A 3D Network Based Shape Prior for Automatic Myocardial Disease Segmentation in Delayed-Enhancement MRI

K. Brahim, A. Qayyum, A. Lalande, A. Boucher, A. Sakly, F. Meriaudeau

Version of Record: https://www.sciencedirect.com/science/article/pii/S1959031821000269
Manuscript_7332aa59aacbd6fc1fec6be2656d8d72

A 3D Network based Shape Prior for automatic myocardial disease segmentation in Delayed-Enhancement MRI[⋆,⋆⋆]

Khawla Brahim[a,b,c,*], Abdul Qayyum[b], Alain Lalande[b], Arnaud Boucher[b], Anis Sakly[c] and Fabrice Meriaudeau[b]

[a]National Engineering School of Sousse, University of Sousse, Tunisia
[b]ImViA EA 7535 laboratory, University of Burgundy, France
[c]LASEE laboratory, National Engineering School of Monastir, University of Monastir, Tunisia

ARTICLE INFO

Keywords:
Myocardial infarction segmentation
LGE-MRI
Microvascular-obstructed regions

ABSTRACT

Objectives: In this work, a new deep learning model for relevant myocardial infarction segmentation from Late Gadolinium Enhancement (LGE)-MRI is proposed. Moreover, our novel segmentation method aims to detect microvascular-obstructed regions accurately. Material and methods: We first segment the anatomical structures, i.e., the left ventricular cavity and the myocardium, to achieve a preliminary segmentation. Then, a shape prior based framework that fuses the 3D U-Net architecture with 3D Autoencoder segmentation framework to constrain the segmentation process of pathological tissues is applied. Results: The proposed network reached outstanding myocardial segmentation compared with the human-level performance with the average Dice score of '0.9507' for myocardium, '0.7656' for scar, and '0.8377' for MVO on the validation set consisting of 16 DE-MRI volumes selected from the training EMIDEC dataset. Conclusion: It is concluded that our approach's extensive validation and comprehensive comparison against existing state-of-the-art deep learning models on three annotated datasets, including healthy and diseased exams, make this proposal a reliable tool to enhance MI diagnosis.

1. Introduction

Cardiovascular disease is a worldwide health emergency of which myocardial infarction (MI) is the most frequent. Myocardial ischemic necrosis is caused by coronary artery occlusion, resulting in blockage of flowing blood to the myocardium [1]. However, patients with acute MI may contain MVO region, which is characterized by deficient hypo-perfusion due to decreased blood flow [2, 3].

DE-MRI is a sophisticated clinical MRI examination for assessing myocardial viability and quantification of myocardial scar. Myocardium zones, including infarcted and MVO tissues, can be highly identified in LGE-MR images, which are typically acquired after several minutes of a gadolinium-based contrast agent injection. Furthermore, accurate segmentation of the clinically significant myocardial damaged areas (scar and MVO tissues) provides a significant predictive value for MI diagnosis and therapeutic strategies such as revascularization [4, 5, 6, 7].

Fig.1 illustrates two slices from two myocardial infraction subjects, the corresponding left ventricular cavity, healthy myocardium, scar, and MVO annotations. Both over-enhanced (infarcted myocardium) and hypo-enhanced (no-reflow) regions have low contrast and ambiguous contours.

Manual annotation is mainly a time-consuming, qualified-reliant task and subject to inter-and intra-experts heterogeneity. Therefore, developing an accurate and automatic myocardial segmentation network is essentially required to assist medical experts; however, it is still a challenging task due to bad delineations regions contrast, image acquisition artifacts and respiratory motion.

Recently, deep learning-based networks have achieved state-of-the-art results for different clinical applications and have successfully shown great potential in various medical image segmentation tasks. In this work, we propose a deep learning algorithm for automatic and accurate MI segmentation from LGE-MRI. The EMIDEC segmentation challenge [8] supplies a large dataset to evaluate different methods. We used the EMIDEC STACOM 2020 challenge dataset to validate our model.

*Corresponding author
Khawla.Brahim@u-bourgogne.fr (K. Brahim); Abdul.Qayyum@u-bourgogne.fr (A. Qayyum); Alain.Lalande@u-bourgogne.fr (A. Lalande); Arnaud.Boucher@u-bourgogne.fr (A. Boucher); sakly_anis@yahoo.fr (A. Sakly); Fabrice.Meriaudeau@u-bourgogne.fr (F. Meriaudeau); Fabrice.Meriaudeau@u-bourgogne.fr (F. Meriaudeau)

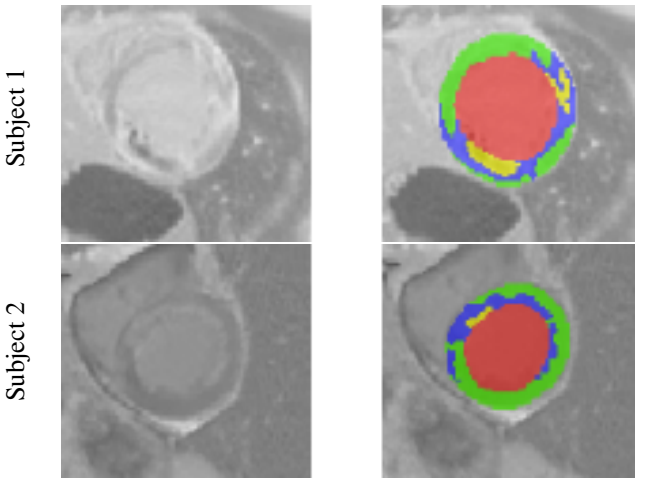


Figure 1: Visual examples of two slices from two different cases. From left to right: DE-MR images and annotations. In the 2nd column, the red, green, blue, and yellow colors represent the left ventricular cavity, healthy myocardium, scar, and MVO.

This paper is organized as follows: Related Works are described in Sect. 2. Material and Method are explained in Sect. 3. Comparative results are outlined in Sect. 4 and Sect. 5 concludes the paper.

2. Related work

In medical imaging, deep learning-based models have received researchers' interest because of their feature extraction efficiency. Among the developed models for medical image segmentation, U-Net [9] is the most popular. It is based on an encoder-decoder pipeline with intermediary skip connections, including low-level spatial information for definite delineation. U-Net-based models performed state-of-the-art results on different segmentation challenges [10, 11]. This popular architecture has been adjusted to various tasks focusing on cardiac magnetic resonance imaging [12, 13, 14, 15]. Poudel et al. [16] proposed a 2.5-dimensional network based on a recurrent fully convolutional network used during learning and retaining inter-slice information across a gated recurrent unit at the bottom of the algorithm pipeline. This network produces a coherent spatial boundary of the left ventricle cavity. Oda et al. [17] introduced a new deep learning model called Boundary-Enhanced Segmentation Network for the semantic segmentation of cells on histopathological images. As compared to U-Net, the proposed applied two decoding paths for restoring the original image resolution. Results on ganglion cells reveal that the mean Dice score representing segmentation accuracy achieves 74.0%.

As medical images are volumetric data, Deep learning models were further extended from 2D to 2.5D and then to 3D architectures. These methods, which served as an insight for our approach, were trained to provide a segmentation mask, contouring the whole image's relevant structures. Various 3D FCN demonstrate superior performance in the segmentation of cardiovascular volumes [18, 19]. 3D U-Net [20] was developed to extract volumetric spatial and intra-slice information using 3D convolutions efficiently. The proposed model demonstrates remarkable results in comparison with an equivalent 2D implementation performance for the segmentation of the highly variable structures of the Xenopus kidney. Yang et al. [21] used a 3D U-Net with pre-trained parameters brought from a C3D temporal transfer learning framework [22] to produce higher performance and fast convergence of their method. The authors have validated their approach using the Automated Cardiac Diagnosis Challenge 2017 datasets for MR image segmentation. The proposed network is a powerful deep CNNs. Milletari et al. [23] proposed V-Net deep-based model. As compared to U-Net architecture, V-Net counts more on the convolutional process and evenly moves to a deeper stage by joining a small learnable function at every stage.

