Four Convolutional Layer Model

Layer	Type	Kernal Size	No. of Kernel	Stride	Output size
0	Input		-	-	320x320
1	Convl	3x3	8	1	320x320
	ReLu	-	-	-	320x320
	Pooll	2	8	2	160x160
2	Conv2	3x3	16	1	160x160
	ReLu	-	-	-	160x160
	Pool2	2	16	2	80x80
3	Conv3	3x3	32	1	80x80
	ReLu	-	-	-	80x80
	Pool3	2	32	2	40x40
4	Conv4	3x3	64	1	40x40
	ReLu	-	-	-	40x40
	Pool4	2	32	2	20x20
5	FC	-	-	-	2
	Softmax	-	-	-	2

Batch normalization is not typically included by default in basic DCNN architectures, as its effects can vary across different network configurations [125], [126] .This study investigates incorporating batch normalization into the best-performing DCNN model from previous experiments to evaluate its impact on classification performance. The batch normalization layer was placed after each convolutional layer, just before the activation function, to standardize activations before they pass through non-linear functions. Given that batch normalization can interact differently with various optimizers, further experiments were conducted to evaluate the model's performance with and without batch normalization under different training conditions. Specifically, three learning rates (0.0001, 0.001, and 0.01) were tested with two types of optimizers, SGDM and Adam, to determine the optimal settings. This experimental setup allows for a comprehensive assessment of how batch normalization influences model stability and accuracy across a range of configurations. The placement of the batch normalization layers within the convolutional block is illustrated in Figure 3.12.

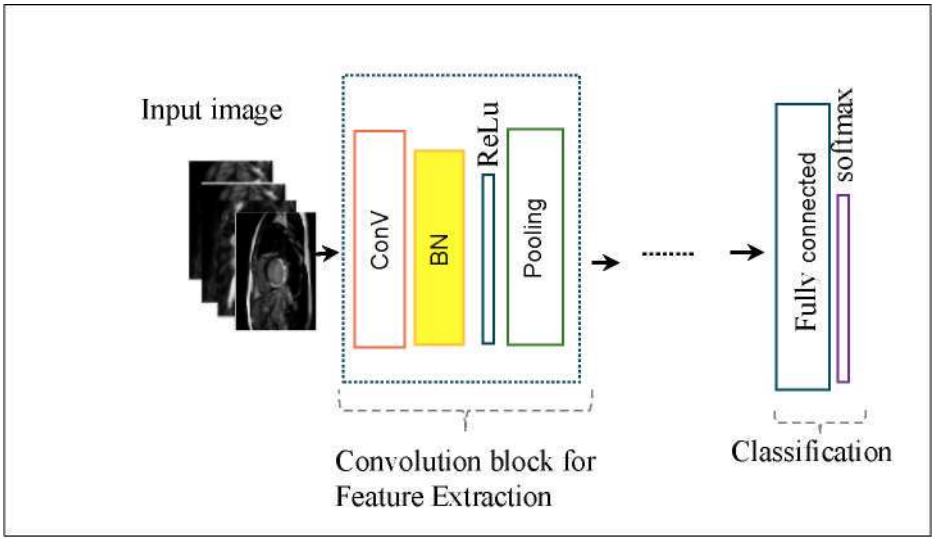


Figure 3.12 Placement of Batch Normalization Layers Within the Convolutional Block

3.4.2 Pre-trained DCNN Model

In addition to the S-DCNN model trained from scratch, this study also explores customized pre-trained DCNN architectures, with modified classification layers, as an alternative network. Pretrained networks offer an effective solution for small image data samples [127]. These models have been previously trained on extensive datasets,

learning to recognize general image features like edges, textures, and shapes. This version of DCNN can also be deployed using an approach known as transfer learning [128], [129] with a pre-trained model. By leveraging this pre-existing knowledge, pre-trained models can be fine-tuned for specific tasks, such as classifying LV and non-LV regions in LGE-MRI images. This study opted to concentrate on three well-known CNN models, namely Alexnet, GoogLeNet, and SqueezeNet, where each of these architectures has its own strengths in performing the classification task. AlexNet by Alex Krizhevsky et. al [130] was the first prominent model to in-corporate the convolutional layers that define a CNN. Its architecture consists of five convolutional layers followed by three fully connected layers, with an input size fixed at 227x227x3. Notably, AlexNet introduced Rectified Linear Units (ReLU) as the activation function, which improved learning efficiency compared to the previously used sigmoid function. GoogLeNet [131] introduced a more complex architecture by incorporating nine inception modules, each designed to perform multilevel feature extraction using 1x1, 3x3, and 5x5 filters. The network consists of 27 layers of deep convolutional layers. It requires input images sized at 224x224x3. The inception modules enable GoogLeNet to capture both local and global features effectively. SqueezeNet [132] introduced a highly compact architecture with nine fire modules, designed to reduce the model's parameters while maintaining classification accuracy. It has 18 layers and requires the same input size as AlexNet. Each fire module uses 1x1 convolution filters to act as bottleneck layers, reducing the computational load before passing data to the 3x3 filters. All three pre-trained models were imported from the MATLAB Deep Learning Toolbox.

The input images were resized to match each model's expected input dimensions, and the final classification layer of each model was customized to produce two output classes for binary classification of LV and non-LV regions. For a fair comparison with our S-DCNN model, all customized pre-trained networks were configured with the same learning rate, batch size, training epochs and optimizer settings as Table 3.6 below.

Table 3.6
Hyperparameter Setting Details for DCNNs Training

DCNN Parameter	Default Value
Learning Rate	0.0001
MaxEpochs	70
Validation Frequency	32 iterations
Mini Batch Size	16
Solver	Adam

3.4.3 Performance Evaluation

This study investigates the proposed method for detecting LGE-MRI images with LV region by classifying into LV and non-LV regions using four S-DCNN models and three pre-trained models. The LV classification comprised images spanning from the basal (base) to the apical regions of the heart, while the non-LV classification included images located below the apical and above the basal regions. To ensure consistency in model comparisons, this classification task performance was tested using 12^{th} Gen Intel 17-12700H processor, NVIDIA GeForce RTX 3050 Laptop GPU and 16GB of RAM. The performance was assessed using classification metrics derived from the Confusion Matrix (CM), including true positive (TP), true negative (TN), false positive (FP), and false negative (FN) outcomes. These metrics provide a quantitative basis for evaluating the effectiveness of each model in accurately classifying images. The definition for each CM value is described below [133]:

TP - The number of correctly classified LV regions
TN - The number of correctly classified non-LV regions
FP - The number of incorrectly classified as LV regions
FN - The number of incorrectly classified as non-LV regions

Below are the lists of performance metrics using the CM components to evaluate the proposed method.

Accuracy (Ace) - represents the ratio of the number of correct classifications (both LV and non-LV) to the total number of images/cases considered. It is calculated

as the proportion of correctly classified samples (true positives and true negatives) to the total number of samples. The formula for accuracy is given by Equation 3.1:

$$\text{ccuracy} \quad \frac{TP+TN}{T_p + F_N + T_N + F_p} \tag{3.1}$$

Recall/Sensitivity (SN) - the ratio of the number of correctly classified LV regions to the total actual LV regions. The percentage of sensitivity is given by Equation 3.2:

$$\text{Recall/sensitivity}=\frac{TP}{(TP+FN)} \tag{3.2}$$

Precision - the ratio of correctly classified LV regions to the total number of predicted LV regions, given by Equation 3.3:

$$\text{Precision}=\frac{TP}{(TP+FP)} \tag{3.3}$$

Fl-score - harmonic mean of precision and recall, providing a balance between these two metrics. It gives a single measure of the model's predictive performance, particularly useful when there is an uneven class distribution, given by Equation 3.4:

$$\text{Fl-Score}=2\frac{\text{Precision x Recall}}{\text{Precision+ Recall}} \tag{3.4}$$

The next stage utilized images containing only the LV region, as classified by the best-performing automated S-DCNN, for the tasks of LV region and scar segmentation.

3.5 Phase 3: New Dual-Stages DeeplabV3+-Based Myocardial Scar Segmentation

As discussed in Chapter 2, Section 2.2, the presence of myocardial scar tissue in LGE-CMRI is a key indicator of MI. This section presents a comprehensive methodology for developing a new automatic dual-stage myocardial scar segmentation model using LGE-CMR images that contain only the LV region obtained from the previous phase. The proposed approach comprises two segmentation stages, each designed to perform a distinct task tailored to its specific segmentation objective. Motivated by the DeepLabV3+ encoder-decoder architecture and its successful application in image segmentation studies [78], [79], [134], this study enhances the encoder architecture at each stage to improve segmentation performance. DeepLabV3+ is well-suited for these segmentation tasks due to its ability to extract multi-scale contextual features, making it effective in handling LV anatomical variations and detecting abnormalities of different scar sizes. The first stage focuses on developing the DLT-LV network for precise segmentation of the LV region, where scars are likely to be present. The customized network architecture is detailed in Subsection 3.5.1.1. This stage also includes a post-processing procedure following the segmentation of the LV region. This procedure addresses potential issues such as small gaps within the LV region and isolated mis-segmentations produced by the DLT-LV network. This procedure is discussed in detail at the end of the first stage, as presented in Subsection 3.5.1.2.

Subsequently, Subsection 3.5.2 turns its attention to the development of the DLT-Scar network, which focuses on the segmentation of scar tissue within the identified LV region. The same model architecture used in the first stage was applied in this second stage, but the selection of the backbone network was determined through the same experimental evaluation conducted in the first stage. Given that myocardial scars typically occupy a very small proportion of the LGE-CMR image relative to the background, the segmentation task is inherently affected by class imbalance, which can significantly degrade model performance if not properly addressed [37], [87], [88]. To overcome this, the study investigates several loss functions suited for imbalanced segmentation tasks and evaluates their impact on the DLT-Scar network's performance. Additionally, a novel approach utilizing the PSO algorithm is proposed to optimize this model by automatically searching for the optimal values of Tversky loss function

hyperparameters, as detailed in section 3.5.2.2. To the best of our knowledge, this is the first investigation to automate the hyperparameter optimization of the Tversky loss function using the PSO-DLT-Scar segmentation model. Figure 3.13 shows the proposed framework for the Dual Stage Myocardial Scar segmentation model, and the flowchart of the proposed two-stage model is demonstrated in Figure 3.14. The offline mode involved the development, training, and testing of the DCNN, while the online mode focused on developing a fully automated myocardial scar segmentation and detection model, utilizing all trained DCNN models.

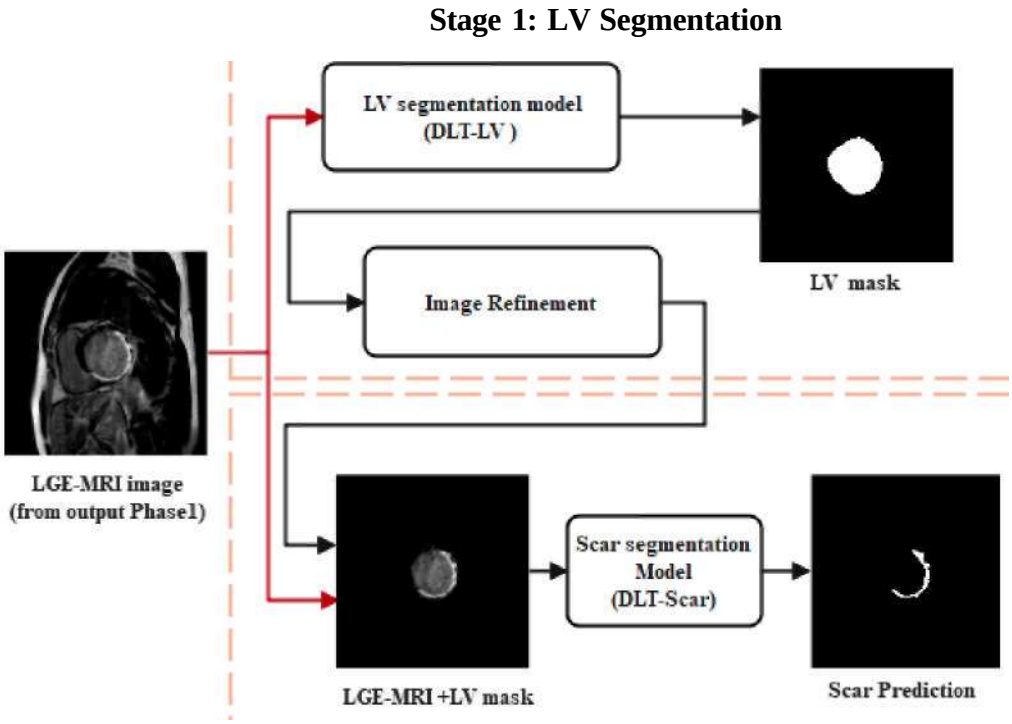


Figure 3.13 Proposed Framework for Dual-Stage Myocardial Scar Segmentation

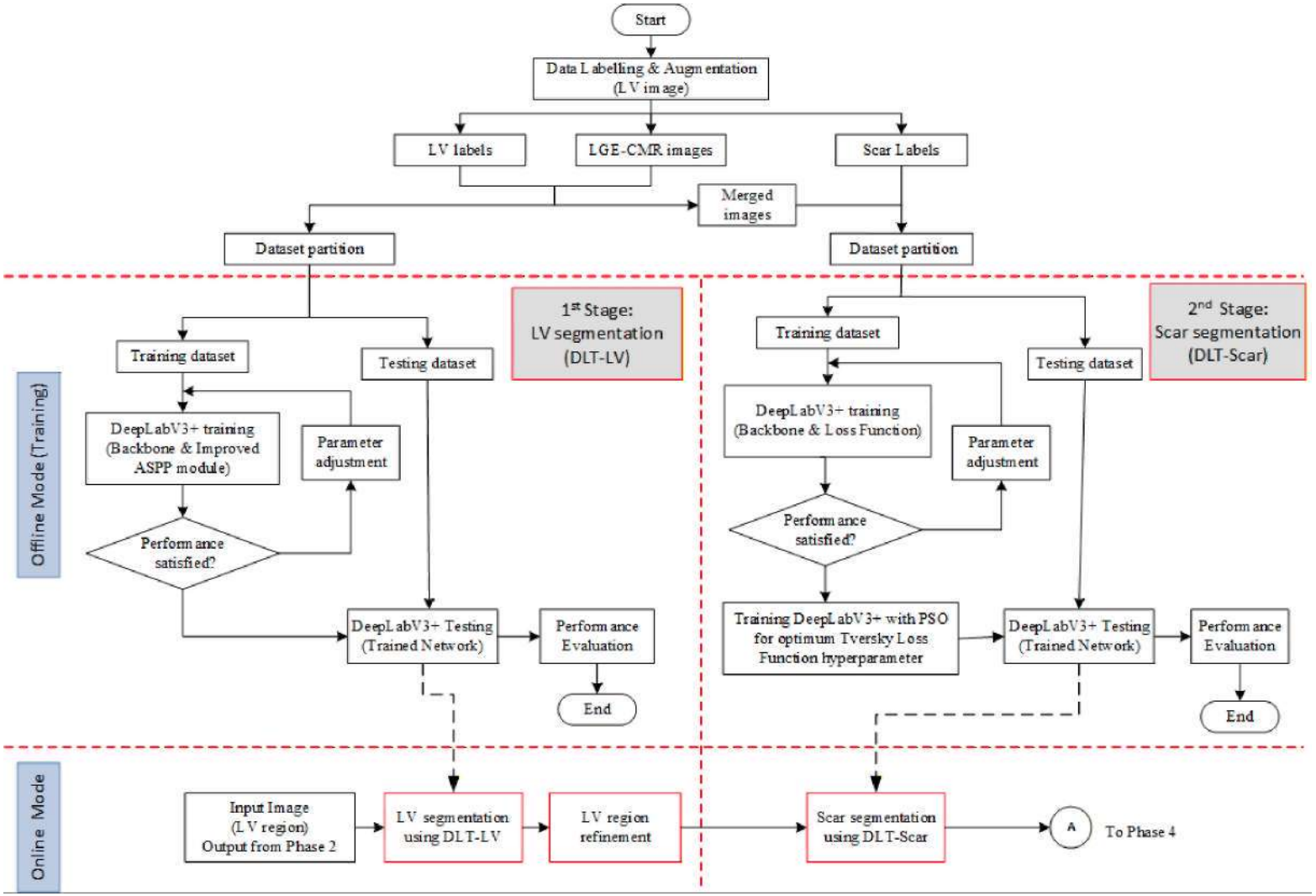


Figure 3.14 Flowchart of Dual-Stage Myocardial Scar Segmentation

3.5.1 First Stage: LV Region Segmentation Model

This subsection focuses on the first stage of the segmentation process, which targets the delineation of the LV region, as illustrated in Figure 3.14. Figure 3.15 illustrates the multiscale morphology of the LV region, with representative short-axis LGE-MRI slices captured at three anatomical levels: (a) apical, (b) mid-ventricular (middle), and (c) basal sections. This stage is critical, as its segmentation output will directly influence the accuracy and reliability of the subsequent scar segmentation. The proposed LV segmentation approach combines a modified DeepLabV3+ network with a post-processing procedure to achieve refined LV region segmentation. The architectural modifications are detailed in Subsection 3.5.1.1, while the post-processing procedure for segmentation refinement is presented in Subsection 3.5.1.2.

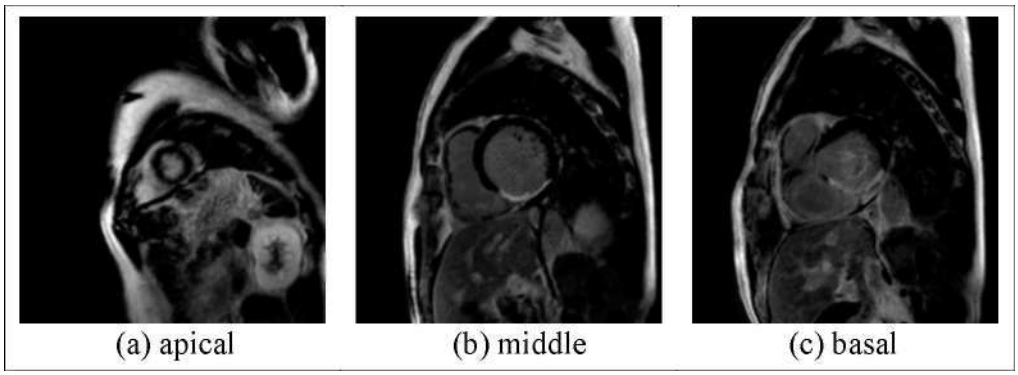


Figure 3.15 The Example of LV Region in LGE-CMR Image

3.5.1.1 DeepLabV3+ Tailored to LV (DLT-LV) Network

DeepLabV3+, introduced by Chen et al. [79], is an advanced semantic segmentation network that integrates both an encoder and a decoder. The decoder in DeepLabV3+ consists of a backbone architecture that serves as the feature extractor, along with the atrous spatial pyramid pooling (ASPP) module, a crucial component for capturing multi-scale contextual information. The integration of both the backbone and ASPP module within the encoder of DeepLabV3+ results in a highly effective semantic segmentation framework. A key aspect of the modifications in this study involved adjusting the encoder of DeepLabV3+, as it is essential for effective feature extraction in image segmentation tasks. These modified elements are highlighted in the encoder part of the DeepLabV3+ structure, as shown in Figure 3.16.

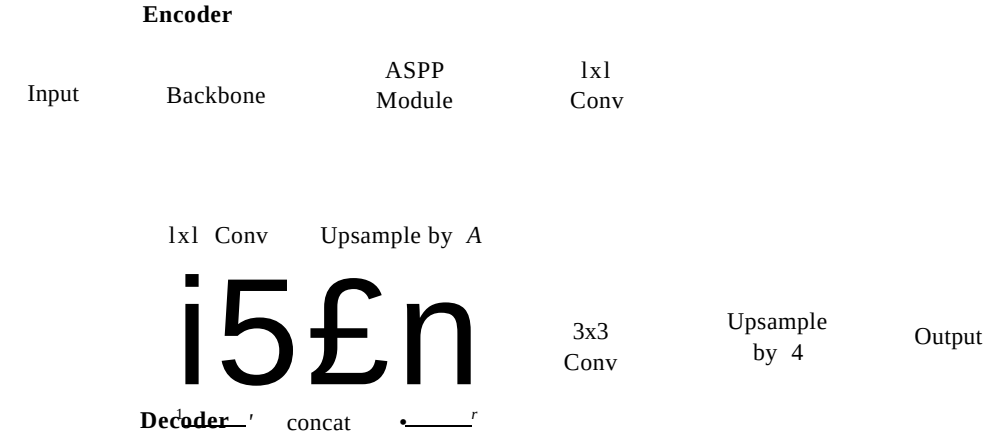


Figure 3.16 The Encoder-Decoder Structure of the DeepLabv3+ Model

Network backbone

The backbone network is crucial in feature extraction, capturing essential information from the input image within the DeepLabV3+ model providing high-quality feature maps to subsequent modules, such as the Atrous Spatial Pyramid Pooling (ASPP) to achieve high performance in semantic segmentation tasks. This study compared three prominent backbones networks, each offering unique strengths for the task: ResNet50, Xception, and MobileNetV2. Xception has been commonly utilized in prior research as the default backbone for DeepLabV3+ [79]. MobileNetV2 [135] was chosen for its lightweight design, which enhances computational efficiency, making it ideal for applications with limited resources or real-time processing needs. Its streamlined structure reduces the model size while preserving accuracy, making it well-suited for extensive datasets such as LGE-MRI scans. Alternatively, ResNet50 [136] is particularly effective at segmenting small or complex anatomical features due to its residual connections. Comparing these backbones helps determine which model provides the most balanced performance in terms of accuracy, efficiency, and adaptability for myocardial scar segmentation. This study used networks pre-trained on over a million images from the ImageNet database, which were then retrained with a short-axis LGE-CMR dataset. Pre-trained networks are known for their efficiency and effectiveness, especially when applied to medical imaging tasks with limited training data.

Improved ASPP Module

In DeepLabV3+ encoder, the ASPP is utilized for resampling feature maps created from the backbone architecture at different atrous rates. The ASPP module includes four parallel atrous convolutions with different dilation rates, allowing the network to simultaneously learn features at various scales. These features are concatenated and passed through a 1 x 1 convolution to create a rich representation of contextual information at multiple scales before being applied to the encoder's feature maps. As the dilation rate increases, the receptive field grows, allowing the network to capture broader contextual information from the input. The proposed module includes extending four parallel branches of the atrous depthwise separable convolution proposed in DeepLabV3+ to five parallel branches with dilation rates of 1,2, 6, 12, and 18 to enhance multi-scale feature extraction. The selected dilation rates are lower than the default values (1,6, 12, and 18) used in the original ASPP module, as this approach is particularly effective in enhancing the representation of small-scale features, as demonstrated in [137], [138]. Additionally, a set of batch normalization and ReLU layers are applied after each pointwise convolution. Batch normalization following each convolution stabilizes training by normalizing the input distributions, which has been shown to improve network convergence and generalization in deep learning models, especially when applied to complex medical imaging tasks [139]. The inclusion of ReLU activation also enhances non-linearity, allowing the ASPP module to better capture complex patterns in myocardial images. Together, the modification ASPP architecture is tailored to address specific challenges in the segmentation of cardiac structures, where capturing fine details at is crucial for accurate LV delineation. The original ASPP module in the DeepLabV3+ model and the improved ASPP module used in the DLT-LV model are displayed for comparison in Figure 3.17. The three backbones mentioned above were tested on the modified ASPP module and compared to the original ASPP. The best-performing backbone was selected as this study's automated DeepLabV3+-based model for LV segmentation, referred as DLT-LV, which will be applied in the subsequent refinement procedure described in the following section. In this section, all training parameters, including the learning rate and number of epochs, were carefully selected to promote stability throughout the training process until the learning progress reached a steady optimal point. This choice of parameters facilitated

balanced convergence toward the highest possible model performance, as detailed in Table 3.7.

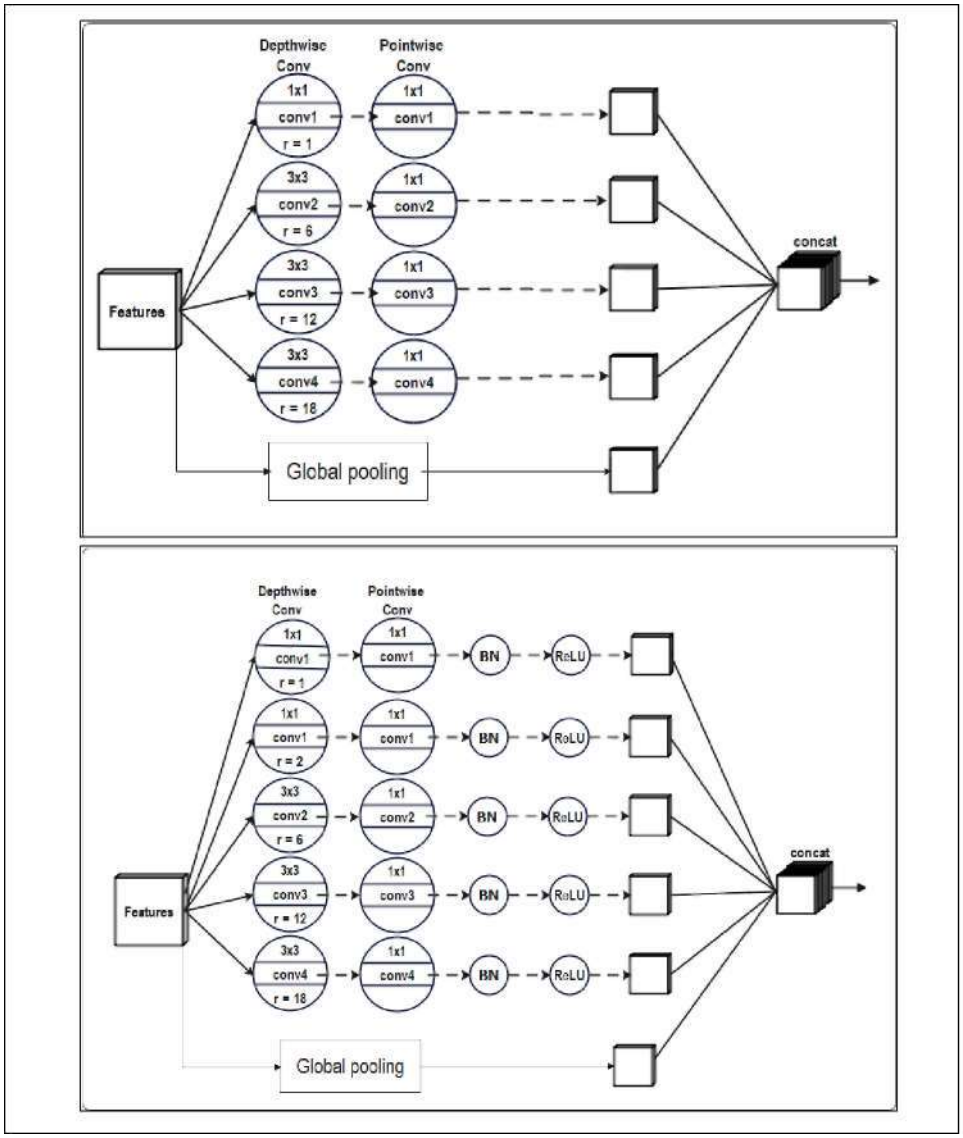


Figure 3.17 The Original ASPP Module in the DeepLabV3+ (Top) and the Proposed Improved ASPP Module in LV Segmentation Model (DLT-LV) (Bottom). The 'r' Denotes the Atrous Rate

Table 3.7
Second Stage: Scar Segmentation Settings for Semantic Segmentation

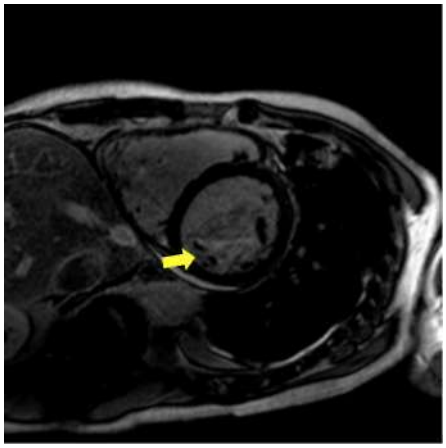
Parameter	**Settings**
Initial Learning Rate	0.0001
Max Epochs	100
Mini Batch	8

Parameter	Settings
Function loss	Generalized Dice Loss Function
Network Optimizer	ADAM

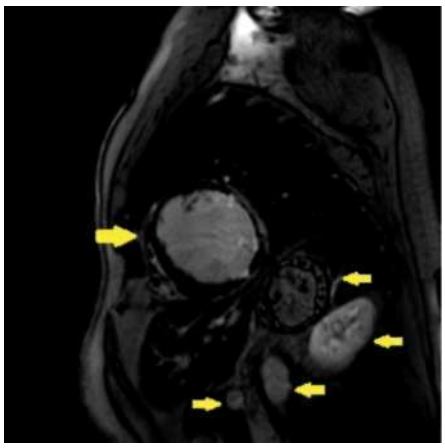
3.5.1.2 DLT-LV Refinement Using Morphological Procedure

The segmentation of the LV chamber in short-axis LGE-CMR images presents several challenges due to the inherent complexity of the imaging modality. These challenges include intensity inconsistencies within the LV cavity, along with the influence of papillary muscles and trabeculations. Additionally, background artifacts and surrounding anatomical structures that closely resemble the LV further complicate segmentation. As shown in Figure 3.18, these factors may cause the network to misinterpret structures, leading to small gaps within the segmented LV chamber and isolated mis-segmentations outside the LV region. Such issues can negatively affect segmentation accuracy, emphasizing the need for additional refinement techniques. To minimize incorrect interpretations for the next stage, the output from the DLT-LV network requires additional refinement and filtering strategies to ensure robust and reliable results. Therefore, this involves a morphological procedure to further refine the segmentation and address potential issues such as small gaps within the LV chamber and isolated mis-segmentations. The steps involved in this procedure are as follows:

i) For a given binary image;
   a) Define holes as sets of background pixels surrounded by object foreground pixels.
   b) Fill all holes with the value of adjacent pixels in the horizontal or vertical direction to ensure that any small gaps within the segmented LV region are closed.
ii) Measure the area/size of each object in the binary image. Then, identify the object with the largest area.
iii) Retain only the largest object and remove any smaller objects. This step ensures that isolated mis-segmentations are discarded.



