



Article

Automatic Left Ventricle Segmentation from Short-Axis Cardiac MRI Images Based on Fully Convolutional Neural Network

Zakarya Farea Shaaf^{1,*}, Muhammad Mahadi Abdul Jamil^{1,*}, Radzi Ambar¹ , Ahmed Abdu Alattab^{2,3}, Anwar Ali Yahya^{3,4} and Yousef Asiri⁴ 

¹ Faculty of Electrical and Electronic Engineering, Universiti Tun Hussein Onn Malaysia, Parit Raja, Batu Pahat 86400, Johor, Malaysia; aradzi@uthm.edu.my

² Department of Computer Science, College of Science and Arts, Sharurah, Najran University, Najran 61441, Saudi Arabia; aalattab@nu.edu.sa

³ Department of Computer Science, Faculty of Computer Science and Information Systems, Thamar University, Dhamar 87246, Yemen; aaesmail@nu.edu.sa

⁴ Department of Computer Science, College of Computer Science and Information Systems, Najran University, Najran 61441, Saudi Arabia; yasiri@nu.edu.sa

* Correspondence: zakaryashaaf@gmail.com (Z.F.S.); mahadi@uthm.edu.my (M.M.A.J.)

Abstract: Background: Left ventricle (LV) segmentation using a cardiac magnetic resonance imaging (MRI) dataset is critical for evaluating global and regional cardiac functions and diagnosing cardiovascular diseases. LV clinical metrics such as LV volume, LV mass and ejection fraction (EF) are frequently extracted based on the LV segmentation from short-axis MRI images. Manual segmentation to assess such functions is tedious and time-consuming for medical experts to diagnose cardiac pathologies. Therefore, a fully automated LV segmentation technique is required to assist medical experts in working more efficiently. Method: This paper proposes a fully convolutional network (FCN) architecture for automatic LV segmentation from short-axis MRI images. Several experiments were conducted in the training phase to compare the performance of the network and the U-Net model with various hyper-parameters, including optimization algorithms, epochs, learning rate, and mini-batch size. In addition, a class weighting method was introduced to avoid having a high imbalance of pixels in the classes of image's labels since the number of background pixels was significantly higher than the number of LV and myocardium pixels. Furthermore, effective image conversion with pixel normalization was applied to obtain exact features representing target organs (LV and myocardium). The segmentation models were trained and tested on a public dataset, namely the evaluation of myocardial infarction from the delayed-enhancement cardiac MRI (EMIDEC) dataset. Results: The dice metric, Jaccard index, sensitivity, and specificity were used to evaluate the network's performance, with values of 0.93, 0.87, 0.98, and 0.94, respectively. Based on the experimental results, the proposed network outperforms the standard U-Net model and is an advanced fully automated method in terms of segmentation performance. Conclusion: This proposed method is applicable in clinical practice for doctors to diagnose cardiac diseases from short-axis MRI images.

Keywords: left ventricle segmentation; cardiac short-axis MRI; fully convolutional network; pixel weights balancing; medical image processing



Citation: Shaaf, Z.F.; Jamil, M.M.A.; Ambar, R.; Alattab, A.A.; Yahya, A.A.; Asiri, Y. Automatic Left Ventricle Segmentation from Short-Axis Cardiac MRI Images Based on Fully Convolutional Neural Network. *Diagnostics* **2022**, *12*, 414. <https://doi.org/10.3390/diagnostics12020414>

Academic Editor: Laura Burattini

Received: 8 December 2021

Accepted: 16 January 2022

Published: 5 February 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Cardiovascular disease is regarded as one of the most severe threats to human health, and it has contributed to an increase in the global mortality rate. According to the World Health Organization, 17.9 million people died from cardiovascular disease in 2016, accounting for 31% of worldwide deaths [1]. As a result, there is a growing emphasis on research and technologies that can effectively improve the diagnosis of cardiovascular diseases while also lowering the mortality rate caused by those diseases. In recent years, the diagnosis of cardiovascular diseases has become more accessible thanks to advancements in medical

imaging techniques such as computed tomography (CT) and cardiac magnetic resonance imaging (CMRI).

MRI is one of the most regularly utilized medical imaging modalities for diagnosing cardiovascular disease because it is non-invasive and produces high-resolution images. Segmentation of cardiac short-axis MRI is critical for quantifying cardiac function by analyzing clinical metrics such as ventricular volumes, stroke volumes, and myocardium in the early detection of cardiovascular diseases. Segmentation of the LV is vital for accurate assessments of cardiac function indicators such as ejection fraction, LV volume, and LV mass, all of which are important in diagnosing cardiovascular diseases [2–4]. To better understand the LV segmentation task, Figure 1 shows short-axis (SAX) images of the LV at the basal, middle, and apical slices along with their corresponding ground truth (labels). The primary goal of LV segmentation is to delineate the LV's contours (epicardium and endocardium). On the other hand, manual LV segmentation is a time-consuming and error-prone task for medical experts. Therefore, a fully automated LV segmentation method from the short axis is urgently needed.

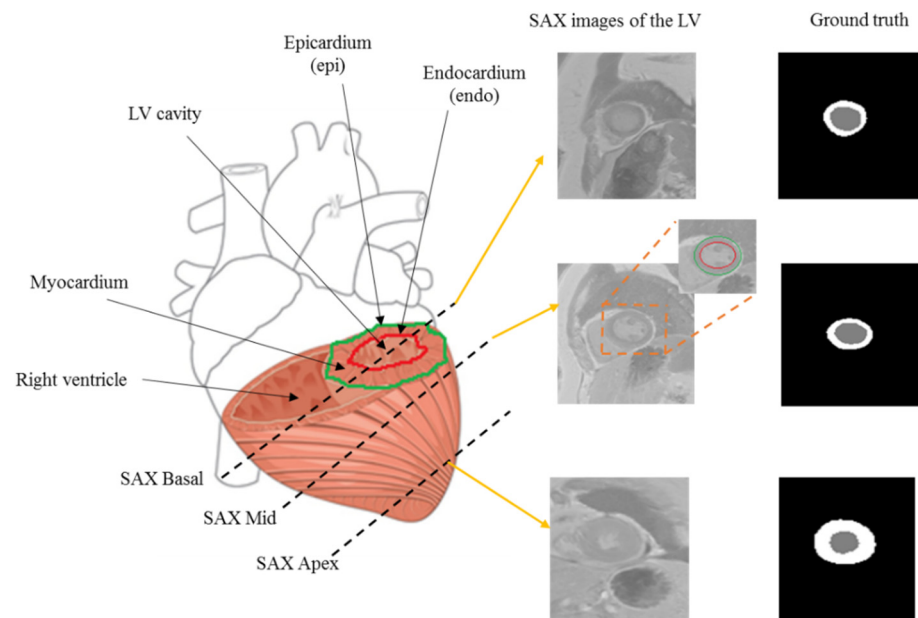


Figure 1. LV short-axis MRI and corresponding ground truth (labels).

For LV segmentation, a variety of techniques have been proposed. Active contour, level set, and graph cut are examples of model-based approaches. These models, on the other hand, are semi-automatic and rely heavily on a successful initialization step. Deep learning-based algorithms have become frequently used in medical image segmentation due to rapid advancements in computer hardware and the availability of massive training data. Convolutional neural networks (CNNs), a standard deep learning-based method, have recently achieved excellent results in various computer vision fields, including object detection [5], image classification [6], and image segmentation [7]. Following this trend, several CNN-based techniques for LV segmentation have been proposed [8–15] and have shown promising results in clinical practice. However, accurate segmentation of the LV and myocardium from cardiac MRI remains a challenge in clinical practice for several reasons, including changes in the LV morphology across slices, an imbalance in pixels between the LV area and the background, and incorrect pixel representation for the target area. Furthermore, Xiong et al. [16] complained that deep learning methods are data-driven and need a massive amount of data for training, and the available labeled dataset of the LV is small. Thus, a small labeled data results in poorer performance when utilizing deep learning approaches.