As organs have a constrained shape (specific anatomy), a constrained position (relative to other organs), and a constrained topology, various 3D-U-Net models were developed based on these "a priori". Zotti et al. [14, 24] developed a novel CNN method called grid-Net for MRI cardiac anatomy segmentation. A modified U-Net architecture with convolution kernels on stridden functions for each depth level between encoding and decoding paths, whence

results were compared and constrained with a volumetric probability map illustrating the entire heart shape. It used a shape prior, which forces the algorithm to provide anatomically accurate results. Oktay et al. [25] presented a generic training framework that includes prior anatomical knowledge into CNNs as a regulariser model, making predictions attending the anatomical features (e.g., shape, tissue structure) of the anatomy structure via the learned shape. The latest models incorporate individual blocks such as the inception module, attention module, inception residual block, and convolutional block attention module (CBAM). Oktay et al. [26] proposed attention module in U-Net network that integrates attention module in skip connection from the encoder to the decoder path to extract spatial localization. The attention module does not require multiple training models that need more training parameters. The attention gating (AG) module progressively overwhelms the feature response in the irrelevant background regions and reinforces the foreground's learning capability. The attention module improved segmentation accuracy by driving attention coefficients. These coefficients are obtained by computing and combining "rich feature maps with low spatial information" attained from the 2D upsample decoder layers with the corresponding encoder layer's high-level semantic outputs. Shankaranarayana et al. [27] proposed a retinal depth estimation model based on a dilated residual inception (DRI) block for multi-scale features extraction. Ibtehaz and Rahman et al. [28] introduced a U-Net with residual inception modules on various multimodal medical imaging datasets.

Several medical imaging attempts based on deep learning models have also been included in the 2020 EMIDEC MICCAI challenge to support the clinicians in the early detection of MI and treatment. Zhang [29] proposed a cascaded convolutional neural framework which firstly used a 2D U-Net to extract the intra-slice information, and then a 3D U-Net to focus on the volumetric spatial information for final segmentation. Feng et al.[30] presented a rotation-based augmentation approach and proved its effectiveness. 2D dilated U-Net was applied as the backbone framework structure. Yang and Wang [31] developed a hybrid modified U-Net architecture which uses the squeeze-and-excitation residual (SE-Res) module and selective kernel (SK) block in the encoding and decoding paths, respectively producing robust segmentation performance. Huellebrand et al. [32] proposed to compare a hybrid mixture model based on [33] and a CNN for myocardial Pathologies' segmentation in DE-MRI. Zhou et al. [34] developed anatomy prior based network, which fuses the U-Net segmentation method with attention blocks. Their approach proposed a neighborhood penalty technique to measure the inclusion relationship between the different myocardial tissues. Results prove the validity of the proposed network in pathology segmentation.

Compared to these methods, the proposed anatomical model consists of encoders and decoders based on expansion, depth-wise, squeeze and excitation blocks, and projection layers with the progressive pattern. We used various depths at each encoder block to get multiscale features at a different place within the encoder path. Our anatomical proposed model is designed based on a proposed kernel-based atrous spatial pyramid pooling module that caters to the encoder side's information. Our pathology approach's novelty has incorporated prior knowledge to 3D U-Net, and optimized loss function adapted to the anatomy structures. The final objective function fuses Jaccard based loss function with a shape constraint loss to significantly re-weight the training for all tissues. The proposed 3D Autoencoder regulariser represents a process-specific training goal. It encodes the probability of a voxel being part of a particular tissue. More importantly, our approach learns high-level features (useful for differentiating anatomical structures) and low-level features (helpful for producing refine segmentation results). Extensive results demonstrated the promising performance of our method.

3. Material and Method

3.1. Clinical Images and Metadata

The dataset for EMIDEC challenge [8] provides 150 exams, including LGE-MRI associated with 12 clinical physiology characteristics. The image data was obtained on Siemens MRI scanners on 1.5T and 3T. The training set (67 pathological cases, 33 normal cases) and the testing set (33 pathological cases, 17 normal cases) consist of 1/3 normal and 2/3 of myocardial damage subjects. There are no overlapped exams in the training and testing sets. Each exam includes a series of 5-10 short-axis slices covering the whole left ventricle myocardium from the base to the apex. Manual delineations (contours of the relevant cardiac regions, including left ventricular cavity, healthy myocardium, MI, and MVO) are given to challengers for the training set and drawn by a biophysicist with more than 15 years of skill in medical imaging. More details of EMIDEC challenge dataset can be found in http://www.emidec.com/ and [8].

3.2. Image Pre-processing

LGE-MRI in the EMIDEC Challenge has an irregular shape. To surmount the shape mismatch and to remove irrelevant anatomical structures, original volumes were cropped to a normalized set whose center was the centroid of the left ventricular cavity annotation. Hence, all present subjects are reshaped to $96 \times 96 \times 16$ by adding empty slices [35]. A standard adaptive histogram equalization algorithm is further adjusted slice-wise to enhance the local image contrast [36, 37]. The noise was reduced using the non-local mean denoising algorithm that achieved good results when applied to previous images [38]. Finally, 16 testing cases were randomly chosen from the training dataset of EMIDEC challenge to validate our approach (Table 1).

Table 1: Stratification of the EMIDEC dataset, including pathological and normal cases used for myocardial segmentation in LGE-MRI. An example is given for one fold as 5 fold-cross validation is done.

EMIDEC dataset (n=100)	Healthy Cases	Pathological Cases	
		Infarcted Cases	Infarcted + MVO (a subclass of MI) Cases
Training dataset (n=68)	29	12	27
Validation dataset (n=16)	2	6	8
Testing dataset (n=16)	2	9	5

3.3. Method

In this section, we propose our shape prior based approach for myocardial pathology segmentation. As shown in Fig. 2, our proposal consists of two stages. Firstly, we introduce the architecture of the anatomical network to segment the myocardium and left-ventricular cavity structures. Secondly, the 3D pre-trained Autoencoder network and the 3D U-Net model were merged to yield the final myocardial segmentation output, including damaged areas (MI and no-reflow). Details of each step are presented in the following paragraphs.

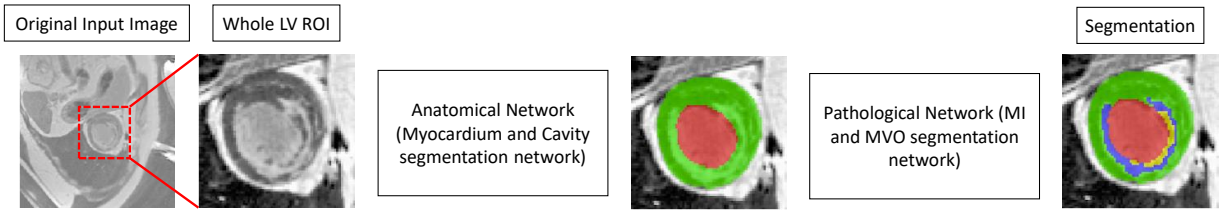


Figure 2: Pipeline of the proposed segmentation network. We first crop the LV region of interest (ROI). A new anatomical network is then used to segment the left ventricular cavity and the myocardium areas from the ROI images. Finally, a pathological network is used to detect pathological tissues (MI and MVO). The red, green, blue, and yellow colors present the left ventricular cavity, the myocardium, MI, and MVO.

3.3.1. Anatomical Network

The proposed model is built based on the architecture of encoder-decoder with skip connections. We have introduced the proposed Inception residual block with CBAM in the encoding path and the proposed EDP (expansion, depth-wise, and projection layers) block module after 2D upsampling layer in the decoding path. The attention module has been used in skip connections, which had been constructed between the analysis and synthesis paths in the same stage. Channels' number is doubled at each Inception residual block, and feature maps' input size is divided by two using depth-wise convolution layer in encoding path.

A modified inception module has been proposed on the encoder side of our anatomical network. In Inception residual block, the features maps are aggregated from different branches using kernels of various sizes. The residual connections supply smooth learning concerning the input feature maps, instead of learning an unreferenced function [39]. The proposed network is shown in Fig. 3.