Intensity inconsistencies and Papillary Muscles within LV cavity



Background Artifacts and Similar Anatomical Structures

Figure 3.18 Examples of Segmentation Challenges in Short-Axis LGE-CMR Images

3.5.2 Second Stage: Scar Segmentation Model

After the image refinement process, the resulting LV mask images are merged with the corresponding LGE-CMR images and subsequently used as input for the second stage of the modified DeepLabV3+ network, referred to as DLT-Scar. This network retains the same encoder-decoder architecture as the DLT-LV model from the previous stage, incorporating the modified ASPP module. The backbone for this stage was selected based on the same comparative experiment involving three backbones (MobileNetV2, ResNet50 and Xception) as described in Section 3.5.1.1, with the best-performing architecture adopted to ensure optimal feature extraction. This section focuses on the second stage of the proposed Dual-Stage DeepLabV3+-based Myocardial Scar Segmentation Model, specifically targeting the selection of the most suitable loss function for the model's classification layer to address class imbalance

caused by the small foreground region which is the scar. As mentioned in Section 2.3.3, during the training of semantic segmentation, class imbalance due to small target regions can cause the network to become biased toward the dominant class (background pixels). This often results in poor segmentation performance for the minority class (scar regions), as the model tends to overlook critical areas, leading to higher rates of false negatives and under-segmentation. Loss functions play a crucial role in guiding the training process of deep learning models for segmentation. By addressing class imbalance in the classification layer, these functions help the model effectively learn features from both the foreground and background classes. This section discusses the second stage of the proposed Dual-Stage approach, which focuses on scar tissue segmentation, with a particular emphasis on optimizing the loss functions to improve performance. Experiments were conducted on the DLT-Scar model using four state-of-the-art loss functions, as described in Subsection 3.5.2.1, to determine the most adaptable loss function for achieving optimal segmentation performance.

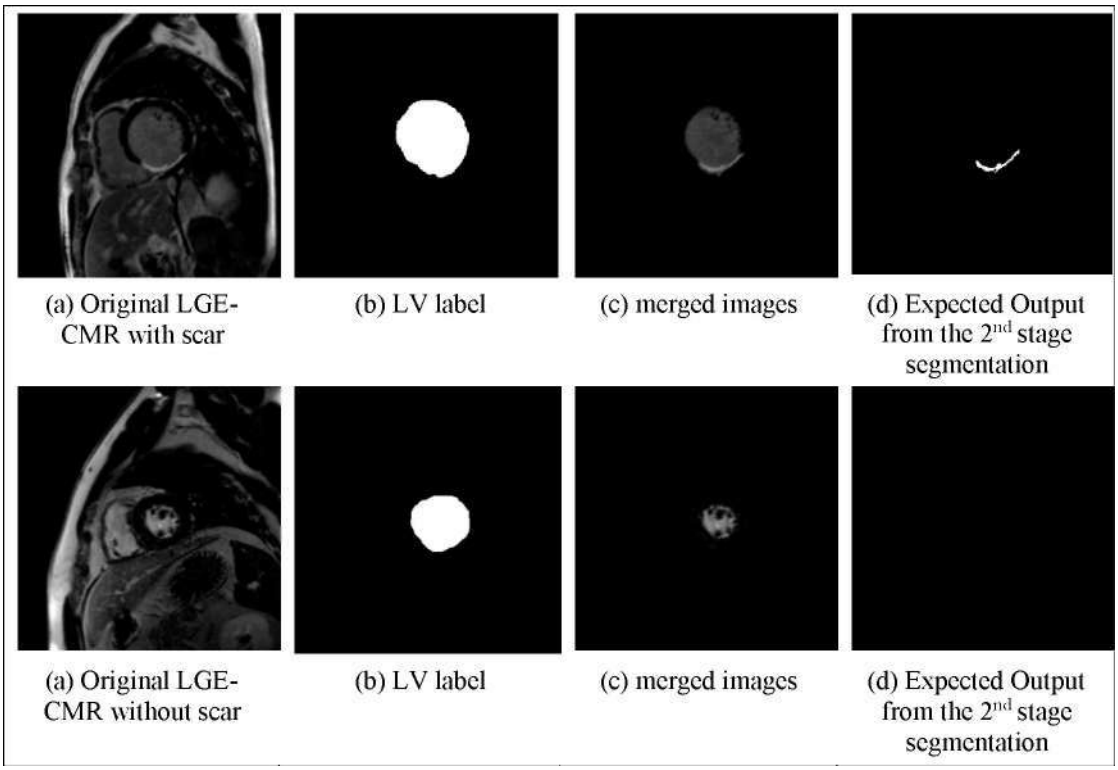


Figure 3.19 Examples of Merged Images, the Input Images to the Second Stage and Its Expected Output From the 2nd Stage Segmentation

Finally, to further enhance segmentation performance in scar tissue detection, Subsection 3.5.2.2 introduces a novel automated method for determining the optimal

hyperparameter values of the Tversky loss function. This approach utilizes PSO to fine-tune the loss function automatically, aiming to improve the model's sensitivity to small scar regions while simplifying the hyperparameter search process. As illustrated in Figure 3.19, the input image used in this stage is trained against manually segmented ground truth scar labels, enabling the model to effectively learn and distinguish scar tissue during segmentation.

3.5.2.1 Loss Fu nction Selection for Imbalance Class

Loss functions provide feedback to the model while training by measuring how far its predictions deviate from the actual ground truth. Minimizing the loss allows the model to adjust its parameters, gradually improving its ability to produce accurate segmentations. Each loss function handles class imbalance in different ways, making it important to choose the most suitable one for tasks involving small and underrepresented regions. In this study, four loss functions were selected based on their ability to effectively manage class imbalance and enhance segmentation performance for small regions: the Generalized Dice loss function, Weighted Cross-Entropy loss function , Focal loss function and Tversky loss function. These functions were compared to determine the best approach for addressing the challenges of class imbalance in this context with same training settings used across the scar segmentation framework as in Table 3.8.

Table 3.8
DCNN Layer Second Stage: Scar Segmentation Settings for Semantic Segmentation

Parameter	**Settings**
Initial Learning Rate	0.0001
Max Epochs	100
Mini Batch	8
network Optimizer	ADAM

3.5.2.2 Proposed Automated Tversky loss Function Hyperparameters Optimization Using PSO

Tversky loss function relies on two independent hyperparameters: a and P, to control the relative penalties on false positives and false negatives. While researchers often employ P=0.7 and a=0.3 [37], [88] it is important to note that the optimal values of these parameters can vary depending on the specific tasks and datasets [36]. While manual random search can be a straightforward method for hyperparameter tuning, it is inefficient due to the extensive manual adjustments required, which are time-consuming for complex CNNs and lack the systematic exploration needed to optimize parameters effectively [140]. To address this limitation, it is necessary to explore automated methodologies for defining parameter spaces to achieve higher performance. Hence, this study further investigated the use of PSO to solve the problem of locating the optimal Tversky loss hyperparameters automatically ideally for LGE-CMR images for segmenting myocardial scars.

This study proposes a modification to the existing pathway by implementing PSO optimization in the scar segmentation, DLT-Scar model training flow. In the default CNN model training process prior to optimization, feature segmentation is straightforward, without any automated feedback mechanism to refine the final segmentation features based on the manual selection of Tversky loss hyperparameters. The proposed PSO-based optimization specifically targets the Tversky loss function in the classification layer of the DLT-Scar architecture. Additionally, the two hyperparameters a and P were simplified to a single parameter search by exploiting their mathematical dependency (where P = 1 - a) [36], [38]. This approach merges the control of false positives and false negatives into a single entity, thereby reducing the hyperparameter search space. The choice to optimize only a or P doesn't make a major difference, because once a is optimized, P's effect is already reflected.

To automatically determine the optimal value of a, the DLT-Scar model was integrated with PSO. In this implementation, the original Tversky loss function originally parameterized by both a and P was simplified to a single tunable parameter by customizing the MATLAB classification layer to include only one user-defined variable for the a value within the DLT-Scar model. In this study, the initial value of a was set to 0.3. The PSO algorithm was guided by a fitness function defined using the Dice Similarity Coefficient (Dice score), which was computed on a testing dataset after

training the DLT-Scar model with each candidate a value. The optimization process began with a population of particles or swarms, each representing a potential a value randomly initialized within a defined range between 0.1 and 0.5 [36], and assigned initial velocities. At every iteration, the objective function was evaluated for each particle, and the particles updated their velocities and positions based on both their personal best (pbest) and the global best (gbest) solutions found so far. These updates were modulated by a linearly decreasing inertia weight to balance exploration and exploitation of the search space. The algorithm proceeded iteratively until a stopping criterion was met, defined by either convergence of the solution or reaching a maximum of 50 iterations. Upon completion, the optimal a value was extracted from the final gbest, representing the most effective hyperparameter configuration found during the search. The corresponding pseudo-code implementation of this optimization strategy is detailed in Figure 3.20.

PSO DLT-Scar optimization algorithm

```
1   Initialization PSO parameter : Swam size (S), a boundary conditions (nlow,nmax),
    max iterations, dataset (train,val,test) ,Number of training epochs
    Output: a (gbest), Highest Dice score achieved with that a (gbestfitness)
2   fori=l to S do
3       Fitness evaluation <— calculate PositionFitness (train DLT_Scar(a))
4       find gbest
5   end
6   for i < iteration do        <— Particle Updates
7       fori =1 to S
8       update velocity & position for each particle
9       CalculatePositionFitness
10      If fitness(i) > particle.pbestfitness
11          update pbest fitness & position
12            if fitness(i) > gbestfitness
13                update gbest fitness & position
14            end
15      end
16  end
17  Return gbest, gbestfitness
```

Figure 3.20 Pseudocode for PSO-DLT-Scar Training For Tversky Hyperparameter Searching

3.5.3 Performance Evaluation of LV and Myocardial Scar Segmentation

The evaluation of the proposed method in the development of Dual-Stages DeepLabV3+-based Myocardial Infarction segmentation model focused on accurately segmenting two key target classes: the left ventricle (LV) region in the first stage and the infarcted tissue or scar in the second stage. All models were trained and tested using a computer equipped with a NVIDIA GeForce RTX 4070Ti GPU and 8GB of RAM NVIDIA. For LV segmentation tasks, this study utilized two similarity metrics: the Dice Similarity Coefficient (Dice Score) and the Jaccard Index, also known as Intersection over Union (IoU). These metrics are commonly used in image segmentation to quantify how well the predicted segmentation matches the ground truth. The Dice Score is particularly useful for evaluating overlap between two regions, while the IoU Index provides a ratio of intersection over the union of the predicted and ground truth regions. Equation (3.10) demonstrates the calculation of the Dice Score and IoU in Boolean representation between the ground truth (GT Area) provided during manual delineation by radiologists and the resulting segmented area (Predicted Area). In this context, AA represents the ground truth region and BB represents the predicted region. Both metrics range from 0 to 1, with a value of 1 indicating perfect overlap. While both assess similarity, the Dice Score tends to emphasize overlap more due to the factor of 2 in its numerator, making it more sensitive to small, well-segmented regions. On the other hand, the IoU dividing the intersection by the total union of the two regions, which is particularly useful for larger regions with fewer discrepancies. Figure 3.21 provides an example of the LV region segmentation result as the target class, showing the binary image predicted from the segmentation process as the Predicted Area (green), the GT Area provided by a radiologist (magenta), and their overlapping region (white). The overlapping region is used to calculate the IoU and assess the segmentation process's performance. Both the Dice Score and IoU metrics complement each other, offering a balanced evaluation of segmentation performance, especially in cases with varying region sizes in LV scar segmentation.

$$\text{Dice Similarity Index (Dice score)} \sim \frac{2\,(\text{Area of Overlap})}{\text{GT Area}+\text{Predicted Area}} \tag{3.5}$$

$$\text{Jaccard Index (IoU)} = \frac{\text{Area of Overlap}}{\text{GT Area+Predicted Area-|AA n BB|}} \quad (3.6)$$

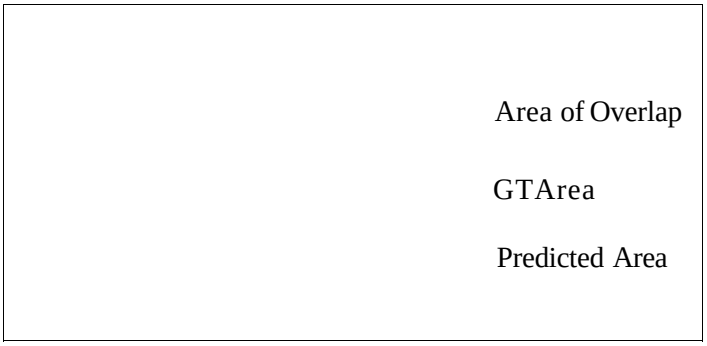


Figure 3.21 Example of the LV Region Segmentation Result as the Target Class, Including Area of Overlap, Ground Truth (GT) Area and Predicted Area (Right)

In the second segmentation task, the Dice coefficient was employed as the primary evaluation metric as it is the most widely used metric for evaluating scar and tumor segmentation in medical imaging due to its sensitivity to small, well-segmented regions as mentioned above. In addition to the Dice Score, the F2-Score was also used to evaluate the second-stage scar segmentation. The F2-Score, which emphasizes recall over precision, has been frequently employed as metric in medical applications [36] as it gives more weight to correctly identifying positive cases (scar tissue). Identifying all relevant target areas is more critical than missed target regions. F2-Score formulated in Equation 3.7

$$\text{F2-Score} = \frac{5\,(\text{Precision x Recall})}{4^{*}\,\text{Precisions- Recall}} \quad \textbf{(3.7)}$$

Recall/Sensitivity (SN) - the ratio of the number of correctly classified Scar class to the total actual Scar class. The percentage of sensitivity is given by Equation 3.8:

$$\text{Recall/sensitivity} = \frac{TP}{(TP+FN)} \quad \textbf{(3.8)}$$

Precision - the ratio of correctly classified Scar class to the total number of predicted Scar class, given by Equation 3.9:

$$\text{Precision} = \frac{}{(TP+FP)}$$

Where,

TP = The number of correctly predicted scar class (True Positive),

FP = The number of incorrectly predicted as scar class (False Positive),

FN = The number of incorrectly predicted as background classes (False Negative)

3.6 Phase 4: The Proposed MI Detection using Scar Sequential Slice Reconstruction (3SR) and DCNN

This section explains the final phase of the proposed automated MI detection. A patient confirmed to have myocardial scars on LGE-CMR is diagnosed with a history of MI. In previous steps, Section 3.5, not all segmented outputs correctly represent scars as there are some scar-like objects obtained due to the similar pixel intensity between scar tissue and artifacts from motion and magnetic field inhomogeneities within the LGE-CMR images. These artifacts can create false hyperintensities that may be incorrectly classified as scar tissue by segmentation algorithms. False segmentation can mistakenly classify healthy tissue slices as infarcted, leading to an inaccurate diagnosis.

In short-axis LGE MRI, the heart is typically imaged in cross-sectional slices from base to apex. Observations in this study reveal that myocardial scars typically exhibit a gradual change in scar size, shape and continuity across sequential slices as illustrated in Figure 3.22. The scar initially appears in the first slices as a small area at any location within the myocardium and generally enlarges in the same anatomical location in the following slices. Although it may change in size or thickness due to natural tapering, the scar should typically remain consistent with the same region. Discontinuity in the pattern may indicate the presence of imaging artifacts as shown in Figure 3.23.

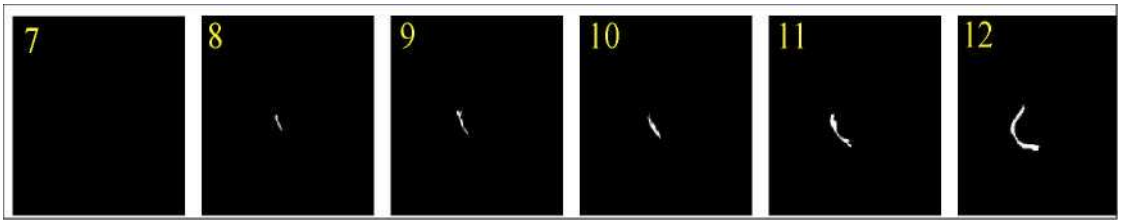


Figure 3.22 Sequential Scar Progression: An Example of the Ground Truth from Slices 7 to 12 of a Patient with an Infarcted Myocardial Scar. The Yellow Numbers Indicate the Number of the Slice and the White Regions Represent the Segmented Scar

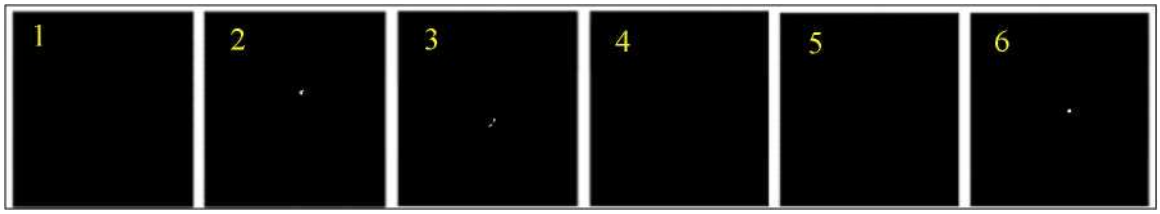


Figure 3.23 An Example Discontinuity in the Pattern in Sequential Scar Slices. The Yellow Numbers Indicate the Number of the Slice

Most previous studies have relied on a single CMR segmentation to predict images with scar regions. However, single-segmented images often lack the contextual information necessary for accurately identifying true scars. To address this limitation, the final phase of this study aims to enhance the detection of MI by identifying true scar tissue through the continuity of sequential images, combined with a DCNN model, to accurately differentiate between true scar tissue indicative of MI and healthy tissue classified as non-MI. This method aligns with Objective 4 and Phase 4 outlined in Figure 3.1, proposing a novel strategy for detecting accurate scar images that contribute to MI diagnosis by leveraging the sequence of scar continuity in images and a DCNN.

The proposed MI detection process begins by identifying a reference image containing the largest segmented scar tissue in a patient, along with sequential slices of the same image. These slices are used to determine potential MI regions through scar sequential slice reconstruction, which captures and consolidates pattern information across images. Once the reconstructed slice is generated, the S-DCNN model, as proposed in Section 3.4.1, is employed. This model has been trained on a newly defined task to analyze and identify image patterns, determining whether the patient has healthy tissue, which is classified as non-MI or has suffered from an MI. Figure 3.24 illustrates the workflow for the MI detection method. To validate the proposed method, a comparative analysis was conducted between the proposed scar image reconstruction

and single-segmented using the same S-DCNN. The following subsection provides a step-by-step explanation of the scar reconstruction process for potential MI images used as input for S-DCNN.

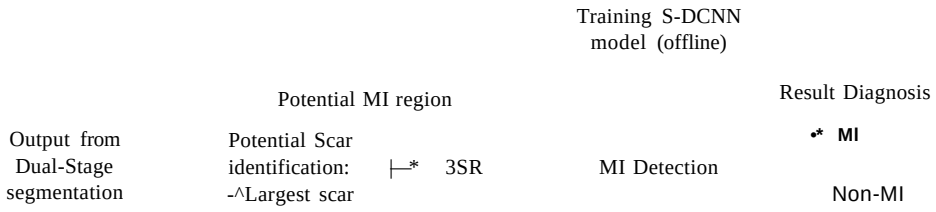


Figure 3.24 Workflow for Automated Detection of Myocardial Infarction

3.6.1 Scar Sequential Slice Reconstruction Method

As previously mentioned, myocardial scars indicate areas of the heart affected by infarction. In practice, the radiologist must compare the LGE-MRI images before and after to confirm true scar images. True scar tissue typically shows persistent enhancement. On the other hand, artifacts or temporary changes (e.g., due to motion, poor contrast timing) may appear differently. This is an important characteristic of myocardial scars [141], [142]. To address this, the proposed Scar Sequential Slice Reconstruction (3SR) method generates three-channel RGB composite images based on the spatial position of a reference slice within a patient's short-axis LGE-CMR sequence. The reference slice is defined as the slice containing the largest segmented scar, as this region generally represents the most clinically significant manifestation of myocardial infarction.

To ensure the continuity of slices is preserved at the edges of the images, the RGB composite construction process in 3 SR is determined by the position of the reference slice within the short-axis LGE-CMR sequence. Let the total number of slices in the sequence be denoted as P, and the reference slice index as i, where $1 < i < P$. The RGB channels are assigned based on three consecutive slices according to the following cases, and the pseudocode algorithm for this Scar Sequential Slice Reconstruction is shown in Figure 3.25.

1. Apical case (reference slice is the first slice, $i = 1$): The composite image is formed using the first slice and the next two consecutive slices. 3SR= Slice (1) +Slice (2) +Slice (3)
2. Basal case (reference slice is the last slice, $i = P$): The composite image is constructed using the last slice and the two preceding slices. 3 SR= Slice (P — 2) + Slice (P - 1) + Slice (P)
3. Intermediate case (reference slice is neither the first nor last slice $1 < i < P$): the composite includes the preceding slice, the reference slice, and the following slice. 3SR= Slice $(i - 1)$ + Slice (i) + Slice $(i + 1)$

Proposed Scar Sequential Slice Reconstruction algorithm

```
Initialization variables: Largest Segmented Scar Area (L)
input: total patient images (P)
Output: A new RGB image
fori=ltoPdo
Read image segmented LGE
Convert RGB to binary image
Find the largest area of scar in the sequential of patient slices as reference image:
Calculate total white area (segmented as scar) = S
    if S > L
        update area image L = S
        update reference image = i
    end
end
Load the reference image (i)
Check if the reference image is the first, last, or middle in the sequence
if i == 1:
Generate RGB images using the reference image and two slices immediately following it.
if i == P: (Last image)
Generate RGB images using the reference slice and two slices immediately preceding it.
else
Generate RGB images using the reference slice and one slice before and after it.
```

Figure 3.25 Pseudocode Algorithm for This Scar Sequential Slice Reconstruction

After reconstructing the potential MI images, this image will be fed to the DCNN model to predict whether the image is diagnosed as MI or healthy tissue labeled as non-MI. The S-DCNN architecture model proposed in subsection 3.4 is employed

for this task. Prior to that, the S-DCNN model has to be trained with a new task, which is to detect MI or non-MI.to distinguish between MI and non-MI patterns. For this purpose, a total of 1083 images were prepared for training and testing, comprising both MI and non-MI cases. Both the training and testing datasets were derived from the same dataset generated in Section 3.5, which consists of segmented images reconstructed into three-channel composites using the Scar Sequential Slice Reconstruction (3SR) strategy. Figure 3.26 displays examples of MI and non-MI patterns used during the training process.

To evaluate the effectiveness of the 3SR approach, the S-DCNN model was trained and tested using two different input strategies, as illustrated in Figure 3.27. In the first approach, shown in Figure 3.27(a), images were reconstructed using the 3SR method from a sequence of consecutive slices was input into the model. In the second (baseline) strategy, shown in Figure 3.27(b), individual segmented slices were used without reconstruction. To ensure a fair comparison, a mini-batch size of 16, a maximum number of epochs of 70, and the Adam optimizer were selected for the S-DCNN model during training on both single images and scar sequential slice reconstruction images.

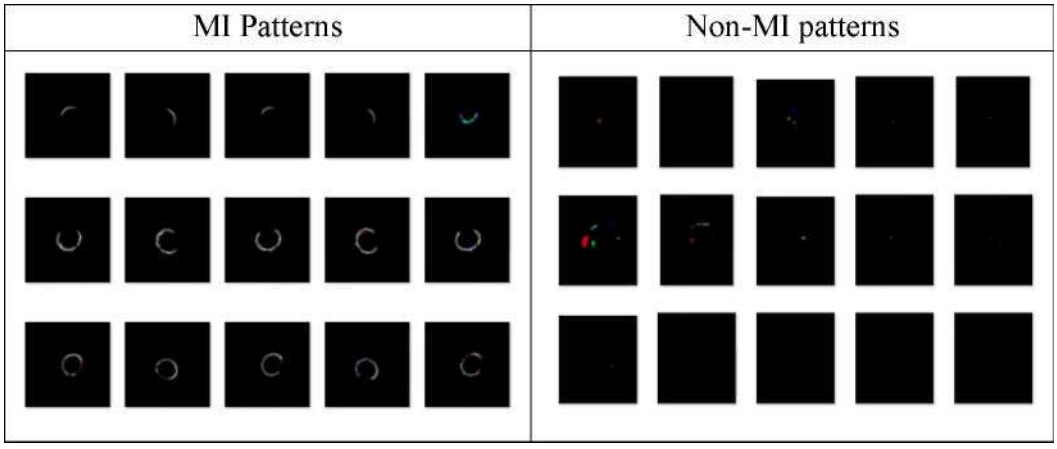
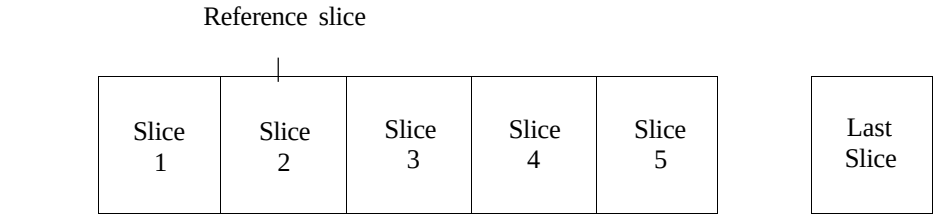
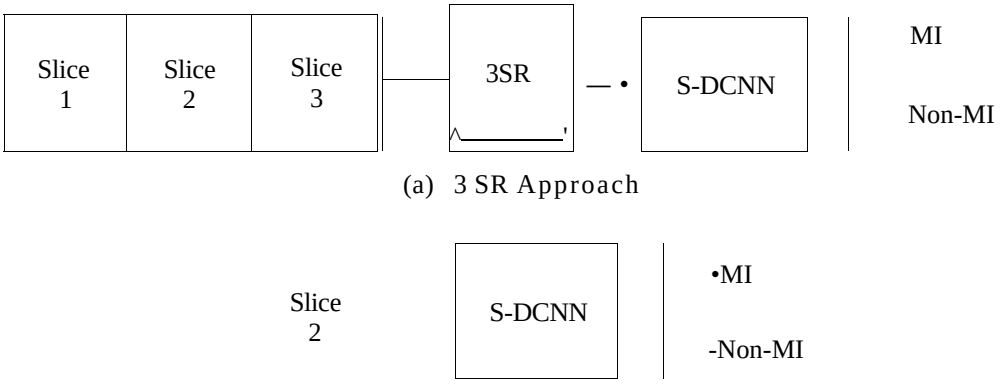


Figure 3.26 Displays Examples of MI and Non-MI Pattern Samples From 3SR Method Used in Training Dataset



The output segmented slices for a given patient.