The ability to develop and compare the performances of FCN models is based on several conditions, including input data normalization, CNN layer selection, pixel balancing in input image labels, and fine-tuning the model's training options. As a result, the goal of this paper was to design an FCN-based segmentation model for the LV from short-axis MR images, which includes the following contributions:

1. Compare various optimization algorithms and select the most reliable one to train the proposed model;
2. Class weighting method to avoid high imbalance of pixels between object and background classes in image's labels;
3. Pixel normalization of labels to allow the model to learn and extract features from input images accurately;
4. Achieve state-of-the-art results for automatic LV segmentation.

The rest of this paper is organized as follows: Section 2 describes the related works, the materials and methods are introduced in Section 3, and the experimental results and discussion are presented in Section 4, followed by a conclusion in Section 5.

2. Related Works

In recent years, segmentation and quantification of the LV from cardiac MRI images have received much attention to diagnose cardiovascular disease. Many studies have proposed semi-automatic segmentation methods to delineate the LV borders, such as active contour [17,18], level set [19–21], graph cut [22], dynamic programming, and atlas-based models. These traditional segmentation methods necessitate user intervention, which is a time-consuming and tedious task. The difference between semi-automatic and fully automatic segmentation is that the latter is better suited to process large batches of cardiac MRI images.

For segmenting the LV and myocardium from CMR images, CNNs in various orders have been proposed. Dangi et al. [23] created a CNN-based multi-task learning (MTL) model for simultaneous LV segmentation and quantification. They used the U-net architecture [24], separating segmentation and regression at the final upsampling layer. This network is capable of learning feature representation while also improving generalization. Moradi et al. [25] developed a deep-learning-based method called MFP-U-net for LV segmentation from echocardiography images, and they designed a network with a feature pyramid that can detect and recognize the LV in MRI. Wu et al. [26] proposed an automatic segmentation model for the LV from cardiac MRI. They used a CNN model to locate the LV and the U-net model to segment it. Abdeltawab et al. [10] devised a framework that begins with FCN-based localization of the LV and extraction of the heart section's ROI. The extracted ROIs are then fed into the FCN2 network, which segments the LV cavity and myocardium. Dong et al. [27] proposed a CNN-based model with two parallel subnetworks to detect endocardium and epicardium contours of the LV, incorporating the MTL concept. The FCN [28] is a CNN expansion with different last layers used for different tasks. Traditional CNN methods, for example, use fully connected layers for image classification to predict objects, whereas an FCN applies a deconvolution (transposed) layer instead of a fully connected layer in semantic segmentation. Several FCN-based models have been used to improve LV segmentation performance [29–31]. The network proposed by Cui et al. [32] was an attention U-Net model based on an FCN structure for cardiac short-axis MRI segmentation. U-Net [24] has been commonly applied in medical image segmentation, particularly in the segmentation of cardiac images [25,33,34].

Some researchers used a hybrid model that combined deep learning methods with traditional models to achieve an optimal LV segmentation performance from short-axis cardiac MRI images. For example, Ngo et al. [35] used a deep learning model combined with a level set for automatic LV segmentation. Avendi et al. [36] developed a fully automatic segmentation model for the LV using deep learning algorithms and deformable models. Due to the strong correlation between sequential frames during the cardiac cycle, a 3D model with a recurrent neural network (RNN) has been proposed. Long

short-term memory (LSTM) is a popular RNN [37] technique for detecting heart motion using spatiotemporal dynamics. Zhang et al. [38] created a multi-level LSTM model for LV segmentation that used low-resolution level features to train one model and high-resolution level features to train another. Additionally, due to the large slice thickness, Baumgartner et al. [39] found that segmentation by 2D CNN performed better than 3D CNN. Furthermore, due to significant morphological differences in LV shape across slices caused by heart movement, RNN models reproduce incorrect features and require high computational costs. Bernard et al. [40] conducted a benchmark study and discovered that FCNs are used in most advanced algorithms for LV segmentation from short-axis MRI images.

In recent years, researchers have been paying more attention to the segmentation of LV boundaries (endo- and epicardium) from short-axis MRI images. Table 1 summarizes the most recent studies in LV segmentation from short-axis MRI using deep learning models. Furthermore, the LV segmentation challenges [40–42] and benchmark datasets with ground truth contours are provided. Deep learning methods have lately obtained excellent results in the segmentation of medical images. CNN is one of the most widely used methods in medical image analysis [23,43] among these approaches. Medical images are segmented at the pixel level, as opposed to image-level classification [27]. Traditional CNN methods must be improved in order to achieve robust semantic segmentation. Furthermore, according to recent research, image pixel class imbalance can affect CNN performance during classification and segmentation [44]. Buda et al. [45] provided a thorough analysis of the CNN class imbalance problem. Data-level methods and classifier methods are two types of solutions to this problem. Oversampling [46] and data augmentation [47] are data-level methods that work with training datasets, whereas classifier-level methods such as cost-sensitive learning [48], hard mining [49], and loss function work with model training options.

Table 1. Current studies in LV segmentation from cardiac MRI using deep learning algorithms.

Author/Year	Dataset	Subjects No.	Data Preparation	Deep Learning Model
Cui et al. [32]/2021	LVSC	200	- Cropping using multi-scale methods - Pixel normalization	Attention U-net architecture
Tan et al. [50]/2017	LVSC	200	- Resampling pixels using linear interpolation - Cropping - Pixel normalization - Augmentation (during training)	CNR
Tran et al. [51]/2016	SCD, LVSC and RVSC	45, 200 and 48	- Cropping using multi-resolution approach - Augmentation	FCN
Khend et al. [8]/2019	ACDC-2017, LVSC and Kaggle	150, 200 and 500	- Cropping - Augmentation	FCN DenseNet
Wang et al. [15]/2020	CAP	450	- Cropping - Augmentation - Pixel normalization	FCN
Wu et al. [26]/2020	SCD	45	- Image filtering - Cropping/downsampling	CNN + U-Net
Wu et al. [52]/2021	SCD	45	- Augmentation	GAN
Dong et al. [27]/2020	MICCAI 2019	56	- Pixel normalization	Parallel CNNs
Du et al. [53]/2019	2900 collected images	156	- Cropping - Normalization	Multi-task CNR + RNN
Abdeltawab et al. [10]/2020	ACDC-2017	150	- Cropping	Two FCNs

Pixel imbalance between the target class and the background class has a significant effect on segmentation performance, which requires an effective solution. Hence, various methods have been proposed to deal with this issue; for example, the focal Tversky loss function (FTL) was introduced by Cui et al. [32], Dong et al. [27] applied cross-entropy loss function instead of the dice loss function, and Wang et al. [15] used dynamic pixel-wise (PW) weighting. In addition, the authors normalized the pixel intensity of the input images to improve the learning ability of the models. Cui et al. [32] used mean–variance normalization (MVN) to normalize the pixel intensity on an input image by subtracting the difference

from its average value and dividing by its standard deviation, and Wang et al. [15] used min-max normalization. Based on the above literature, in this study we created a 2D FCN technique with fewer parameters for accurately segmenting the LV and myocardium from short-axis MRI images. After using appropriate normalization and conversion techniques, the input images were used to extract pixels. The 2D PNG images have some advantages compared with NIfTI images, such as flexible image visualization, augmentation (rotation, cropping, and rescaling), and efficient exclusion of unwanted images.

3. Materials and Methods

3.1. Task Description

The procedures of the proposed system for LV segmentation are shown in Figure 2. The steps of the system are as follows: (i) preparation of the MRI images, including resizing and pixel normalization; (ii) training the FCN model with a comparison between three optimization algorithms, namely stochastic gradient descent (SGD), adaptive moment estimation (Adam), and root mean square propagation (RMSProp); and (iii) testing the trained model for extraction of ROI features and segmentation to delineate the LV contours.