The Fig. 4a presents the proposed Inception residual block. Compared with the original Inception residual module, Batch Normalization (BN) layer has been used after each convolutional layer except for bottleneck layers. We applied

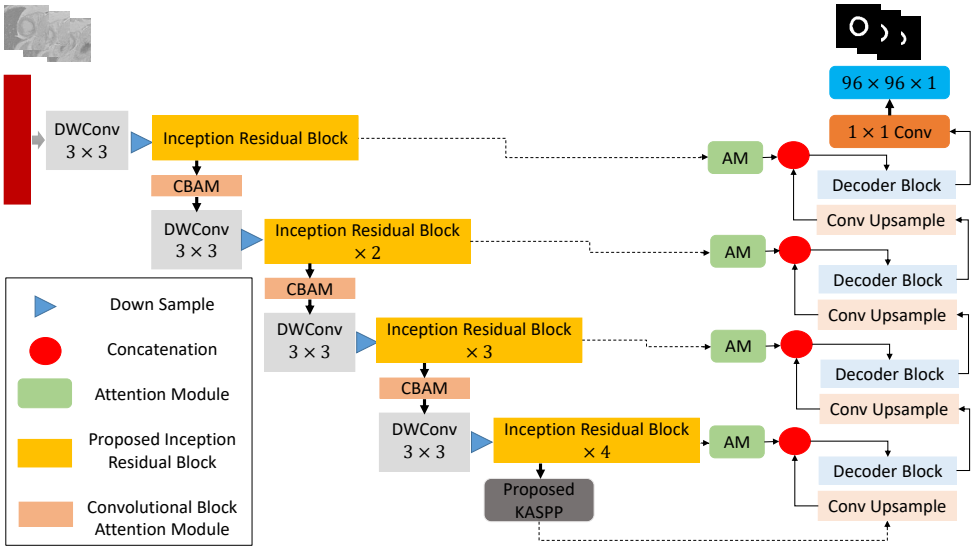


Figure 3: Anatomical network based on Inception residual, CBAM, and Decoder (EDP) blocks.

1×1 and 3×3 kernel, and also proposed 5×5 kernel branch as inspired by the DeepLab architecture [40]. BN layer yields smooth training and may avert gradient vanishing while retaining convolutional layers. The feature maps are aggregated by convolving with three kernels, namely 1×1, 3×3, and 5×5. The 3×3 and 5×5 kernels are further reduced into 1×3, 3×1, 1×5 and 5×1 to reduce the number of parameters.

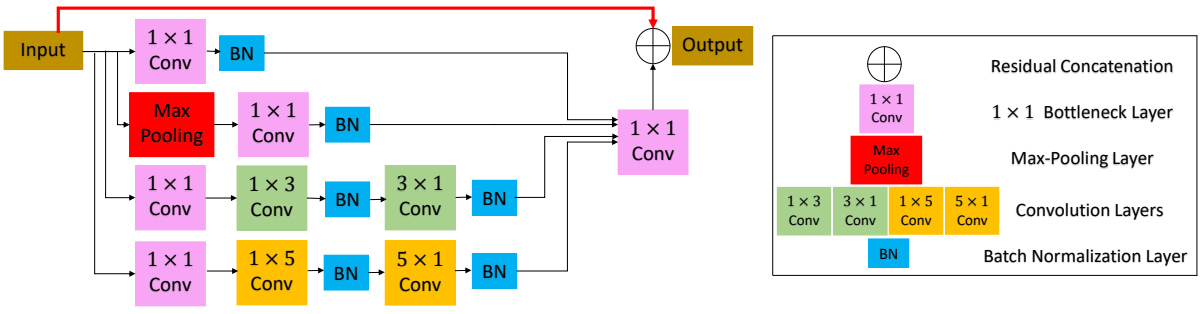
Assuming that x_l is the output of the l^{th} layer, $c_{(n\times n)}(.)$ is a $n \times n$ kernel convolutional layer, $c_b(.)$ denotes the BN layer, and 1×1 Conv designates the bottleneck layer. The output of each Inception residual block is given in Eq. 1.

$$x_{l+1} = c_{1\times1}(c_{1\times1}(x_l).c_b(c_{3\times3}(c_{1\times1}(x_l))).c_b(c_{3\times3}(c_b(c_{3\times3}(c_{1\times1}(x_l))))).c_b(c_{5\times5}(c_b(c_{5\times5}(c_{1\times1}(x_l))))) + x_l \quad (1)$$

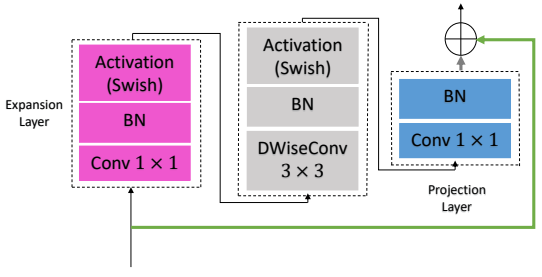
In the synthesis path, we have introduced EDP block to extract relevant semantic information. The entire layer structure for the decoder block is presented in Fig. 4b.

Sanghyun Woo et al. [41] presented the CBAM that extracts attention maps and multiplied them with input feature maps to obtain adaptive refinement features. The channel attention map exploits the inter-channel relationship of features, and the feature map obtained from this is considered a feature detector. The spatial attention module focuses on what is meaningful, given an input image through the benefit of the combination of average-pooling and max-pooling. In the spatial attention module, the features were refined channel-wise based on spatial attention. The spatial attention map is generated by exploiting the inter-spatial relationship of features. The feature descriptors are extracted using average-pooling and max-pooling operations along the channel axis and concatenated. The pooling operations along the channel axis is supposed to be effective in informative prominence regions. The 1×1 2D convolutional layer is applied on the 2D descriptor to get the raw attention map, and the batch-normalization layer produced an adequate response. After the batch normalization layer, the swish based sigmoid function is used on the final attention map. The proposed two attention modules can be placed in a parallel or sequential way. The sequential arrangement produced better performance. In our design approach, the spatial attention module used before the channel attention module provided a slightly better response. We inserted the CBAM module at the encoder side to extract delicate features from every Inception residual block stage. The proposed CBAM block is shown in Fig. 5.

Ye Huang et al. [42] proposed a kernel-sharing atrous convolutional (KSAC) layer in atrous spatial pyramid pooling (ASPP) module. The 3×3 kernel is shared with atrous convolutional layers with various dilation rates. In this work, we have extended KSAC based ASPP module and fused different features extracted from the down-sampling path with the various scale features (five scales) in the KASPP as illustrated in Fig. 6. The proposed KASPP module extracts multiscale contextual information from the encoder side of the anatomical network.



(a) Proposed Inception residual block.



(b) EDP block.

Figure 4: Proposed Inception residual block and EDP block.

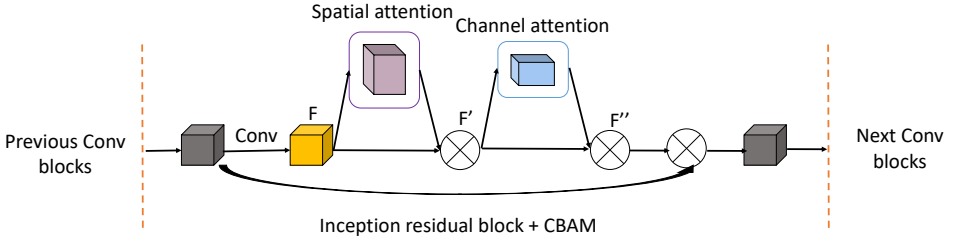


Figure 5: Proposed Convolutional block attention module (CBAM).

3.3.2. Pathological Network

The experiments revealed that extracting volume patches of size $12 \times 12 \times 12$ pixels 3 from the EMIDEC training dataset achieves excellent results for diseased myocardial tissues' segmentation.