(b) Single Approach

Figure 3.27 Integration of Scar Sequential Slice Reconstruction (3SR) with the S-DCNN Model Compared to the Single-Slice Approach

3.6.2 Performance Evaluation of 3SR and S-DCNN

The detection of true image with MI and vs. Non-MI that indicates truly scar appearance in the image is evaluated by measuring the percentages of True Positive (TP) and True Negative (TN) according to the CM components , with detection accuracy as defined in Equation 3.10.

$$\text{Accuracy} \quad \frac{\text{Total Number of correctly images classified}}{\text{Total Number of images}} \qquad (3.10)$$

To further evaluate the model's precision in detecting scarred regions, sensitivity (also known as recall) was used. Sensitivity focuses specifically on the model's ability to correctly identify MI images. The percentage of sensitivity is given by Equation 3.11:

$$\text{Recall/sensitivity} = \frac{TP}{(TP+FN)} \quad (3\ 11)$$

On the other hand, specificity measures the model's ability to correctly identify non-Mil images, reducing false positives by ensuring healthy tissues are not misclassified as MI. The percentage of specificity is given by Equation 3.12:

$$\text{Specificity} = \frac{TN}{(TN+FP)} \quad (3\ 12)$$

The definition for each CM value is described below:

TP - The number of correctly classified MI

TN - The number of correctly classified non-MI

FP - The number of incorrectly classified as MI

FN - The number of incorrectly classified as non-MI

3.7 Summary

This chapter discusses the overall proposed method for fully automatic myocardial infarction (MI) detection in short-axis LGE-MRI images. Currently, in Malaysia, manual detection remains the most commonly used method for MI diagnosis. This process involves direct human intervention, requiring radiologists to visually inspect the images and identify myocardial scars indicative of MI. The accuracy of this method heavily depends on the radiologist's knowledge and experience. The proposed method is developed based on radiologists' manual interpretation of the images during clinical evaluation. The automated approach begins by processing the images received by the radiologist for MI detection evaluation.

In this study, a dataset comprising short-axis LGE-CMRI image sequences was collected locally from the Advanced Medical and Dental Institute (AMDI), Universiti Sains Malaysia (USM), Malaysia. This dataset was utilized to develop and test the proposed method at all stages, as described in Section 3.3. In alignment with research objective 1, a new automated procedure for MI detection was developed, focusing on

LV classification since the LGE-CMRI images contain both LV and non-LV regions, as detailed in Section 3.4. By utilizing a DCNN to classify LV and non-LV regions, this study introduces an innovative approach that eliminates the need for manual intervention, streamlines the process and facilitates LV segmentation for subsequent MI detection. Moreover, the proposed classification method reduces analysis time by excluding non-LV regions from further evaluation, ensuring that only relevant regions are analyzed for MI detection.

Next, Section 3.5 addresses Research Objective 2 by introducing a novel dual-stage segmentation framework specifically designed to optimize myocardial scar segmentation, with each stage tailored to its respective segmentation task. The first-stage network, DLT-LV, is designed to segment the target LV from complex backgrounds. In this stage, the original backbone of the DeepLabV3+ model was replaced with alternative network architectures, accompanied by enhancements to the Atrous Spatial Pyramid Pooling (ASPP) module. These modifications were specifically tailored to improve the segmentation of LV regions, building on the foundation of the original DeepLabV3+ model. To address variations in LV size, the ASPP module was extended with smaller dilated rate tuning to extract multiscale features with greater distinguishability. Three pretrained architectures were evaluated as backbones to assess the performance of the proposed ASPP module. Additionally, a post-processing procedure was introduced to refine the segmentation outputs of the DLT-LV network. The second-stage network, DLT-Scar, is designed to segment myocardial scars and employs the same model architecture as DLT-LV, with the exception of the backbone. The optimal backbone for each stage was chosen based on its task-specific performance. In this stage, an investigation into four different loss functions was conducted for scar segmentation, aiming to address the challenges of class imbalance in medical imaging. Among the evaluated loss functions, the Tversky loss function demonstrated the best performance, making it the ideal choice for the task. To further enhance the model's effectiveness as in line with Research Objective 3, a novel method was proposed to automate the optimization of the Tversky loss function's hyperparameters within the DLT-Scar model. This optimization was achieved by using PSO to fine-tune the hyperparameters and improve the segmentation accuracy. Furthermore, the hyperparameters of the Tversky loss function were simplified to a single parameter, a, by recognizing the redundancy between the two original hyperparameters, a and p. This simplification significantly reduced the hyperparameter search space, making the

optimization process more efficient while maintaining model performance. Overall, the proposed dual-stage segmentation framework enhances myocardial scar detection by optimizing both the LV segmentation and scar segmentation stages, offering a robust solution for myocardial scar detection in complex and imbalanced medical imaging scenarios.

In medical diagnosis, higher recall is often prioritized to ensure that all potential cases are identified. However, increasing recall also makes the model more susceptible to false positives. This limitation is particularly challenging in LGE-MRI, where scar-like artifacts can closely resemble true myocardial scars, potentially leading to false diagnoses, particularly in healthy tissue. To address this issue, and in line with the radiologist's approach of reviewing images before and after to identify true scars, the Scar Sequential Slice Reconstruction method was introduced following the segmentation process. This method is consistent with Research Objective 4 aims to improve the identification of true scars in short-axis LGE-CMRI, thereby ensuring more accurate MI detection. Section 3.6 details the proposed automated method for true scar identification, which further classifies MI and healthy tissue using the DCNN-S model. The following chapter presents the results and discussion of the proposed methodology.

CHAPTER 4
RESULTS AND DISCUSSION

4.1 Introduction

This chapter continues the findings of the study by presenting and discussing the results. It is divided into four main parts. The first part focuses on the results and discussion of the proposed method for classifying LV and non-LV regions. The study provides a detailed analysis of the shallow deep convolutional neural network (S-DCNN) model, developed from scratch, by evaluating different network layers and optimizing the model using batch normalization, two network optimizers, and learning rates. Additionally, a comparison is made between the results of this S-DCNN model and several well-known pre-trained DCNN architectures, which are considered alternative methods.

The second part presents an in-depth analysis of the newly proposed Dual-stage DeepLabV3+-based segmentation model on LGE-MRI images containing only the LV region. The first stage involves LV region segmentation, while the second stage focuses on myocardial scar segmentation. The analysis is presented separately for each stage, providing a comprehensive evaluation of the results. The first stage examines the performance of the proposed DLT-LV architecture, followed by a morphological procedure to refine the LV segmentation by eliminating isolated mis-segmentations and small gaps. The second stage focuses on analyzing the proposed DLT-Scar segmentation model. The focus then shifts toward optimizing the DLT-Scar segmentation model, where a comparison of different loss functions is conducted to address class imbalance more effectively, ultimately leading to an effective solution for the task. Then, the analysis explores the outcomes of automated hyperparameter searching for the Tversky loss function using the PSO profiling approach to achieve optimal performance.

The third part of this chapter analyzes the performance of a newly proposed method for predicting the presence of true scar tissue in LGE-CMR segmented images, which is essential for identifying MI. This method focuses on detecting myocardial scar tissue, a key indicator of MI, by analyzing its appearance across a sequence of LGE-CMR images. In this approach, the sequential features of the scar regions are

reconstructed to capture the continuity and progression of scarring across three adjacent image slices. Once these sequential features are restructured, the DCNN-S model, introduced in Section 3.4, is applied to classify the images and determine whether the patient has normal myocardial tissue or a history of myocardial infarction. This part of the study evaluates the effectiveness of the sequential reconstruction method in distinguishing true scar images compared to using a single image for MI classification.

Finally, the last part of this chapter describes the proposed fully automated myocardial infarction segmentation and detection framework, evaluated against the radiologist's manual interpretation. This system aims to assist radiologists in diagnosing the presence of myocardial scar associated with myocardial infarction in LGE-MRI images through a quick and straightforward process.

4.2 Analysis of LV Classification Using DCNN

This section presents the analysis of the proposed automatic method for detecting images containing the LV region by classifying images with LV and non-LV regions using DCNN models. The purpose of this method is to identify the LV region as the initial step for the subsequent segmentation process, where only images with the LV region will be used in the segmentation stage. As outlined in Subsection 3.4, the DCNN model was designed specifically to classify images based on the presence of LV and non-LV regions in sequential LGE-MRI. Section 4.2.1 presents the results and analysis of the new proposed S-DCNN architecture, while Section 4.2.2 presents the results obtained using three well-known DCNN architectures, namely AlexNet, GoogLeNet, and SqueezeNet, employing pre-trained models as alternative classifiers. The performance of these models is compared to the proposed S-DCNN through both qualitative and quantitative analyses.

4.2.1 Analysis of S-DCNN Model

To analyse the performance of the LV and non-LV classification using S-DCNN model, several parameters such as the number of convolution layers, batch normalization layer and training algorithm are investigated. For the initial step, the effect of increasing the number of convolution layers from two (2) to four (4) in building S-DCNN model was investigated. Adam optimizer algorithm was adopted to train with

an initial learning rate of 0.0001. The main purpose of this observation is to analyse the effect of adding a convolution layer to the network performance in terms of its prediction accuracy. Table 4.1 presents the precision, recall, F1-score, and accuracy results for different numbers of convolutional layers (Conv Layer) used in the S-DCNN model. The results show that increasing the number of convolutional layers generally leads to improved performance up to a point. With one convolutional layer, the model achieves an accuracy of 88.14% and an F1-score of 87.79%. As additional layers are introduced, the model's ability to extract hierarchical and more abstract features improves, leading to better discrimination between LV and non-LV regions. This is evident as the performance steadily increases, with three convolutional layers providing the best performance, achieving an accuracy of 91.89% and an F1-score of 91.80%. Precision and recall follow a similar upward trend, improving as more layers are introduced. However, when the model uses four convolutional layers, the performance slightly decreases, with accuracy dropping to 89.18% and F1-score to 88.80%. This reduction can be attributed to two primary factors. First, adding excessive layers introduces more parameters, increasing model complexity and making it more prone to overfitting, especially when training on a limited dataset. The model starts memorizing noise rather than learning meaningful patterns, leading to reduced generalization to unseen data. Second, deeper networks require more computational resources and may suffer from vanishing gradients, where information from the earlier layers is weakened as it propagates through the network, hindering effective feature learning.

Table 4.1
Precision, Recall, F1-Score, Accuracy Using Different Convolutional Layers

Conv Layer	Precision	Recall	F-Score	Accuracy
1	90.52	85.21	87.79	88.14
2	92.24	85.79	88.90	89.29
3	**92.90**	**90.71**	**91.80**	**91.89**
4	92.03	85.79	88.80	89.18

Next, the investigation focuses on the incorporation of batch normalization layers into the highest-scoring model from the previous experiment, the S-DCNN with three (3) convolutional layers. As batch normalization can interact with different

optimizers during training, this study further evaluated the model with and without batch normalization, using three different learning rates (0.0001, 0.001, 0.01) and two optimizers: SGDM and Adam. Figure 4.1 illustrates the impact of incorporating batch normalization for the Adam and SGDM optimizers across three different learning rates. The results consistently showed significant performance improvements with BN across all configurations. At a learning rate of 0.0001, the Adam optimizer without BN achieved 91.89%, but with BN, it increased substantially to 97.36%. SGDM at the same learning rate followed a similar trend, improving from 90.16% without BN to 94.26% with BN. This suggests that batch normalization stabilizes training and enhances performance at lower learning rates. For a learning rate of 0.001, SGDM showed a dramatic improvement from 52.26% to 94.96% when BN was applied. Likewise, Adam also benefited from BN, improving from 88.49% to 91.96%. At a higher learning rate of 0.01, both Adam and SGDM performed poorly without BN, achieving 42.09% and 42.26%, respectively, but with BN, both improved significantly to 82.56% and 84.46%, respectively. These findings reinforce the crucial role of BN in mitigating internal covariate shift and allowing the network to train effectively across different learning rates. Overall, the Adam optimizer with a learning rate of 0.0001 and batch normalization delivered the best performance, achieving 97.36% test accuracy, making it the optimal combination for this model and task. This underscores the importance of BN in deep learning architectures, particularly in stabilizing training and improving model robustness under various hyperparameter settings.

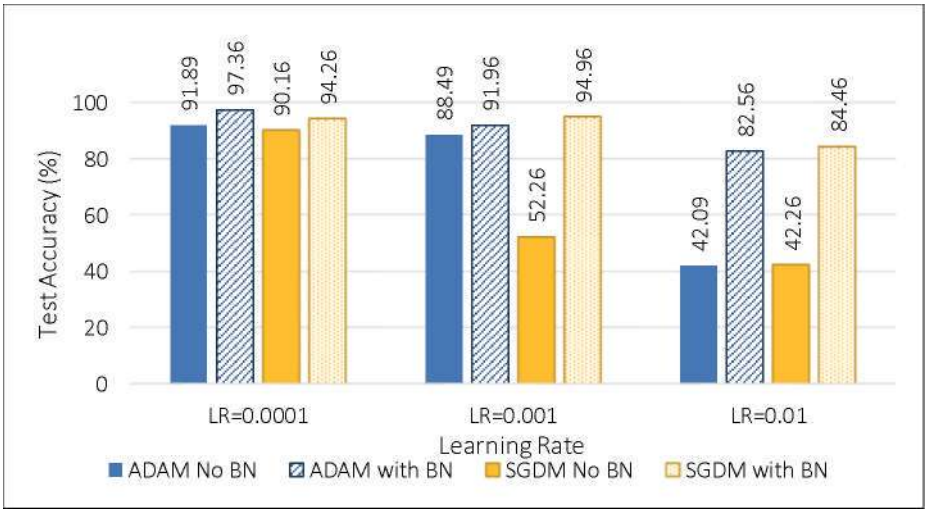


Figure 4.1 Test Accuracy Results for the Adam and SGDM Optimizers with and Without BN Across Three Different Learning Rates

4.2.2 Analysis of Pre-Trained DCNN Model

This subsection compares the S-DCNN model against three customized well-known DCNN architectures: AlexNet, SqueezeNet, and GoogleNet. The results in Table 4.2 demonstrate the effectiveness of the proposed method using DCNN models in terms of depth, accuracy metrics, and training time. In this context, the depth refers to the total number of layers within the respective DCNN models, reflecting their architectural complexity. The proposed S-DCNN model, designed with only 4 layers, stands out for its lightweight structure. In comparison, AlexNet possesses a moderate depth of 8 layers, while SqueezeNet and GoogleNet are considerably deeper with 18 and 22 layers, respectively. All models demonstrated strong performance in classifying LV regions, achieving accuracy rates above 90%. The S-DCNN with 3 convolutional layers achieves an accuracy of 97.36% slightly outperforms AlexNet (96.43%), but falling just below the performance of GoogleNet (98.75%) and SqueezeNet (98.39%) by approximately 2%. GoogleNet exhibits the highest overall performance with a precision of 99.21%, recall of 98.29%, F-Score of 98.74%, and accuracy of 98.75%, indicating its superior capability in accurate classification. SqueezeNet closely follows, with precision and recall values of 99.06% and 97.71%, respectively, resulting in an F-Score of 98.38% and an accuracy of 98.39%. These metrics highlight SqueezeNet's robustness, making it a strong alternative to GoogleNet. Notably, the S-DCNN model excels in computational efficiency, with a training time of just 20 minutes and 47 seconds, significantly shorter than the 2 hours and 27 minutes required by GoogleNet, 1 hour and 7 minutes for SqueezeNet, and 1 hour and 6 minutes for AlexNet. While S-DCNN model exhibits slightly lower accuracy compared to pre-trained models like SqueezeNet and GoogleNet offers significant advantages in simplicity and computational efficiency, making it well-suited for integration as a computer-aided system with other networks in subsequent processes. Its less complex architecture translates to faster training times and reduced resource requirements, making it a practical choice for resource-constrained environments. Furthermore, unlike pre-trained models that require fixed input image sizes, the S-DCNN model offering greater adaptability to different image sizes datasets and potential future applications without the need for image resizing, which can introduce noise and potentially degrade performance. These strengths in simplicity and efficiency make the S-DCNN as a

compelling alternative to pre-trained models, particularly when computational resources are limited.

Table 4.2
Comparison of the Training Time and Accuracy S-DCNN and Pre-Trained AlexNet, SqueezeNet and GoogleNet

DCNN Model	**Depth (layers)**	**Precision (%)**	**Recall (%)**	**F-Score (%)**	**Accuracy (%)**	**Training Time (h:m:s)**
S-DCNN (3 conv)	4	97.77	96.93	97.35	97.36	0:20:47
Alexnet	8	97.03	95.79	96.41	96.43	1:06:35
SqueezeNet	18	99.06	97.71	98.38	98.39	1:07:05
GoogleNet	22	99.21	98.29	98.74	98.75	2:27:03

The details of the percentage of each output class in the confusion matrix are illustrated in Figure 4.2 (a) to (d), which shows the testing classification performance for each output class. All models exhibit a similar pattern of achieving high accuracy in classifying non-LV images while encountering relatively greater challenges in classifying LV images, as evidenced by higher false negative rates. Additionally, the models demonstrate balanced performance with a strong emphasis on minimizing misclassifications. These similarities highlight the proposed method for LV and non-LV classification is robust and capable of maintaining consistent accuracy across different architectures.

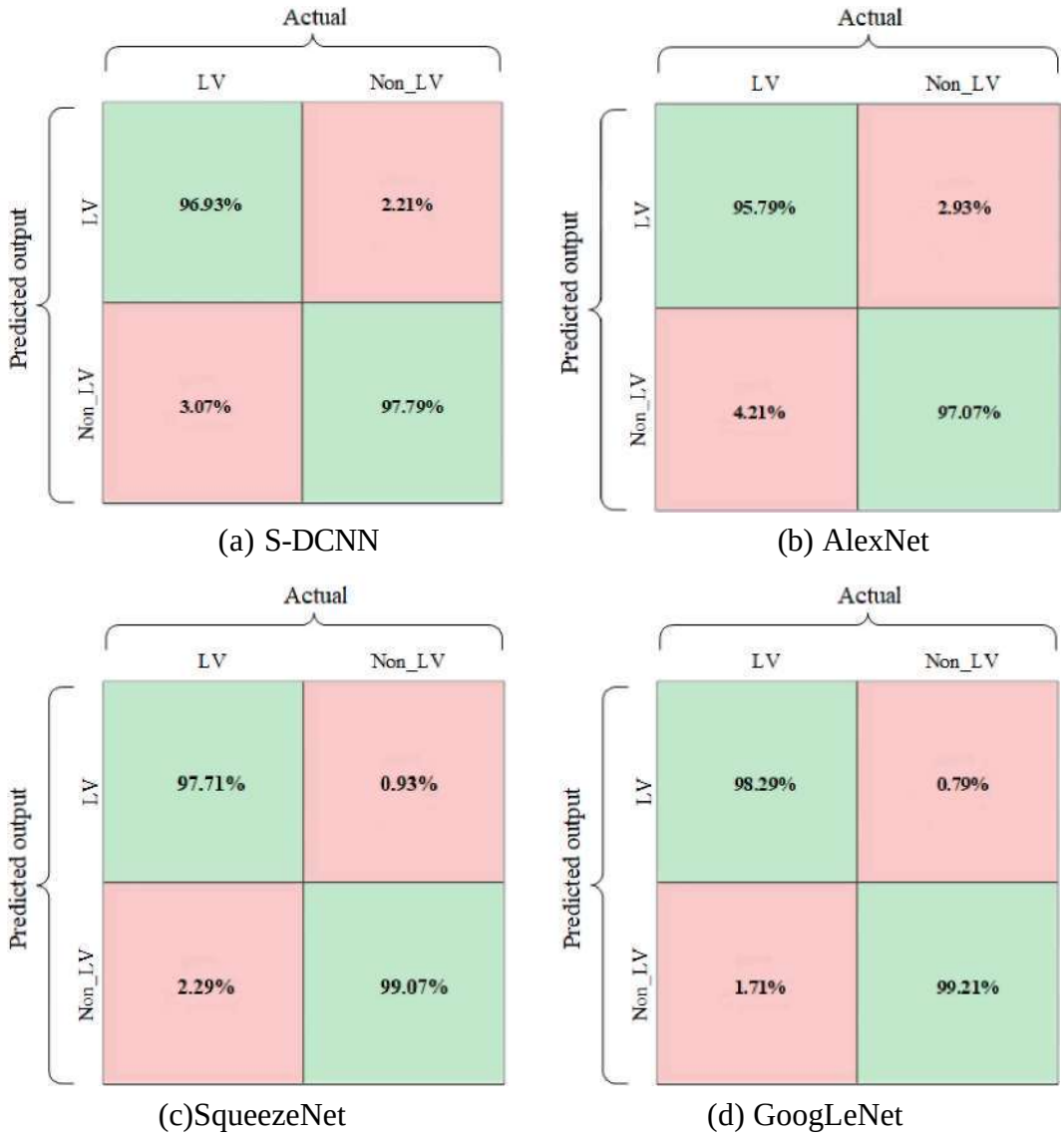


Figure 4.2 Confusion Matrix Illustrating the Classification Performance Across Different DCNN Models

4.3 Analysis of New Dual-Stages DeeplabV3+-Based Myocardial Scar Segmentation

This section presents the analysis of the newly proposed fully automated dual-stage myocardial scar segmentation model, applied to the original LGE-MRI images identified in Section 4.2 as containing the LV region.

The segmentation pipeline consists of two stages: the first involves LV region segmentation using the DLT-LV model, and the second focuses on myocardial scar segmentation using the DLT-Scar model. In the first stage, the LV segmentation results using DLT-LV was benchmarked against the standard DeepLabV3+ architecture. This

was followed by an analysis incorporating post-processing refinement techniques to enhance the segmentation output by addressing issues such as small gaps within the LV region and isolated mis-segmentation. The second stage of this section presents the analysis of myocardial scar segmentation using the DLT-Scar network, with particular emphasis on evaluating different loss functions and optimizing the Tversky loss function through the integration of the PSO algorithm. The performance of the PSO-optimized configuration is compared against the standard Tversky parameter settings to assess its effectiveness in improving segmentation accuracy.

4.3.1 Analysis for First Stage: LV Region Segmentation Model

This subsection presents the analysis of the proposed DLT-LV model for LV region segmentation, which forms the first stage of the dual-stage DeepLabV3+-based myocardial scar segmentation framework. It consists of two key analyses: the first focuses on evaluating the proposed DLT-LV architecture, beginning with architectural experiments involving three different backbone networks and a modified ASPP module. The second analysis assesses a post-processing refinement procedure applied to the DLT-LV segmentation output to address minor segmentation errors. This refinement is crucial for generating accurate input for the second-stage scar segmentation and is aimed at improving LV segmentation accuracy.

4.3.1.1 Analysis of DLT-LV Network

This study began by investigating the backbone of the DeepLabV3+ for the LV segmentation task before further enhancement with the proposed ASPP module as described in Section 3.5.1.1. Figures 4.3, 4.4, and 4.5 present the segmentation results for the upper basal, middle, and lower apical segments of the LV region on short-axis LGE-CMRI images, respectively, comparing the outputs from the DeepLabV3+ model with three different backbone networks, namely MobileNetV2, ResNet50, and Xception. The purpose of this finding is to select the ideal backbone for LV region segmentation for further improvement. The results reveal distinct strengths and weaknesses for each backbone, with all three backbones successfully segmenting the middle LV region but encountering significant challenges in segmenting the apical and basal segments of the LV region.

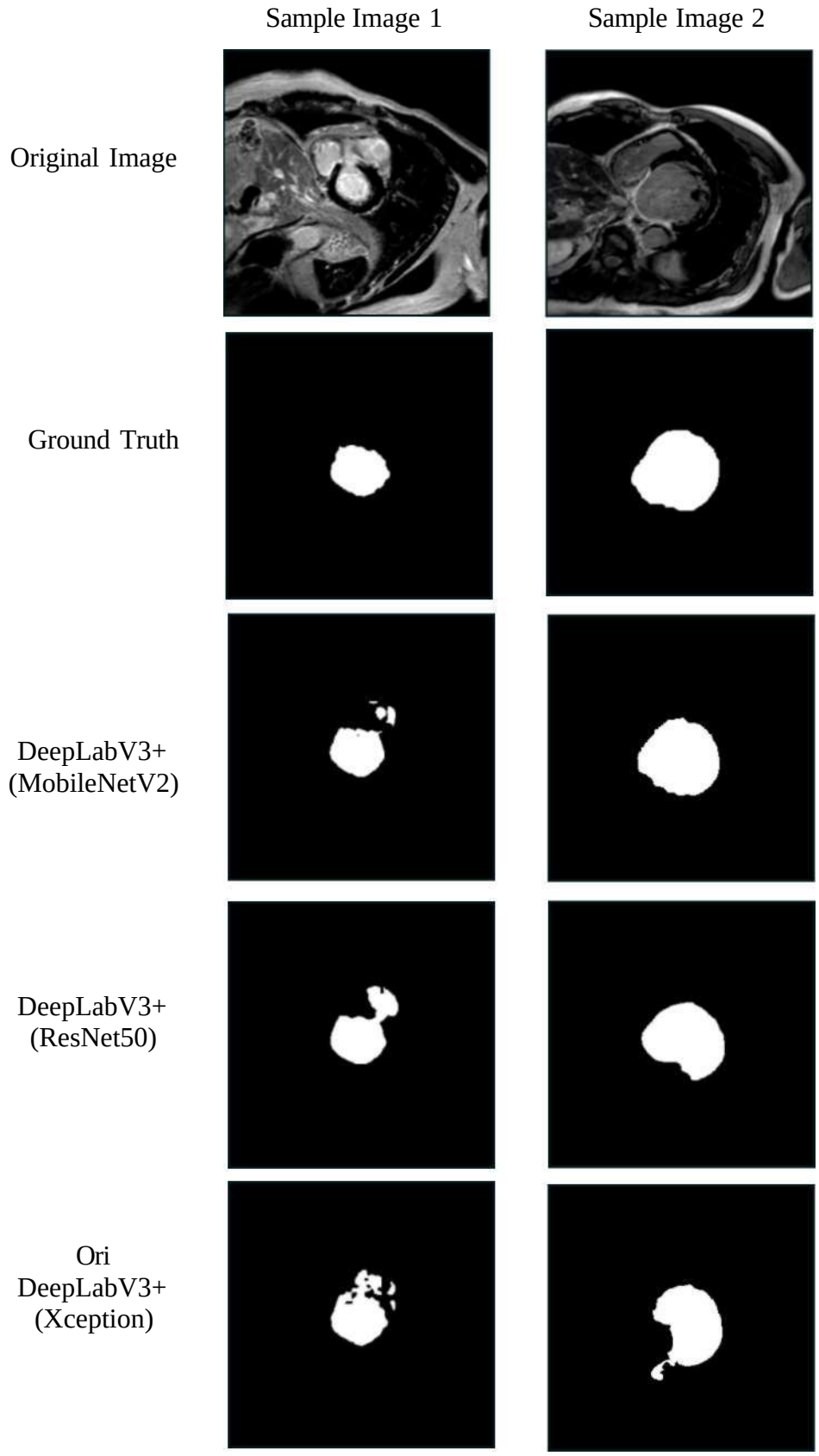


Figure 4.3 Example of Visual Segmentation Results of the Top Basal LV Chamber on LGE-CMRI and its Ground Truths

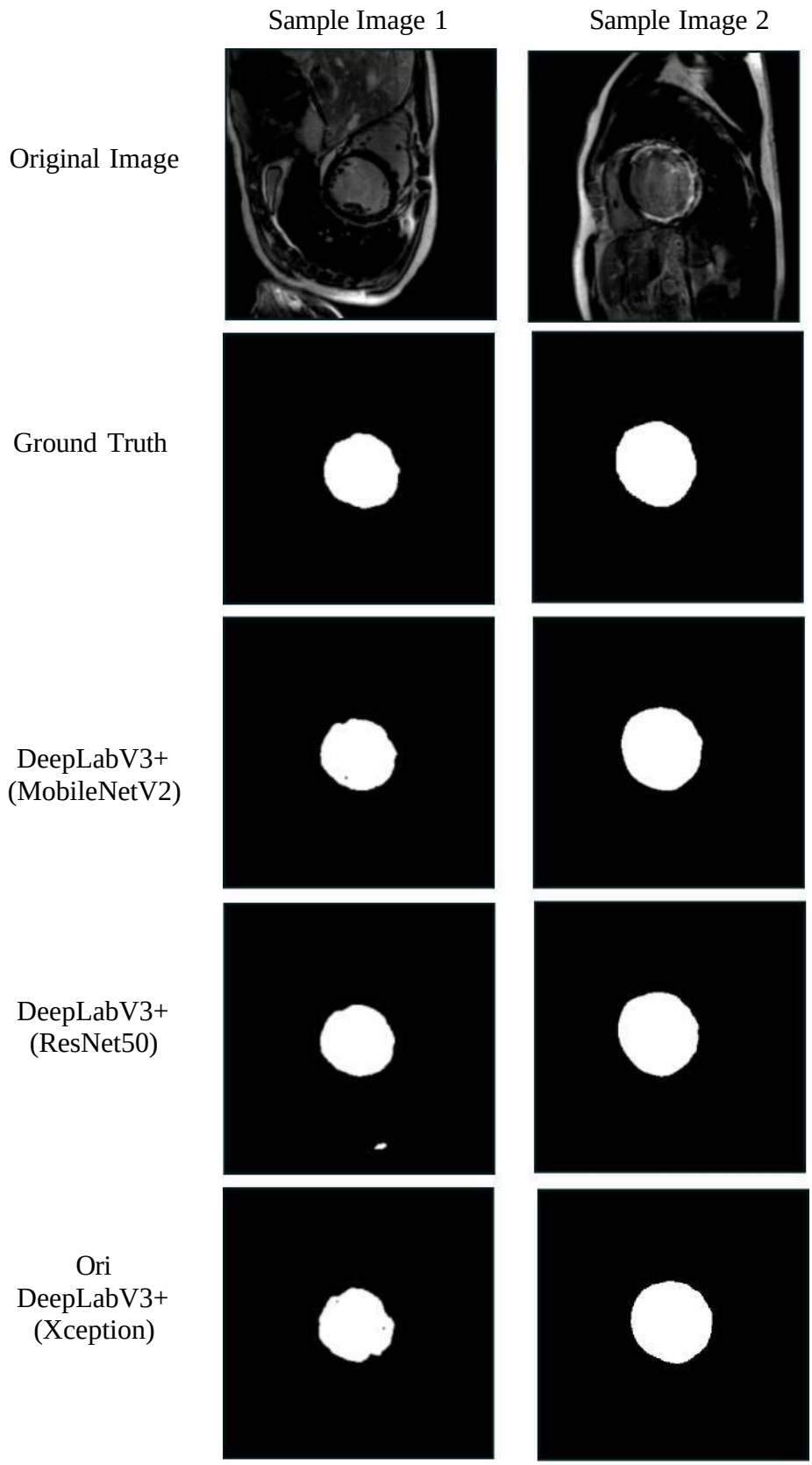


Figure 4.4 Examples of Visual Segmentation Results of the Middle LV Chamber on LGE-CMRI and Its Ground Truths

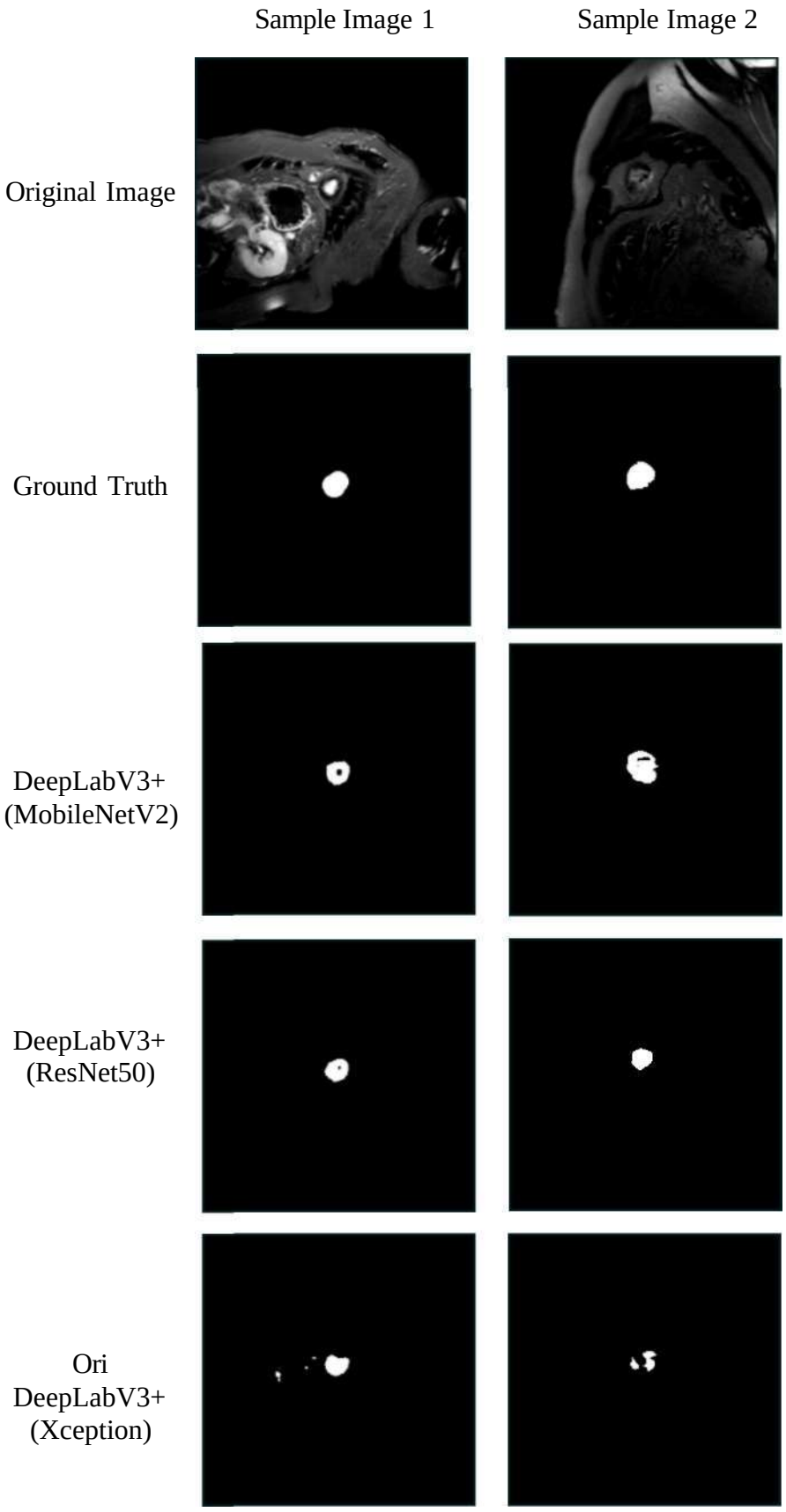


Figure 4.5 Visual Segmentation Results of the Lowest Apical LV Chamber on LGE-CMRI and Its Ground Truths

As illustrated in Figures 4.3 and 4.5, which correspond to the basal and apical LV regions, the DeepLabV3+ model using the MobileNetV2 backbone produces outputs that show comparatively better alignment with the ground truth than those generated by ResNet50 and Xception. However, all models still exhibit noticeable inaccuracies in segmenting these regions, largely due to the tapering geometry of the heart at the apical end and the structural complexity present in the basal region. In addition, inter-subject anatomical variability, such as differences in heart size, further contributes to these challenges. Collectively, these factors result in a multi-scale representation of the LV region across slices, which complicates consistent boundary delineation during segmentation. Following the qualitative observations, this issue is further investigated through quantitative analysis using the proposed modified ASPP module, which is designed to enhance the model's ability to capture multi-scale contextual information and improve segmentation accuracy in complex regions. The ASPP module in DeepLabV3+ is designed to capture multi-scale contextual information by applying dilated convolutions at varying rates, allowing the model to effectively manage variations in object size and spatial resolution. Subsection 3.5.1.1 details the architectural modifications introduced in the proposed ASPP module aimed at enhancing feature extraction for LV segmentation. To evaluate the effectiveness of this enhancement, all three backbone networks mentioned earlier were integrated into the DeepLabV3+ architecture to perform LV segmentation using the proposed ASPP module. This experimental provides a comprehensive assessment of the proposed design relative to the original DeepLabV3+ configuration, which employs the Xception backbone with the standard ASPP module.

Table 4.3 presents the quantitative performance comparison of the original and proposed ASPP modules across three DeepLab V3+ backbone networks (MobileNetV2, ResNet50, and Xception) for LV segmentation. The performance is evaluated using mean IoU and mean Dice Score. Using the original ASPP module, the results indicate that MobileNetV2 achieved the highest mean Dice Score of 96.12%, followed by ResNet50 at 96.09%. Xception, which serves as the benchmark backbone in the original DeepLabV3+ architecture, however yielded the lowest Dice Score of 95.47%. A similar trend is observed in the mean IoU, where MobileNetV2 again achieved the highest performance at 94.91%, followed closely by ResNet50 at 94.78%, while Xception recorded the lowest at 94.44%. These results highlightthat all three backbones achieved strong LV segmentation performance, with both mean Dice scores and IoU values

exceeding 90%. Among them, MobileNetV2 emerged as the most effective backbone when paired with the standard ASPP configuration. When the proposed modified ASPP module was integrated, performance improvements were observed across all three backbones. The MobileNetV2 backbone again achieved the highest overall performance, with a Dice score of 97.02% and an IoU of 95.60%. Similarly, ResNet50 and Xception showed performance gains, reaching Dice scores of 96.96% and 96.02%, respectively. The consistent improvements across all backbones highlight the effectiveness of the proposed modified ASPP module in enhancing contextual feature representation and segmentation performance.

Table 4.3
Quantitative Comparison of Original and Improved ASPP Modules Across Three Backbone Networks

Backbone	**Original ASPP**		**Improved ASPP module**	
	IoU	**Dice Score (%)**	**IoU**	**Dice Score (%)**
MobileNetV2	94.91	96.12	**95.60**	**97.02**
ResNet50	94.78	96.09	95.32	96.96
Xception	94.44	95.47	94.91	96.02

Figure 4.6 compares the performance of all three backbone networks in terms of mean Dice Score and training time, based on segmentation results using the proposed modified ASPP module. As illustrated, MobileNetV2 achieves the highest Dice Score of 97.02% while also recording the shortest training time, approximately 540 seconds, demonstrating both superior accuracy and computational efficiency. The effectiveness of MobileNetV2 can be attributed to its lightweight architecture, which enables faster convergence during training process. In contrast, ResNet50 achieves a comparable Dice Score of 96.96% but requires a significantly longer training time of about approximately 940 seconds. Xception which serves as the benchmark backbone in the original DeepLabV3+ architecture, shows the lowest performance with a Dice Score of 96.02% and the longest training time at approximately 1,020 seconds. These results highlight the practicality of the MobileNetV2 as the backbone in DeeplabV3+ architecture for LV segmentation task particularly in applications where computational resources are limited or real-time segmentation is required.

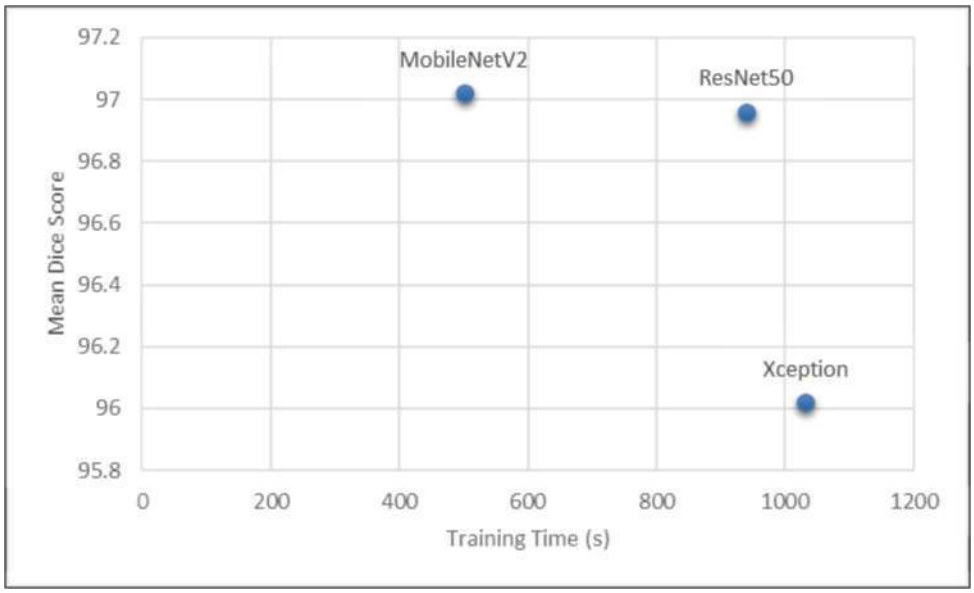


Figure 4.6 Correlation Analysis of the Mean Dice Score and Training Time for the Three DeeplabV3+ Backbones

Following the identification of MobileNetV2 as the most suitable backbone for LV segmentation, further analysis was performed to visually assess the performance of the proposed DLT-LV model. Figure 4.7 presents a side-by-side comparison of segmentation outputs generated by the baseline DeepLabV3+ model using the standard ASPP module and the proposed DLT-LV model, both incorporating the MobileNetV2 backbone. The results are displayed across three representative LV slices: (a) upper basal, (b) middle, and (c) lowest apical regions. Across all anatomical levels, the DLT-LV model demonstrates improved segmentation performance compared to the baseline model. Although minor boundary misalignments are still observed, the segmentation output adequately captures the LV region, ensuring its reliable inclusion for the subsequent scar segmentation stage. Quantitatively, the DLT-LV model demonstrated an increment in Dice and IoU scores in each region, with particularly notable improvements in the challenging basal and apical slices. For example, in the apical slice, the Dice Score increased from 77.06% to 79.10% and the IoU improved from 62.68% to 65.42%. While the segmentation of small or ambiguous LV regions remains a challenge, the proposed ASPP module incorporating adjusted dilation rates, batch normalization, and an additional convolutional branch demonstrates improved multi-scale contextual learning by more effectively capturing spatial features across varying resolutions.

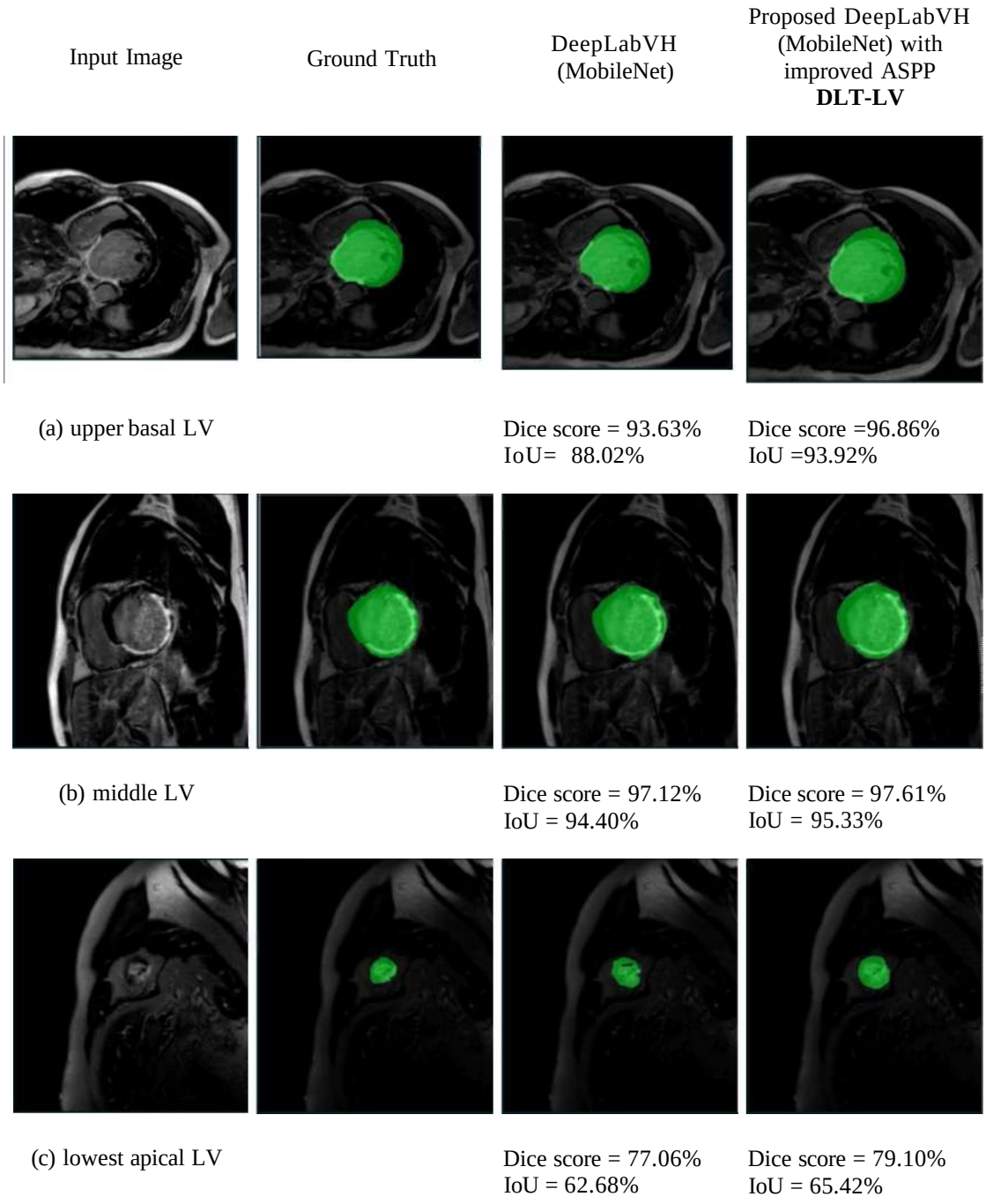


Figure 4.7 Qualitative and Quantitative Comparison of LV Segmentation Using Baseline model and Proposed DLT-LV Models Across Basal, Middle, and Apical Slices

4.3.1.2 Analysis of DLT-LV Refinement Using Morphological Procedure

At the end of the LV segmentation stage, this study proposed morphological procedures to correct and remove segmentation errors, such as small gaps within the LV region and isolated mis-segmentations, that occurred after DLT-LV model

segmentation. By doing so, it is expected that the DLT-LV model will correctly segment the LV region, ensuring that the input image fed to the second stage is accurate. For this analysis, the segmented outputs of all testing images from DLV-LV were refined using a morphological algorithm. Figure 4.8 presents a visual analysis of the image refinement process, comparing segmentation outputs before and after applying the proposed algorithm. As shown in Figure 4.8(a), the initial output from the DLT-LV model contains segmentation errors, including spurious regions and internal gaps. Following the application of the targeted morphological algorithm, these errors are effectively eliminated, as illustrated in Figure 4.8(b).

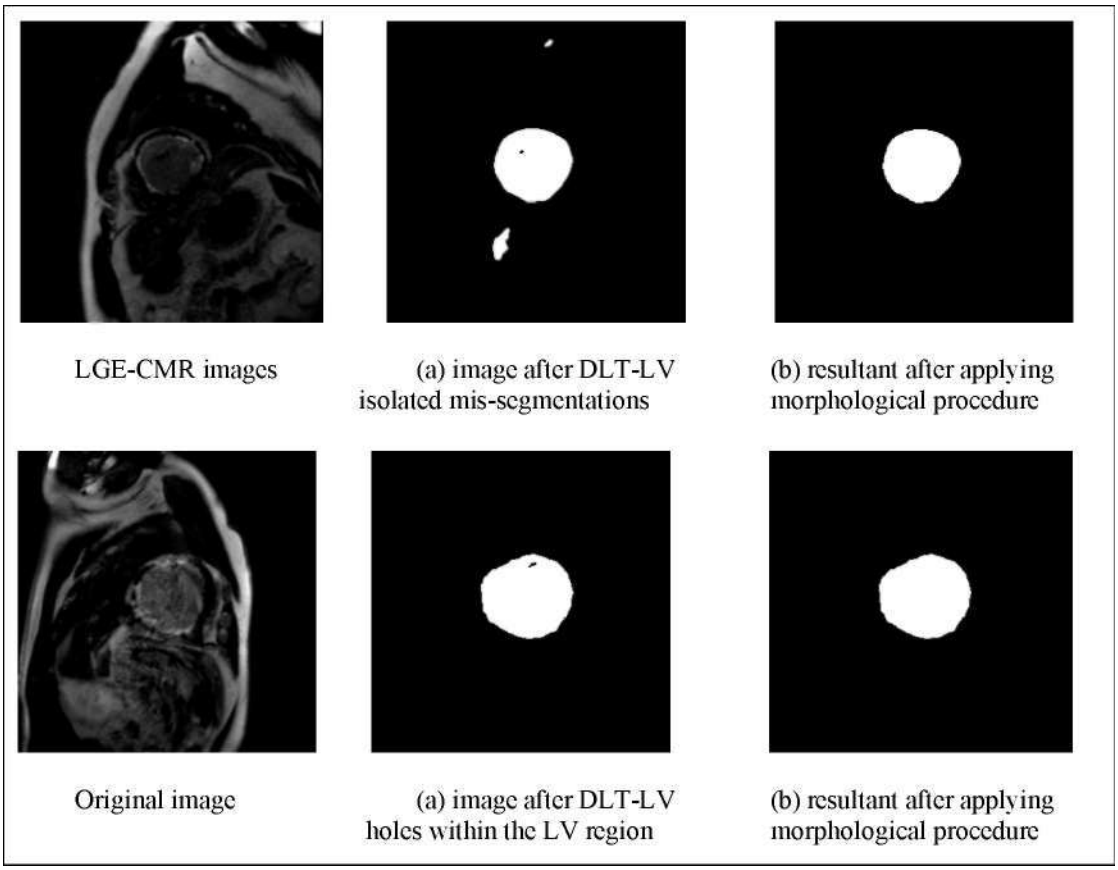


Figure 4.8 Example of Segmented Image Before and After Image Refinement Using the Morphological Procedure

The quantitative assessment before and after applying the image refinement performance is presented in Table 4.4. Following the application of the refinement algorithm, the results show a slight improvement in both the Dice score and IoU around 0.78% and 0.37%, respectively. Small increment in IoU relative to the Dice score because the Dice score is generally more sensitive to small changes in overlap because

it gives more weight to the intersection. Although the results show only a small increment in performance, this post-segmentation procedure is significant as it addresses critical segmentation errors by correcting false negatives (isolated mis-segmentations) and false positives (closing gaps within the LV). Additionally, it acts as a precautionary measure to prevent segmentation errors that might mistakenly remove or add regions resembling scars in other areas. By improving the overall delineation of the LV region, it provides a reliable foundation for the subsequent scar segmentation task to be performed accurately.

Table 4.4
Performance Comparison of DLT-LV Segmentation Output with and Without the Morphological Procedure.

LV Output	**Mean Dice Score (%)**	**Mean IoU (%)**
Without morphological procedure	97.02	95.60
With morphological procedure	**97.80**	**95.97**

4.3.2 Analysis for Stage 2: Scar Segmentation Model

This subsection analyzes the performance of the DLT-Scar model, which shares the same architecture as DLT-LV used previously for LV segmentation, with the exception of the backbone network. The optimal backbone for myocardial scar segmentation was selected from MobileNetV2, ResNet50, and Xception based on experimental results demonstrating the highest segmentation accuracy. Initially, the optimal backbone for myocardial scar segmentation was selected from MobileNetV2, ResNet50, and Xception, based on experimental results demonstrating the highest segmentation accuracy. Following this, a comparative analysis was conducted to evaluate the effectiveness of various loss functions employed in the classification layer of the DLT-Scar network. Subsequently, PSO was applied to fine-tune the hyperparameters of the best-performing loss function, and the corresponding results are discussed. This network uses the merged images from LGE-CMR and LV mask labels from the previous stage as input to perform scar segmentation. During training, the network was trained on manually segmented scar images to accurately classify the scar tissue.

4.3.2.1 Analysis of DLT-Scar Backbone Selection

Table 4.5 presents the comparative performance analysis of three backbone networks (MobileNetV2, ResNet50, and Xception) integrated into the DeepLabV3+ architecture for the task of myocardial scar segmentation, from LGE-CMR images. It is noteworthy that this backbone selection experiment was also conducted in the first-stage segmentation task (LV segmentation), under the same training settings, as detailed in Section 3.5.1.1. Among the evaluated backbones, ResNet50 outperformed the others, achieving the highest scar mean Dice Score of 71.98%, precision of 0.6821, and recall of 0.7822, resulting in the best overall F2-Score of 0.7599. MobileNetV2, despite being computationally efficient, reported the lowest scar Dice Score (67.39%) and F2-score (0.6948), highlighting potential trade-offs between model size and segmentation performance. Xception, while slightly outperforming MobileNetV2 in scar Dice Score and F2-Score, did not surpass ResNet50 on any performance metric. The results highlight ResNet50's effectiveness in capturing fine-grained features essential for segmenting small and heterogeneous regions such as myocardial scars. This is likely due to its residual connections, which preserve lower-level details and support richer deep feature representation. Given its strong performance, ResNet50 has been selected as the backbone architecture for myocardial scar segmentation in this study.

Table 4.5
Comparison of the Different DeepLabV3+ Backbone for Scar Segmentation

Backbone	**Mean Dice Score (%)**		**Precision**	**Recall**	F2-Score
	Background	Scar			
MobileNetV2	99.88	67.39	0.6413	0.7096	0.6948
ResNet50	99.88	71.98	**0.6821**	0.7822	0.7599
Xception	99.87	68.17	0.6557	0.7096	0.69811

While the background segmentation yielded a Dice Score exceeding 99%, the scar Dice Score remained substantially lower at 71.98% emphasizing the challenges of detecting and delineating myocardial scars due to their small size, weak contrast, and highly imbalanced representation in the dataset compared to the background. This significant performance gap underscores the need for further optimization, particularly in enhancing the model's sensitivity to small and irregular scar regions. To address this,

the next objective of the study focuses on analyzing the effect of different loss functions specifically within the DLT-Scar model, which utilizes ResNet50 as its backbone.

4.3.2.2 Analysis of Function Loss Selection for Imbalance Class

A comparative analysis of four state-of-the-art loss functions was carried out to address class imbalance and improve the model's effectiveness in accurately segmenting myocardial scars, particularly by adjusting how the network penalizes misclassifications involving the underrepresented scar class. Figure 4.9 demonstrates the segmentation results obtained using the proposed DLT-Scar model, highlighting the impact of different loss functions namely Weighted Cross-Entropy (CE), Weighted Dice, Focal Loss, and the Tversky Loss. The input LGE-CMR images and corresponding ground truth masks are shown for four representative cases. Visually, Tversky Loss and Weighted Dice Loss produce segmentation masks that most closely resemble the ground truth in terms of shape across all examples. However, Tversky Loss demonstrates superior capability in accurately capturing both the shape and completeness of scar regions, particularly in Image 2 and Image 3, where Dice Loss produces narrower and less continuous segmentations compared to the ground truth, thereby introducing false negatives. In contrast, the segmentation results using Weighted CE exhibit noticeable over-segmentation and thickened boundaries, indicating that the model struggles to accurately distinguish scar tissue from non-scarred areas. Conversely, the segmentation results using Focal Loss tend to under-segment, resulting in the exclusion of scar areas that are less apparent.

Table 4.6 presents the quantitative performance analysis of the proposed DLT-Scar model trained using the four previously mentioned loss functions. The evaluation was conducted on the test set using standard metrics widely adopted in medical image segmentation, namely mean Dice Score, Precision, Recall, and F2-Score. Among the evaluated loss functions, the Tversky Loss achieved the best overall performance, with a scar Dice Score of 72.71%, a Precision of 0.6891, Recall of 0.7903, and the highest F2-Score of 0.7677. The Weighted Dice loss also performed well, producing a scar Dice Score of 71.98%, Precision of 0.6821, Recall of 0.7822, and an F2-Score of 0.7599, slightly lower than those obtained with Tversky Loss but still competitive. In contrast, the Weighted CE loss function yielded the highest Recall of 0.7943, indicating high sensitivity to scar regions. However, this came at the cost of low Precision (0.4339) and

the lowest Dice Score of 56.12%, suggesting a high rate of false positives. The Focal Loss produced the highest Precision of 0.8097 but suffered from a low Recall of 0.5252 and a moderate Dice Score of 60.71%, indicating a tendency to under-segment and miss scar regions. These findings clearly demonstrate that the choice of loss function has a significant impact on segmentation performance. Overall, Tversky Loss shows as the most effective or this myocardial scar segmentation task, primarily due to its ability to handle class imbalance by offering more flexibility and its superior balance between detecting scar tissue accurately and minimizing false positives.

Table 4.6
Result of Performance Analysis on the Test Set for Four Different Loss Functions Used in Training for Scar Segmentation

Loss function	**Mean Dice Score (%)**		**Precision**	**Recall**	**F2-Score**
	Background	**Scar**			
Weighted CE	99.75	56.12	0.4339	**0.7943**	0.6811
Weighted Dice	99.88	71.98	0.6821	0.7822	0.7599
Focal Loss	99.88	60.71	**0.8097**	0.5252	0.5649
Tversky Loss (a =0.3, (3=0.7)	**99.89**	**72.71**	0.6891	0.7903	**0.7677**

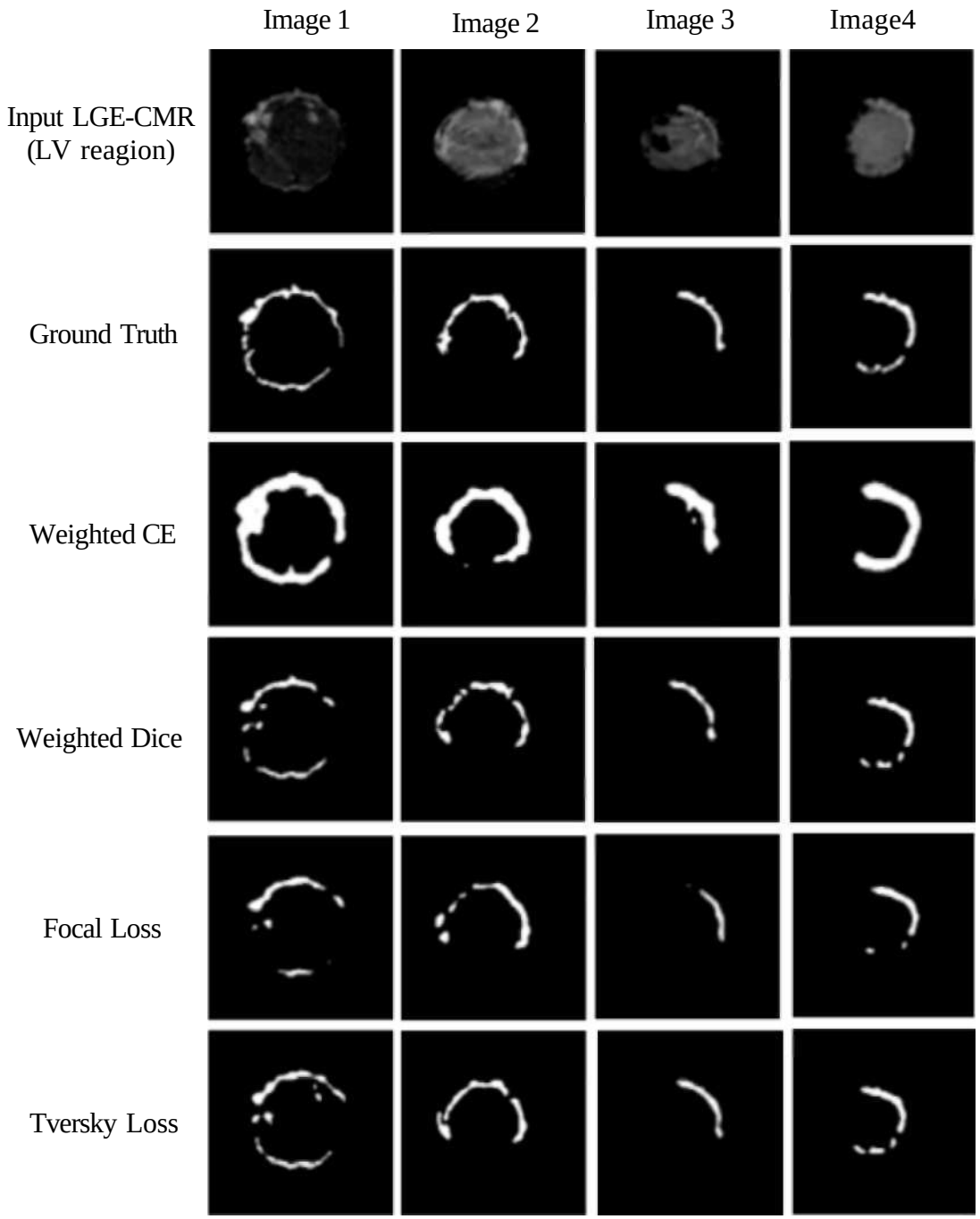


Figure 4.9 Visual Comparison of Myocardial Scar Segmentation Results Using Different Loss Functions in the DLT-Scar Model

4.3.2.3 Analysis of the Proposed Automated Tversky Loss Function Hyperparameters Optimization Using PSOfor Scar Segmentation Model

Building upon the promising results obtained using the Tversky Loss, this subsection focuses on further improving the model's quantitative performance by automatically optimizing its hyperparameters, a and P, which control the trade-off

between false positives and false negatives. Rather than relying on fixed, manually selected values, this study applies the PSO algorithm to automatically search for the optimal Tversky loss parameter (a), based on the simplified formulation described in Chapter 3, Subsection 3.5.2.2.

The objective of the PSO algorithm in this study is to automatically optimize the Tversky loss function hyperparameter (a) in the DLT-Scar model, with the Dice Score defined as the fitness function. The value of this function directly reflects the segmentation accuracy of the model, where a higher Dice Score indicates better scar delineation. The PSO is configured with 5 particles or also known as swarm and a maximum of 50 iterations. Figure 4.10 show the convergence behavior of the PSO algorithm applied to optimize the hyperparameter of the Tversky loss function. The y-axis represents the function value and the iteration count on the x-axis. The sharp decrease in the early iterations, with significant improvements in the objective function, indicates that the PSO algorithm quickly identified better hyperparameter value for the Tversky loss function during the initial search phase. By iteration 20, the function value begins to stabilize, suggesting that the algorithm had converged on an optimal solution. This value remained consistent as the optimization process stopped at iteration 38, indicating that no further improvements were possible, and the algorithm had effectively converged on the optimal hyperparameter value for the Tversky loss function. The best function value achieved, as shown in the graph, is 0.74314, which corresponds to a Tversky loss function hyperparameter (a) value of 0.2141.

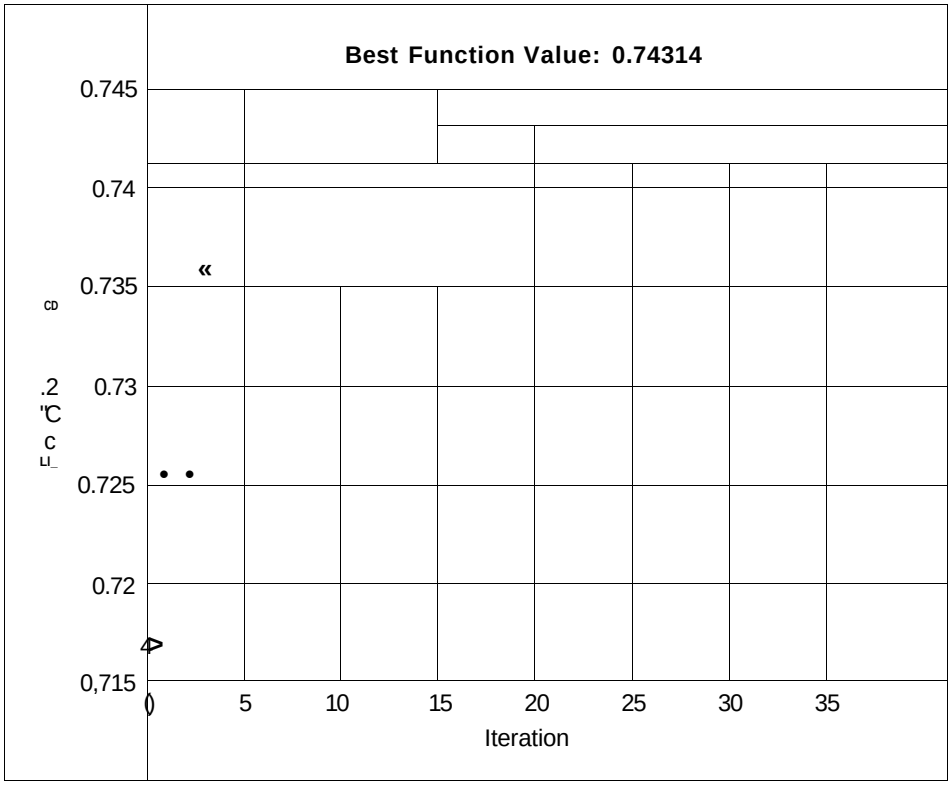


Figure 4.10 Convergence Behavior of the PSO Algorithm Applied to Optimize the Hyperparameter of the Tversky Loss Function in DLT-Scar Model

Table 4.7
Result of Performance of DLT-Scar Model with Default and PSO-optimized Tversky Loss Parameter.

Tversky Loss hyperparameters	**Mean Dice Score (%)**		**Precision**	**Recall**	**F2-Score**
	Background	**Scar**			
a = 0.3, (3=0.7	99.89	72.71	0.6891	0.7903	0.7677
PSO (a= 0.2141)	99.90	**74.31**	**0.7219**	**0.8092**	**0.7893**

The quantitative analysis of test set performance across default Tversky loss hyperparameters, (a = 0.3, P = 0.7) with the optimized hyperparameters obtained through PSO (a = 0.2141) is presented in Table 4.7. While background segmentation remained consistently excellent across all configurations, the PSO-optimized a value of 0.2141 produced a distinct advantage in terms of scar segmentation. Notably, it achieved the highest Dice Score and the F2-Score improved by 1.6% (72.71% to 74.31%) and 2.16% (0.7677 to 0.7893), respectively.

In summary, these results highlight that the automated PSO optimization significantly improves scar segmentation by optimizing a to 0.2141, the loss function

assigns relatively more weight to false negatives (FN) and less weight to false positives (FP). This adjustment shifts the model's focus toward improving recall without drastically reducing precision. This confirms the value of PSO as an effective strategy for fine-tuning the Tversky loss and addressing the class imbalance challenges inherent in medical image segmentation, where maximizing recall is often prioritized to ensure no potential signs of the condition are missed.

4.4 Analysis of the Proposed MI Detection Using 3SR and DCNN

The previous results in Section 4.3.2 indicate that the proposed Dual-Stage-DeeplabV3+based segmentation model has successfully predicted scar images. However, the model also identified scar-like objects in some images, which could result in false diagnoses especially for healthy tissue. Therefore, a single image is often insufficient for reliably identifying the presence of a true scar in short-axis LGE MRI, which strongly suggests a history MI and plays a critical role in accurate MI detection. This section representing the final stage of this study, provides a comprehensive description and analysis of the results obtained from the MI detection method, employing the proposed scar sequential slice reconstruction with the S-DCNN model. The scar sequential slice reconstruction algorithm, along with its training parameter setup, is discussed in detail in Section 3.6.1 of Chapter 3, while the S-DCNN network architecture is comprehensively covered in Section 3.4.1. A comparison was conducted between the S-DCNN model's performance on the same segmented images, which were produced from the predicted range of images containing scars as in Section 4.3.3. This evaluation included both single segmented images and reconstructed sequential images to assess the impact of scar sequential slice reconstruction on true scar detection, thus indicating MI detection accuracy.

4.4.1 Description of Dataset for Training and Testing the MI Detection

Figures 4.1 l(a-c) to 4.12 (a-c) and Figures 4.13 to 4.14 (a-c) present examples of a sequential segmented image from a non-MI subject and a subject with MI, respectively. Scar segmentation was performed using the previous method described in Section 3.5.2. The ground truth for each image was adopted from the manual segmentation performed by radiologists in the previous section, where images

containing scar tissue were labelled as MI and those without scar tissue were labelled as non-MI. Figures 4.11(d) and 4.12(d) present the results of scar sequential slice reconstruction obtained from a healthy subject, while Figures 4.13(d) and 4.14(d) display the results of scar sequential slice reconstruction from an MI subject. The RGB composite visually represents the continuity of the scar pattern across consecutive slices. As observed in Figures 4.11(d) to 4.12(d), false scar-like segmentations frequently exhibit not only discontinuity but also a scattered pattern in the sequential slice reconstructions. These discontinuities and scattered patterns may suggest the presence of imaging artifacts. In contrast, the true scar patterns as in Figures 4.13(d) and 4.14(d) exhibit a continuous pattern across adjacent slices, maintaining their position within the same anatomical region. Despite variations in size or thickness, the circumferential placement of the scar remains relatively consistent across slices.

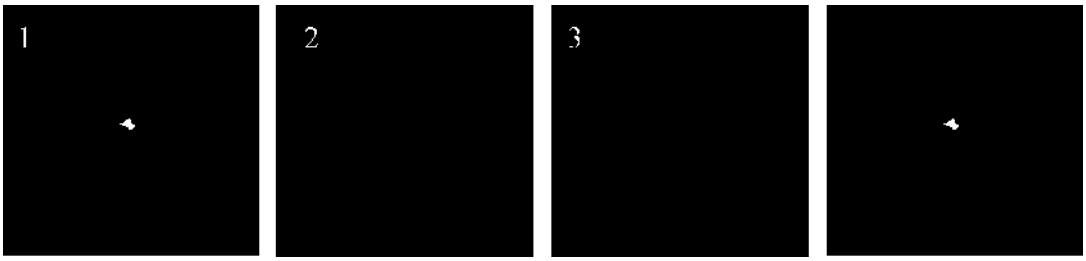


Figure 4.11 (a-c) and (d) Present Examples of Sequential Slices Showing Discontinuity of the Scar From the Segmented of a Healthy Subject and the Corresponding Results of Scar Sequential Slice Reconstruction, Respectively

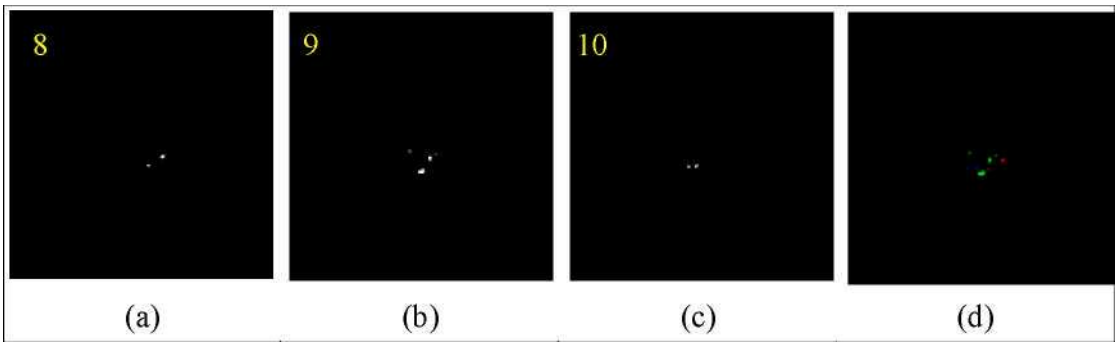
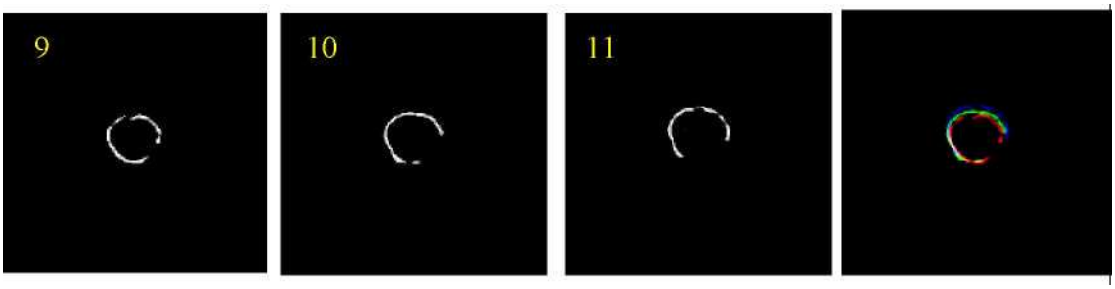
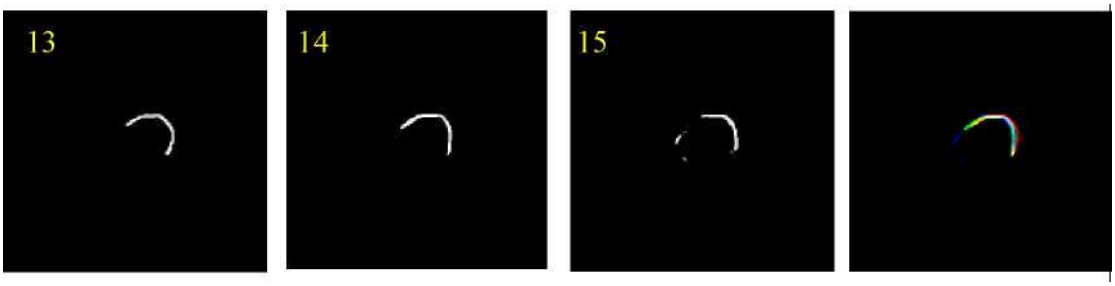


Figure 4.12 (a-c) and (d) Present Examples of Scattered Pattern in Sequential Slices From the Segmented of a Healthy Subject and the Corresponding Results of Scar Sequential Slice Reconstruction, Respectively



(a) (b) (c) (d)

Figure 4.13 (a-c) and (d) Present Examples of Sequential Slices Showing Continuity of the Scar From the Segmented of a MI Subject and the Corresponding Results of Scar Sequential Slice Reconstruction, Respectively



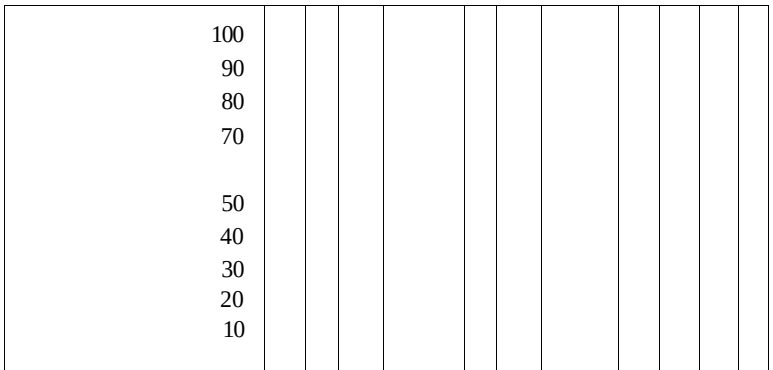
(a) (b) (c) (d)

Figure 4.14 (a-c) and (d) Present Examples of Sequential Slices Showing Continuity of the Scar From the Segmented of a MI Subject and the Corresponding Results of Scar Sequential Slice Reconstruction, Respectively

4.4.2 Analysis of Scar Sequential Slice Reconstruction Method

This section presents the results and performance of the proposed MI detection method. The goal is to assess the impact of the sequential image patterns on MI detection compared to single segmented images using the same S-DCNN model. The analysis began by evaluating the S-DCNN using scar segmentation images obtained from the segmentation result described in Section 4.3.2 without featuring the sequential image reconstruction. The results were then compared to the output of the S-DCNN when fed with an image generated through the proposed sequential image reconstruction applied to the segmented predicted scar. Figure 4.15 shows the performance evaluation bar chart comparing the S-DCNN model's ability to detect MI using single segmented images versus reconstructed sequential images. The accuracy of the S-DCNN model increased significantly from 73.66% to 94.30%, demonstrating the effectiveness of incorporating sequential image patterns in capturing the continuity of myocardial structures, which is often missed in single-segmented images. Similarly,

specificity improved from 49.46% to 92.50%, reflecting the model's enhanced ability to identify non-MI cases and reduce false alarms. Although sensitivity slightly decreased from 97.85% to 96.15%, it remains high, maintaining overall classification reliability. Insights from the confusion matrix as in Figure 4.16, further support these findings, showing a marked reduction in false positives while explaining a slight increase in false negatives. The false positive rate was significantly reduced from 50.54% to 7.50%, highlighting the model's improved precision and its ability to reduce unnecessary false alarms. However, the false negative rate increased slightly from 2.15% to 3.85%. A detailed analysis, as visually demonstrated in Figure 4.17(a), revealed that the increase was due to large anatomical gaps between slices in the sequential reconstruction image. Variations in cardiac size, related to the thickness of the LGE-MRI slices, may have contributed to these large gaps, resulting in discontinuities. Additionally, Figure 4.17 (b) illustrates the misclassifications, where false positives occurred because the sequential pattern of healthy tissue resembled true scar tissue. The proposed method introduces a novel contribution to MI detection by leveraging sequential image reconstruction to address the limitations of single segmented images. This approach provides a more comprehensive analysis of myocardial structures and significantly enhances diagnostic accuracy.



	Accuracy	Sensitivity	Specificity
Single Image	73.66	97.85	49.46
Sequential images	94.30	96.15	92.50

Figure 4.15 Accuracy, Sensitivity and Specificity Performance of the S-DCNN Model for MI Detection: Comparison Between Single Segmented Images and Reconstructed Sequential Images

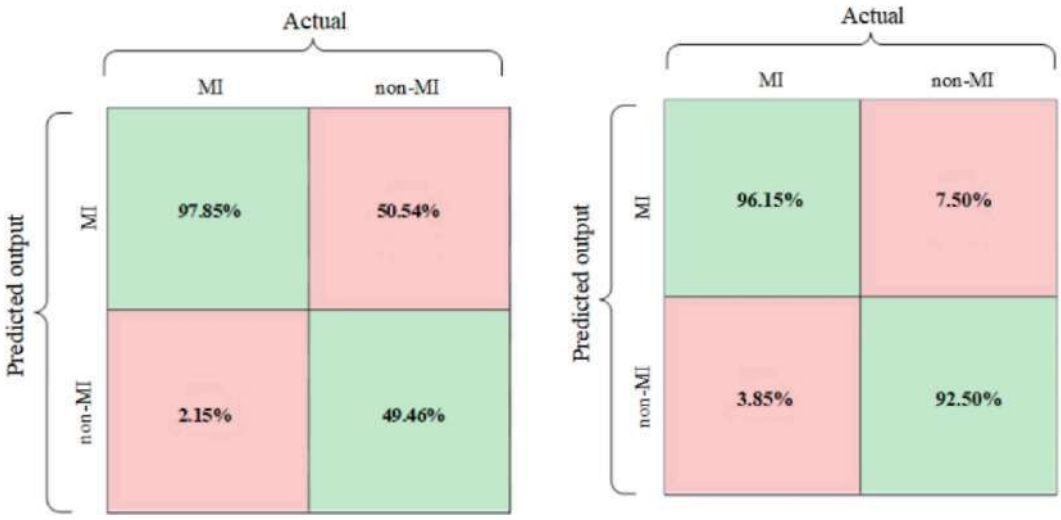


Figure 4.16 Confusion Matrix Plot for MI Detection Using the S-DCNN Classifier: (a) Single Image, (b) Sequential Image Reconstruction

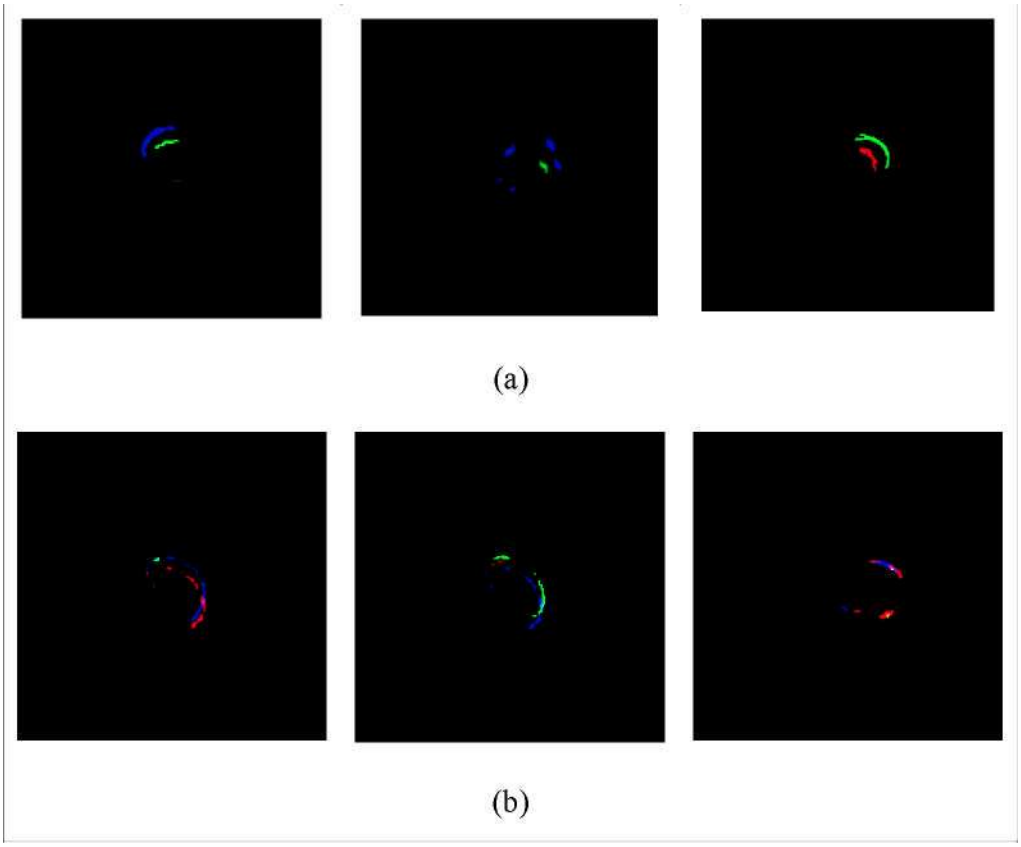


Figure 4.17 Analysis of Misclassifications in the Sequential Reconstruction of LGE-MRI Segmented Image:(a) Misclassifications of False Negative Caused By Large Anatomical Gaps Between Slices Resulting in Discontinuities and (b) Misclassifications of False Positives Caused by the Sequential Pattern of Healthy Tissue Resembling True Scar Tissue

4.5 Fully Automated Myocardial Infarction Detection

This study presents a novel fully automated framework for the detection of MI in late gadolinium enhancement cardiac magnetic resonance (LGE-CMR) images. The proposed system is designed to assist radiologists in the diagnosis of MI by providing a fast, reliable, and fully automated analysis pipeline. The first step involves an automated LV region classification, which filters out images that do not contain the LV, thereby allowing the subsequent segmentation process to focus exclusively on clinically relevant slices. The myocardial scar segmentation is performed using the proposed dual-stage DeepLabV3+-based segmentation model, specifically tailored for accurate scar delineation. Detailed training results and performance analyses of this segmentation model are provided in Section 4.3. Following segmentation, the detection of myocardial infarction begins by identifying the image slice containing the largest segmented scar tissue for each patient. This slice serves as a reference point for the proposed 3 SR, which consolidates scar continuity across sequential slices to enhance infarct pattern recognition. Once the reconstructed representation is formed, it is passed through the proposed S-DCNN model for classification. Further training and analysis of this classification method are discussed in Section 4.4.

To validate the proposed system, 30 subjects were used, including 10 patients with confirmed MI and 20 healthy control subjects. Table 4.8 presents a validation analysis comparing the predictions made by the proposed fully automated MI detection against radiologist reports to assess the system's diagnostic performance. The confusion matrix in Figure 4.18 provides a detailed breakdown of the classification results. Overall, the model correctly classified 29 out of 30 patients, resulting in a total classification accuracy of 96.7%. This high accuracy demonstrates the reliability of the proposed system in distinguishing between pathological and non-pathological cases. The system successfully identified all 10 MI cases, resulting in a True Positive Rate (Sensitivity) of 100%. This indicates the model's strong ability to detect infarcted cases, a crucial factor in clinical applications where missing an MI diagnosis can have severe implications. The model also correctly classified 19 out of 20 Non-MI patients, achieving a True Negative Rate (Specificity) of 95%. Only one healthy patient (Patient 18) was incorrectly classified as having MI, representing a False Positive. Importantly, the absence of any False Negatives (FN = 0) highlights the model's clinical safety in avoiding missed MI diagnoses. These validation results underscore the effectiveness

and clinical potential of the proposed end-to-end MI detection framework, showing strong agreement with radiologist-reported findings.

Table 4.8
Validation Analysis of a Fully Automated MI Detection Outputs with Radiologist Reports

No Patient	Patients' Radiologist Report	Proposed Fully Automated MI Detection Prediction
1	MI	MI
2	MI	MI
3	MI	MI
4	MI	MI
5	MI	MI
6	MI	MI
7	MI	MI
8	MI	MI
9	MI	MI
10	MI	MI
11	Non-MI	Non-MI
12	Non-MI	Non-MI
13	Non-MI	Non-MI
14	Non-MI	Non-MI
15	Non-MI	Non-MI
16	Non-MI	Non-MI
17	Non-MI	Non-MI
18	Non-MI	MI
19	Non-MI	Non-MI
20	Non-MI	Non-MI
21	Non-MI	Non-MI
22	Non-MI	Non-MI
23	Non-MI	Non-MI

No Patient	Patients' Radiologist Report	Proposed Fully Automated MI Detection Prediction
24	Non-MI	Non-MI
25	Non-MI	Non-MI
26	Non-MI	Non-MI
27	Non-MI	Non-MI
28	Non-MI	Non-MI
29	Non-MI	Non-MI
30	Non-MI	Non-MI
	Correctly classified:	29/30 = 96.7%

		PREDICTED	
	CLASSES	MI	Non-MI
3 HO<	MI	10	0
	Non-MI	1	19

Figure 4.18 Classification Outcomes for MI and Non-MI Cases Fully Automated Myocardial Infarction Detection

4.5.1 Comparison of Fully Automated MI Detection with Other Methods

In this section, the performance of the proposed fully automated MI Segmentation and detection model will be benchmarked against several recent studies that utilized the DCNN approach. Table 4.9 shows the comparison of the proposed method with five different methods. Moccia et al. [28] developed a CNN-based model for myocardial scar segmentation using LGE-CMR images, achieving a Dice score of approximately 55%. Aside from the low performance in scar segmentation, their approach also relied on manually cropped myocardial regions and lacked classification components. Similarly, Chen et al. [23] reported segmentation Dice scores 67% using an ensemble of U-Net-based architectures and their results are also confined to middle slices only. Their method also excluded MI detection, making it unsuitable for end-to-end automated and clinical diagnostic coverage. Dan et al. [112]introduced ACSNet for

anatomically constrained scar segmentation across apical, middle, and basal LV levels. However, the method reported a relatively low scar segmentation Dice score of only 57% and it did not include any classification component. Abdulkareem et al. reported significantly lower detection performance (68%) using LGE-CMR and pre-contrast cine CMR, highlighting the effectiveness of our approach that integrates both segmentation and detection in a sequential learning framework. Abdulkareem et al. [108] proposed a complicated classification-only approach using pre-contrast cine CMR and ResNet50, reporting an MI detection accuracy of just 68%. Their method lacked segmentation and was limited by the absence of contrast-enhanced visibility and structural context.

A more recent study by De la Rosa et al. [115] proposed a cascaded approach in which a CNN is first used to detect infarcted slices, followed by a rule-based morphological segmentation and final refinement using a DCNN ensemble. Despite achieving a high scar segmentation Dice score of 77.22%, their rule-based morphological segmentation method relies on the manual selection of infarcted images, which limits its automation and generalizability. In contrast, our proposed method achieved a comparable Dice score of 74.31% while delivering a higher MI detection accuracy of 96.67%, outperforming all related works.

Table 4.9
Comparison of the Proposed Dual-Stage Model with Related Works in Myocardial Scar Segmentation

Model	**LV region Classification (Accuracy %)**	**Myocardial Scar Segmentation (mean Dice Score %)**	**Myocardial Infarction Detection (Accuracy %)**
Moccia et al. [28]	-	55.00	-
Chen et al. [23]	-	67.00	-
Dan et al. [112]	-	57.00	-
Abdulkareem et al. [108]	-	-	68.00
De la Rosa et al. [115]	-	**77.22**	90.63
Proposed Method	**97.36**	74.31	**96.67**

The proposed model addresses the limitations of previous works by using an end-to-end automated pipeline that integrates classification, segmentation, and

detection, without manually selecting slices or cropping the LV region. Furthermore, the dataset encompasses the full range of LGE-CMR slices, from the lowest epicardial to the upper basal regions, as typically received by radiologists. This approach makes it more practical for real-world clinical applications. Although direct comparisons with previous studies may be limited by differences in datasets, it is important to note that the results fall within the expected performance range. Variations in dataset characteristics such as anatomical complexity and image quality can significantly influence segmentation outcomes.

4.6 Conclusion

In this chapter, the results of the proposed methodology for fully automated segmentation and detection of MI in LGE-MRI images have been thoroughly discussed. The results were analyzed based on the study's objectives, starting with the new proposed classification method for identifying images containing LV and non-LV regions using a DCNN model. Section 4.2 presents two viable options for automated image classification with LV and non-LV regions: using the proposed S-DCNN with a three-convolutional layer model or the customized pre-trained network, GoogLeNet. While the GoogLeNet model achieved a higher accuracy of 98.75%, the S-DCNN demonstrated a reliable accuracy of 97.36% with notable advantages in simplicity and computational efficiency. The promising results of this research highlight the potential of both automated approaches. The output of this work, which identifies the LV region, can automate the manual process, serving as input for the next assessment stage: myocardial scar segmentation for MI detection.

Section 4.3 elaborates analysis of the novel Dual stage DeeplabV3+based myocardial scar segmentation model by using the original DeepLabv3+ as the benchmark. Subsection 4.3.1 evaluates the performance of the proposed DLT-LV network model for first-stage LV region segmentation, incorporating MobileNetV2 as the backbone and a morphological refinement step. The analysis demonstrates that the proposed DLT-LV model (Dice: 97.02%), IoU: 95.60%>) significantly outperforms both the original DeepLabV3+ model with Xception backbone (Dice: 95.47%, IoU: 94.44%) and the baseline DeepLab V3+ model with MobileNetV2 backbone (Dice: 96.12%, IoU: 94.9P/o). The inclusion of the morphological refinement in the post-processing step further enhances segmentation quality by eliminating isolated mis-segmentations and

filling internal gaps, thus improving the structural integrity of the LV mask. Quantitatively, the application of the refinement procedure improved the mean Dice score from 97.02% to 97.80% and the mean IoU from 95.60% to 95.97%. These results demonstrate that the proposed approach not only outperforms the baseline model but also provides a more reliable foundation for the subsequent scar segmentation stage.

Subsection 4.3.2 presents a comprehensive evaluation of myocardial scar segmentation using the proposed DLT-Scar model, with particular focus on backbone selection, loss function analysis, and automated loss function optimization to address class imbalance in LGE-MRI images. Initially, three backbone networks (MobileNetV2, Xception, and ResNet50) were tested under the same training conditions. ResNet50 outperformed the others, achieving the highest scar segmentation accuracy with a mean Dice Score of 71.98%), Precision of 0.6821, Recall of 0.7822, and F2-Score of 0.7599. This strong performance is likely due to its residual connections, which help retain low-level features and improve detection of small and heterogeneous scar regions. As a result, ResNet50 was chosen as the backbone for the DLT-Scar model. To further address the class imbalance between the small scar regions and the large background, a comparative analysis of four loss functions was conducted (Weighted Cross-Entropy, Weighted Dice, Focal Loss, and Tversky Loss). The analysis highlights the highest performance of the Tversky loss function (a = 0.3, P = 0.7) achieving a scar Dice Score of 72.71%, Precision of 0.6891, Recall of 0.7903, and F2-Score of 0.7677, outperforming the other loss functions, making it the ideal choice for the task. To further optimize segmentation performance, the study introduced an automated tuning strategy for the Tversky loss hyperparameters using Particle Swarm Optimization (PSO). By simplifying the Tversky formulation to a single tunable parameter (a), the PSO algorithm identified an optimal a value of 0.2141, which led to further improvements in segmentation performance. The PSO-optimized model achieved a Dice Score of 74.31%, Precision of 0.7219, Recall of 0.8092, and the highest F2-Score of 0.7893, representing an increase of 1.6% in Dice Score and 2.16% in F2-Score compared to the default Tversky configuration. This subsection highlights that the integration of ResNet50 as the backbone, combined with the Tversky Loss function and PSO-based hyperparameter optimization, significantly enhances the model's ability to segment myocardial scars in LGE-MRI images. The refined architecture and loss function effectively address the challenges of class imbalance and small target region

detection, thereby improving the reliability and clinical applicability of the proposed segmentation framework.

Based on the findings in Section 4.3.2, increasing recall enhances the model's sensitivity to potential cases but also leads to a higher susceptibility to false positives. Moreover, Scar-like artifacts can appear similar to true myocardial scars, which could result in false diagnoses especially for healthy tissue. To cater this issue, the new 3SR method in combination with the S-DCNN was introduced following the segmentation process to improve the identification of true scars in short-axis LGE-MRI, ensuring more accurate MI detection. A comparative evaluation was performed using the same S-DCNN classifier to detect MI from single-segmented images and reconstructed sequential images produced by the 3SR method. The results clearly show that the 3SR method significantly improves MI detection, with accuracy rising to 94.30%), specificity to 92.50%), and sensitivity remaining high at 96.15%), compared to the use of single segmented images, which achieved only 73.66%> accuracy, 97.85%> sensitivity, and 49.46% specificity.

Finally, Section 4.5 validates the overall proposed fully automated myocardial infarction (MI) detection system against radiologist reports. The system correctly classified 29 out of 30 patients, achieving an overall accuracy of 96.7%. This section also provides a comparison with existing methods, highlighting the advantages of the proposed approach in terms of automation, accuracy, and clinical reliability.

In summary, this end-to-end framework offers accurate, efficient, and clinically relevant MI detection, demonstrating state-of-the-art performance and strong potential for real-world application in LGE-CMR analysis. Future work will focus on extending this framework to include scar quantification, paving the way for more effective MI screening and diagnosis.

CHAPTER 5
CONCLUSION

5.1 Conclusion

This study presented a fully automated, end-to-end framework for segmentation and detection of myocardial infarction (MI) in short-axis late gadolinium enhancement-cardiac magnetic resonance (LGE-CMR) images. The system was successfully developed using deep convolutional neural networks (DCNNs) approaches to assist radiologists by providing a fully automated system for the analysis and identification of myocardial Infarction.

The framework begins with a newly proposed automated classification method for identifying images containing the left ventricular (LV) region, addressing Objective 1. This method classifies LGE-CMR slices into LV and non-LV regions. A novel shallow deep convolutional neural network (S-DCNN) was introduced and benchmarked against three pre-trained networks: GoogleNet, AlexNet, and SqueezeNet. The S-DCNN achieved a reliable classification accuracy of 97.36%, which was slightly higher than AlexNet (96.43%) but moderately lower than GoogleNet (98.75%) and SqueezeNet (98.39%). However, the S-DCNN offers notable advantages in terms of architecture simplicity and computational efficiency. The output from this classification stage eliminates the need for manual screening and serves as the input for the next stage: myocardial scar segmentation for MI detection.

The second objective of this study (Objective 2) represents a key contribution, focusing on the development of a segmentation framework to localize myocardial scar regions associated with infarction in LGE-CMR images. To address this, a novel dual-stage segmentation architecture was proposed using a customized DeepLabV3+ network comprising two task-specific stages: DeepLabV3+ Tailored to LV (DLT-LV) for segmenting the LV region and DeepLabV3+ Tailor to Scar (DLT-Scar) for identifying myocardial scar within the LV region. The DLT-LV model, developed using a MobileNetV2 backbone and a modified Atrous Spatial Pyramid Pooling (ASPP) module, achieved a Dice Score of 97.02% and an IoU of 95.60%. This performance surpassed the benchmark model, the original DeepLabV3+, with Xception backbone and standard ASPP configuration by approximately 2% Dice Score. A morphological

post-processing procedure was introduced to refine the LV segmentation masks by correcting small gaps within the LV chamber and eliminating isolated mis-segmentations. Although the improvement in Dice Score was minimal, with an increase of approximately 1%, this refinement step is essential to ensure that accurate and anatomically consistent LV masks are provided to the subsequent scar segmentation stage. For the second stage, the DLT-Scar model employed the same architecture as DLT-LV, with the exception of using a ResNet50 backbone. To address the challenge of class imbalance, specifically the small size of scar regions relative to the background, the Tversky loss function with standard hyperparameters ($a = 0.3$, $P = 0.7$) was utilized. This configuration yielded a scar segmentation mean Dice Score of 72.71% and an F2-Score of 0.7677, demonstrating the model's effectiveness in accurately segmenting myocardial scars in challenging LGE-CMR data.

Building on the proposed dual-stage segmentation framework, the third objective (Objective 3) introduced an optimization strategy by integrating the DLT-Scar segmentation network with the Particle Swarm Optimization (PSO) algorithm. This method was designed to automatically fine-tune the Tversky loss function, which was mathematically simplified to a single hyperparameter (a). By embedding PSO within the training process of the DLT-Scar model, the framework eliminated the need for manual hyperparameter selection, enabling adaptive loss optimization tailored to the segmentation task. This PSO-DLT-Scar method optimized the a value (0.2141), resulting in a further improvement in scar segmentation performance, with the mean Dice Score increasing to 74.31% and the F2-Score to 0.7893.

The output of myocardial scar segmentation obtained from the previous stage was further analyzed using a novel MI detection approach (Objective 4), referred to as Scar Sequential Slice Reconstruction (3SR) and DCNN classification. This method leverages the anatomical continuity of the left ventricle across consecutive short-axis LGE-CMR slices by reconstructing a volumetric representation based on the largest detected scar region. The resulting reconstructed image serves as the input for a trained S-DCNN model classifier, which performs binary classification to detect the presence of myocardial infarction as part of radiologist's report. The final implementation of this automated detection system is presented in Section 4.6. The proposed method achieved a myocardial infarction detection accuracy of 96.67%, successfully identifying all 10 MI cases and 19 out of 20 non-MI cases. These findings present a scalable, fully

automated approach that enhances diagnostic consistency and operational efficiency within clinical cardiac imaging workflows.

5.2 Research Contribution

This research contributes to the overall process of automated Myocardial Infarction segmentation and detection using image processing and artificial intelligence techniques. Based on the objectives of this study, the contribution of this research is listed as follows:

i. A new automated method for classifying LV regions in a series of LGE-CMR images using a novel shallow deep convolutional neural network (S-DCNN) model, developed to operate efficiently in resource-constrained settings while maintaining high classification accuracy. Since LGE-CMR scans include both LV and non-LV slices, this method addresses the gap in existing studies by automating the classification process, where LV slices are typically identified through manual screening by radiologists or manually selected in most studies. In addition, this study also explores an alternative classification architecture utilizing a pre-trained network, providing a flexible and scalable solution for improved cardiac assessment. This approach streamlines the subsequent myocardial scar segmentation process and eliminates manual workload.

ii. A novel, fully automated dual-stage segmentation framework based on DeepLabV3+, designed to independently handle left ventricular (LV) and myocardial scar segmentation as two task-specific stages, namely DLT-LV and DLT-Scar. To strengthen inter-stage consistency and improve overall segmentation reliability, the framework incorporates a custom morphological post-processing procedure between the two stages. This refinement step corrects common LV segmentation errors, such as gaps within the LV region and isolated mis-segmentations, ensuring that a cleaner and more anatomically accurate mask is passed to the scar segmentation stage. In conjunction with the framework design, this study also contributes a set of architectural enhancements to the DeepLabV3+ model, including a modified Atrous Spatial Pyramid Pooling (ASPP) module with adjusted dilation rates, batch normalization, and an added parallel convolutional branch to improve multi-

scale contextual learning. In addition, task-specific backbone networks were employed (MobileNetV2 for LV segmentation and ResNet50 for scar segmentation) each selected based on empirical performance and suitability for their respective segmentation objectives.

iii. A new optimization strategy is introduced for tuning the Tversky loss function hyperparameter used in training the DLT-Scar model, employing Particle Swarm Optimization (PSO) to automatically identify the optimal a value. To reduce computational complexity, the loss formulation is simplified by optimizing only a single parameter a, with P implicitly defined as (1 - a). This simplification narrows the search space while maintaining effective control over the trade-off between false positives and false negatives. By emphasizing false negative reduction, this strategy addresses the class imbalance commonly found in medical imaging where small target regions present major challenges.

iv. A new method integrating Scar Sequential Slice Reconstruction (3SR) is introduced to enhance the reliability of prior myocardial infarction (MI) detection. Unlike conventional single-image segmentation, this approach utilizes the spatial continuity of scars across segmented sequential LGE-CMR slices to distinguish true infarcted regions from imaging artifacts. A reference image containing the largest segmented scar is established to assess spatial consistency across slices, effectively reducing false positives caused by motion artifacts and magnetic field inhomogeneities, particularly in healthy subjects undergoing LGE-CMR scans. Furthermore, incorporating this method with the previously proposed S-DCNN model enables automated detection of subjects with and without MI

v. This study's contribution includes demonstrating that the fully automated system developed based on the research objectives shows improved consistency and reduced variability in readings with high accuracy compared to radiologists' report. This highlights the potential to improve diagnostic reliability in medical imaging by providing more standardized, time-saving, and reproducible results.

5.3 Recommendation for Future Work

This study successfully developed an end-to-end framework for fully automated MI segmentation and detection using DCNNs in short-axis LGE-CMR images. However, several potential directions remain for further enhancement and exploration.

A promising extension of this research would be the development of a real-time automated system that facilitates integration and control between the cardiac MRI acquisition process and the end-to-end segmentation-detection pipeline. Embedding this framework within the clinical imaging workflow could enable instantaneous processing and interpretation of cardiac scans, supporting radiologists in making faster, data-driven decisions at the point of care.

The proposed dual-stage DeepLabV3+ segmentation model represents a valuable step toward automated myocardial scar segmentation, with significant potential for clinical extension. Future work could expand this framework to include the quantification of infarct burden and transmurality, including the classification of infarcts into subendocardial, mid-myocardial, subepicardial, and transmural categories for clinical interpretation and treatment planning.

Given that this study focused exclusively on short-axis LGE-CMR images due to their widespread clinical use and availability, future work could explore the incorporation of additional imaging planes, such as long-axis views (two-, three-, and four-chamber) within the same LGE-CMR modality. Leveraging diverse anatomical perspectives would enable more comprehensive myocardial coverage, enhance spatial context, and potentially improve the robustness of scar localization, particularly in complex or ambiguous infarct patterns. Moreover, building on this multi-plane imaging approach, the method could also be extended and adapted to detect other myocardial pathologies, such as non-ischemic cardiomyopathies, where scar morphology and distribution differ significantly from ischemic patterns. Tailoring the segmentation and classification strategies to accommodate these conditions would further broaden the clinical applicability and impact of the proposed system.

Scar regions in LGE-CMR images often exhibit ambiguous intensity profiles, which can hinder accurate boundary delineation. Incorporating pre-segmentation image enhancement techniques, such as denoising autoencoder, attention-guided enhancement networks or transformer-based restoration models, offers promising capabilities for improving visual clarity and emphasizing clinically relevant features. These approaches

may contribute to more accurate and reliable segmentation performance, particularly in cases with low scar-to-background contrast or diffuse scar distribution.

Although the sequential slice reconstruction (3SR) method demonstrated promising results, inter-slice discontinuities were still observed in some infarcted cases. These inconsistencies are likely attributed to variations in cardiac size and non-uniform slice thickness during LGE-CMR acquisition. Future research could explore the integration of motion-aware techniques, such as optical flow estimation or temporal interpolation networks. Combining these methods with the 3 SR framework may result in more anatomically coherent and clinically reliable infarct detection.

REFERENCES

[1] P. Rajiah, M. Y. Desai, D. Kwon, and S. D. Flamm, "MR Imaging of Myocardial Infarction," *RadioGraphics*, vol. 33, no. 5, pp. 1383-1412, Sep. 2013, doi: 10.1148/rg.335125722.

[2] D. M. Papetti *et al.*, "An accurate and time-efficient deep learning-based system for automated segmentation and reporting of cardiac magnetic resonance-detected ischemic scar," *Comput Methods Programs Biomed*, vol. 229, p. 107321, Feb. 2023, doi: 10.1016/j.cmpb.2022.107321.

[3] O. W. Health, "Cardiovascular diseases (CVDs)," 11 June 2021. Accessed: Nov. 29, 2021. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds)

[4] Y. Xie, E. Xu, B. Bowe, and Z. Al-Aly, "Long-term cardiovascular outcomes of COVID-19," *Nat Med*, vol. 28, no. 3, pp. 583-590, Mar. 2022, doi: 10.1038/s41591-022-01689-3.

[5] "Department of Statistics Malaysia." Accessed: Apr. 24, 2025. [Online]. Available: https://www.dosm.gov.my/portal-main/release-content/statistics-on-causes-of-death-malaysia-2023

[6] "INTERACTIVE: Heart disease is top cause of early deaths in Malaysia | The Star." Accessed: Apr. 24, 2025. [Online]. Available: https://www.thestar.com.my/news/nation/2023/ll/29/interactive-heart-disease-is-top-cause-of-early-deaths-in-malaysia

[7] L. Huang *etal.*, "Cardiac Involvement in Patients Recovered From COVID-2019 Identified Using Magnetic Resonance Imaging," *JACC Cardiovasc Imaging*, vol. 13, no. 11, pp. 2330-2339, 2020, doi: 10.1016/j.jcmg.2020.05.004.

[8] A. Di Toro *et al*, "Long COVID: long-term effects?," *European Heart Journal Supplements*, vol. 23, no. SupplementE, pp. E1-E5, Oct. 2021, doi: 10.1093/eurheartj/suab080.

[9] S. M. Sidik, "Heart-disease risk soars after COVID — even with a mild case," *Nature*, vol. 602, no. 7898, pp. 560-560, Feb. 2022, doi: 10.1038/d41586-022-00403-0.

[10] G. Li, W. Zhang, D. Jia, J. Rong, Z. Yu, and D. Wu, "Epidemic of the SARS-CoV-2 Omicron variant in Shanghai, China in 2022: Transient and persistent

effects on Out-of-hospital cardiac arrests," *Resuscitation,* vol. 186, p. 109722, May 2023, doi: 10.1016/j.resuscitation.2023.109722.

[11] O. P. Simonetti *et al,* "An Improved MR Imaging Technique for the Visualization of Myocardial Infarction," *Radiology,* vol. 218, no. 1, pp. 215-223, Jan. 2001, doi: 10.1148/radiology.218.1.r01ja50215.

[12] R. C. Hendel *etal,* "ACCF/ACR/SCCT/SCMR/ASNC/NASCI/SCAI/SIR 2006 appropriateness criteria for cardiac computed tomography and cardiac magnetic resonance imaging: a report of the American College of Cardiology Foundation Quality Strategic Directions Committee Appropriateness C," *J Am Coll Cardiol,* vol. 48, no. 7, pp. 1475-97, Oct. 2006, doi: 10.1016/j.jacc.2006.07.003.

[13] E. Wu *et al,* "Infarct size by contrast enhanced cardiac magnetic resonance is a stronger predictor of outcomes than left ventricular ejection fraction or end-systolic volume index: prospective cohort study," *Heart,* vol. 94, no. 6, pp. 730-736, Jun. 2008, doi: 10.1136/hrt.2007.122622.

[14] A. Busse *et al,* "Cardiac MRI—Update 2020," *Radiologe,* vol. 60, no. SI, pp. 33-40, Nov. 2020, doi: 10.1007/sOOl 17-020-00687-1.

[15] 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, doi: 10.1161/CIRCULATIONAHA.106.653568/FORMAT/EPUB.

[16] G. Ocak *et al,* "Trends in Mortality Due to Myocardial Infarction, Stroke, and Pulmonary Embolism in Patients Receiving Dialysis," *JAMA Netw Open,* vol. 5, no. 4, p. e227624, Apr. 2022, doi: 10.1001/jamanetworkopen.2022.7624.

[17] R. Karim *et al,* "Evaluation of current algorithms for segmentation of scar tissue from late Gadolinium enhancement cardiovascular magnetic resonance of the left atrium: An open-access grand challenge," *Journal of Cardiovascular Magnetic Resonance,* vol. 15, no. l,pp. 1-17, Dec. 2013, doi: 10.1186/1532-429X-15-105.

[18] C. Chen *et al,* "Deep Learning for Cardiac Image Segmentation: A Review," *Front Cardiovasc Med,* vol. 7, no. March, Mar. 2020, doi: 10.3389/fcvm.2020.00025.

[19] Y. Wu, Z. Tang, B. Li, D. Firmin, and G Yang, "Recent Advances in Fibrosis and Scar Segmentation From Cardiac MRI: A State-of-the-Art Review and Future Perspectives," *Front Physiol,* vol. 12, Aug. 2021, doi: 10.3 3 89/fphys.2021.709230.

[20] Y. Liu, M. Zhang, Q. Zhan, D. Gu, and G. Liu, "Two-Stage Method for Segmentation of the Myocardial Scars and Edema on Multi-sequence Cardiac Magnetic Resonance," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 12554 LNCS, pp. 26-36, 2020, doi: 10.1007/978-3-030-6565l-5_3/COVER/.

[21] L. Cai, J. Gao, and D. Zhao, "A review of the application of deep learning in medical image classification and segmentation," *Ann TranslMed*, vol. 8, no. 11, pp. 713-713, 2020, doi: 10.21037/atm.2020.02.44.

[22] Y. Xu, R. Quan, W. Xu, Y. Huang, X. Chen, and F. Liu, "Advances in Medical Image Segmentation: A Comprehensive Review of Traditional, Deep Learning and Hybrid Approaches," *Bioengineering*, vol. 11, no. 10, p. 1034, Oct. 2024, doi: 10.3390/bioengineeringl 1101034.

[23] Z. Chen *et ah,* "Automatic deep learning-based myocardial infarction segmentation from delayed enhancement MRI," *Computerized Medical Imaging and Graphics*, vol. 95, p. 102014, Jan. 2022, doi: 10.1016/j.compmedimag.2021.102014.

[24] F. Zabihollahy, S. Rajan, and E. Ukwatta, "Machine Learning-Based Segmentation of Left Ventricular Myocardial Fibrosis from Magnetic Resonance Imaging," *Curr Cardiol Rep*, vol. 22, no. 8, p. 65, Aug. 2020, doi: 10.1007/sl 1886-020-01321-1.

[25] D. J. Im *et ah,* "Guidelines for cardiovascular magnetic resonance imaging from the Korean society of cardiovascular imaging—part 3: Perfusion, delayed enhancement, and tl-and t2 mapping," *Korean J Radiol*, vol. 20, no. 12, pp. 1562-1582, 2019, doi: 10.3348/kjr.2019.0411.

[26] F. Zabihollahy, J. A. White, and E. Ukwatta, "Myocardial scar segmentation from magnetic resonance images using convolutional neural network," in *Medical Imaging 2018: Computer-Aided Diagnosis*, K. Mori and N. Petrick, Eds., SPIE, Feb. 2018, p. 106. doi: 10.1117/12.2293518.

[27] A. S. Fahmy, E. J. Rowin, R. H. Chan, W. J. Manning, M. S. Maron, and R. Nezafat, "Improved Quantification of Myocardium Scar in Late Gadolinium Enhancement Images: Deep Learning Based Image Fusion Approach.," *JMagn Re son Imaging*, vol. 54, no. 1, pp. 303-312, Jul. 2021, doi: 10.1002/jmri.27555.

[28] S. Moccia *et ah,* "Automated Scar Segmentation From Cardiac Magnetic Resonance-Late Gadolinium Enhancement Images Using a Deep-Learning

Approach," in *Medical physics*, Dec. 2018, pp. 1740-1751. doi: 10.22489/CinC.2018.278.

[29] F. Lau, T. Hendriks, J. Lieman-Sifry, S. Sail, and D. Golden, "ScarGAN: Chained Generative Adversarial Networks to Simulate Pathological Tissue on Cardiovascular MR Scans," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 11045 LNCS, Springer International Publishing, 2018, pp. 343-350. doi: 10.1007/978-3-030-00889-5_39.

[30] K. Brahim, A. Qayyum, A. Lalande, A. Boucher, A. Sakly, and F. Meriaudeau, "A 3D Network Based Shape Prior for Automatic Myocardial Disease Segmentation in Delayed-Enhancement MRI," *IRBM*, vol. 1, no. Mi, pp. 1-11, Feb. 2021, doi: 10.1016/j.irbm.2021.02.005.

[31] Z. F. Shaaf, M. M. A. Jamil, R. Ambar, A. A. Alattab, A. A. Yahya, and Y. Asiri, "Automatic Left Ventricle Segmentation from Short-Axis Cardiac MRI Images Based on Fully Convolutional Neural Network," *Diagnostics*, vol. 12, no. 2, p. 414, Feb. 2022, doi: 10.3390/diagnosticsl2020414.

[32] V. P. Jani, M. Ostovaneh, E. Chamera, Y. Kato, J. A. C. Lima, and B. Ambale-Venkatesh, "Deep learning for automatic volumetric segmentation of left ventricular myocardium and ischaemic scar from multi-slice late gadolinium enhancement cardiovascular magnetic resonance," *Eur Heart J Cardiovasc Imaging*, vol. 25, no. 6, pp. 829-838, May 2024, doi: 10.1093/ehjci/jeae022.

[33] S. Moccia *et ah*, "Development and testing of a deep learning-based strategy for scar segmentation on CMR-LGE images," *Magnetic Resonance Materials in Physics, Biology and Medicine*, vol. 32, no. 2, pp. 187-195, Apr. 2019, doi: 10.1007/sl0334-018-0718-4.

[34] R. Sharma, C. F. Eick, and N. V. Tsekos, "Myocardial Infarction Segmentation in Late Gadolinium Enhanced MRI Images using Data Augmentation and Chaining Multiple U-Net," in *2020 IEEE 20th International Conference on Bioinformatics andBioengineering (BIBE)*, IEEE, Oct. 2020, pp. 975-980. doi: 10.1109/BIBE50027.2020.00165.

[35] R. Wang, T. C. R. Z. Lei, Bingtao Meng, Hongying Nandi, and Asoke K, "Medical image segmentation using deep learning: A survey," *IET Image Process*, vol. 16, no. 5, pp. 1243-1267, 2022.

[36] S. S. M. Salehi, D. Erdogmus, and A. Gholipour, "Tversky Loss Function for Image Segmentation Using 3D Fully Convolutional Deep Networks," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 10541 LNCS, Springer Verlag, 2017, pp. 379-387. doi: 10.1007/978-3-319-67389-9_44.

[37] N. Nasalwai, N. S. Punn, S. K. Sonbhadra, and S. Agarwal, "Addressing the Class Imbalance Problem in Medical Image Segmentation via Accelerated Tversky Loss Function," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 12714 LNAI, Springer, Cham, 2021, pp. 390-402. doi: 10.1007/978-3 -030-75768-7_31.

[38] N. Abraham and N. M. Khan, "A Novel Focal Tversky Loss Function With Improved Attention U-Net for Lesion Segmentation," in *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI2019)*, IEEE, Apr. 2019, pp. 683-687. doi: 10.1109/ISBI.2019.8759329.

[39] E. B. Schelbert *et ah*, "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, Nov. 2010, doi: 10.1161/CIRCFMAGING108.835793.

[40] R. Kirschner *et ah*, "Quantification of myocardial viability distribution with Gd(DTPA) bolus-enhanced, signal intensity-based percent infarct mapping," *Magn Reson Imaging*, vol. 29, no. 5, pp. 650-658, Jun. 2011, doi: 10.1016/j.mri.2011.02.010.

[41] A. Ammari, R. Mahmoudi, B. Hmida, R. Saouli, and M. Hedi Bedoui, "Deep-active-learning approach towards accurate right ventricular segmentation using a two-level uncertainty estimation," *Computerized Medical Imaging and Graphics*, vol. 104, Mar. 2023, doi: 10.1016/j.compmedimag.2022.102168.

[42] R. Alizadehsani *et ah*, "Machine learning-based coronary artery disease diagnosis: A comprehensive review," *Comput Biol Med*, vol. 111, no. June, p. 103346, Aug. 2019, doi: 10.1016/j.compbiomed.2019.103346.

[43] W. J. Richardson, S. A. Clarke, T. A. Quinn, and J. W. Holmes, "Physiological Implications of Myocardial Scar Structure," in *Comprehensive Physiology*, vol. 5, no. 4, Wiley, 2015, pp. 1877-1909. doi: 10.1002/cphy.cl40067.

[44] K. Thygesen *et ah,* "Fourth Universal Definition of Myocardial Infarction (2018)," *J Am Coll Cardiol,* vol. 72, no. 18, pp. 2231-2264, Oct. 2018, doi: 10.1016/j.jacc.2018.08.1038.

[45] "Symptoms of Heart Attack in Men | The Harley Street." Accessed: Jan. 06, 2025. [Online]. Available: https://www.harleystreet.sg/blog/symptoms-of-heart-attack-in-men/

[46] H. K. Awada, M. P. Hwang, and Y. Wang, "Towards comprehensive cardiac repair and regeneration after myocardial infarction: Aspects to consider and proteins to deliver," *Biomaterials,* vol. 82, pp. 94-112, Mar. 2016, doi: 10.1016/j.biomaterials.2015.12.025.

[47] M. B. Stokes and R. Roberts-Thomson, "Diagnostic tests: The role of cardiac imaging in clinical practice," ^4w^ Prefer, vol.40, no. 4, pp. 151-155, Aug. 2017, doi: 10.18773/austprescr.2017.045.

[48] R. E. Thornhill, F. S. Prato, and G. Wisenberg, "The Assessment of Myocardial Viability: A Review of Current Diagnostic Imaging Approaches," *Journal of Cardiovascular Magnetic Resonance,* vol. 4, no. 3, pp. 381-410, 2002, doi: 10.1081/JCMR-120013301.

[49] R. M. Lang *et ah,* "Recommendations for Cardiac Chamber Quantification by Echocardiography in Adults: An Update from the American Society of Echocardiography and the European Association of Cardiovascular Imaging," *Journal of the American Society of Echocardiography,* vol. 28, no. 1, pp. 1-39.el4, Jan. 2015, doi: 10.1016/j.echo.2014.10.003.

[50] L. D. Weberling, D. Lossnitzer, N. Frey, and F. Andre, "Coronary Computed Tomography vs. Cardiac Magnetic Resonance Imaging in the Evaluation of Coronary Artery Disease," *Diagnostics,* vol. 13, no. 1, p. 125, Dec. 2022, doi: 10.3390/diagnosticsl3010125.

[51] A. Baggiano *et ah,* "STress computed tomogRaphy perfusion and stress cArdiac magnetic resonance for ThE manaGement of suspected or known coronarY artery disease: resources and outcomes impact," *J Cardiovasc Comput Tomogr,* vol. 18, no. 6, pp. 553-558, Nov. 2024, doi: 10.1016/j.jcct.2024.08.001.

[52] 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, Apr. 2016, doi: 10.1007/sl0334-015-0521-4.

[53] M. Penso *et ah,* "Automated left and right ventricular chamber segmentation in cardiac magnetic resonance images using dense fully convolutional neural network," *Comput Methods Programs Biomed,* vol. 204, p. 106059, Jun. 2021, doi: 10.1016/j.cmpb.2021.106059.

[54] M. A. Al-antari *et ah,* "Deep learning myocardial infarction segmentation framework from cardiac magnetic resonance images," *Biomed Signal Process Control,* vol. 89, p. 105710, Mar. 2024, doi: 10.1016/j.bspc.2023.105710.

[55] S. Whiteman, Y. Alimi, M. Carrasco, J. Gielecki, A. Zurada, and M. Loukas, "Anatomy of the cardiac chambers: A review of the left ventricle," *Translational Research in Anatomy,* vol. 23, no. October 2020, p. 100095, 2021, doi: 10.1016/j.tria.2020.100095.

[56] J. Brownstein *et ah,* "Method and Atlas to Enable Targeting for Cardiac Radioablation Employing the American Heart Association Segmented Model.," *Int JRadiat Oncol Biol Phys,* vol. 111, no. 1, pp. 178-185, Sep. 2021, doi: 10.1016/j.ijrobp.2021.03.051.

[57] J. Schulz-Menger *et ah,* " Standardized image interpretation and post-processing in cardiovascular magnetic resonance - 2020 update: Society for Cardiovascular Magnetic Resonance (SCMR): Board of Trustees Task Force on Standardized Post-Processing," *Journal of Cardiovascular Magnetic Resonance,* vol. 22, no. 1,pp. 1-22, Mar. 2020, doi: 10.1186/S12968-020-00610-6/FIGURES/10.

[58] N. Ho and Y.-C. Kim, "Evaluation of transfer learning in deep convolutional neural network models for cardiac short axis slice classification," *Sci Rep,* vol. 11, no. 1,p. 1839, Dec. 2021, doi: 10.1038/s41598-021-81525-9.

[59] A. S. Flett *et ah,* "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, Feb. 2011, doi: 10.1016/j.jcmg.2010.11.015.

[60] C. B. Monti *et ah,* "Image quality of late gadolinium enhancement in cardiac magnetic resonance with different doses of contrast material in patients with chronic myocardial infarction," *Eur Radiol Exp,* vol. 4, no. 1, p. 21, Dec. 2020, doi: 10.1186/s41747-020-00149-2.

[61] M. Tang *et al,* "Assessment of Pulmonary Arteries Hemodynamics and Its Relationship With Cardiac Remodeling and Myocardial Fibrosis in Athletes With Four-Dimensional Flow MRI," *Journal of Magnetic Resonance Imaging,* vol. 60, no. 1, pp. 377-387, Jul. 2024, doi: 10.1002/jmri.29048.

[62] R. Shah *et al,* "Quality assessment of cardiac magnetic resonance myocardial scar imaging prior to ventricular arrhythmia ablation," *Int JCardiovasc Imaging,* vol. 39, no. 2, pp. 411-421, Nov. 2022, doi: 10.1007/sl0554-022-02734-5.

[63] A. K. Attili, A. Schuster, E. Nagel, J. H. C. Reiber, and R. J. van der Geest, "Quantification in cardiac MRI: advances in image acquisition and processing," *Int J Cardiovasc Imaging,* vol. 26, no. SI, pp. 27-40, Feb. 2010, doi: 10.1007/sl0554-009-9571-x.

[64] O. Bernard *etal,* "Deep Learning Techniques for Automatic MRI Cardiac Multi-Structures Segmentation and Diagnosis: Is the Problem Solved?," *IEEE Trans Med Imaging,* vol. 37, no. 11, pp. 2514-2525, Nov. 2018, doi: 10.1109/TMI.2018.2837502.

[65] Y. Wu, Z. Tang, B. Li, D. Firmin, and G. Yang, "Recent Advances in Fibrosis and Scar Segmentation From Cardiac MRI: A State-of-the-Art Review and Future Perspectives," *Front Physiol,* vol. 12, p. 709230, Aug. 2021, doi: 10.3 3 89/fphys.2021.709230.

[66] Q. Tao, B. P. F. Lelieveldt, and R. J. van der Geest, "Deep Learning for Quantitative Cardiac MRI," *American Journal of Roentgenology,* vol. 214, no. 3, pp. 529-535, Mar. 2020, doi: 10.2214/AJR.19.21927.

[67] M. Habijan *et al,* "Overview of the Whole Heart and Heart Chamber Segmentation Methods," *Cardiovasc Eng Technol,* vol. 11, no. 6, pp. *125-1 Al,* Dec. 2020, doi: 10.1007/sl3239-020-00494-8.

[68] B. M. Jepson *et al,* "Proposed competencies for the performance of cardiovascular computed tomography in pediatric and adult congenital heart disease," *J Cardiovasc Comput Tomogr,* vol. 17, no. 5, pp. 295-301, Sep. 2023, doi: 10.1016/j.jcct.2023.08.002.

[69] F. Catapano *et al,* "Competence of radiologists in cardiac CT and MR imaging in Europe: insights from the ESCR Registry," *Eur Radiol,* vol. 34, no. 9, pp. 5666-5677, Feb. 2024, doi: 10.1007/s00330-024-10644-4.

[70] T. C. Kwee and R. M. Kwee, "Workload of diagnostic radiologists in the foreseeable future based on recent scientific advances: growth expectations and

role of artificial intelligence," *Insights Imaging*, vol. 12, no. 1, p. 88, Dec. 2021, doi: 10.1186/sl3244-021-01031-4.

[71] L. Alzubaidi *et al*, "Review of deep learning: concepts, CNN architectures, challenges, applications, future directions," *J Big Data*, vol. 8, no. 1, p. 53, Mar. 2021, doi: 10.1186/s40537-021-00444-8.

[72] R. Yamashita, M. Nishio, R. K. G. Do, and K. Togashi, "Convolutional neural networks: an overview and application in radiology," *Insights Imaging*, vol. 9, no. 4, pp. 611-629, Aug. 2018, doi: 10.1007/sl3244-018-0639-9.

[73] Q.-Y. Xiang *et al*, "Performance analysis of quantum convolutional layers for image classification," *Phys Scr*, vol. 99, no. 12, p. 125120, Dec. 2024, doi: 10.1088/1402-4896/ad8dl7.

[74] G. Huang, Z. Liu, G. Pleiss, L. van der Maaten, and K. Q. Weinberger, "Convolutional Networks with Dense Connectivity," *IEEE Trans Pattern Anal Mach Intell*, vol. 44, no. 12, pp. 8704-8716, Dec. 2022, doi: 10.1109/TPAMI.2019.2918284.

[75] H. Zhang and J. Hui, "A New Semantic Segmentation Network with FPFN and Dense ASPP," in *10th International Conference on Control, Automation and Information Sciences, ICCAIS 2021 - Proceedings*, Institute of Electrical and Electronics Engineers Inc., 2021, pp. 799-804. doi: 10.1109/ICCAIS52680.2021.9624497.

[76] W. Qiu, L. Gu, F. Gao, and T. Jiang, "Building Extraction From Very High-Resolution Remote Sensing Images Using Refine-UNet," *IEEE Geoscience and Remote Sensing Letters*, vol. 20, 2023, doi: 10.1109/LGRS.2023.3243609.

[77] S. Guo and C. Zhu, "Cascaded ASPP and Attention Mechanism-based Deeplabv3+ Semantic Segmentation Model," in *Proceedings of 2022 8th IEEE International Conference on Cloud Computing and Intelligence Systems, CCIS 2022*, Institute of Electrical and Electronics Engineers Inc., 2022, pp. 315-318. doi: 10.1109/CCIS57298.2022.10016433.

[78] G. Sunandini, R. Sivanpillai, V. Sowmya, and V. V. Sajith Variyar, "Significance of Atrous Spatial Pyramid Pooling (ASPP) in Deeplabv3+ for Water Body Segmentation," in *Proceedings of the 10th International Conference on Signal Processing and Integrated Networks, SPIN 2023*, Institute of Electrical and Electronics Engineers Inc., 2023, pp. 744-749. doi: 10.1109/SPIN5 7001.2023.10116882.

[79] L.-C. Chen, Y. Zhu, G. Papandreou, F. Schroff, and H. Adam, "Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation," in *Computer Science*, vol. 11211, Springer, Cham, 2018, pp. 833-851. doi: 10.1007/978-3-030-01234-2_49.

[80] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, "DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs," Jun. 2016, [Online]. Available: http://arxiv.org/abs/1606.00915

[81] L.-C. Chen, G. Papandreou, F. Schroff, and H. Adam, "Rethinking Atrous Convolution for Semantic Image Segmentation," Jun. 2017, [Online]. Available: http://arxiv.org/abs/1706.05587

[82] A. Aqthobirrobbany, R. D. H. Al-Fahsi, I. Soesanti, and H. A. Nugroho, "Enhanced U-Net architecture with CNN backbone for accurate segmentation of skin lesions in dermoscopic images," *International Journal of Advances in Intelligent Informatics*, vol. 10, no. 3, p. 490, Aug. 2024, doi: 10.26555/ijain.vl0i3.1379.

[83] M. Bal-Ghaoui, M. H. El Yousfi Alaoui, A. Jilbab, and A. Bourouhou, "U-Net transfer learning backbones for lesions segmentation in breast ultrasound images," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 13, no. 5, p. 5747, Oct. 2023, doi: 10.11591/ijece.vl3i5.pp5747-5754.

[84] N. Das and S. Das, "Attention-UNet architectures with pretrained backbones for multi-class cardiac MR image segmentation," *Curr Probi Cardiol*, vol. 49, no. 1, p. 102129, Jan. 2024, doi: 10.1016/j.cpcardiol.2023.102129.

[85] J. Liu, J. He, J. Zhang, J. S. Ren, and H. Li, "EfficientFCN: Holistically-Guided Decoding for Semantic Segmentation," 2020, pp. 1-17. doi: 10.1007/978-3-030-58574-7J.

[86] V. Badrinarayanan, A. Kendall, and R. Cipolla, "SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation," *IEEE Trans Pattern Anal Mach Intell*, vol. 39, no. 12, pp. 2481-2495, Dec. 2017, doi: 10.1109/TPAMI.2016.2644615.

[87] M. Hammad, M. H. Alkinani, B. B. Gupta, and A. A. Abd El-Latif, "Myocardial infarction detection based on deep neural network on imbalanced data,"

Multimed Syst, vol. 28, no. 4, pp. 1373-1385, Aug. 2022, doi: 10.1007/s00530-020-00728-8.

[88] M. Yeung, E. Sala, C.-B. Schonlieb, and L. Rundo, "Unified Focal loss: Generalising Dice and cross entropy-based losses to handle class imbalanced medical image segmentation," *Computerized Medical Imaging and Graphics*, vol. 95, p. 102026, Jan. 2022, doi: 10.1016/j.compmedimag.2021.102026.

[89] P. O. Bressan *et al.*, "Semantic segmentation with labeling uncertainty and class imbalance applied to vegetation mapping," *International Journal of Applied Earth Observation and Geoinformation*, vol. 108, p. 102690, Apr. 2022, doi: 10.1016/j.jag.2022.102690.

[90] P. O. Bressan *et al.*, "Semantic Segmentation with Labeling Uncertainty and Class Imbalance," 2021, [Online]. Available: http://arxiv.org/abs/2102.04566

[91] Y. Liu *et al.*, "Exploring Uncertainty Measures in Bayesian Deep Attentive Neural Networks for Prostate Zonal Segmentation," *IEEE Access*, vol. 8, pp. 151817-151828, 2020, doi: 10.1109/ACCESS.2020.3017168.

[92] C. H. Sudre, W. Li, T. Vercauteren, S. Ourselin, and M. Jorge Cardoso, "Generalised Dice Overlap as a Deep Learning Loss Function for Highly Unbalanced Segmentations," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 10553 LNCS, Springer Verlag, 2017, pp. 240-248. doi: 10.1007/978-3-319-67558-9_28.

[93] P. Goyal, "Shallow SegNet with bilinear interpolation and weighted cross-entropy loss for Semantic segmentation of brain tissue," in *2022 IEEE International Conference on Signal Processing, Informatics, Communication and Energy Systems (SPICES)*, IEEE, Mar. 2022, pp. 361-365. doi: 10.1109/SPICES52834.2022.9774193.

[94] J. Ni, J. Wu, J. Tong, Z. Chen, and J. Zhao, "GC-Net: Global context network for medical image segmentation," *ComputMethods Programs Biomed*, vol. 190, p. 105121, Jul. 2020, doi: 10.1016/j.cmpb.2019.105121.

[95] T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollar, "Focal Loss for Dense Object Detection," *IEEE Trans Pattern AnalMach Intell*, vol. 42, no. 2, pp. 318-327, Feb. 2020, doi: 10.1109/TPAMI.2018.2858826.

[96] H. Hui, X. Zhang, Z. Wu, and F. Li, "Dual-Path Attention Compensation U-Net for Stroke Lesion Segmentation," *ComputIntellNeurosci*, vol. 2021, no. 1, Jan. 2021, doi: 10.1155/2021/7552185.

[97] S.-W. Yoo *et al*, "CACSNet for automatic robust classification and segmentation of carotid artery calcification on panoramic radiographs using a cascaded deep learning network," *Sci Rep*, vol. 14, no. 1, p. 13894, Jun. 2024, doi: 10.1038/s41598-024-64265-4.

[98] F. E. Fernandes Junior and G. G. Yen, "Particle swarm optimization of deep neural networks architectures for image classification," *Swarm Evol Comput*, vol. 49, pp. 62-74, Sep. 2019, doi: 10.1016/j.swevo.2019.05.010.

[99] Y. Ru *et al*, "Accurate prediction of the rheological behavior of MWCNT-A1203/water-ethylene glycol nanofluid with metaheuristic-optimized machine learning models," *International Journal of Thermal Sciences*, vol. 211, p. 109691, May 2025, doi: 10.1016/j.ijthermalsci.2025.109691.

[100] D. Sahu *et al*, "Beyond boundaries a hybrid cellular potts and particle swarm optimization model for energy and latency optimization in edge computing," *Sci Rep*, vol. 15, no. 1, p. 6266, Feb. 2025, doi: 10.1038/s41598-025-90348-x.

[101] X. Zhou, X. Wang, and R. Guo, "Assessment model of ozone pollution based on SHAP-IPSO-CNN and its application," *Sci Rep*, vol. 15, no. 1, p. 3404, Jan. 2025, doi: 10.1038/s41598-025-87702-4.

[102] S. Gangwar, R. Devi, and N. A. Mat Isa, "Optimized exposer region-based modified adaptive histogram equalization method for contrast enhancement in CXR imaging," Sczitep, vol. 15, no. 1, p. 6693, Feb. 2025, doi: 10.1038/s41598-025-90876-6.

[103] A. Maqbool, A. U. Rehman, A. Arshad, K. Mahmoud, and M. Lehtonen, "Hybrid metaheuristic optimization based DSM approach towards effective energy recommender system," *Electric Power Systems Research*, vol. 246, p. 111645, Sep. 2025, doi: 10.1016/j.epsr.2025.111645.

[104] B. Revathi, S. K. K. Elizabeth, P. Nagaraj, S. S. Birunda, and D. Nithya, "Particle Swarm Optimization based Detection of Diabetic Retinopathy using a Novel Deep CNN," in *2023 Third International Conference on Artificial Intelligence and Smart Energy (ICAIS)*, IEEE, Feb. 2023, pp. 998-1003. doi: 10.1109/ICAIS56108.2023.10073926.

[105] L. C. Amado *et al,* "Accurate and objective infarct sizing by contrast-enhanced magnetic resonance imaging in a canine myocardial infarction model," *J Am Coll Cardiol,* vol. 44, no. 12, pp. 2383-2389, Dec. 2004, doi: 10.1016/j.jacc.2004.09.020.

[106] A. Kolipaka, G. P. Chatzimavroudis, R. D. White, T. P. O'Donnell, and R. M. Setser, "Segmentation of non-viable myocardium in delayed enhancement magnetic resonance images," *Int J Cardiovasc Imaging,* vol. 21, no. 2-3, pp. 303-311, Apr. 2005, doi: 10.1007/sl0554-004-5806-z.

[107] 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, May 2016, doi: 10.1016/j.media.2016.01.004.

[108] M. Abdulkareem *etal,* "Predicting post-contrast information from contrast agent free cardiac MRI using machine learning: Challenges and methods," *Front Cardiovasc Med,* vol. 9, Jul. 2022, doi: 10.3389/fcvm.2022.894503.

[109] W. Ding *et al,* "Aligning Multi-Sequence CMR Towards Fully Automated Myocardial Pathology Segmentation," *IEEE Trans Med Imaging,* vol. 42, no. 12, pp. 3474-3486, Dec. 2023, doi: 10.1109/TMI.2023.3288046.

[110] M. Chen, L. Fang, Q. Zhuang, and H. Liu, "Deep Learning Assessment of Myocardial Infarction From MR Image Sequences," *IEEE Access,* vol. 7, pp. 5438-5446, 2019, doi: 10.1109/ACCESS.2018.2889744.

[Ill] A. S. Fahmy *etal,* "Automated Cardiac MR Scar Quantification in Hypertrophic Cardiomyopathy Using Deep Convolutional Neural Networks," *JACC Cardiovasc Imaging,* vol. 11, no. 12, pp. 1917-1918, Dec. 2018, doi: 10.1016/j.jcmg.2018.04.030.

[112] D. M. Popescu *et al,* "Anatomically informed deep learning on contrast-enhanced cardiac magnetic resonance imaging for scar segmentation and clinical feature extraction," *Cardiovasc Digit Health J,* vol. 3, no. 1, pp. 2-13, Feb. 2022, doi: 10.1016/j.cvdhj.2021.11.007.

[113] S. Wang, A. M. S. E. K. Abdelaty, K. Parke, J. R. Arnold, G. P. McCann, and I. Y. Tyukin, "Myl-Net: Fully Automatic Detection and Quantification of Myocardial Infarction from Cardiovascular MRI Images," *Entropy,* vol. 25, no. 3, p. 431, Feb. 2023, doi: 10.3390/e25030431.

[114] J. Ma, "Cascaded Framework for Automatic Evaluation of Myocardial Infarction from Delayed-Enhancement Cardiac MRI," *arXivpreprint*, pp. 1-7, Dec. 2020, Accessed: Apr. 09, 2025. [Online]. Available: http://arxiv.org/abs/2012.14556

[115] E. de la Rosa, D. Sidibe, T. Decourselle, T. Leclercq, A. Cochet, and A. Lalande, "Myocardial Infarction Quantification from Late Gadolinium Enhancement MRI Using Top-Hat Transforms and Neural Networks," *Algorithms*, vol. 14, no. 8, p. 249, Aug. 2021, doi: 10.3390/al4080249.

[116] F. Zabihollahy, J. A. White, and E. Ukwatta, "Convolutional neural network-based approach for segmentation of left ventricle myocardial scar from 3D late gadolinium enhancement MR images.," *MedPhys*, vol. 46, no. 4, pp. 1740-1751, Apr. 2019, doi: 10.1002/mp. 13436.

[117] F. Zabihollahy, M. Rajchl, J. A. White, and E. Ukwatta, "Fully automated segmentation of left ventricular scar from 3D late gadolinium enhancement magnetic resonance imaging using a cascaded multi-planar U-Net (CMPU-Net)," *MedPhys*, vol. 47, no. 4, pp. 1645-1655, 2020, doi: 10.1002/mp. 14022.

[118] A. S. Fahmy *et ah*, "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, Jan. 2020, doi: 10.1148/radiol.2019190737.

[119] K. Brahim, A. Qayyum, A. Lalande, A. Boucher, A. Sakly, and F. Meriaudeau, "A deep learning approach for the segmentation of myocardial diseases," in *2020 25th International Conference on Pattern Recognition (ICPR)*, IEEE, Jan. 2021, pp. 4544-4551. doi: 10.1109/ICPR48806.2021.9412793.

[120] Y. EL kati, S.-L. Wang, M. M. Taresh, and T. A. A. Ali, "Enhancing breast cancer diagnosis using deep learning and gradient multi-verse optimizer: a robust biomedical data analysis approach," *PeerJ Comput Sci*, vol. 10, p. e2578, Dec. 2024, doi: 10.7717/peerj-cs.2578.

[121] P. Nagpal, S. A. Bhinge, and A. Shitole, "A Comparative Analysis of ResNet Architectures," in *2022 International Conference on Smart Generation Computing, Communication and Networking (SMART GENCON)*, IEEE, Dec. 2022, pp. 1-8. doi: 10.1109/SMARTGENCON56628.2022.10083966.

[122] I. Kandel and M. Castelli, "The effect of batch size on the generahzabihty of the convolutional neural networks on a histopathology dataset," *ICTExpress*, vol. 6, no. 4, pp. 312-315, Dec. 2020, doi: 10.1016/j.icte.2020.04.010.

[123] H.-C. Shin *et al*, "Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics and Transfer Learning," *IEEE Trans Med Imaging*, vol. 35, no. 5, pp. 1285-1298, May 2016, doi: 10.1109/TMI.2016.2528162.

[124] Y. Singh, M. Saini, and Savita, "Impact and Performance Analysis of Various Activation Functions for Classification Problems," in *2023 IEEE International Conference on Contemporary Computing and Communications (InC4)*, IEEE, Apr. 2023, pp. 1-7. doi: 10.1109/InC457730.2023.10263129.

[125] M. Hasani and H. Khotanlou, "An Empirical Study on Position of the Batch Normalization Layer in Convolutional Neural Networks," in *2019 5th Iranian Conference on Signal Processing and Intelligent Systems (ICSPIS)*, IEEE, Dec. 2019, pp. 1-4. doi: 10.1109/ICSPIS48872.2019.9066113.

[126] C. Garbin, X. Zhu, and O. Marques, "Dropout vs. batch normalization: an empirical study of their impact to deep learning," *Multimed Tools Appl*, vol. 79, no. 19-20, pp. 12777-12815, 2020, doi: 10.1007/sll042-019-08453-9.

[127] P. Sahu, A. Chug, A. P. Singh, D. Singh, and R. P. Singh, "Implementation of CNNs for Crop Diseases Classification: A Comparison of Pre-trained Model and Training from Scratch," *Inter nationalJournal ofComputer Science and Network Security*, vol. 20, no. 10, pp. 206-215, 2020.

[128] C. Oksiiz and M. K. Gullii, "Deep Feature Extraction Based Fine-Tuning," in *2020 28th Signal Processing and Communications Applications Conference (SIU)*, IEEE, Oct. 2020, pp. 1-4. doi: 10.1109/SIU49456.2020.9302108.

[129] J. Margeta, A. Criminisi, R. Cabrera Lozoya, D. C. Lee, and N. Ayache, "Fine-tuned convolutional neural nets for cardiac MRI acquisition plane recognition," *ComputMethods Biomech BiomedEngImaging Vis*, vol. 5, no. 5, pp. 339-349, Sep. 2017, doi: 10.1080/21681163.2015.1061448.

[130] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Commun ACM*, vol. 60, no. 6, pp. 84-90, 2017, doi: 10.1145/3065386.

[131] C. Szegedy *et al*, "Going deeper with convolutions," in *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, Jun. 2015, pp. 1-9. doi: 10.1109/CVPR.2015.7298594.

[132] F. N. Iandola, S. Han, M. W. Moskewicz, K. Ashraf, W. J. Dally, andK. Keutzer, "SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5MB

model size," vol. 4, p. 370, Feb. 2016, [Online]. Available: http://arxiv.org/abs/1602.07360

[133] M. F. Abdullah, S. N. Sulaiman, M. K. Osman, N. K. A. Karim, I. S. Isa, and I. L. Shuaib, "Designation of Thorax and Non-Thorax Regions for Lung Cancer Detection in CT Scan Images using Deep Learning," *Journal of Electrical & Electronic Systems Research*, vol. 17, no. DEC 2020, pp. 41-49, Dec. 2020, doi: 10.24191/jeesr.vl7il.006.

[134] C. Wang, P. Du, H. Wu, J. Li, C. Zhao, and H. Zhu, "A cucumber leaf disease severity classification method based on the fusion of DeepLabV3+ and U-Net," *Comput Electron Agric*, vol. 189, Oct. 2021, doi: 10.1016/j.compag.2021.106373.

[135] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, andL.-C. Chen, "MobileNetV2: Inverted Residuals and Linear Bottlenecks," in *2018IEEE/CVF Conference on Computer Vision and Pattern Recognition*, IEEE, Jun. 2018, pp. 4510^520. doi: 10.1109/CVPR.2018.00474.

[136] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, Jun. 2016, pp. 770-778. doi: 10.1109/CVPR.2016.90.

[137] H. Polat, "A modified DeepLabV3+ based semantic segmentation of chest computed tomography images for COVID-19 lung infections," *Int J Imaging Syst Technol*, vol. 32, no. 5, pp. 1481-1495, Sep. 2022, doi: 10.1002/ima.22772.

[138] Y. Zhang and T. Shen, "Small Object Detection with Multiple Receptive Fields," *IOP ConfSer Earth Environ Sci*, vol. 440, no. 3, p. 032093, Feb. 2020, doi: 10.1088/1755-1315/440/3/032093.

[139] R. Davuluri and R. Rengaswamy, "Improved Classification Model using CNN for Detection of Alzheimer's Disease," *Journal of Computer Science*, vol. 18, no. 5, pp. 415-425, May 2022, doi: 10.3844/jcssp.2022.415.425.

[140] Monica and P. Agrawal, "A Survey on Hyperparameter Optimization of Machine Learning Models," in *2024 2nd International Conference on Disruptive Technologies (ICDT)*, IEEE, Mar. 2024, pp. 11-15. doi: 10.1109/ICDT61202.2024.10489732.

[141] P. Gut *et ah*, "Improved myocardial scar visualization using free-breathing motion-corrected wideband black-blood late gadolinium enhancement imaging

in patients with implantable cardiac devices," *Diagn Interv Imaging*, vol. 106, no. 5, pp. 169-182, May 2025, doi: 10.1016/j.diii.2024.12.001.

[142] S. Sridi *et ah*, "Improved myocardial scar visualization with fast free-breathing motion-compensated black-blood Tl-rho-prepared late gadolinium enhancement MRI," *Diagn Interv Imaging*, vol. 103, no. 12, pp. 607-617, Dec. 2022, doi: 10.1016/j.diii.2022.07.003.

APPENDICES

APPENDIX 1

Ethical Approval and Extension Approval

Jawatankuasa Etika
Penyelidikan Manusia USM (JEPeM)

Human Research Ethics Committee USM (HREC)

Universiti Sains Malaysia
Kampus Kesihatan
16150 KuDang Kenan, Kelantan, Malaysia.
Tel. : +609 - 767 3000/2354/2362
Fax. : + 609 -767 2351
Email: jepem(S>usm.my
La man Web : www.jepem.kk.usm.my
www.usm.my

22nd December 2023

Assoc. Prof. Ir. Ts. Dr. Siti Noraini Sulaiman
School of Electrical Engineering
Universiti Teknologi Mara Cawangan Pulau Pinang
Kampus Permatang Pauh
13500 Permatang Pauh, Pulau Pinang

JEPeM USM Code: USM/JEPeM/21090623
Study Protocol Title: An Automated Segmentation of Left Ventricular Scar from Cardiac MR Images Based on Deep Convolutional Neural Networks for Myocardial Infarction Diagnosis

Dear Assoc. Prof. **Ir.** Ts. Dr. Siti Noraini,

We wish to inform you that the Jawatankuasa Etika Penyelidikan Manusia, Universiti Sains Malaysia (JEPeM-USM) acknowledged receipt of the Continuing Review Application dated 13* December 2023.

Upon review of JEPeM-USM Form 3(B) 2021: Continuing Review Application Form, the committee **AGREED** for the **EXTENSION OF APPROVAL (commencing from 5th February 2024 to 4'h February 2025).** The document is included in the protocol file.

Thank you for your continuing compliance with the requirements of JEPeM-USM.

"MALAYSIA MADANI"

"BERKHIDMAT UNTUK NEGARA"

Sincerely,

ASSOC. PROF. DR. HASLINA HAROON
Deputy Chairperson
Jawatankuasa Etika Penyelidikan (Manusia), JEPeM
Universiti Sains Malaysia

c.c Secretary
Jawatankuasa Etika Penyelidikan (Manusia), JEPeM
Universiti Sains Malaysia

JAWAIANKUASA ETIKA
PtNYEUDIKAN MANUSIA

JiwBtmkuBii Etlki
Penydidikan Minus*m* USM (JEPeM]

Human Research Ethkx QnmniUBe USM (MEC)

Jni.?-ji:i ^ains Ma eyjie
^HT.aus ^ejihatar.
1615D 'iLsair. Ke-iarv Kelantnri, Malnysss.
Tel. : -509 - 767 30«y2551;2362
fiK. :4HS-76723J1
Email : jaaeT^LZin.niy
UiTiari Web : www.jec^n.kk. jsi^.rny
www. jETi.ny

5* February 2023

Assoc. Prof. Ir. Ts. Dr. Siti Noraini Sulaiman
School of Electrical Engineering
Universiti Teknologi Mara Cawangan Pulau Pinang
Kampus Permatang Pauh
13500 Permatang Pauh, Pulau Pinang

JEPeM USM Code: USMUEPeM/21090623

Study Protocol Title: An Automated Segmentation of Left Ventricular Scarfrom Cardiac MR Images Based on Deep Convolutional Neural Networks For Myocardial Infarction Diagnosis

Dear Prof.,

We wish to inform you that Jawatankuasa Etika Penyelidikan Manusia, Universiti Sains Malaysia (JEPeM-USM) acknowledged receipt your application for continuing review dated 20[11] January 2023.

Upon review of JEPeM-USM Form 3(B) 2021: Continuing Review Application Form, the committee AGREED for the EXTENSION OF APPROVAL (commencing on 5[1]* February 2023 to 4* February 2024). The report is included in the protocol fie.

JEPeM-USM has noted that there is no research activity conducted during the period of 22[1u3] November 2022 until 4[m] February 2023 and this has been included in the protocol fie.

Principle Investigator {PI) should be aware and concerned about the ethical expiration of the study in the Future.

Thank you for your continuing compliance with the requirements of the JEPeM-USM.

"BERKHIDMAT UNTUK NEGARA"

Sincerely,

(ASSOC. PROF. DR. SHAHROM MAHMUD)
Deputy Chairperson
Jawatankuasa Etika Penyelidikan (Manusia), JEPeM
Universiti Sains Malaysia

c.c Secretary
Jawatankuasa Etika Penyelidikan (Manusia), JEPeM
Universiti Sains Malaysia

£ ^
JAWSTikKUISS ETIKA
PIKHELUHKINMAhUSIS.

APPENDIX 2

Flowchart for Applying JEPeM USM Ethical Approval

ACTIVITY			Responsibility
			JEPeM-USM Secretariat
OR Research Committee			
Yes (3 copies]		No	
Send to JEPeM-USM Chair		Send back to applicant J,	
		Vetted/Presented at PTJ	Applicant/ Supervisor
i	1	Resubmit	
Eaempted	Aalgn Protocol for Expedited Oft FuHboard Review		JEPeM-USM
Send *cknoiMlwfa«fiwnt kcft« to *npik«4	Forward Protocol for Enpedited 00 Fuiiboird Review *(trfrr to Fknnhort 6 or 7 SOP II)*		JEPeM-USM Secretariat

APPENDIX 3

Anonymized Cardiac MRI Radiology Report Samples from AMDI (IPPT)

PPUSMB

Patient Name	MRN #	Access. #	DOB	Age	Sex	Date	Mod	Study Description	Ordered by
ANONYMIZED			ANONYMIZED					Cardiac Without & Witt Contrast	

Radiology Report

Cardiac MRI (Viability protocol) performed on 17/4/2023 at Philips MRI 3T.

Sequences:
Axial BB and WB
Cine: 4Ch, LVOT, *2Ch*. SA stack
EGE: 4Ch_ LVOT, and 2Ch stacks
LGE: 4Ch. LVOT *2Ch* SA
Tl native and post-contrast: SA (1 SL)

"Weight: 75 kg
Height: l i S m
Body surface area: 1.77 *m-* (Du Bois)
BP: 173/79 mruHg
Haematocnt: 44_S% on 17.04.2023.

Findings:
Left ventricle (LV):
Increased indexed end diastolic volume and indexed end systolic volume with reduced ejection fraction.
Global hypokinesia with thin akinetic basal anterolateral wall: and basal to mid lateral and inferior walls.
Thickened basal-iiud anteroseptal wall, maximum 1.2S cm.
Increased indexed mass.
No aneurysmal formation.

Right ventricle (RV):
Normal indexed end diastolic: volume and end systolic volume with low normal ejection fraction.
Normal tricuspid annular plane systolic excursion (TAPSE). measuring 2.4 cm.
No regional wall motion abnormality.
No aneurysmal formation.

Atria:
Marginally dilated tight atrium - indexed area 15.4 cmVm- (Normal <15.0cm-/m^2)
Mildly dilated left atrium - indexed area 17 cm^2/ui' (Normal < 15 Ocni'/m-)
No interatrial septal wall thickening.

Valves:
Mitral valve regurgitation (MR), likely mild "by visual assessment under scan condition.

Post-gadoliniLim:
Early phase:
No intraluminal thrombus or mass. No early myocardial enhancement.

Late phase:
Transmural hypereiihancement of basal anterolateral wall; and basal to mid lateral and inferior walls.

Thoracic great vessels:
The thoracic aorta is normal in size.
Mildly dilated pulmonary trunk measuring 3.1 cm.
Normal systemic and pulmonary venous returns.

Impressions:
1. Dilated LV with reduced ejection fraction, LVEF 27 %
2. Normal RV size with low normal election fraction (RVEF *45* %)
3. Transmural myocardial infarction of basal antero lateral wall; and basal to mid lateral and inferior walls. MI
4. No myocardial inflammation or infiltration
5. Basal-mid anteroseptal wall hypertrophy, likely secondary to hypertension.
6. Mild MR. Suggest ECHO correlation for proper valve assessment.
7. Mild bi-atnal dilatation.

Features of ischemic dilated cardiomyopathy.
LAD: All segments are viable.
LCx: 3 of 5 segments are non-viable
RCA : 2 of 5 segments are non-viable.
See myocardial viability map for summary.

Reported with Dr ,Dr ,Dr. , Dr
Reporting team: Dr *J* Dr / Dr / Dr .' Dr / Dr ' Dr

Reported by: DR ANONYMIZED ,DR! ,DR
JJR
09/05/2023 _

PPUSMB

Patient Name	MRN # Access. #	DOB	Age	Sex	Dale	Mod	Study Description	Ordered by
ANONYM IZED							Cardiac Without Contrast	

Radiology Report

Cardiac MRI performed on 20/2/2023 at Philips Acliieva 3.0T scanner

Sequences:
Axial EB and WB
Cine: 4Ch, LVOT, 2Ch, SA stack
STIR: *4Ch*.. LVOT, 2Ch, SA stack
EGE: 4Ch, LVOT, and 2Ch
LGE: 4Ch, LVOT, 2Ch, SA
T2 map SA base, mid and apical
Tl native and post-contrast: SA (1 SL)

Weight: 58 kg
Height: 1.54 m
Body surface area: 1.58 m^2 (Du Bois)
BP 133/76 inmHg
Haernatocnt (20/2/2023) :40.0%

Findings:
LefWentncle £LV):
High normal indexed end diastolic volume and increased indexed end systolic volume with low normal ejection fraction
Global hypokinesia especially from basal to mid segments, more notable at the basal septum.
Normal wall thickness and indexed mass.
No aneurysmal formation.

Right ventricle (RV):
Normal indexed end diastolic volume and indexed end systolic volume with normal ejection fraction
Normal tricuspid annular plane systolic excursion (TAPSE). measuring 2.0 cm.
Slightly hypokinetic apical free wall.
No aneurysmal formation

Atria:
Normal right atrial size, area 14 cm^2.
Normal left atrium size, area 18 cm^2.
No interatrial septal wall thickening.

Valves:
Suspicious of trivial mitral (MR) and aortic (AR) regurgitation by visual assessment under scan condition.

STIR:
No significant increase in myocardial signal intensity.

Parametric mapping:
Relatively increased mean native Tl relaxation time and extracellular volume fraction at basal septum.
Suboptimal assessment for myocardial T2 values due to respiratory motion artefact.

Post-gadolinium:
Early phase: No intraluminal thrombus or mass.
Late phase: Enhancement of the inferior RV insertion point.

Thoracic great vessels:
The thoracic aorta and pulmonary' artery are normal in size.
Normal systemic and pulmonary venous returns.

Impressions:
1. High normal LV size with low normal ejection fraction. LVEF 52%.
2. Normal RV size with normal ejection fraction, RVEF 57%.
3. Relative increase in Tl values and ECV at "basal septum. Limited assessment of T2 map.
4. Fibrosis at inferior RV insertion points.
5_ No evidence of myocardial infarction. Non-MI
6. Suspicious trivial MR & AR. Suggest ECHO correlation for proper valve evaluation.

Correlating w'ith clinical history, improvement of LVEF from 23% Jin previous echocardiography) to 52% and evidence of interstitial fibrosis (relative increase in Tl values/ECV), this could indicate a resolving myocarditis.
The fibrosis at the inferior RV insertion site is usually due to the myocardial strain The remaining LV wall hypokinesia is likely a sequelae of the myocarditis. However, given the diabetic status and mild coronary artery disease on angiography, there is also a possibility of hibernating myocardium as a result of microvascular dysfunction. Consider stress myocardial perfusion after 6 months.

Reported with Dr Numl Ain Bt Mat Idns (Radiologist), Dr Siti Aishah Binti Ahmad Maulana (Radiologist) and Dr Khairil Amir Sayuti (Consultant Radiologist)
Dr Saifi.il / Dr Qistina / Dr Thiban / Dr Atifah .' Dr Jaya

Reported by: AMONVHIZED DR ,DR

AUTHOR'S PROFILE

Dayang Suhaida Awang Damit received her Bachelor of Engineering (Electrical-Telecommunications) with Honours from Universiti Teknologi Malaysia, Skudai, Johor in 2006. She subsequently earned a Master of Science in Electronic Design from the School of Electrical and Electronics Engineering, Universiti Sains Malaysia in 2010. Currently, she is a lecturer at the Electrical Engineering Studies Department, Universiti Teknologi MARA (UiTM) Pulau Pinang. Her research interests encompass wireless communication, image processing, artificial intelligence, and biomedical engineering.

LIST OF PUBLICATIONS:

Grant

1. Title: *New Convolutional Neural Networks - Autoencoder Model with Fusion Correlation Layer for Left Ventricle Classification and Scar Tissue Segmentation in Cardiac Magnetic Resonance Images of Myocardial Infarction Disease.*
   Code/Amount: FRGS/1/2023/SKK06/UITM/5/3Q33/2023) - (RM 96,231.00).
   From: Fundamental Research Grant Scheme (FRGS), Ministry of Higher Education, Malaysia. Muhammad Khusairi bin Osman, Siti Noraini Sulaiman, **Dayang Suhaida Binti Awang Damit,** Noor Khairiah A. Karim, Khairil Amir Bin Sayuti. Duration: 1 October 2023 - 30 September 2025,

Refereed Journal

1. **Dayang Suhaida Awang Damit,** Siti Noraini Sulaiman, Muhammad Khusain Osman, Noor Khairiah A. Karim, Samsul Setumin, "Automated DeepLabV3+ based model for left ventricle segmentation on short-axis late gadolinium enhancement-magnetic cardiac resonance imaging images", International Journal of Electrical and Computer Engineering (IJECE). Vol. 14, No. 3, June 2024, **Scopus indexed,** Status: **Published.**
2. **Dayang Suhaida Awang Damit,** Siti Noraini Sulaiman, Muhammad Khusain Osman, Noor Khairiah A. Karim, Samsul Setumin, "Better Network Efficiency Through Batch Normalization and Optimal Learning Rate in Left Ventricle Chamber Classification", Pertanika Journal of Science and Technology. Volume 33, Issue 2, March 2025 **Scopus & WoS indexed,** Status: **Published.**
3. **Dayang Suhaida Awang Damit,** Siti Noraini Sulaiman, Muhammad Khusain Osman, Noor Khairiah A. Karim, Samsul Setumin, "Classification of Left Ventricle and Non- Left Ventricle Segment for Cardiac Assessment Using Deep Convolutional Neural Network", Journal of Electrical and Electronics Systems Research (JEESR) Vol. 21, pp. 31-38 (October 2022). **MyCite indexed,** Status: **Published.**

List of Proceedings

1. **Dayang Suhaida Awang Damit,** A'isyah Nadiyah Hilmi, Siti Noraini Sulaiman, Muhammad Khusairi Osman, Noor Khairiah A. Karim, Nor Adni Mat Leh, "Impact of Class Labeling on Myocardium Segmentation using Cascaded Deep Learning for Improved Myocardial Infarction Segmentation", in 14th IEEE International Conference on Control System, Computing and Engineering (ICCSCE24), 25 Aug- 26 Aug 2024, **Scopus indexed,** Status: **Published.**
2. **Dayang Suhaida Awang Damit,** Siti Noraini Sulaiman, Muhammad Khusairi Osman, Noor Khairiah A. Karim, Noor Fadzilah Razali, Mohd Ikmal Fitri Marzuki "Navigating Tversky Loss Function Hyperparameter Spaces using Particle Swarm Optimization for Myocardial Scar Segmentation", 20th IEEE International

Colloquium on Signal Processing & Its Application (CSPA 2024), 1 -2 March 2024, **Scopus indexed,** Status: **Published.**

3. **Dayang Suhaida Awang Damit,** Siti Noraini Sulaiman, Muhammad Khusairi Osman, Noor Khairiah A. Karim, Mohd Firdaus Abdullah, Belinda Chong Chiew Meng, "Optimizing Network Classification Performance by Geometric Transformations on Delayed Enhancement Cardiac Magnetic Resonance Imaging", in 2023 13th IEEE International Conference on Control System, Computing and Engineering (ICCSCE), 25 Aug- 26 Aug 2023, **Scopus indexed,** Status: **Published.**

Book Chapter

1. **Dayang Suhaida Awang Damit,** Siti Noraini Sulaiman, Muhammad Khusairi Osman, Noor Khairiah A. Karim, Belinda Chong Chiew Meng, "Performance Improvement with Optimization Algorithm in Isolating Left Ventricle and Non-Left Ventricle Cardiac" for Intelligent Multimedia Signal Processing for Smart Ecosystems: Springer Book Chapter. **Scopus indexed,** Status: **Published (October 2023).**

Copyright/Patent

1. Automated Classification of Left Ventricle and Non-Left Ventricle Segment in Late Gadolinium Enhancement Cardiac MRI using Deep Convolutional Neural Network. - No. LY2022P05193 (2022), Intellectual Property Corporation of Malaysia (MylPO).
2. DLT-LV: A DeeplabV3+-based Tailored Model For Automated Left Ventricle Region Delineation. - No. LY2023P06250 (2023), Intellectual Property Corporation of Malaysia (My IPO).
3. 3SR-DCNN: Scar Sequential Slice Reconstruction with DCNN Classification Method for Automated Myocardial Infarction Detection. - No. LY2025P02797 (2025), Intellectual Property Corporation of Malaysia (MylPO).

Awards

1. **Dayang Suhaida Awang Damit,** Siti Noraini Sulaiman, Muhammad Khusain Osman, Noor Khairiah A. Karim, Noor Fadzilah Razali, Mohd Ikmal Fitri Marzuki, "Navigating Tversky Loss Function Hyperparameter Spaces using Particle Swarm Optimization for Myocardial Scar Segmentation," in 20th IEEE International Colloquium on Signal Processing & Its Application (CSPA 2024), 1 -2 March 2024, **Best Paper Award.**

2. **Dayang Suhaida Awang Damit,** Siti Noraini Sulaiman, Muhammad Khusairi Osman, Noor Khairiah A. Karim, Samsul Setumin, "AI-Driven Left Ventricle Detection in LGE-Cardiac MRI Via Lightweight DCNN Architecture", in Negeri Sembilan International Exposition (NSIEx2024), 10 Oct 2024, **Gold Medal.**

3. **Dayang Suhaida Awang Damit,** Siti Noraini Sulaiman, Muhammad Khusairi Osman, Noor Khairiah A. Karim, Samsul Setumin, "Left Ventricle Segment Segregation Technique Using Deep Convolutional Neural Network for Cardiac MRI Images", in Penang International Invention, Innovation and Design (PIID2023), 16-18 May 2023, **Silver Medal.**

4. **Dayang Suhaida Awang Damit,** Siti Noraini Sulaiman, Muhammad Khusairi Osman, Noor Khairiah A. Karim, Mohd Firdaus Abdullah, Syafiqah Aqilah Saifudin, "LV Quest: An Automated Model for Classify the Left Ventricle Chamber," in SIPJM Invention, Innovation & Technology Expo (SI2TE 2023), 10-12 July 2023, **Bronze Medal**