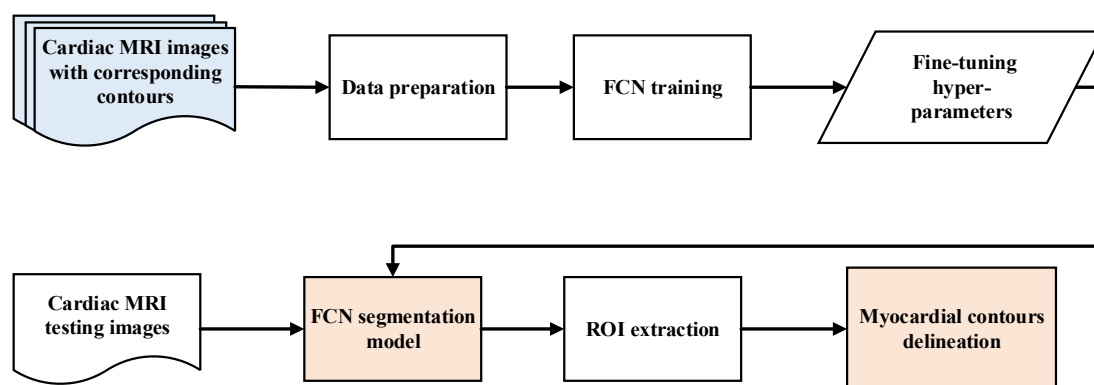


Figure 2. Diagram of the proposed model.

3.2. Data Description

The dataset for this study was acquired at the University Hospital of Dijon in France and was provided from the automatic evaluation of myocardial infarction from delayed-enhancement cardiac MRI (EMIDEC) [42] during the MICCAI conference 2020. This dataset contains sequences of short-axis MRI images with ground truth for 150 patients (100 for training and 50 for testing). Each case has a text file with clinical information, a neuroimaging informatics technology initiative (NIfTI) file with the short-axis images of the LV, and an NIfTI file with the labeled masks. The masks consist of four different pixels for each area, which are 0, 1, 2, 3, and 4, representing the background, LV cavity, normal myocardium (NM), myocardial infarction (MI), and no-reflow (NREFLOW), respectively, as shown in Figure 3. The dataset can be downloaded from the website (<http://emidec.com/> (accessed on 1 December 2021)).

3.3. Data Preparation

Medical images are commonly stored in NIfTI or DICOM format after being acquired from medical imaging modalities. Although these formats have high precision for images, they provide volumetric (voxels, height, and depth) data with unequal depth in various slices/series. Furthermore, image preprocessing steps such as augmentation and excluding unwanted images from apical slices are quite tedious with volumetric data (3D). Thus, in this study, the NIfTI data were converted to 2D images (PNG) as inputs to train the proposed 2D FCN.

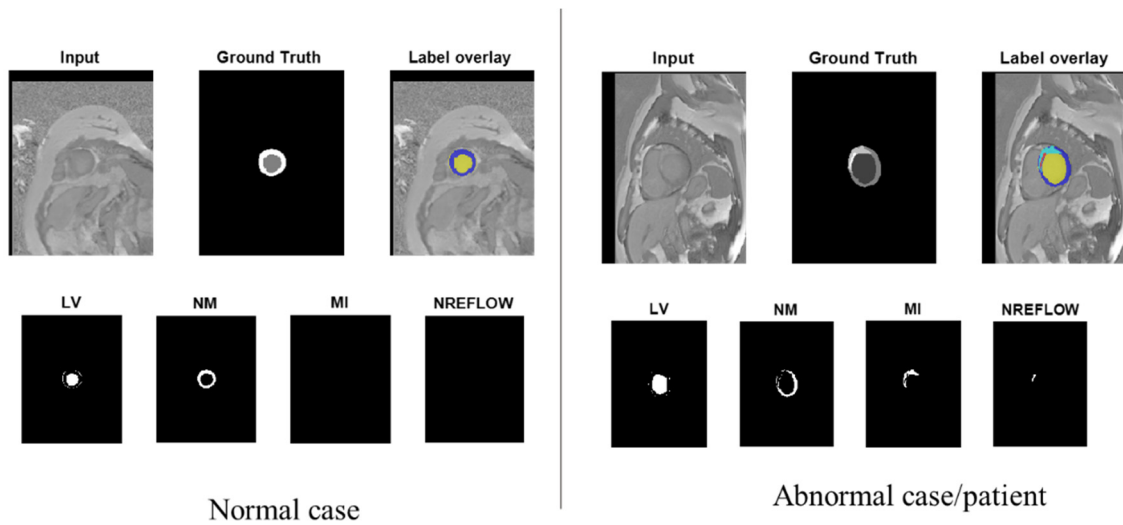


Figure 3. MR images' visualization from EMIDEC dataset.

3.3.1. Data Conversion and Normalization

Data were converted from NifTI to PNG images with extensive consideration of the pixels representation. An open-source toolkit named XMedCon was used for medical image conversion [54]. This platform is a graphical user interface (GUI) that gives immediate visual control on selected options with various features, including simple image processing, volume manipulation, pixel values support, and supporting image formats for all medical modalities. The principle of this toolkit is to preserve data and assure that the default output represents the pixel data as retrieved from the original study. This initial step has achieved a more desirable performance for adequate pixel representation than direct conversion by programming code in MATLAB or Python.

The pixels for the LV and myocardium after image conversion and pixel normalization are depicted in Figure 4, showing the robustness of the conversion step in the right image, whereas the left image represents image pixels using the usual conversion method. The size of all input images was 256×192 pixels, with normalized pixel intensities from [1, 2] to [128, 255] for the LV and myocardium, respectively. The pixel normalization (N_P) was applied using the following equation:

$$N_P = 255 \times mat2gray(N_{ori}) \tag{1}$$

where N_{ori} represents the pixels matrix of the original image.

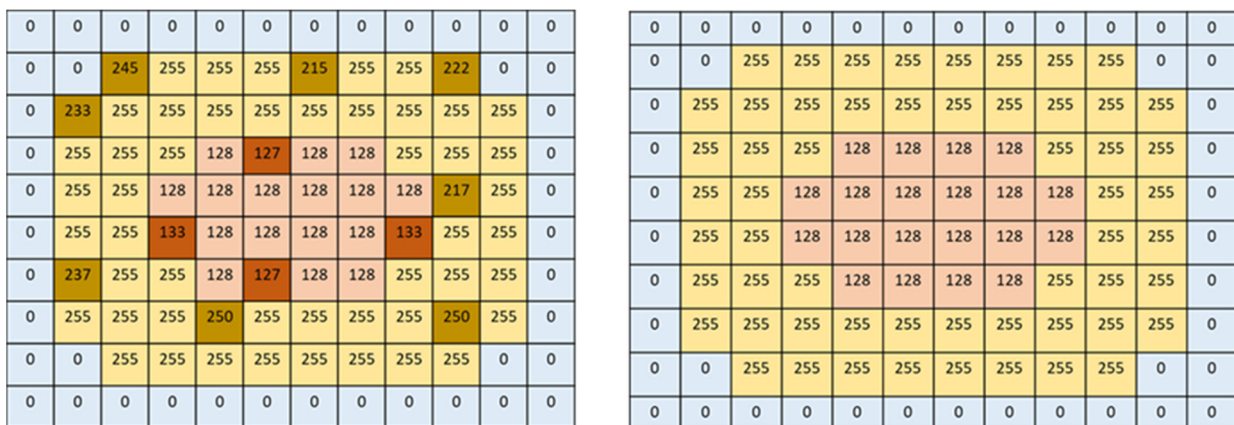


Figure 4. Pixels of image labeled by usual conversion (left) and conversion by XMedCon (right).

3.3.2. Balancing of Class Weight Pixels

Most pixels in the ground truth (labels) are for background, leading to class imbalance. During the learning process, network biases to learning the dominant class (background) result in weak segmentation performance. The balancing of pixels for three classes, including background (BG), LV, and myocardium (Myo), is presented. The class weighting method was used to compute class weights, namely inverse frequency weighting, where the weights of classes are the inverse of the class frequencies.

3.4. Network Architecture

The first step in creating a new FCN is to define and select suitable layers. Figure 5 depicts the proposed FCN architecture with input and output images. This network takes the principle of U-Net architecture, which has encoder (contraction/downsampling) and decoder (expansion/upsampling) paths, as shown in Figure 6. The contraction path extracts local features and restores feature maps in the expansion path of the network. The network is designed to train a few samples. The layers of the downsampling path are 3×3 convolutions, batch normalization, and a rectified linear unit (ReLU) as an activation function with padding to keep the output size of the convolution layer the same as its input. Then, there is downsampling followed by max-pooling operation with a size of 2×2 and stride of 2 to reduce the input size. The 4×4 transpose convolution (deconvolution) and convolution layers are applied in the upsampling path followed by a pixel classification layer with a softmax layer to predict the output image. The cross-entropy term is used as a loss function in this network.

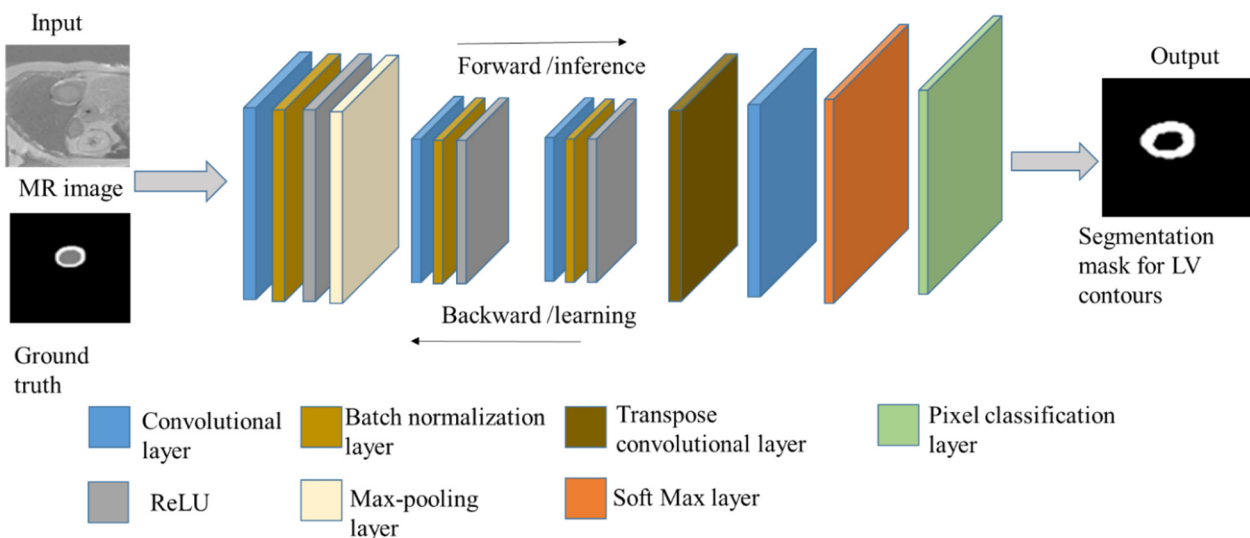


Figure 5. Architecture layers of the proposed FCN.

After defining the layers of the FCN model, the next step is setting up the training options for the network by specifying some parameters such as the solver, the maximum number of epochs, and the learning rate. Solvers such as SGD, Adam, and RMSProp update the network parameters using a subset of data at each step called a mini-batch to minimize the loss function. The parameter updating is named an iteration while the epoch passes through the entire data during network training. The learning rate is a crucial parameter for network training that can shorten the training time and minimize the loss in training progress.

During training, the network performs a forward pass, where each layer takes the output from the previous layer as the input and then outputs the results to the next layer, and a backward pass, where each layer takes the derivative of the loss concerning the layer’s outputs and computes it to the inputs to propagate the results. At the end of the forward pass, the network output layer calculates the loss L between the target T and the prediction

Y. Table 2 illustrates the layer types of the FCN model with their kernel sizes and learnable parameters, such as bias and weights. The advantage of the proposed FCN is that its training time is faster and it requires less memory space than U-Net, which requires much time for training and consists of many parameters that need a high computational cost.

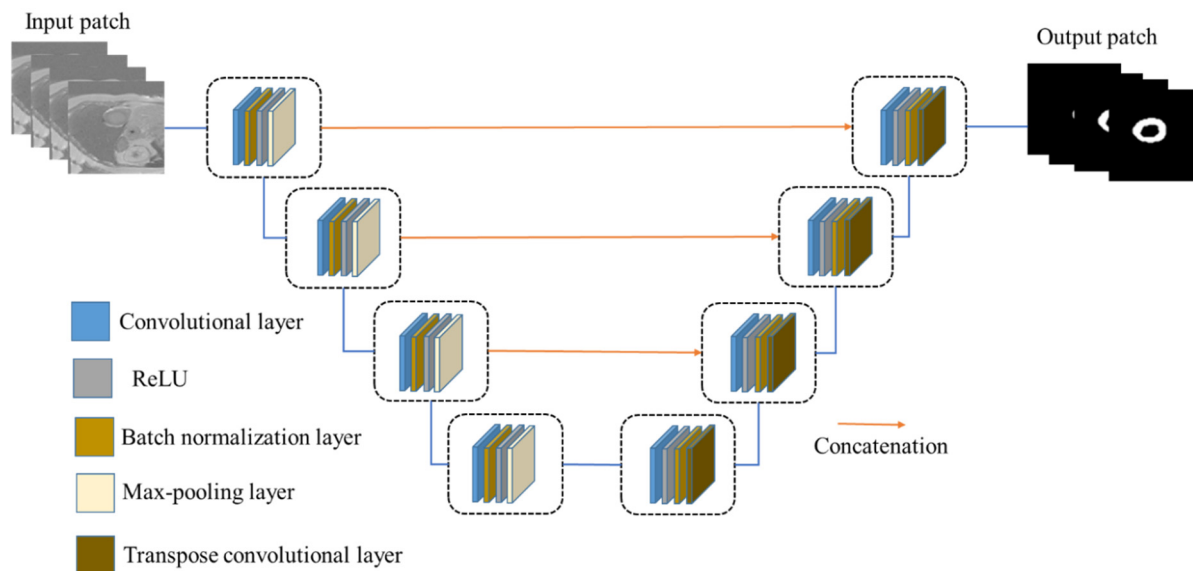


Figure 6. Schematic architecture of the U-Net model.

In this work, the proposed network was trained with three different optimization algorithms to select one after comparing their performances and efficiency. Moreover, the hyper-parameters of the training options, such as epochs, mini-batch size, and learning rate, were fine-tuned through sequent experiments to select desirable parameters for the proposed network. Based on the experiments, the optimal mini-batch sizes for normalized images were 4 and 8 due to the data size and to lessen the memory space. Thus, the selection of mini-batch size was decided, which evaluated the gradient of the loss function and updated the weights significantly, resulting in a considerable performance of the network. The performance evaluation of this work was determined using metrics such as sensitivity, specificity, negative predictive value (NPV) and positive predictive value (PPV), Jaccard index, and dice score coefficient (DSC).

Table 2. Analyzing layers of the proposed FCN.

Layer Number	Layer Type	Kernel Size	Learnable
1	Image input	$256 \times 192 \times 1$	-
2	Convolution	$256 \times 192 \times 16$	Weights $3 \times 3 \times 1 \times 16$ Bias $1 \times 1 \times 16$
3	Batch normalization	$256 \times 192 \times 16$	Offset $1 \times 1 \times 16$ Scale $1 \times 1 \times 16$
4	ReLU	$256 \times 192 \times 16$	-
5	Max pooling	$128 \times 96 \times 16$	-
6	Convolution	$128 \times 96 \times 32$	Weights $3 \times 3 \times 16 \times 32$ Bias $1 \times 1 \times 32$
7	Batch normalization	$128 \times 96 \times 32$	Offset $1 \times 1 \times 32$ Scale $1 \times 1 \times 32$
8	ReLU	$128 \times 96 \times 32$	-
9	Convolution	$128 \times 96 \times 64$	Weights $3 \times 3 \times 32 \times 64$ Bias $1 \times 1 \times 64$

Table 2. *Cont.*

Layer Number	Layer Type	Kernel Size	Learnable
10	Batch normalization	$128 \times 96 \times 64$	Offset $1 \times 1 \times 64$ Scale $1 \times 1 \times 64$
11	ReLU	$128 \times 96 \times 64$	-
12	Transpose convolutional layer	$128 \times 96 \times 16$	Weights $4 \times 4 \times 16 \times 64$ Bias $1 \times 1 \times 16$
13	Convolution	$256 \times 192 \times 2$	Offset $3 \times 3 \times 16 \times 2$ Scale $1 \times 1 \times 2$
14	Softmax	$256 \times 192 \times 2$	-
15	Pixel classification layer	-	-

4. Results and Discussion

The proposed method was implemented using MATLAB software (version r2020b) with an Intel (R) i7-3770 central processing unit (CPU), 20 GB DDR3 random access memory (RAM), and Nvidia GeForce GTX 1050 Ti. The initial experiments were conducted to select the hyper-parameters used for the training of the proposed network. After that, the network was trained to segment the LV from MRI images based on the proposed conversion method for the images.

4.1. Hyper-Parameters Selection

These experiments aimed to select hyper-parameters for the network training, including optimization algorithms (SGDM, Adam, and RMSProp), learning rate, epochs, and mini-batch size. The algorithms' performance was compared using learning rates of 0.01 and 0.001 over 30, 50, 100, and 150 epochs at mini-batch sizes of 4 and 8. As shown in Tables 3 and 4, the Adam algorithm achieved the highest performance using a learning rate of 0.001, 150 epochs, and a mini-batch size of 4. Based on the hyper-parameters selection, the proposed FCN was trained and gained a mini-batch accuracy of 91.18% and mini-batch loss of 0.005, as illustrated in Figure 7. Hence, it is evident that the network's segmentation performance can be improved by using fewer mini-batches and increasing the number of epochs.

Table 3. Optimization algorithms' performance in the trained network at mini-batch size 4.

Learning Rate	Epochs	SGDM %	ADAM %	RMSProp %
0.01	30	74.67	60.18	50.25
	50	76.41	54.39	54.85
	100	82.45	51.34	65.11
	150	76.81	60.69	47.22
0.001	30	70.68	78.91	81.76
	50	74.18	81.27	83.82
	100	77.27	87.14	85.65
	150	76.92	91.18	89.20

Table 4. Optimization algorithms' performance in the trained network at mini-batch size 8.

Learning Rate	Epochs	SGDM %	ADAM %	RMSProp %
0.001	30	60.63	79.85	78.00
	50	69.00	81.17	81.69
	100	70.65	87.60	81.96
	150	60.37	90.42	87.44

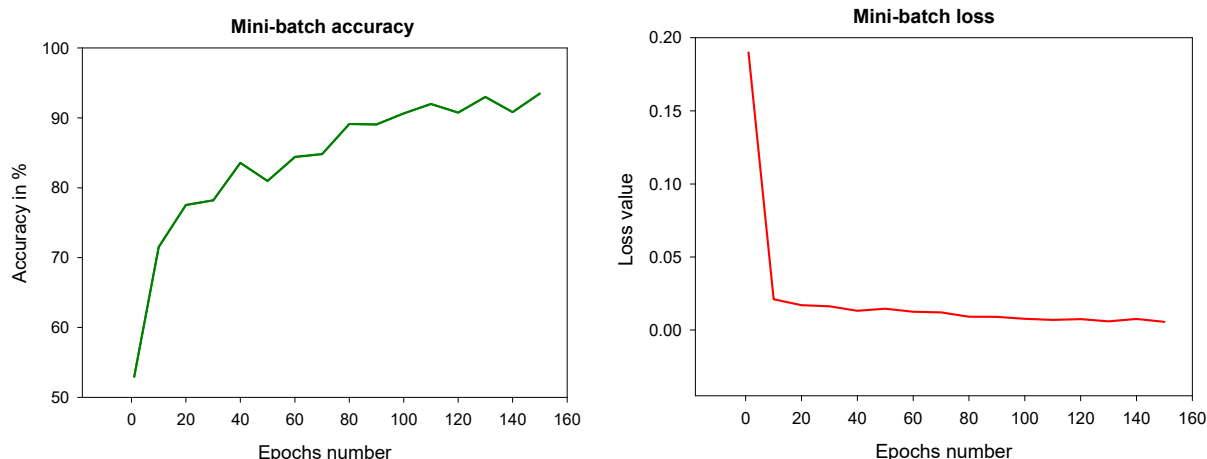


Figure 7. The mini-batch accuracy and mini-batch loss when training the proposed network.

4.2. Network Performance

The image conversion using the XMedCon toolkit and pixel normalization with pixels weight balancing resulted in an improved performance. The comparison of evaluation metrics between the FCN model and U-Net models based on the proposed image conversion and normalization is shown in Table 5. The proposed FCN model outperformed U-Net models in terms of Jaccard index, sensitivity, NPV, and dice similarity with scores of 0.87, 0.98, 0.99, and 0.93, respectively. The U-Net model with the Sgdm solver performed well only in specificity and PPV, obtaining 0.98 for both. The performance of U-Net without data conversion by the XMedCon toolkit was the worst among the four models, with less efficiency in minimizing loss function based on data features. Furthermore, as depicted in Table 6, the proposed model outperformed the other trained models in terms of global and mean accuracies, mean intersection over union (IoU), weighted IoU, and mean boundary F1 (BF) score, with values of 0.95, 0.96, 0.90, 0.91, and 0.89, respectively. The results shown in Tables 5 and 6 prove that models using images converted by the XMedCon toolkit as input perform well. Although, the exact features representation of pixels after conversion, an imbalance of target and background pixel classes was found. Thus, pixel weight balancing of the background, LV, and myocardium (Myo) classes was applied and achieved high balancing of the pixels as shown in Figure 8. The confusion matrices used to refine the trained and proposed models are depicted in Figure 9 with prediction of classes' pixels of the LV and myocardium.

Table 5. Comparison of evaluation metrics between trained models and the proposed FCN model (✓ represents conversion by XMedCon and × represents conversion by coding).

Model	Conversion by XMedCon	Jaccard Index	Sensitivity	Specificity	PPV	NPV	DSC
U-Net + sgdm	✓	0.84	0.86	0.98	0.98	0.89	0.91
U-Net	×	0.60	0.78	0.85	0.72	0.89	0.74
U-Net + adam	✓	0.84	0.93	0.91	0.89	0.94	0.91
The proposed FCN	✓	0.87	0.98	0.94	0.89	0.99	0.93

Table 6. Performance results of the trained models (✓ represents conversion by XMedCon and × represents conversion by coding).

Model	Conversion by XMedCon	Global Accuracy	Mean Accuracy	Mean IoU	Weighted IoU	Mean BF-Score
U-Net + adam	✓	0.93	0.92	0.86	0.86	0.89
U-Net	×	0.83	0.82	0.69	0.71	0.67
U-Net + sgdm	✓	0.92	0.92	0.85	0.85	0.85
Our FCN	✓	0.95	0.96	0.90	0.91	0.89

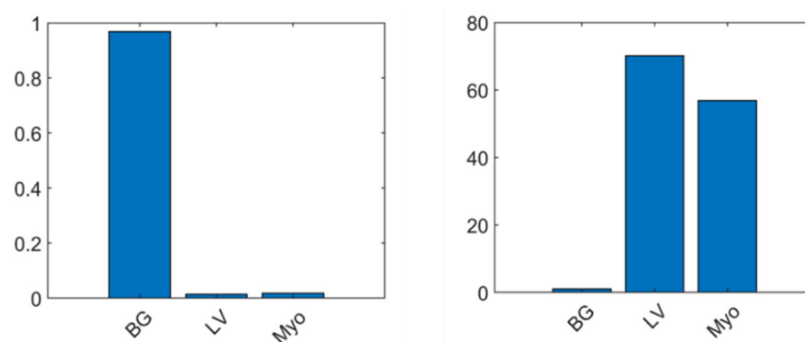


Figure 8. Pixels of classes before balancing (left) and after balancing (right).

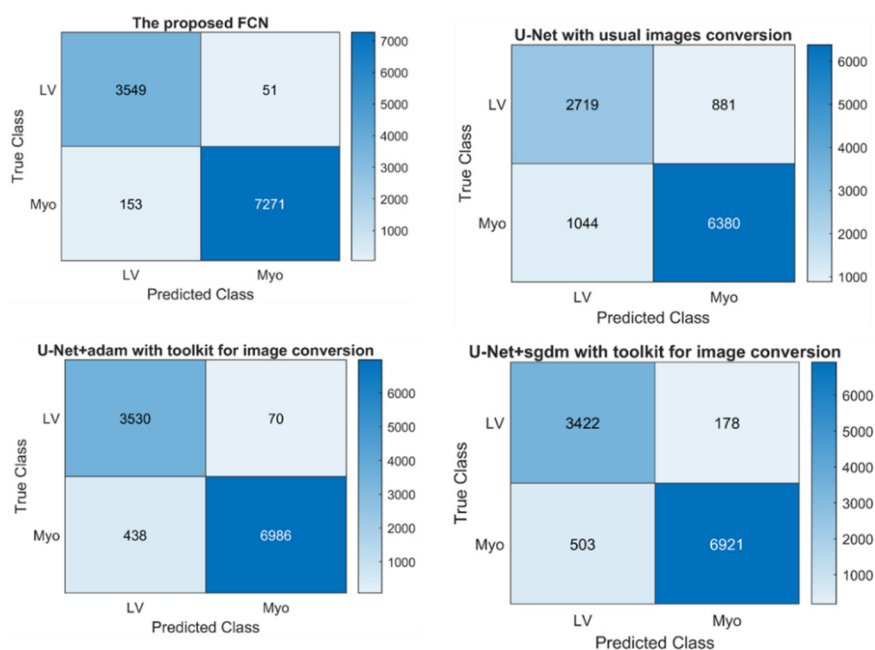


Figure 9. Confusion matrices of the trained models.

Figure 10 shows a visual comparison of the output for LV labels using the selected model (with LR = 0.001, epoch = 150) and other models with the referenced ground truths. The proposed model's segmentation results are very close to the expert-provided LV boundaries. Furthermore, the figure compares the segmentation results obtained from the proposed model with pre-trained U-Net models under various conditions. The proposed FCN model outperformed other methods for delineating LV contours.

4.3. Comparison with Recent Methods

The quantitative comparison of the LV segmentation results between the proposed model and other advanced methods is depicted in Table 7. These methods include the attention U-Net architecture [32], convolutional neural network regression (CNR) method [50], FCN method [51], multi-scale FCN DenseNet [8], and a dynamic pixel-wise weighting-based FCN [15]. The detailed datasets and data preparation steps for these models are presented under the related work sections in Table 1. It can be observed that the proposed method achieved a robust performance compared to other published methods. For most evaluation metrics, including the Jaccard index, sensitivity, PPV, NPV, and DSC, the proposed FCN model outperformed other methods, except for the specificity of the method proposed by Wang et al. [15]. It is important to know that Wang's method involves a dynamic pixel-wise weighting technique to adjust the pixel's weight according to the upper layer's segmentation accuracy and forces the pixel classifier to consider the misclassified ones. The network's performance was based on specific data of a hundred images that

underwent normalization and manipulation to be suited for the trained network. The network was tested for ten normal subjects on the same training dataset and had the most significant advantage of being trained using a small normalized dataset.

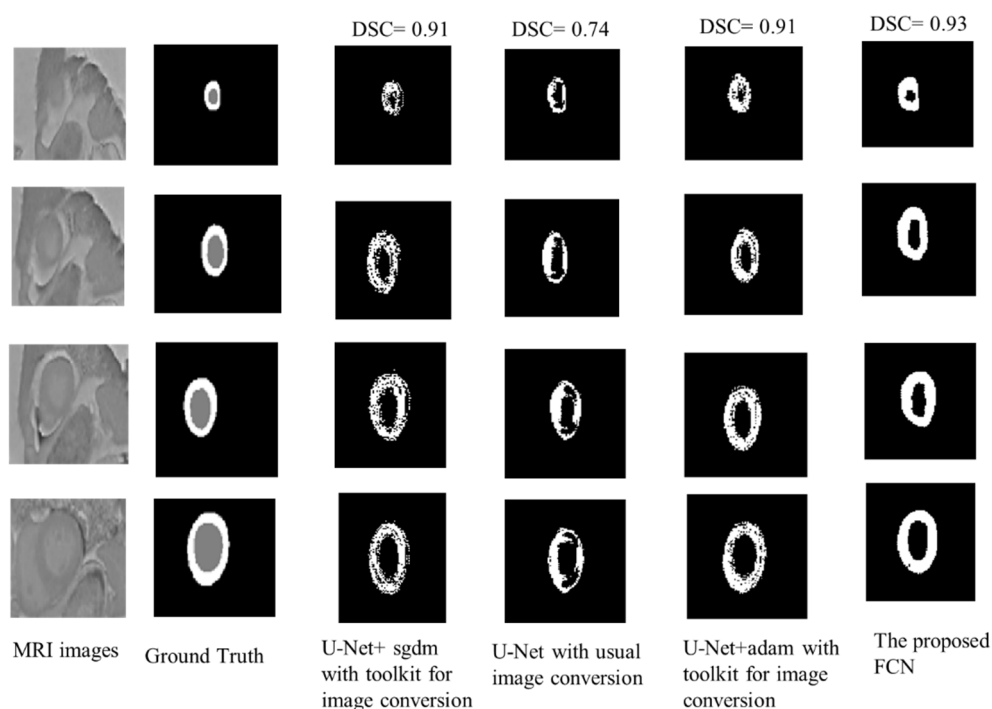


Figure 10. Segmentation results from comparison between U-Net models under various conditions and the proposed FCN.

Table 7. Performance comparison between the proposed model and other state-of-the-art models in automatic LV segmentation.

Method	Jaccard Index	Sensitivity	Specificity	PPV	NPV	Dice
Cui et al. [32]	0.75	0.87	0.92	0.87	0.93	-
Tan et al. [50]	0.77	0.88	0.95	0.86	0.96	-
Tran et al. [51]	0.74	0.83	0.96	0.86	0.95	-
Khend et al. [8]	0.74	0.84	0.96	0.87	0.95	0.84
Wang et al. [15]	0.70	0.90	0.99	0.77	0.99	0.80
The proposed FCN	0.87	0.98	0.94	0.89	0.99	0.93

To the best of our knowledge, LV segmentation is essential to evaluate cardiac function by measuring parameters such as LV volume, LV mass, and ejection fraction. The results show that the performance of the proposed method to delineate the LV contours is very close to the ground truth provided by clinical experts. Thus, on high-contrast images, the proposed network obtains intelligible results allowing doctors to detect cardiac diseases such as myocardial infarction precisely based on automatic LV segmentation.

4.4. Limitation of the Study

The size of the endocardial and epicardial regions from LV segmentation in apical slices was not always accurate compared with basal and middle slices. The main limitation in this study is the number of datasets that need to be enlarged and appropriately configured to train FCN models. In addition, the setting up of significant parameters for network training requires more training data with augmentation.

5. Conclusions

In this paper, an FCN was proposed for LV segmentation from short-axis MRI. The selection of training hyper-parameters, such as optimization algorithm, epoch's number, learning rate, and mini-batch size, was based on multiple experiments training various model structures. The input images used for the model were initially converted using a toolkit that keeps the feature representation of pixels the same as the original data. Data normalization in this study performed well and allowed the network to learn feature extraction accurately. In addition, pixel weighting was introduced to avoid an imbalance in target class and background class pixels. Overall, the proposed network has achieved a robust performance in terms of Jaccard index, dice metric, sensitivity, specificity, PPV, and NPV, which is a significant step towards reducing manual segmentation by clinical experts during the diagnosis of cardiac diseases.

Author Contributions: Conceptualization, Z.F.S., M.M.A.J. and R.A.; methodology, Z.F.S., M.M.A.J. and R.A.; software, Z.F.S.; validation, Z.F.S., M.M.A.J. and R.A.; formal analysis, Z.F.S., A.A.A., A.A.Y. and Y.A.; investigation, Z.F.S., A.A.A., A.A.Y. and Y.A.; resources, Z.F.S., M.M.A.J. and R.A.; data curation, Z.F.S.; writing—original draft preparation, Z.F.S., M.M.A.J. and R.A.; writing—review and editing, Z.F.S., A.A.A., A.A.Y. and Y.A.; visualization, Z.F.S., M.M.A.J., R.A., A.A.A., A.A.Y. and Y.A.; supervision, M.M.A.J. and R.A.; project administration, Z.F.S., M.M.A.J. and R.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the Ministry of Higher Education (MOHE) through the Fundamental Research Grant Scheme (FRGS) (FRGS/1/2020/TK0/UTHM/02/16) and the Universiti Tun Hussein Onn Malaysia (UTHM) through an FRGS Research Grant (Vot K304).

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Mozaffarian, D.; Benjamin, E.J.; Go, A.S.; Arnett, D.K.; Blaha, M.J.; Cushman, M.; de Ferranti, S.; Després, J.-P.; Fullerton, H.J.; Howard, V.J.; et al. Heart Disease and Stroke Statistics—2015 Update. *Circulation* **2015**, *131*, e29–e322. [[CrossRef](#)] [[PubMed](#)]
2. Zhang, N.; Yang, G.; Gao, Z.; Xu, C.; Zhang, Y.; Shi, R.; Keegan, J.; Xu, L.; Zhang, H.; Fan, Z.; et al. Deep learning for diagnosis of chronic myocardial infarction on nonenhanced cardiac cine MRI. *Radiology* **2019**, *291*, 606–607. [[CrossRef](#)] [[PubMed](#)]
3. Chen, M.; Fang, L.; Zhuang, Q.; Liu, H. Deep Learning Assessment of Myocardial Infarction from MR Image Sequences. *IEEE Access* **2019**, *7*, 5438–5446. [[CrossRef](#)]
4. Suinesiaputra, A.; Dhooge, J.; Duchateau, N.; Ehrhardt, J.; Frangi, A.F.; Gooya, A.; Grau, V.; Lekadir, K.; Lu, A.; Mukhopadhyay, A.; et al. Statistical Shape Modeling of the Left Ventricle: Myocardial Infarct Classification Challenge. *IEEE J. Biomed. Health Inform.* **2018**, *22*, 503–515. [[CrossRef](#)]
5. Ren, S.; He, K.; Girshick, R.; Sun, J. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *39*, 1137–1149. [[CrossRef](#)]
6. Baloglu, U.B.; Talo, M.; Yildirim, O.; Tan, R.S.; Acharya, U.R. Classification of myocardial infarction with multi-lead ECG signals and deep CNN. *Pattern Recognit. Lett.* **2019**, *122*, 23–30. [[CrossRef](#)]
7. Rizwan, I.; Haque, I.; Neubert, J. Deep learning approaches to biomedical image segmentation. *Inform. Med. Unlocked* **2020**, *18*, 100297. [[CrossRef](#)]
8. Khened, M.; Kollerathu, V.A.; Krishnamurthi, G. Fully convolutional multi-scale residual DenseNets for cardiac segmentation and automated cardiac diagnosis using ensemble of classifiers. *Med. Image Anal.* **2019**, *51*, 21–45. [[CrossRef](#)]
9. Hu, H.; Pan, N.; Wang, J.; Yin, T.; Ye, R. Automatic segmentation of left ventricle from cardiac MRI via deep learning and region constrained dynamic programming. *Neurocomputing* **2019**, *347*, 139–148. [[CrossRef](#)]
10. Abdeltawab, H.; Khalifa, F.; Taher, F.; Alghamdi, N.S.; Ghazal, M.; Beache, G.; Mohamed, T.; Keynton, R.; El-Baz, A. A deep learning-based approach for automatic segmentation and quantification of the left ventricle from cardiac cine MR images. *Comput. Med. Imaging Graph.* **2020**, *81*, 101717. [[CrossRef](#)]
11. Wu, J.; Gan, Z.; Guo, W.; Yang, X.; Lin, A. A fully convolutional network feature descriptor: Application to left ventricle motion estimation based on graph matching in short-axis MRI. *Neurocomputing* **2020**, *392*, 196–208. [[CrossRef](#)]
12. Budai, A.; Suhai, F.I.; Csorba, K.; Toth, A.; Szabo, L.; Vago, H.; Merkely, B. Fully automatic segmentation of right and left ventricle on short-axis cardiac MRI images. *Comput. Med. Imaging Graph.* **2020**, *85*, 101786. [[CrossRef](#)] [[PubMed](#)]
13. Wang, X.; Zhai, S.; Niu, Y. Left ventricle landmark localization and identification in cardiac MRI by deep metric learning-assisted CNN regression. *Neurocomputing* **2020**, *399*, 153–170. [[CrossRef](#)]

14. Penso, M.; Moccia, S.; Scafuri, S.; Muscogiuri, G.; Pontone, G.; Pepi, M.; Caiani, E.G. Automated left and right ventricular chamber segmentation in cardiac magnetic resonance images using dense fully convolutional neural network. *Comput. Methods Programs Biomed.* **2021**, *204*, 106059. [[CrossRef](#)] [[PubMed](#)]
15. Wang, Z.; Xie, L.; Qi, J. Dynamic pixel-wise weighting-based fully convolutional neural networks for left ventricle segmentation in short-axis MRI. *Magn. Reson. Imaging* **2020**, *66*, 131–140. [[CrossRef](#)]
16. Xiong, J.; Po, L.M.; Cheung, K.W.; Xian, P.; Zhao, Y.; Rehman, Y.A.U.; Zhang, Y. Edge-sensitive left ventricle segmentation using deep reinforcement learning. *Sensors* **2021**, *21*, 2375. [[CrossRef](#)]
17. Zhang, Z.; Duan, C.; Lin, T.; Zhou, S.; Wang, Y.; Gao, X. GVFOF: A novel external force for active contour based image segmentation. *Inf. Sci.* **2020**, *506*, 1–18. [[CrossRef](#)]
18. Wu, Y.; Wang, Y.; Jia, Y. Segmentation of the left ventricle in cardiac cine MRI using a shape-constrained snake model. *Comput. Vis. Image Underst.* **2013**, *117*, 990–1003. [[CrossRef](#)]
19. Feng, C.; Zhang, S.; Zhao, D.; Li, C. Simultaneous extraction of endocardial and epicardial contours of the left ventricle by distance regularized level sets. *Med. Phys.* **2016**, *43*, 2741–2755. [[CrossRef](#)]
20. Liu, Y.; Captur, G.; Moon, J.C.; Guo, S.; Yang, X.; Zhang, S.; Li, C. Distance regularized two level sets for segmentation of left and right ventricles from cine-MRI. *Magn. Reson. Imaging* **2016**, *34*, 699–706. [[CrossRef](#)]
21. Yang, C.; Wu, W.; Su, Y.; Zhang, S. Left ventricle segmentation via two-layer level sets with circular shape constraint. *Magn. Reson. Imaging* **2017**, *38*, 202–213. [[CrossRef](#)]
22. Uzunbaş, M.G.; Zhang, S.; Pohl, K.M.; Metaxas, D.; Axel, L. Segmentation of myocardium using deformable regions and graph cuts. In Proceedings of the 2012 9th IEEE International Symposium on Biomedical Imaging (ISBI), Barcelona, Spain, 2–5 May 2012; pp. 254–257. [[CrossRef](#)]
23. Dangi, S.; Yaniv, Z.; Linte, C.A. Left Ventricle Segmentation and Quantification from Cardiac Cine MR Images via Multi-task Learning. In *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*; Springer: Berlin/Heidelberg, Germany, 2019; pp. 21–31; ISBN 9783030120283.
24. Ronneberger, O.; Fischer, P.; Brox, T. U-net: Convolutional networks for biomedical image segmentation. In Proceedings of the Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics); Springer: Berlin/Heidelberg, Germany, 2015; Volume 9351, pp. 234–241.
25. Moradi, S.; Oghli, M.G.; Alizadehasl, A.; Shiri, I.; Oveisi, N.; Oveisi, M.; Maleki, M.; Dhooge, J. MFP-Unet: A novel deep learning based approach for left ventricle segmentation in echocardiography. *Phys. Med.* **2019**, *67*, 58–69. [[CrossRef](#)] [[PubMed](#)]
26. Wu, B.; Fang, Y.; Lai, X. Left ventricle automatic segmentation in cardiac MRI using a combined CNN and U-net approach. *Comput. Med. Imaging Graph.* **2020**, *82*, 101719. [[CrossRef](#)] [[PubMed](#)]
27. Dong, Z.; Du, X.; Liu, Y. Automatic segmentation of left ventricle using parallel end–end deep convolutional neural networks framework. *Knowl.-Based Syst.* **2020**, *204*, 106210. [[CrossRef](#)]
28. Shelhamer, E.; Long, J.; Darrell, T. Fully Convolutional Networks for Semantic Segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *39*, 640–651. [[CrossRef](#)]
29. Li, J.; Yu, Z.L.; Gu, Z.; Liu, H.; Li, Y. Dilated-Inception Net: Multi-Scale Feature Aggregation for Cardiac Right Ventricle Segmentation. *IEEE Trans. Biomed. Eng.* **2019**, *66*, 3499–3508. [[CrossRef](#)]
30. Zhou, X.-Y.; Yang, G.-Z. Normalization in Training U-Net for 2-D Biomedical Semantic Segmentation. *IEEE Robot. Autom. Lett.* **2019**, *4*, 1792–1799. [[CrossRef](#)]
31. Fahmy, A.S.; El-Rewaidy, H.; Nezafat, M.; Nakamori, S.; Nezafat, R. Automated analysis of cardiovascular magnetic resonance myocardial native T1 mapping images using fully convolutional neural networks. *J. Cardiovasc. Magn. Reson.* **2019**, *21*, 7. [[CrossRef](#)] [[PubMed](#)]
32. Cui, H.; Yuwen, C.; Jiang, L.; Xia, Y.; Zhang, Y. Multiscale attention guided U-Net architecture for cardiac segmentation in short-axis MRI images. *Comput. Methods Programs Biomed.* **2021**, *206*, 106142. [[CrossRef](#)] [[PubMed](#)]
33. Yang, Y.; Feng, C.; Wang, R. Automatic segmentation model combining U-Net and level set method for medical images. *Expert Syst. Appl.* **2020**, *153*, 113419. [[CrossRef](#)]
34. Baldeon-Calisto, M.; Lai-Yuen, S.K. AdaResU-Net: Multiobjective adaptive convolutional neural network for medical image segmentation. *Neurocomputing* **2020**, *392*, 325–340. [[CrossRef](#)]
35. Ngo, T.A.; Lu, Z.; Carneiro, G. Combining deep learning and level set for the automated segmentation of the left ventricle of the heart from cardiac cine magnetic resonance. *Med. Image Anal.* **2017**, *35*, 159–171. [[CrossRef](#)] [[PubMed](#)]
36. Avendi, M.R.; Kheradvar, A.; Jafarkhani, H. A combined deep-learning and deformable-model approach to fully automatic segmentation of the left ventricle in cardiac MRI. *Med. Image Anal.* **2016**, *30*, 108–119. [[CrossRef](#)] [[PubMed](#)]
37. Liang, X.; Shen, X.; Xiang, D.; Feng, J.; Lin, L.; Yan, S. Semantic Object Parsing with Local-Global Long Short-Term Memory. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; IEEE: Piscataway, NJ, USA, 2016; pp. 3185–3193.
38. Zhang, D.; Icke, I.; Dogdas, B.; Parimal, S.; Sampath, S.; Forbes, J.; Bagchi, A.; Chin, C.-L.; Chen, A. A multi-level convolutional LSTM model for the segmentation of left ventricle myocardium in infarcted porcine cine MR images. In Proceedings of the 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018), Washington, DC, USA, 4–7 April 2018; IEEE: Piscataway, NJ, USA, 2018; pp. 470–473.

39. Baumgartner, C.F.; Koch, L.M.; Pollefeys, M.; Konukoglu, E. An Exploration of 2D and 3D Deep Learning Techniques for Cardiac MR Image Segmentation. In *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*; Springer: Berlin/Heidelberg, Germany, 2018; pp. 111–119; ISBN 9783319755403.
40. Bernard, O.; Lalande, A.; Zotti, C.; Cervenansky, F.; Yang, X.; Heng, P.-A.; Cetin, I.; Lekadir, K.; Camara, O.; Gonzalez Ballester, M.A.; et al. Deep Learning Techniques for Automatic MRI Cardiac Multi-Structures Segmentation and Diagnosis: Is the Problem Solved? *IEEE Trans. Med. Imaging* **2018**, *37*, 2514–2525. [[CrossRef](#)]
41. Suinesiaputra, A.; Cowan, B.R.; Al-Agamy, A.O.; Elattar, M.A.; Ayache, N.; Fahmy, A.S.; Khalifa, A.M.; Medrano-Gracia, P.; Jolly, M.-P.; Kadish, A.H.; et al. A collaborative resource to build consensus for automated left ventricular segmentation of cardiac MR images. *Med. Image Anal.* **2014**, *18*, 50–62. [[CrossRef](#)]
42. Lalande, A.; Chen, Z.; Decourselle, T.; Qayyum, A.; Pommier, T.; Lorgis, L.; de la Rosa, E.; Cochet, A.; Cottin, Y.; Ginjac, D.; et al. Emidec: A Database Usable for the Automatic Evaluation of Myocardial Infarction from Delayed-Enhancement Cardiac MRI. *Data* **2020**, *5*, 89. [[CrossRef](#)]
43. Wen, S.; Wei, H.; Yan, Z.; Guo, Z.; Yang, Y.; Huang, T.; Chen, Y. Memristor-Based Design of Sparse Compact Convolutional Neural Network. *IEEE Trans. Netw. Sci. Eng.* **2020**, *7*, 1431–1440. [[CrossRef](#)]
44. Li, S.; Song, W.; Qin, H.; Hao, A. Deep variance network: An iterative, improved CNN framework for unbalanced training datasets. *Pattern Recognit.* **2018**, *81*, 294–308. [[CrossRef](#)]
45. Buda, M.; Maki, A.; Mazurowski, M.A. A systematic study of the class imbalance problem in convolutional neural networks. *Neural Netw.* **2018**, *106*, 249–259. [[CrossRef](#)] [[PubMed](#)]
46. Haixiang, G.; Yijing, L.; Shang, J.; Mingyun, G.; Yuanyue, H.; Bing, G. Learning from class-imbalanced data: Review of methods and applications. *Expert Syst. Appl.* **2017**, *73*, 220–239. [[CrossRef](#)]
47. Lin, A.; Wu, J.; Yang, X. A data augmentation approach to train fully convolutional networks for left ventricle segmentation. *Magn. Reson. Imaging* **2020**, *66*, 152–164. [[CrossRef](#)] [[PubMed](#)]
48. Jiang, J.; Liu, X.; Zhang, K.; Long, E.; Wang, L.; Li, W.; Liu, L.; Wang, S.; Zhu, M.; Cui, J.; et al. Automatic diagnosis of imbalanced ophthalmic images using a cost-sensitive deep convolutional neural network. *Biomed. Eng. Online* **2017**, *16*, 132. [[CrossRef](#)]
49. Van Grinsven, M.J.J.P.; Van Ginneken, B.; Hoyng, C.B.; Theelen, T.; Sánchez, C.I. Fast Convolutional Neural Network Training Using Selective Data Sampling: Application to Hemorrhage Detection in Color Fundus Images. *IEEE Trans. Med. Imaging* **2016**, *35*, 1273–1284. [[CrossRef](#)] [[PubMed](#)]
50. Tan, L.K.; Liew, Y.M.; Lim, E.; McLaughlin, R.A. Convolutional neural network regression for short-axis left ventricle segmentation in cardiac cine MR sequences. *Med. Image Anal.* **2017**, *39*, 78–86. [[CrossRef](#)]
51. Tran, P.V. A Fully Convolutional Neural Network for Cardiac Segmentation in Short-Axis MRI. *arXiv* **2016**, arXiv:1604.00494.
52. Wu, H.; Lu, X.; Lei, B.; Wen, Z. Automated left ventricular segmentation from cardiac magnetic resonance images via adversarial learning with multi-stage pose estimation network and co-discriminator. *Med. Image Anal.* **2021**, *68*, 101891. [[CrossRef](#)] [[PubMed](#)]
53. Du, X.; Tang, R.; Yin, S.; Zhang, Y.; Li, S. Direct Segmentation-Based Full Quantification for Left Ventricle via Deep Multi-Task Regression Learning Network. *IEEE J. Biomed. Health Inform.* **2019**, *23*, 942–948. [[CrossRef](#)] [[PubMed](#)]
54. Nolf, E.; Voet, T.; Jacobs, F.; Dierckx, R.; Lemahieu, I. (X)MedCon * An OpenSource Medical Image Conversion Toolkit. *Eur. J. Nucl. Med.* **2003**, *30* (Suppl. 2), S246.