



U-Shaped Densely Connected Convolutions for Left Ventricle Segmentation from CMR Images

Khouloud Boukhris^{1,2} , Ramzi Mahmoudi^{2,3} , Asma Ben Abdallah² ,
Mabrouk AbdelAli⁴ , Badii Hmida⁴ , and Mohamed Hédi Bedoui² 

¹ Faculty of Sciences Monastir, University of Monastir, Monastir, Tunisia

² Faculty of Medicine Monastir, Medical Imaging Technology Lab – LTIM-LR12ES06, University of Monastir, Monastir, Tunisia

³ Gaspard-Monge Computer-Science Laboratory, Paris-Est University, Mixed Unit CNRS-UMLV-ESIEE UMR8049, BP99, ESIEE Paris City Descartes, 93162 Noisy Le Grand, France

⁴ Radiology Service- UR12SP40 CHU Fattouma Bourguiba, 5019 Monastir, Tunisia

Abstract. Segmentation of cardiac magnetic resonance images (cMRI) remains a challenging task in the field of scientific research due to its significance in the medical assessment of cardiovascular diseases. Ensuring accurate segmentation of the heart structures, mainly the left ventricle cavity, serves to extract important information and has a major impact on the quantitative analysis of the heart function which helps to conduct the proper diagnosis of doctors. The present paper introduces a simple and efficient U-shaped convolutional neural network aiming to accurately segment the LV from cMR images. We applied our architecture for Left Ventricle (LV) segmentation on cardiac MR images (cMRI), from the Automated Cardiac Diagnosis Challenge (ACDC). Obtained results are promising. This simple model based on CNN has significantly fewer parameters rendering it less demanding in terms of computation. Nevertheless, it has provided accurate segmentation. The tested method achieved LV Dice scores of 0.958 at end-systolic time (ES) and 0.979 at end-diastolic time (ED), which yields a mean Dice score of 0.968 on the ACDC dataset.

Keywords: cMRI Segmentation · CNN · U-Net · Dense net · Convolutions

1 Introduction

Cardiovascular diseases represent the leading cause of death according to the World Health Organization. Therefore, they have become a major healthcare issue over past years worldwide. There are different cardiac imaging techniques for viewing the heart structures that help in making the right diagnosis of these diseases. One of them is Cardiovascular magnetic resonance imaging (cMRI) which represents the current gold standard reference for assessing cardiac function [1]. Indeed, the accurate segmentation of the left ventricle (LV) from these cardiac images is required to retrieve information

on ventricular function, such as left ventricular end-systolic volume (LVESV), the left ventricle end-diastolic volume (LVEDV) and the left ventricle ejection fraction (LVEF) [2]. Consequently, major advances have been made in the field of cardiac image segmentation aiming to evaluate the heart function and establish the right diagnosis and treatment of cardiac diseases.

Before the advent of deep learning, a wealth of techniques had been developed to segment and evaluate the heart function from cardiovascular images including level sets, dynamic programming, active contour, graph cuts, and atlas registration [1, 3, 4]. These early approaches required significant manual intervention by the expert in order to achieve their goals. These first techniques may show promising results on limited datasets, but they generally tend to underperform on large variable datasets. In contrast, deep learning based approaches have proven to be able to overcome these limitations by automatically discovering intricate features from data for object detection and segmentation.

Convolutional neural networks (CNN), which were first introduced by Yann LeCun et al. in 1998 [5], are currently the most widely used techniques in the field of biomedical image classification and segmentation. U-Net [6], which is one of the most remarkable extensions of FCN [7] and therefore of CNN, has proven to be a gold-standard in the field of biomedical segmentation while achieving the highest accuracy [8]. U-Net has received much attention with the field of cardiovascular analysis in the last two years and therefore, several U-shaped architectures have been proposed in the literature for fully automated segmentation of the LV from cine MRI [9–14].

In this paper, we propose a fully automatic deep learning approach for left ventricle LV segmentation in cine MRI. Our proposed method is a U-Net-based architecture using Dense connections [15] in order to achieve fewer parameters while ensuring higher accuracy. This paper is organized as follows. A brief overview of related works is introduced in the next section. Then, the proposed method is presented in Sect. 3. Next, experimental results are provided in Sect. 4. And finally the conclusion and future work are drawn.

2 Related Works

U-Net [6], such as SegNet [16] and PspNet [17], is an encoder-decoder-based architecture that uses skip connections between encoder and decoder blocks. This skip connection consists of concatenating the high-level feature maps from the decoder and the low-level feature maps from the corresponding encoder which have the same spatial resolution (see Fig. 1). In the original U-Net, the encoder is down-sampled in total of 4 times, symmetrically to the decoder which is also up-sampled 4 times. This symmetry enables the model to restore the same size as the input image.

Wenjun Yan et al. [12] proposed a U-net-based method (OF-net) that integrates temporal information from cine MRI into LV segmentation. They incorporated an optical flow (OF) field to capture the cardiac motion towards adding temporal dimension. For this to happen, they used Res-Blocks [18] incrementing, thereby, the number of parameters and so the execution time.

Isensee et al. [11] used a 3D-U-Net inspired architectures for the segmentation of the left and the right ventricles at the end-systolic and the end diastolic time. Zhang

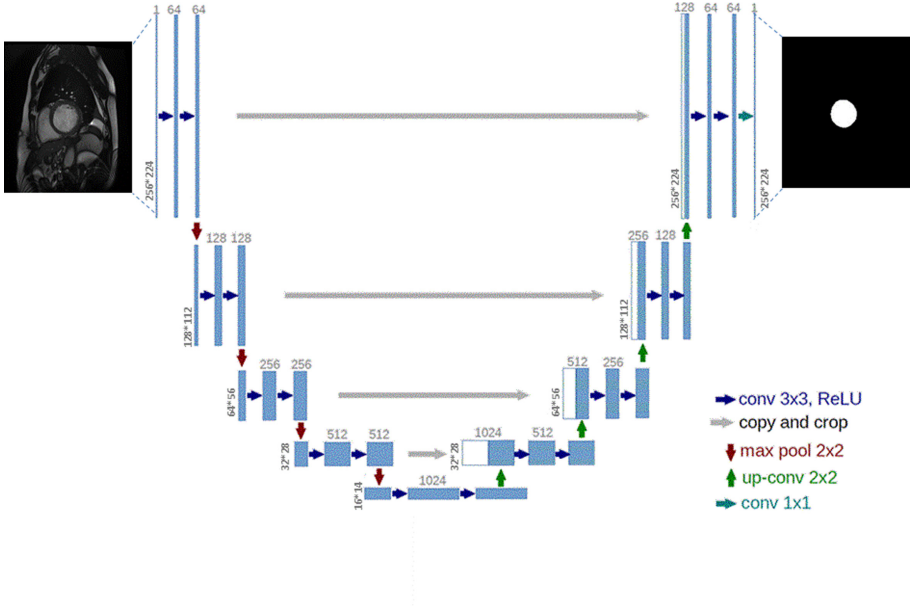


Fig. 1. An illustration of the original U-Net architecture towards LVC segmentation

et al. [19] also combined U-net with SE-Net model in order to reweight the channels of the feature map by giving higher weight to the relevant information and lower weight to the disabled one. Many approaches regarding U-Net have led to good results in LV segmentation from cMR images.

3 Proposed Method

3.1 Dataset

The dataset we adopted in this work is that of The Automated Cardiac Diagnosis Challenge (ACDC). It contains short-axis cMR images along with their corresponding ground truth images of Left Ventricle LV, LV myocardium, and Right Ventricle RV for 100 patients. The ACDC dataset results from clinical examinations acquired at the University Hospital of Dijon France [20].

The 100 patients of the ACDC dataset constitute a total number of 1902 labeled images at both end-systole (ES) and end-diastole (ED) time. In order to enable the evaluation of our method, we divided the labeled data into 80% and 20% which makes 1700 images for the training and 202 for the test. The giving dataset was divided into five subgroups according to the patient's pathology: 20 normal patients, 20 patients with previous myocardial infarction, 20 patients with dilated cardiomyopathy, 20 patients with hypertrophic cardiomyopathy and 20 patients with abnormal right ventricle. The training-test split we have just proposed maintains this subdivision, which means that the 202 test images are composed of four patients from each of these five subgroups.

It is to mention that the standard cMRI acquisition provides 8 to 12 slices from base to apex for each patient.

3.2 Preprocessing

The dataset given by the ACDC challenge has a wide variety of dimensions in the short-axis plane, ranging from 154×224 to 428×512 . Therefore, we resized all the dataset to 256×224 . In addition, the images present a wide range of pixel intensities, which might affect the performance of the segmentation model. To address this issue, we subtracted the mean value from each pixel and divided the result by the standard deviation thus ensuring the data normalization. In addition, as we are interested on segmenting the left ventricle, we applied a simple threshold on the ground truth images to keep only the LV cavity. We finally applied CLAHE [21] Contrast Limited Adaptive Histogram Equalization to enhance the local contrast of the images, which leads to better computational analysis.

3.3 Architecture

In this study we aim to achieve higher accuracy while considerably reducing the number of trainable parameters. For this to happen, we propose a U-shaped model using Dense Blocks for LV segmentation from cMR images. Our architecture is shown in the figure below (Fig. 2).

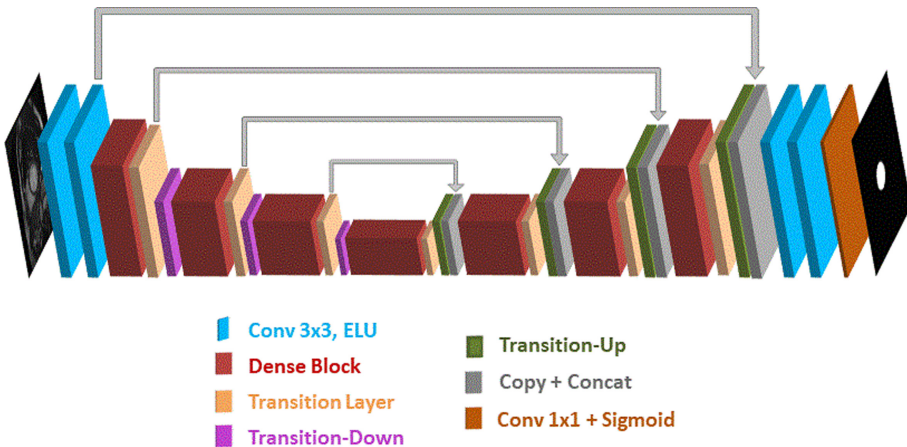


Fig. 2. Illustration of the proposed Dense U-Net

As with U-Net, our architecture is down-sampled then up-sampled symmetrically 4 times. In the first level, the input images are fed into two successive 3×3 unpadded convolutions using Exponential Linear Unit (ELU) and followed by a 2×2 max pooling operation with stride 2 for down-sampling.

The next levels are composed of Dense Blocks followed by Transition layers (same depth) that are down-sampled in the contraction path and up-sampled in the symmetric expanding path. Each Dense Block consists of four consecutive convolution layers having the same resolution, each followed by batch normalization (BN), Exponential Linear Unit (ELU) and a dropout layer of 0.2. The output of each convolution in the dense block is concatenated with the input of the following convolutions. The structure of a Dense Block followed by a Transition-Down is illustrated in the figure below (Fig. 3).

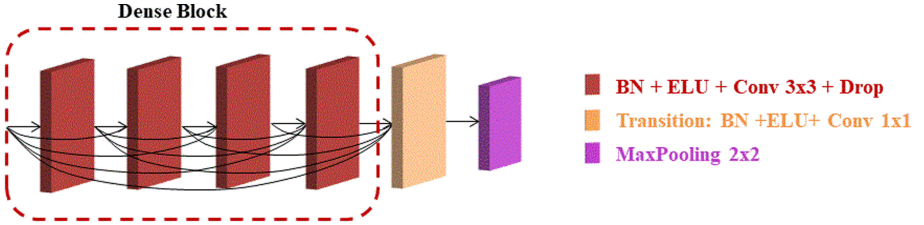


Fig. 3. Illustration of the Dense Block and the Transition-Down

In the contracting path, the filter size of the first dense block starts with 16 and is been duplicated after each down-sampling operation, whilst ensuring symmetry with the expanding path.

Eventually, to obtain the final binary segmentation, the resulting feature maps from the last 3×3 convolution layer of the proposed architecture, are agglomerated and averaged by employing a 1×1 convolution with a sigmoid activation to predict the probability of each output class. In our case, the number of classes is 1, indicating the LV (left Ventricle).

3.4 Post-processing

The resulting masks are resized to their initial dimensions. And no further post-processing is applied to the resulting segmented images.

3.5 Evaluation Metrics

Several metrics were used in order to evaluate the performance of our method, including accuracy, sensitivity, specificity and dice coefficient. To obtain these metrics, we first need to go through the computation of true Positive (TP), True Negative (TN), False Positive (FP) and False Negative (FN).

$$Accuracy = (TN + TP)/(TN + TP + FN + FP) \quad (1)$$

$$Sensitivity = TP/(TP + FN) \quad (2)$$

$$Specificity = TN/(TN + FP) \quad (3)$$

$$Dice\ coefficient = 2TP/(2TP + FP + FN) \quad (4)$$

4 Experiments and Results

The model was trained using binary cross-entropy as loss function and Adam [22] optimizer with its default parameters, starting with its default learning rate which is set to 0.001. We adopted the “reduce learning rate on the plateau” strategy with the aim of automatically reducing the learning rate. The learning rate was reduced by a constant factor of 0.1 when the loss metric has reached a plateau on the validation set, which varied the learning rate from 0.001 to $1e - 6$ over 32 epochs. For model evaluation, we have split the training data into validation and train and tracked binary cross entropy loss and Dice coefficient over the iterations (see Fig. 4). The percentage of the data that was held over validation is 10%.

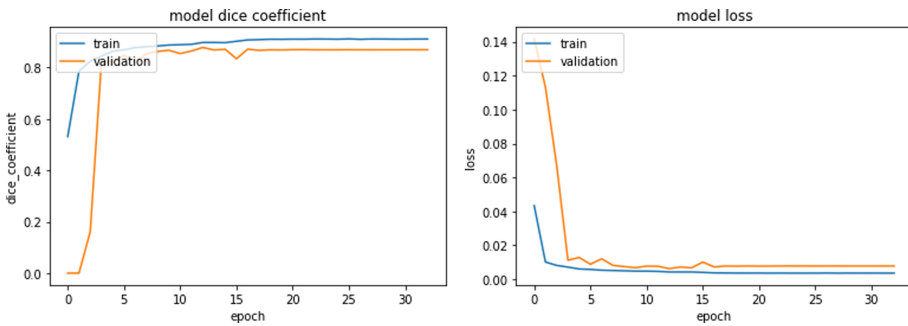


Fig. 4. Visualization of the proposed model history with training and validation

Table 1 presents the evaluation results of the two models (U-Net and the proposed U-shaped densely connected Convolutions) on the previously described test data (202 test images). Both Models were trained using the same preprocessing, the same post-processing and the same hyper parameters including loss function, batch size, learning rate and number of epochs. The U-Net architecture used in this comparison is detailed in the first figure (see Fig. 1).

Table 1. Comparison of LV segmentation performance in terms of Accuracy, Sensitivity, Specificity and Dice coefficient at the end-systolic (ES) and the end-diastolic (ED) time

| Method | Accuracy | | Sensitivity | | Specificity | | Dice | | Parameters |
|-----------------|----------|------|-------------|------|-------------|------|------|------|------------|
| | ES | ED | ES | ED | ES | ED | ES | ED | |
| U-Net | 0.98 | 0.98 | 0.73 | 0.81 | 0.97 | 0.98 | 0.79 | 0.81 | 31M |
| Proposed method | 0.99 | 0.99 | 0.97 | 0.98 | 0.99 | 0.99 | 0.95 | 0.97 | 3M |

The large margin of difference between the proposed networks and U-Net could be explained by the use of dense blocks in the lower levels of U-Net which enables extracting abundant local features via densely connected convolutional layers. This has played a

crucial role in improving the quality of the segmentation especially when dealing with basal and apical slices (see Fig. 5), in cMRI images, that are found to perform poorly with U-Net and other existing methods in the literature. Basal and apical slices have always been challenging in the literature when it comes to left ventricular segmentation. It is worth mentioning that this improvement is achieved despite a reduced number of trainable parameters that is divided by 10 when compared with U-Net.

As it may be observed, the number of trainable parameters has decreased from 31 million parameters with U-Net to only 3 million parameters with the proposed method. Our model is less computationally intensive and therefore helps to gain in terms of time.

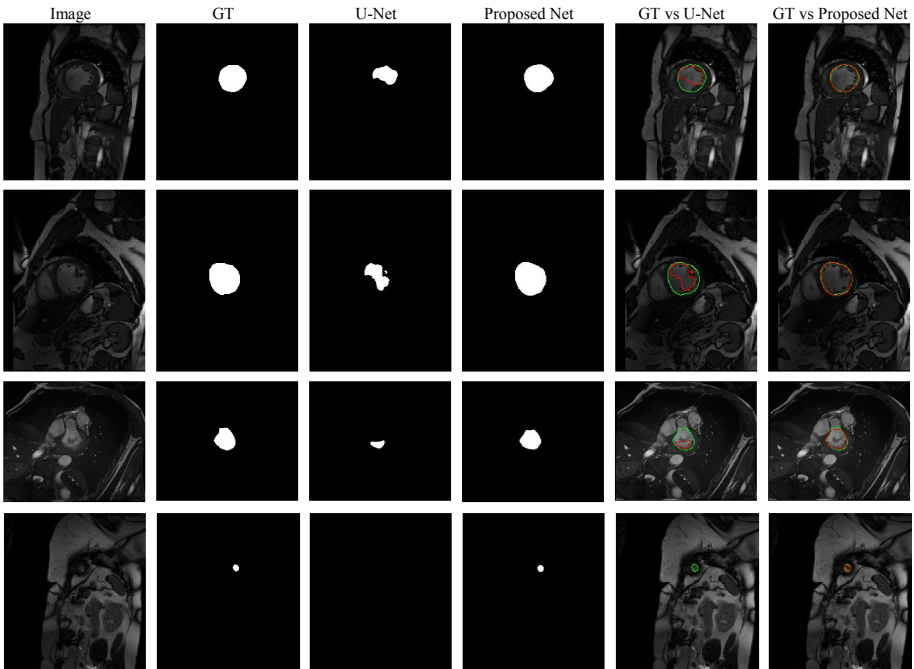


Fig. 5. Qualitative segmentation results of U-Net and the proposed model on the ACDC dataset. The experimental results show that the proposed Dense-U-shaped-Net yields better segmentation masks than the original U-Net especially when dealing with basal and apical slices. The ground truth (GT) is delineated in green color with both comparisons.

Even though we established our test on 20% of the ACDC training data, we conducted a comparison with existing state-of-the-art methods set for the left ventricle (LV). Table 2 shows that our approach outperforms other existing methods.

Table 2. Comparison of LV segmentation performance of the proposed method with the state of the art in terms of Dice coefficient at the end-systolic (ES) and the end-diastolic (ED) time

| Method | Dice | |
|------------------------|------|------|
| | ES | ED |
| Isensee et al. [11] | 0.93 | 0.96 |
| Simantiris et al. [14] | 0.92 | 0.96 |
| Proposed Method | 0.95 | 0.97 |

5 Conclusion

In this paper, a simple efficient method for segmenting LV cMR images is proposed. Experimental results on the ACDC dataset show that our U-shaped method with densely connected Convolutions has proven its ability to enhance the performance of cardiac MRI segmentation compared to other existing methods. The use of dense blocks enables the model extracting abundant features, which led to achieve impressive performance. This improvement is provided with reduced number of trainable parameters compared with other existing approaches that make it less time consuming. The obtained results demonstrated the effectiveness of our proposed method in performing precise LV segmentation, which may help establishing an early diagnosis of heart diseases. Further studies could include combining dilated convolutions and dense connections to learn features at different scales.

References

1. Petitjean, C., Dacher, J.-N.: A review of segmentation methods in short axis cardiac MR images. *Med. Image Anal.* **15**(2), 169–184 (2011). <https://doi.org/10.1016/j.media.2010.12.004>
2. White, H.D., Norris, R.M., Brown, M.A., Brandt, P.W., Whitlock, R.M., Wild, C.J.: Left ventricular end-systolic volume as the major determinant of survival after recovery from myocardial infarction. *Circulation* **76**(1), 44–51 (1987). <https://doi.org/10.1161/01.CIR.76.1.44>
3. Pluempitiwiriyaewej, C., Moura, J.M.F., Lin Wu, Y.-J., Ho, C.: STACS: new active contour scheme for cardiac MR image segmentation. *IEEE Trans. Med. Imaging* **24**(5), 593–603 (2005). <https://doi.org/10.1109/TMI.2005.843740>
4. 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.* **43**(6(Part 1)), 2741–2755 (2016)
5. Lecun, Y., Bottou, L., Bengio, Y., Haffner, P.: Gradient-based learning applied to document recognition. *Proc. IEEE* **86**(11), 2278–2324 (1998). <https://doi.org/10.1109/5.726791>
6. Ronneberger, O., Fischer, P., Brox, T.: U-Net: convolutional networks for biomedical image segmentation. In: Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F. (eds.) *MICCAI 2015*. LNCS, vol. 9351, pp. 234–241. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-24574-4_28

7. Long, J., Shelhamer, E., Darrell, T.: Fully convolutional networks for semantic segmentation, pp. 3431–3440 (2015). Accessed 28 Oct 2020
8. Rizwan, I., Haque, I., Neubert, J.: Deep learning approaches to biomedical image segmentation. *Inform. Med. Unlocked* **18**, 100297 (2020)
9. Jang, Y., Hong, Y., