Segmentation of Late Gadolinium-Enhanced Cardiac Magnetic Resonance Images Based on Residual Attention Mechanisms

Zhaokai Kong[1], Zhenyin Fu[1], Yingyi Geng[1], Ling Xia[1], Ruiqing Dong[2], Jucheng Zhang[3], Dongdong Deng[4*]

[1]Zhejiang University, Hangzhou, China
[2]The Fourth Affiliated Hospital of Soochow University, Suzhou, China
[3]The Second Afiliated Hospital Zhejiang University School of Medicine, Hangzhou, China
[4]Dalian University of Technology, Dalian, China

Abstract

Accurate automated segmentation of late gadolinium-enhanced cardiac magnetic resonance (LGE-CMR) images is crucial for personalized cardiac modeling, particularly for myocardial boundaries. This study aims to validate the accuracy of our proposed deep learning segmentation method for the left ventricular blood cavity (LV) and left ventricular myocardium (Myo). We propose a residual attention-based nnUNetv2 to automatically segment the LV and Myo. The residual attention mechanism enables the network to focus on more informative channels, enhancing discriminative learning and increasing the network's ability to segment the LV and Myo boundary. We also propose an image interpolation method that interpolates the voxel spacing of the original image and increases the number of image layers. Our proposed method achieved a Dice similarity coefficient (DSC) of 85.19% for LV and 72.12% for Myo. The lower DSC score for Myo is primarily due to the poor image quality of the LGE-CMR. The results demonstrate that our proposed residual attention-based nnUNetv2 network can effectively improve the network's ability to segment LV and Myo in LGE-CMR images.

1. Introduction

Sudden cardiac death[1] from ventricular tachycardia (VT)[2,3], often linked to myocardial infarction, represents a significant health issue[4]. Advances in virtual cardiac modeling have improved our understanding and management of VT, enhancing the targeting and effectiveness of treatments like radiofrequency ablation (RFA). Three-dimensional CMR reconstruction is crucial for personalized cardiac modeling, enabling accurate heart anatomy visualization and automated MRI segmentation. This process supports precise patient-specific models that guide clinical interventions and streamline clinical workflows.

Manual segmentation of CMR images in clinical settings is time-consuming and inefficient, with results varying significantly between individuals and lacking reproducibility[5]. Semi-automatic or fully automatic segmentation techniques have long been a focal point in medical image processing research[6]. However, inherent limitations in CMR images present challenges for segmentation: significant variations in heart shape, size, and position among individuals; indistinct boundaries of cardiac substructures due to noise and lesions; and inconsistent image resolution.

Currently, deep learning-based approaches lead the field of medical image segmentation[7]. Zhou et al.[8] have redesigned the skip connections of the U-Net framework, proposing a new architecture, UNet++, that utilizes convolutional neural networks (CNNs) for semantic and instance segmentation. Isensee et al.[9] introduced the nnUNet segmentation method, which automates parameter configuration for data preprocessing, network architecture, training, and post-processing for new tasks without human intervention. Recent studies have employed the Vision Transformer (ViT)[10] to replace the bottleneck layer of U-Net, enhancing segmentation accuracy for abdominal and cardiac images[11]. Other research has utilized the Swin Transformer[12] as the foundational unit for constructing the encoder and decoder of a U-Net, with segmentation experiments showing that this pure Transformer network structure outperforms hybrid CNN-Transformer structures[13]. Huang et al.[14] have developed a feedforward network using convolution operations to replace the multi-layer perceptron (MLP) in classical Transformers, effectively capturing long-range dependencies and local context for medical image segmentation.

To address the issues mentioned, we propose a fully automated segmentation network based on the nnU-Net architecture, incorporating a residual attention mechanism. This network enhances the model's focus on target areas, thereby effectively improving segmentation performance. Additionally, we have developed an image interpolation

 ISSN: 2325-887X DOI: 10.22489/CinC.2024.138

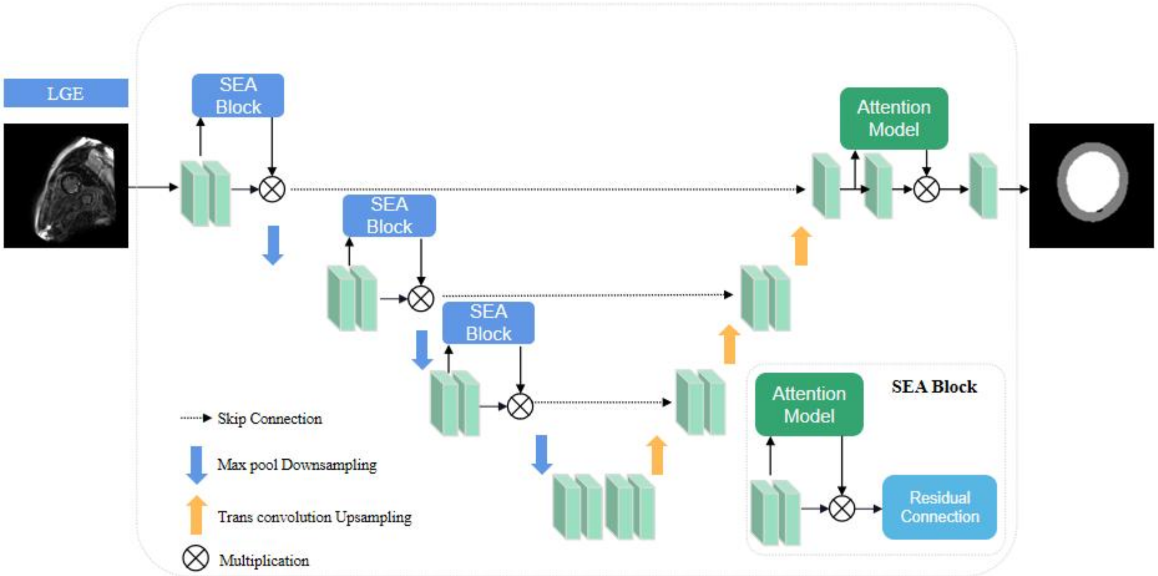

Figure 1. Overview of Segmented Network Architecture.

method that enhances the overall quality of the images.

2. Method

The fully automatic segmentation network in this study is built on the nnUNetv2 framework, on which we propose a residual attention mechanism to enhance the network's performance for LGE-CMR image segmentation. Moreover, we propose an image interpolation method to enhance the image quality.

2.1. nnUNet

nnU-Net is a generic U-Net network framework for automated design, proposed by Isensee et al. in 2019. nnU-Net's core idea is to simplify complex deep learning tasks by automating hyper-parameter searches and network architectural design, and to make it a standardized methodology applicable to different medical image segmentation tasks.

The nnU-Net framework is designed to have good generalization and adaptability. It is able to automatically adapt the model structure and hyperparameters to different types of datasets and tasks. Its performance stems from its systematic optimization and automated configuration of data preprocessing, network architecture, data augmentation, and loss functions.

2.2. Residual Attention Block

Although nnU-Net is more powerful, its segmentation performance for LGE-CMR images is not satisfactory. Therefore, we propose the residual attention mechanism to improve the segmentation performance of the whole network.

We integrate the residual attention module between the encoder and decoder in baseline to enhance the feature representation of the network at different resolutions. The module is able to introduce attentional weights in the multiscale feature maps, thus assigning higher weights to target regions and suppressing interference from irrelevant regions. By introducing the multi-scale attention mechanism, the model can focus on both global and local information.

The residual attention module is mainly composed of three branches. The first is the Feature Extraction Branch, which uses convolutional layers and activation functions for feature mapping. The second part is the Attention Branch, where we use spatial attention and channel attention mechanisms to generate the attention weight maps. The last part is the Residual Connection Branch, which passes the input features directly to the output layer and fuses them with the attention weight processed features.

And in the attention module, we use channel attention module and spatial attention module. The channel attention module is able to adaptively assign weights to different channels by computing attention for each channel of the feature graph. Here we use global average pooling and global maximum pooling and then generate the channel weights through a small multilayer perceptron (MLP) network. And on the spatial attention module, we assign weights to each pixel point of the feature map to focus on important regions in the image. We pool the feature maps over the channel dimension and then use a convolutional layer to generate the spatial attention weight maps.

2.3. Image interpolation on segmentation

Since LGE-CMR images are usually low-resolution, the network outputs the segmentation results with low-resolution output labels. This has a great impact on the subsequent modeling work. So we propose an image interpolation method to improve the overall image quality.

For LGE-CMR images, we use bilinear interpolation. This is a commonly used interpolation method for continuous data, which is able to maintain the smoothness of the image when scaling or resampling. And for labeled images, we use nearest neighbor interpolation. This method is commonly used for discrete categorical data to avoid non-integer errors in the interpolation process for category values.

3. Result

We analyzed LGE-CMR images from 150 patients (100 from EMIDEC, 5 from Anzhen Hospital, and 45 from MS-CMRseg2019). Our training and test set ratio was 7:3, with 105 patients in the training set and 45 patients in the test set. Our proposed method achieved a Dice similarity coefficient (DSC) of 85.19% for LV and 72.12% for Myo.

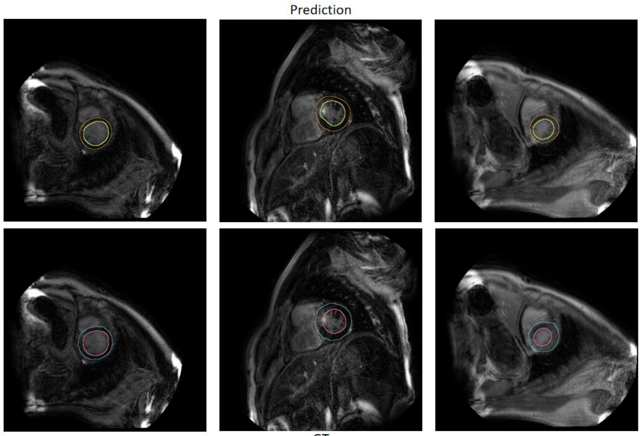

Figure 2. Comparison of network segmentation results with ground truth.

Figure 2 exhibits the segmentation results of our proposed network in comparison with the ground truth. From the results, it can be seen that our proposed network has a greater advantage in for segmenting the boundary of the left ventricular myocardium and can accurately segment the region of the left ventricular blood pool.

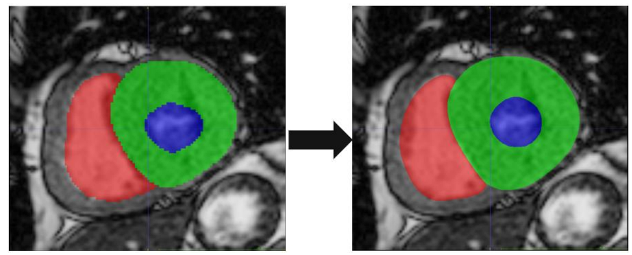

Figure 3. Comparison between original image and segmented labels before and after image interpolation.

From Fig. 3, we can see that the overall image quality has been greatly improved after the image interpolation method. And the boundary of the segmentation result is smoother and clearer. The method can provide great convenience for the subsequent modeling work.

4. Conclusion

In this paper, a fully automated LGE-CMR segmentation network is proposed. The network adds a residual attention mechanism to the framework of nnU-Net to combine global context information and local feature information to achieve more accurate segmentation. And the image interpolation method is introduced to improve the image quality. The results demonstrate that our proposed residual attention-based nnUNetv2 network can effectively improve the network's ability to segment LV and Myo in LGE-CMR images.

Acknowledgments

This research was supported by the Natural Science Foundation of China (62171408) and the the Key Research and Development Program of Zhejiang Province (2023C03088).

References

[1] D. Lloyd-Jones et al., "Heart disease and stroke statistics,"

Circulation, vol. 121, no. 7, pp. e46–e215, Jul. 2010.
[2] V. Delgado et al., "Relative merits of left ventricular dyssynchrony, left ventricular lead position, and myocardial scar to predict long-term survival of ischemic heart failure patients undergoing cardiac resynchronization therapy," Circulation, vol. 123, no. 1, pp. 70–78, Jan. 2011.
[3] D. B, D. S, R. J, S. P, R. P, G. S, A. H. Kadish, and J. J., "Infarct morphology identifies patients with substrate for sustained ventricular tachycardia," Am. Coll. Cardiol, vol. 45, no. 7, pp. 1104–1108, 2005.
[4] D. Deng, H. Arevalo, A. Prakosa, D. Callans, and N. Trayanova, "A feasibility study of arrhythmia risk prediction in patients with myocardial infarction and preserved ejection fraction," Europace, vol. 18, suppl 4, pp. iv60-iv66, 2016.
[5] M. Jacobs et al., "Automated Segmental Analysis of Fully Quantitative Myocardial Blood Flow Maps by First-Pass Perfusion Cardiovascular Magnetic Resonance," IEEE Access, vol. 9, pp. 52796-52811, 2021
[6] B. Hassan, T. Hassan, R. Ahmed, S. Qin and N. Werghi, "Automated Segmentation and Extraction of Posterior Eye Segment using OCT Scans," 2021 International Conference on Robotics and Automation in Industry (ICRAI), Rawalpindi, Pakistan, 2021, pp. 1-5.
[7] E. A. Fouzia, D. Aziz and O. Aziz, "Images Segmentation using Deep Learning Algorithms and Metaheuristics," 2022 8th International Conference on Optimization and Applications (ICOA), Genoa, Italy, 2022, pp. 1-6.
[8] Z. Zhou, M. M. R. Siddiquee, N. Tajbakhsh and J. Liang, "UNet++: Redesigning Skip Connections to Exploit Multiscale Features in Image Segmentation," IEEE Transactions on Medical Imaging, vol. 39, no. 6, pp. 1856-1867, June 2020.
[9] Isensee, F., Jaeger, P.F., Kohl, S.A.A. et al, "nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation," Nat Methods, vol. 18,pp. 203-211, 2021.
[10] Dosovitskiy A , Beyer L , Kolesnikov A ,et al. An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale. 2020.DOI:10.48550/arXiv.2010.11929..
[11] Chen J , Lu Y , Yu Q ,et al. TransUNet: Transformers Make Strong Encoders for Medical Image Segmentation. 2021.DOI:10.48550/arXiv.2102.04306.
[12] Z. Liu et al., "Swin Transformer: Hierarchical Vision Transformer using Shifted Windows," 2021 IEEE/CVF International Conference on Computer Vision (ICCV), Montreal, QC, Canada, 2021, pp. 9992-10002.
[13] Cao H , Wang Y , Chen J ,et al.Swin-Unet: Unet-like Pure Transformer for Medical Image Segmentation. 2021.DOI:10.48550/arXiv.2105.05537.
[14] X. Huang, Z. Deng, D. Li, X. Yuan and Y. Fu, "MISSFormer: An Effective Transformer for 2D Medical Image Segmentation," IEEE Transactions on Medical Imaging, vol. 42, no. 5, pp. 1484-1494, May 2023.

Address for correspondence:

Dongdong Deng.
School of Biomedical Engineering, Dalian University of Technology, Dalian, Liaoning, China.
dengdongdong@dlut.edu.cn