As illustrated in Fig. 7, our 3D pathological network fuses 3D U-Net with a Super-Resolution (SR) module to constrain prior knowledge shape. 3D Autoencoder aims to correctly encode and minimize the original volume that can be reconstructed from the encoded representation. Thus a pre-trained 3D Autoencoder is efficient to regularize the segmentation result into a realistic shape. Pre-trained 3D Autoencoder is linked to the 3D U-Net and takes the segmented volume as input. A regularization term is introduced for restraining the segmentation output.

The 3D U-Net is a classical encoder-decoder segmentation framework in which, firstly, the encoder uses an ensemble of convolution-pooling layers to extract more high-level semantic features. Then the feature maps are up-sampled to recover the localization for every voxel. Shortcut connections of multi-scale feature maps between the encoder and decoder paths in the same stage reach an excellent medical segmentation. Each layer includes two $3 \times 3 \times 3$ convolutions in the encoder path, each attended by a BN, a rectified linear unit (ReLu), and then a $2 \times 2 \times 2$ max pooling with strides of two in each dimension. In the decoder path, each layer includes an up convolution of $2 \times 2 \times 2$ by strides of two in each dimension, attended by two $3 \times 3 \times 3$ convolutions, a BN, and then a ReLu. In the last layer, a $1 \times 1 \times 1$ convolution reduces the output channels' number to the number of predicted classes, which is 5 in our case.

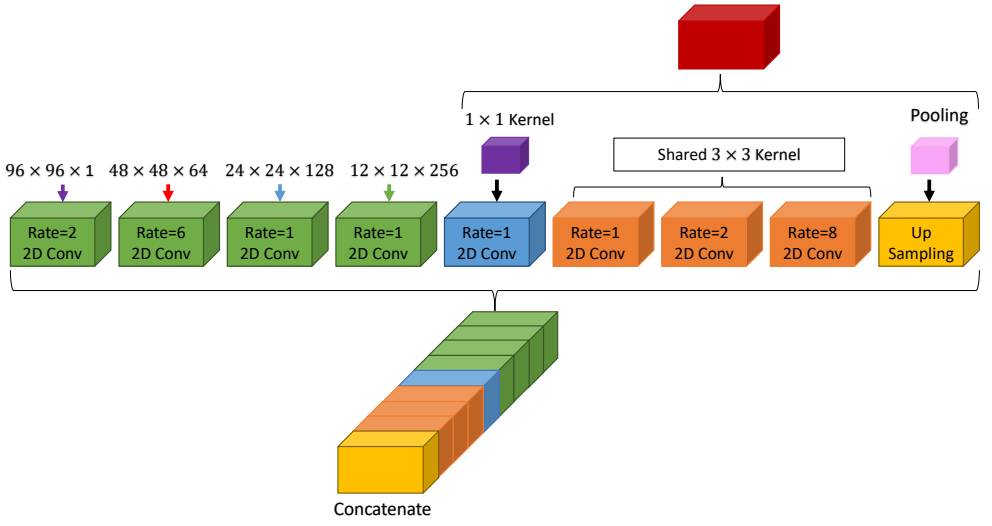


Figure 6: Proposed K-atrous spatial pyramid pooling layer module.

The employed 3D regulariser model has a fully connected layer, including five neurons representing the predicted classes (background, left ventricular cavity, healthy myocardium, scar, and MVO annotations). In the training phase, the weights of the pre-trained 3D Autoencoder model were transferred as a first step, and after that, they were fine-tuned by applying Adam optimizer over the training set. The pre-trained regulariser model has in-depth knowledge about segmenting various feature representations' types, and by fine-tuning its parameters, the 3D Autoencoder model learns the current task feature representations.

The Jaccard loss is derived from the Jaccard metric, which measures the overlap between two objects [43]. A weighted Jaccard (wIOU) and a mean weighted Jaccard (mwIOU) are two loss functions using the class weights and the weight coefficient's ratio to the sum of the weight coefficients. To overcome the class imbalance between relatively small segmentation labels and the extensive background, we optimize the summation of a (wIOU) loss / (mwIOU) loss and a shape constraint loss. The final loss function used for training the pathological network is given in Eq. 2.

$$L_{Final} = L_{Seg} + \lambda_{SR} \times L_{SR} \tag{2}$$

(Where L_{Seg} represents the wIOU or mwIOU based loss function, λ_{SR} denotes the regularization term and L_{SR} indicates the L2 loss function that is defined in Frobenius norm Eq. 3. We choose $\lambda_{SR} = 10^{-2}$.)

$$L_{SR} = \sum_{i=1}^{n} ||RP_i - RG_i||_F^2 \tag{3}$$

(Where n represents the total number of training volumes, RG_i is the reconstructed gold standard, RP_i indicates the reconstructed segmentation results and $||.||_F$ is the Frobenius norm of an $m \times n$ matrix.)

4. Results and discussion

4.1. Evaluation Metrics for Segmentation

For the evaluation of networks segmentation results in 3D, we used some common objective criteria including geometrical metrics such as Dice score (DSC) Eq. 4 for the different areas of interest and Hausdorff Distance (HD) Eq. 5 for the myocardium and clinical metrics including relative volume errors, i.e., Absolute Volume Difference (AVD) Eq. 6 and Absolute Volume Difference Rate according to the volume of the myocardium (AVDR) Eq. 7. [44, 45, 46, 47, 48]. The evaluated metrics were based on the official code https://github.com/EMIDEC-Challenge/Evaluation-metrics.

The DSC computes the spatial overlap of the network-generated boundary and manual contouring and is determined as follows:

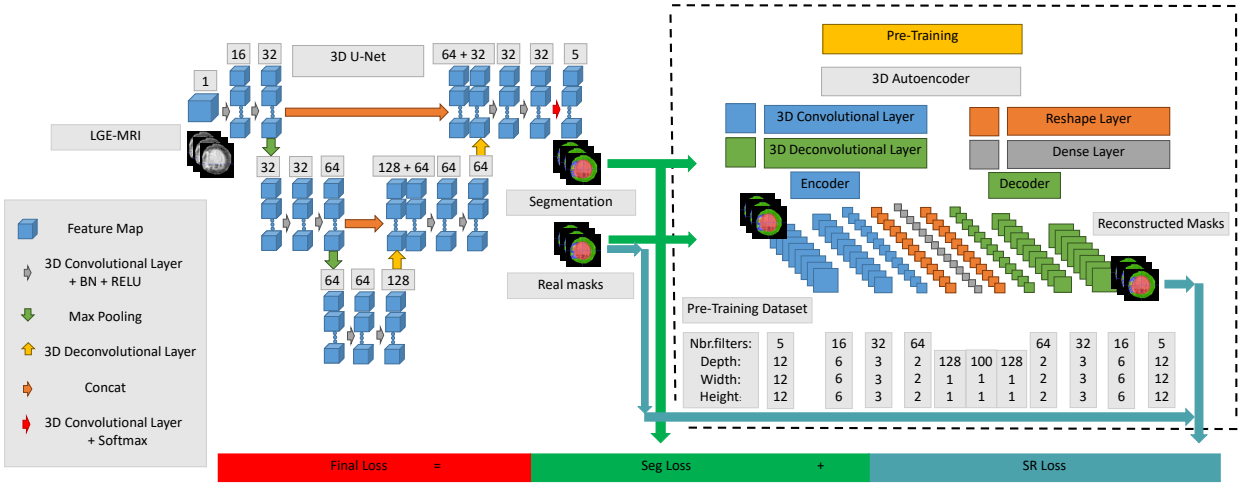


Figure 7: Schematic representation of our pathological approach for myocardial disease segmentation. The number of channels is presented above each feature map.

$$DSC = \frac{2 \times (N \cap M)}{N + M} \tag{4}$$

(Where N and M designate the corresponding network and manually segmented tissues.)

The HD metric computes how far the network-generated delineations and gold standard are from each other and is determined as follows:

$$HD = \max_{x \in N} \min_{y \in M} ||x - y|| \tag{5}$$

$$AVD = |V_N - V_M| \tag{6}$$

$$AVDR = \frac{|V_N - V_M|}{V_{MYO}} \tag{7}$$

(Where V_N is the volume of the network-generated segmentation, V_M is the volume of the manual annotation, and V_{MYO} is the Myocardial volume of the manual annotation.)

4.2. Results

4.2.1. Evaluation results of the proposed model on EMIDEC dataset

There are a total of 100 cases with published labels to train, validate, and test our network's performance. We make random 5-fold cross-validation by shuffling the scan sequence and splitting the EMIDEC dataset into 5 folds using 68 MR scans for training. Table 2 and Table 3 represent the cross-validation results of our segmentation output before the application of post-processing removal of correlated errors in myocardial disease segmentation from LGE-MRI data and two methods segmentation output.

Table 2 gives an extensive evaluation of the conducted experiments' metrics on the 5-fold cross-validation. Our approach proves that with the restricted low number of training samples, it is conceivable to segment LGE-MRI scans. It can be observed that the proposed method can accurately segment the anatomical structures as well as the pathological tissues. For our 5-fold cross-validation results, the Standard Deviation is relatively small, revealing that our approach is stable for the myocardial disease segmentation.

Summary of the quantitative results reported in Table 3 demonstrates that our proposed segmentation model reached the highest DSC and HD compared with two methods included in the 2020 EMIDEC MICCAI challenge. The comparison study shows the effectiveness of our proposed method.

Table 2: Quantitative 5-fold cross-validation results of our myocardial segmentation output in LGE-MRI.

Targets	Metrics	5-fold cross-validation						
		fold0	fold1	fold2	fold3	fold4	Average	Standard Deviation
Myocardium	DSC %	95.27	94.40	95.11	95.23	95.49	95.10	0.37
	AVD mm^3	256.81	343.50	247.88	252.19	231.69	266.41	39.46
	HD mm	4.41	5.29	4.36	4.06	3.89	4.40	0.48
Infarction	DSC %	74.93	77.85	76.37	76.52	75.02	76.14	1.08
	AVD mm^3	342.00	308.69	212.31	176.75	284.81	264.91	61.31
	AVDR %	6.82	6.87	4.09	2.99	5.84	5.32	1.54
MVO	DSC %	71.60	77.03	77.64	70.60	72.06	73.79	2.94
	AVD mm^3	41.06	34.50	60.00	76.31	63.56	55.09	15.28
	AVDR %	0.75	0.75	1.10	1.41	1.24	1.05	0.26

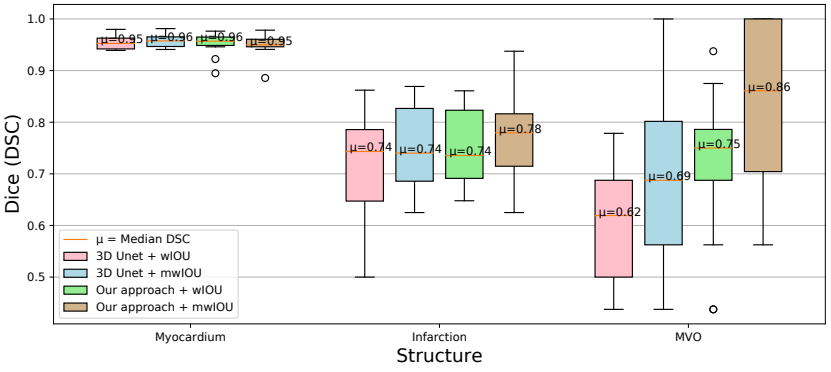
Table 3: Comparison study with previous methods on 5-fold cross-validation of EMIDEC dataset. Best values marked in bold font.

Targets	Metrics	Methods		
		Huellebrand et al. [32]	Zhang [29]	**Our proposed model**
Myocardium	DSC %	81.00	94.40	**95.10**
	AVD mm^3	13655.55	6474.38	**266.41**
	HD mm	16.72	17.21	**4.40**
Infarction	DSC %	36.08	72.08	**76.14**
	AVD mm^3	8980.5	4179.5	**264.91**
	AVDR %	7.07	**3.41**	5.32
MVO	DSC %	54.15	71.01	**73.79**
	AVD mm^3	1501.73	918.69	**55.09**
	AVDR %	1.08	**0.69**	1.05

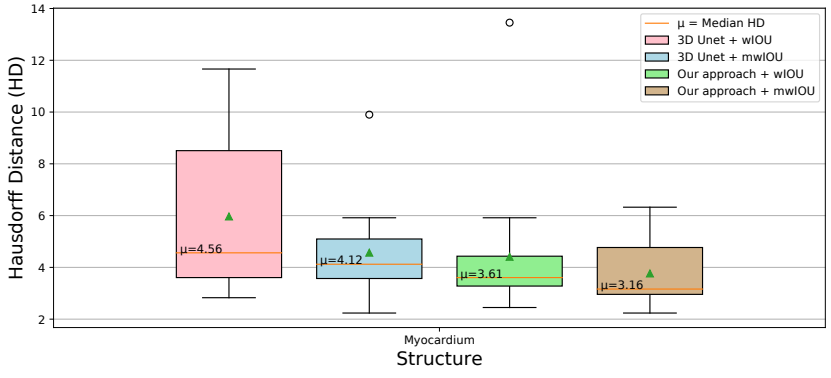
We conduct different experiments for a comparative study of methods based on (wIOU) loss and (mwIOU) loss, which is defined by the ratio of the wIOU to the sum of the weight coefficients in myocardial segmentation. The proposed model is also compared with the baseline 3D U-Net architecture to highlight the importance of shape regularization in improving segmentation. Fig. 8 and Fig. 9 plot quantitative metrics including the Dice score and AVD for each myocardial region for the two methods among all loss functions, showing the pertinence of prior shape and post-processing in making the best trade-off between evaluation metrics. Morphological opening with a kernel of size 3×3 and iterations' number equal to 1 and majority voting method (combining results from different models obtained with varying parameters of the loss function's weighting) as post-processing operations are used to increase sensitivity for segmenting infarct and MVO. Our proposed segmentation method produces excellent results representing top performances for myocardial segmentation in LGE-MRI. For our final segmentation results, the ensembled network achieves an average DSC score of '0.9507', '0.7656', and '0.8377' for myocardium, MI, and MVO, respectively. The proposed network successfully outperforms the baseline 3D U-Net architecture (median HD of '3.16 mm' vs. '4.12 mm').

Fig. 10 illustrates good visualization of our final segmented volumes, 3D U-Net results, and corresponding manual myocardial annotation in our own split testing set. We acquire the final prediction of testing set by gathering the segmentation results of different models using morphological opening and majority voting as post-processing operations. Two axial slice views and a 3D surface rendering with manual myocardial delineation are illustrated on the first row. Segmented images obtained using our 2-stage anatomical and pathological overall's system and 3D U-Net are shown in the matching row. As compared to manual annotations, our proposed network correctly depicted myocardial structures and showed a good agreement with the gold standard. Notably, the proposed shape prior based deep method demonstrated promising performance in segmenting the damaged myocardial areas from DE-MRI.

Table 4 shows how the addition of the pathological network, which is responsible for extracting the damaged areas, affect the segmentation performance of the myocardium and left-ventricular cavity structures obtained from the



(a) Dice score for the myocardium, MI, and MVO areas.



(b) HD for the various segmentation netowrks.

Figure 8: An illustration of the networks results. The proposed network outperformed the 3D U-Net approaches in segmenting myocardial tissues.

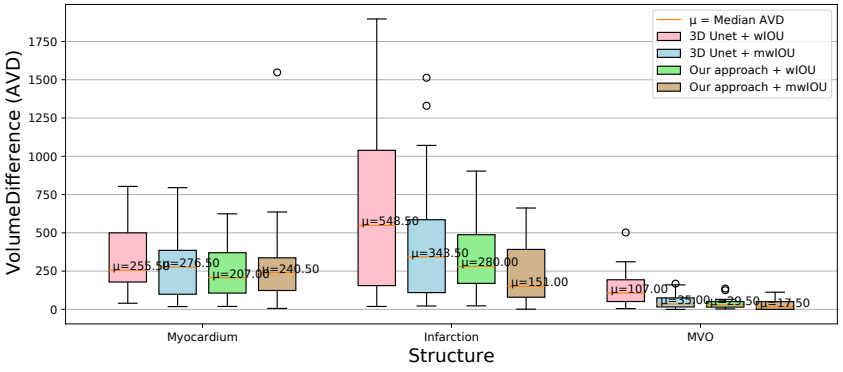
anatomical network. The table quantitatively demonstrates the anatomical segmentation model's performances alone and when adding the shape prior model to it.

Table 4: Performance comparison between the proposed anatomical method and our final proposed model (anatomical + pathological) before and after the application of post-processing on EMIDEC dataset. An example is given for one fold as 5 fold-cross validation is done "fold0".

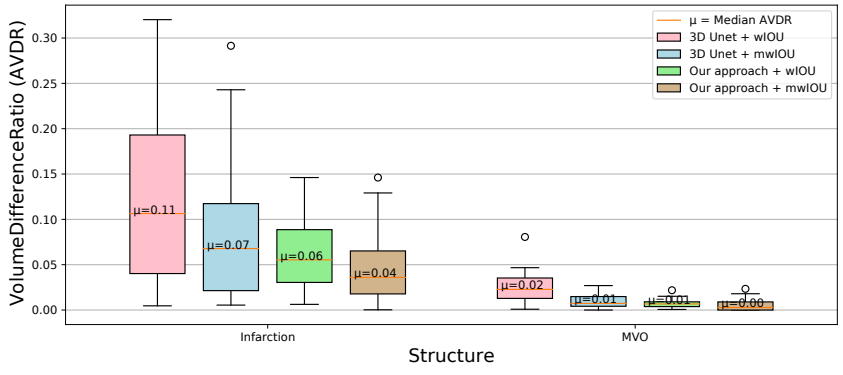
Targets	Metrics	Methods		
		Our proposed anatomical model	Our final proposed model before applying post-processing	Our final proposed model after applying post-processing
Myocardium	DSC %	91.18	95.27	95.07
	AVD mm^3	348.62	256.81	315.5
	HD mm	23.65	4.41	5.02
Infarction	DSC %	–	74.93	76.56
	AVD mm^3	–	342.00	234.12
	AVDR %	–	6.82	4.92
MVO	DSC %	–	71.60	83.77
	AVD mm^3	–	41.06	30.25
	AVDR %	–	0.75	0.60

4.2.2. Validation of the performance of the proposed anatomical model on benchmark datasets

The proposed method (anatomical model) has been evaluated and tested on other publicly available LGE-MRI datasets such as Myocardial pathology segmentation combining multi-sequence CMR (MyoPS) 2020 dataset in MIC-



(a) AVD between proposed networks and the gold standard.



(b) AVDR of MI and MVO in relation to the complete myocardial volume.

Figure 9: An illustration of the networks results. Our approach shows superior AVD and AVDR-values on the MI, and MVO areas.

CAI 2020 challenge and a Multi-sequence Cardiac MR (MS-CMRSeg) 2019 in MICCAI 2019. The data has been preprocessed based on the multivariate mixture model (MvMM) [49]. The performance comparison of our proposed anatomical method and existing deep learning models on MS-CMRSeg 2019 and MyoPS 2020 datasets are shown in Table 5 and Table 6 respectively. These tables prove the reliability of our proposed anatomical model on other clinical datasets. The experiments show that our anatomical approach produced excellent Dice similarity as compared to existing methods.

Table 5: Performance analysis of the proposed anatomical model on MS-CMRSeg 2019 dataset and comparison with existing deep learning-based models. Best values marked in bold font.

Methods	Dataset	DSC (%)	HD (mm)
Our proposed anatomical model	BSSFP+ T1 for training LGE for testing	**81.87**	**9.98**
Sulaiman Vesal [50]	Used different combination for training LGE for testing	74.90	11.35
Víctor M. Campello [51]	Used different combination for training LGE for testing	68.00	12.00

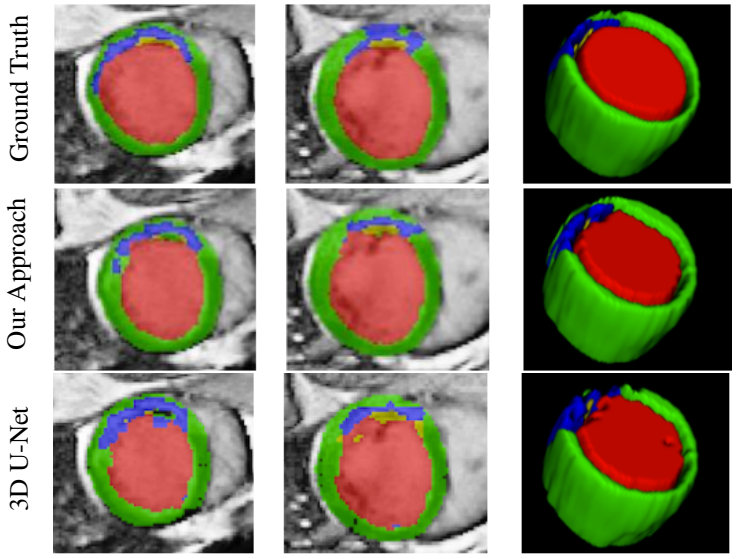


Figure 10: Visual exemplar segmentations for our 2-stage anatomical and pathological overall system and 3D U-Net applied for the same subject on the EMIDEC validation set. Each row displays a case on two different heart locations, followed by a 3D surface rendering. The rows from top to bottom are the ground truth, the proposed method prediction, and 3D U-Net segmentation.

Table 6: Performance comparison between the proposed anatomical method and existing state-of-the-art deep learning models on MyoPS 2020 dataset. Best values marked in bold font.

Models	DSC(%)	HD(mm)	AVD(ml)
Our proposed anatomical model	**0.8469**	**16.3190**	8.6275
DSNet [52]	0.8217	17.9299	**7.3824**
X-Net [53]	0.8005	23.2728	11.6204
U-Net [9]	0.7956	114.7892	26.3757

5. Conclusion

This paper proposes a deep learning-based shape prior model for automatic myocardial infarction segmentation from LGE-MRI. Experiments results have revealed the algorithm's effectiveness in segmenting the anatomical structures. As compared to the manual annotation, our proposed model can find its pertinence in assisting experts in diagnosing different cardiac diseases. In future research, we would fuse features extracted from LGE-MRI and clinical Metadata information to perform the final segmentation.

6. Acknowledgment

This work was partly supported by the French "Investissements d'Avenir" program, ISITE-BFC project (number ANR-15-IDEX-0003).

References

[1] B. Surawicz, T. Knilans, Chou's Electrocardiography in Clinical Practice E-Book: Adult and Pediatric, Elsevier Health Sciences, 2008.
[2] V. Pineda, X. Merino, S. Gispert, P. Mahía, B. Garcia, R. Domínguez-Oronoz, No-reflow phenomenon in cardiac mri: diagnosis and clinical implications, American Journal of Roentgenology 191 (1) (2008) 73–79.
[3] P. Rajiah, M. Y. Desai, D. Kwon, S. D. Flamm, Mr imaging of myocardial infarction, Radiographics 33 (5) (2013) 1383–1412.
[4] R. J. Kim, E. Wu, A. Rafael, E.-L. Chen, M. A. Parker, O. Simonetti, F. J. Klocke, R. O. Bonow, R. M. Judd, The use of contrast-enhanced magnetic resonance imaging to identify reversible myocardial dysfunction, New England Journal of Medicine 343 (20) (2000) 1445–1453.

[5] B. L. Gerber, M. F. Rousseau, S. A. Ahn, J.-B. l. P. de Waroux, A.-C. Pouleur, T. Phlips, D. Vancraeynest, A. Pasquet, J.-L. J. Vanoverschelde, Prognostic value of myocardial viability by delayed-enhanced magnetic resonance in patients with coronary artery disease and low ejection fraction: impact of revascularization therapy, Journal of the American College of Cardiology 59 (9) (2012) 825–835.

[6] R. O. Bonow, G. Maurer, K. L. Lee, T. A. Holly, P. F. Binkley, P. Desvigne-Nickens, J. Drozdz, P. S. Farsky, A. M. Feldman, T. Doenst, et al., Myocardial viability and survival in ischemic left ventricular dysfunction, New England Journal of Medicine 364 (17) (2011) 1617–1625.

[7] K. C. Allman, L. J. Shaw, R. Hachamovitch, J. E. Udelson, Myocardial viability testing and impact of revascularization on prognosis in patients with coronary artery disease and left ventricular dysfunction: a meta-analysis, Journal of the American College of Cardiology 39 (7) (2002) 1151–1158.

[8] A. Lalande, Z. Chen, T. Decourselle, A. Qayyum, T. Pommier, L. Lorgis, E. de la Rosa, A. Cochet, Y. Cottin, D. Ginhac, et al., Emidec: A database usable for the automatic evaluation of myocardial infarction from delayed-enhancement cardiac mri, Data 5 (4) (2020) 89.

[9] O. Ronneberger, P. Fischer, T. Brox, U-net: Convolutional networks for biomedical image segmentation, in: International Conference on Medical image computing and computer-assisted intervention, Springer, 2015, pp. 234–241.

[10] N. Heller, F. Isensee, K. H. Maier-Hein, X. Hou, C. Xie, F. Li, Y. Nan, G. Mu, Z. Lin, M. Han, et al., The state of the art in kidney and kidney tumor segmentation in contrast-enhanced ct imaging: Results of the kits19 challenge, arXiv preprint arXiv:1912.01054 (2019).

[11] F. Isensee, J. Petersen, S. A. Kohl, P. F. Jäger, K. H. Maier-Hein, nnu-net: Breaking the spell on successful medical image segmentation, arXiv preprint arXiv:1904.08128 1 (2019) 1–8.

[12] C. F. Baumgartner, L. M. Koch, M. Pollefeys, E. Konukoglu, An exploration of 2d and 3d deep learning techniques for cardiac mr image segmentation, in: International Workshop on Statistical Atlases and Computational Models of the Heart, Springer, 2017, pp. 111–119.

[13] Q. Zheng, H. Delingette, N. Duchateau, N. Ayache, 3-d consistent and robust segmentation of cardiac images by deep learning with spatial propagation, IEEE transactions on medical imaging 37 (9) (2018) 2137–2148.

[14] C. Zotti, Z. Luo, O. Humbert, A. Lalande, P.-M. Jodoin, Gridnet with automatic shape prior registration for automatic mri cardiac segmentation, in: International Workshop on Statistical Atlases and Computational Models of the Heart, Springer, 2017, pp. 73–81.

[15] J. Zhang, J. Du, H. Liu, X. Hou, Y. Zhao, M. Ding, Lu-net: an improved u-net for ventricular segmentation, IEEE Access 7 (2019) 92539–92546.

[16] R. P. Poudel, P. Lamata, G. Montana, Recurrent fully convolutional neural networks for multi-slice mri cardiac segmentation, in: Reconstruction, segmentation, and analysis of medical images, Springer, 2016, pp. 83–94.

[17] H. Oda, H. R. Roth, K. Chiba, J. Sokolić, T. Kitasaka, M. Oda, A. Hinoki, H. Uchida, J. A. Schnabel, K. Mori, Besnet: boundary-enhanced segmentation of cells in histopathological images, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 2018, pp. 228–236.

[18] Q. Dou, L. Yu, H. Chen, Y. Jin, X. Yang, J. Qin, P.-A. Heng, 3d deeply supervised network for automated segmentation of volumetric medical images, Medical image analysis 41 (2017) 40–54.

[19] L. Yu, X. Yang, J. Qin, P.-A. Heng, 3d fractalnet: dense volumetric segmentation for cardiovascular mri volumes, in: Reconstruction, segmentation, and analysis of medical images, Springer, 2016, pp. 103–110.

[20] Ö. Çiçek, A. Abdulkadir, S. S. Lienkamp, T. Brox, O. Ronneberger, 3d u-net: learning dense volumetric segmentation from sparse annotation, in: International conference on medical image computing and computer-assisted intervention, Springer, 2016, pp. 424–432.

[21] X. Yang, C. Bian, L. Yu, D. Ni, P.-A. Heng, Class-balanced deep neural network for automatic ventricular structure segmentation, in: International workshop on statistical atlases and computational models of the heart, Springer, 2017, pp. 152–160.

[22] D. Tran, L. Bourdev, R. Fergus, L. Torresani, M. Paluri, Learning spatiotemporal features with 3d convolutional networks, in: Proceedings of the IEEE international conference on computer vision, 2015, pp. 4489–4497.

[23] F. Milletari, N. Navab, S.-A. Ahmadi, V-net: Fully convolutional neural networks for volumetric medical image segmentation, in: 2016 fourth international conference on 3D vision (3DV), IEEE, 2016, pp. 565–571.

[24] C. Zotti, Z. Luo, A. Lalande, P.-M. Jodoin, Convolutional neural network with shape prior applied to cardiac mri segmentation, IEEE journal of biomedical and health informatics 23 (3) (2018) 1119–1128.

[25] O. Oktay, E. Ferrante, K. Kamnitsas, M. Heinrich, W. Bai, J. Caballero, S. A. Cook, A. De Marvao, T. Dawes, D. P. O'Regan, et al., Anatomically constrained neural networks (acnns): application to cardiac image enhancement and segmentation, IEEE transactions on medical imaging 37 (2) (2017) 384–395.

[26] O. Oktay, J. Schlemper, L. L. Folgoc, M. Lee, M. Heinrich, K. Misawa, K. Mori, S. McDonagh, N. Y. Hammerla, B. Kainz, et al., Attention u-net: Learning where to look for the pancreas, arXiv preprint arXiv:1804.03999 (2018).

[27] S. M. Shankaranarayana, K. Ram, K. Mitra, M. Sivaprakasam, Fully convolutional networks for monocular retinal depth estimation and optic disc-cup segmentation, IEEE journal of biomedical and health informatics 23 (4) (2019) 1417–1426.

[28] N. Ibtehaz, M. S. Rahman, Multiresunet: Rethinking the u-net architecture for multimodal biomedical image segmentation, Neural Networks 121 (2020) 74–87.

[29] Y. Zhang, Cascaded convolutional neural network for automatic myocardial infarction segmentation from delayed-enhancement cardiac mri.

[30] X. Feng, C. M. Kramer, M. Salerno, C. H. Meyer, Automatic scar segmentation from de-mri using 2d dilated unet with rotation-based augmentation.

[31] S. Yang, X. Wang, A hybrid network for automatic myocardial infarction segmentation in delayed enhancement-mri.

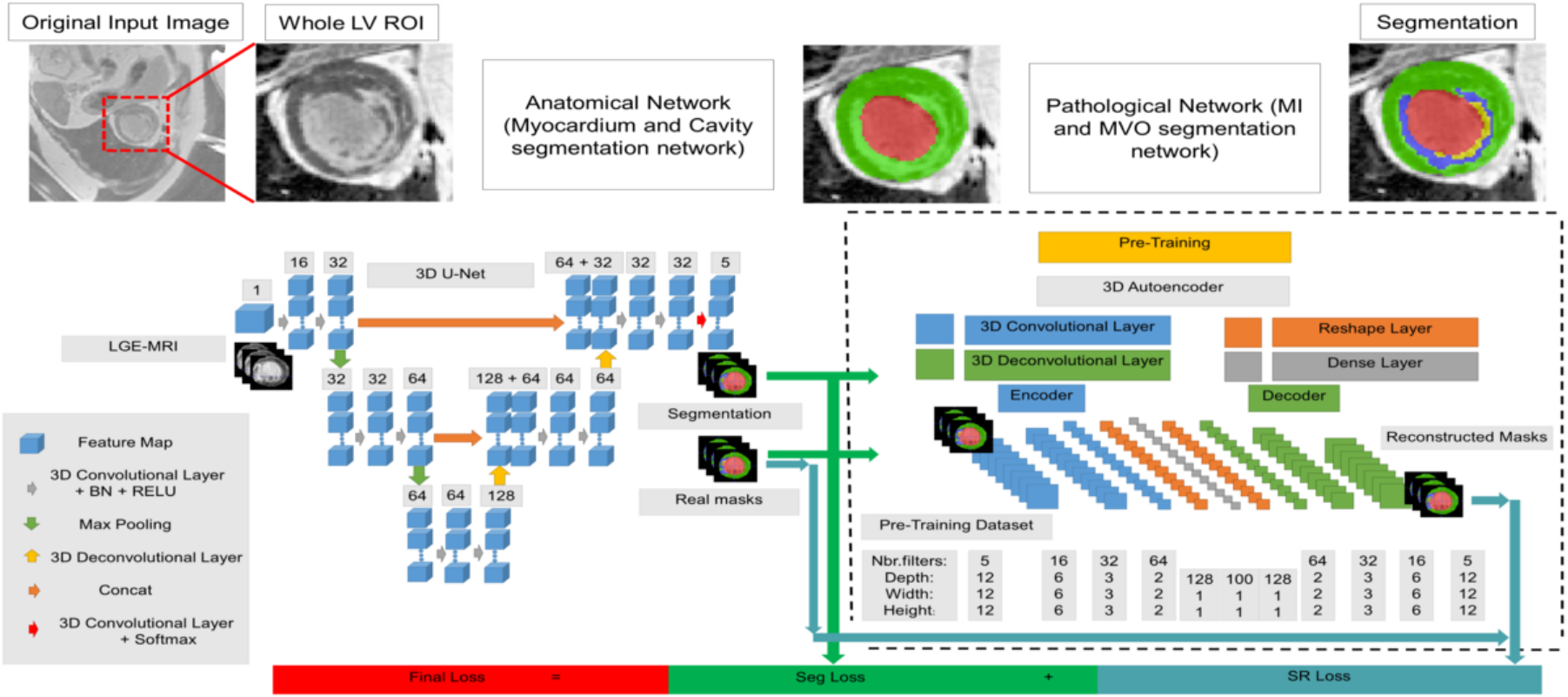
[32] M. Huellebrand, M. Ivantsits, H. Zhang, P. Kohlmann, J.-M. Kuhnigk, T. Kuehne, S. Schönberg, A. Hennemuth, Comparison of a hybrid mixture model and a cnn for the segmentation myocardial pathologies in delayed enhancement mri.

[33] A. Hennemuth, O. Friman, M. Huellebrand, H.-O. Peitgen, Mixture-model-based segmentation of myocardial delayed enhancement mri, in: International Workshop on Statistical Atlases and Computational Models of the Heart, Springer, 2012, pp. 87–96.

[34] Y. Zhou, K. Zhang, X. Luo, S. Wang, X. Zhuang, Anatomy prior based u-net for pathology segmentation with attention, arXiv preprint arXiv:2011.08769 (2020).

[35] A. Qayyum, A. Lalande, F. Meriaudeau, Automatic segmentation of tumors and affected organs in the abdomen using a 3d hybrid model for

computed tomography imaging, Computers in Biology and Medicine 127 (2020) 104097.
[36] K. Zuiderveld, Contrast limited adaptive histogram equalization, Graphics gems (1994) 474–485.
[37] W.-Q. Xie, X.-P. Zhang, X.-M. Yang, Q.-S. Liu, S.-H. Tang, X.-B. Tu, 3d size and shape characterization of natural sand particles using 2d image analysis, Engineering Geology (2020) 105915.
[38] A. Buades, B. Coll, J.-M. Morel, A non-local algorithm for image denoising, in: 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), Vol. 2, IEEE, 2005, pp. 60–65.
[39] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, A. Rabinovich, Going deeper with convolutions, in: Proceedings of the IEEE conference on computer vision and pattern recognition, 2015, pp. 1–9.
[40] L.-C. Chen, Y. Zhu, G. Papandreou, F. Schroff, H. Adam, Encoder-decoder with atrous separable convolution for semantic image segmentation, in: Proceedings of the European conference on computer vision (ECCV), 2018, pp. 801–818.
[41] S. Woo, J. Park, J.-Y. Lee, I. So Kweon, Cbam: Convolutional block attention module, in: Proceedings of the European conference on computer vision (ECCV), 2018, pp. 3–19.
[42] Y. Huang, Q. Wang, W. Jia, X. He, See more than once–kernel-sharing atrous convolution for semantic segmentation, arXiv preprint arXiv:1908.09443 (2019).
[43] Y. Yuan, M. Chao, Y.-C. Lo, Automatic skin lesion segmentation using deep fully convolutional networks with jaccard distance, IEEE transactions on medical imaging 36 (9) (2017) 1876–1886.
[44] W. R. Crum, O. Camara, D. L. Hill, Generalized overlap measures for evaluation and validation in medical image analysis, IEEE transactions on medical imaging 25 (11) (2006) 1451–1461.
[45] A. A. Taha, A. Hanbury, Metrics for evaluating 3d medical image segmentation: analysis, selection, and tool, BMC medical imaging 15 (1) (2015) 29.
[46] K. Kamnitsas, C. Ledig, V. F. Newcombe, J. P. Simpson, A. D. Kane, D. K. Menon, D. Rueckert, B. Glocker, Efficient multi-scale 3d cnn with fully connected crf for accurate brain lesion segmentation, Medical image analysis 36 (2017) 61–78.
[47] K. H. Zou, S. K. Warfield, A. Bharatha, C. M. Tempany, M. R. Kaus, S. J. Haker, W. M. Wells III, F. A. Jolesz, R. Kikinis, Statistical validation of image segmentation quality based on a spatial overlap index1: scientific reports, Academic radiology 11 (2) (2004) 178–189.
[48] D. P. Huttenlocher, G. A. Klanderman, W. J. Rucklidge, Comparing images using the hausdorff distance, IEEE Transactions on pattern analysis and machine intelligence 15 (9) (1993) 850–863.
[49] X. Zhuang, Multivariate mixture model for myocardial segmentation combining multi-source images, IEEE transactions on pattern analysis and machine intelligence 41 (12) (2018) 2933–2946.
[50] S. Vesal, N. Ravikumar, A. Maier, Automated multi-sequence cardiac mri segmentation using supervised domain adaptation, in: International Workshop on Statistical Atlases and Computational Models of the Heart, Springer, 2019, pp. 300–308.
[51] V. M. Campello, C. Martín-Isla, C. Izquierdo, S. E. Petersen, M. A. G. Ballester, K. Lekadir, Combining multi-sequence and synthetic images for improved segmentation of late gadolinium enhancement cardiac mri, in: International Workshop on Statistical Atlases and Computational Models of the Heart, Springer, 2019, pp. 290–299.
[52] M. K. Hasan, L. Dahal, P. N. Samarakoon, F. I. Tushar, R. Martí, Dsnet: Automatic dermoscopic skin lesion segmentation, Computers in Biology and Medicine (2020) 103738.
[53] K. Qi, H. Yang, C. Li, Z. Liu, M. Wang, Q. Liu, S. Wang, X-net: Brain stroke lesion segmentation based on depthwise separable convolution and long-range dependencies, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 2019, pp. 247–255.


Original Input Image
Whole LV ROI
Anatomical Network (Myocardium and Cavity segmentation network)
Pathological Network (MI and MVO segmentation network)
Segmentation
LGE-MRI
3D U-Net
Segmentation
Real masks
Pre-Training
3D Autoencoder
3D Convolutional Layer
3D Deconvolutional Layer
Reshape Layer
Dense Layer
Encoder
Decoder
Reconstructed Masks
Pre-Training Dataset
Feature Map
3D Convolutional Layer + BN + RELU
Max Pooling
3D Deconvolutional Layer
Concat
3D Convolutional Layer + Softmax
Final Loss = Seg Loss + SR Loss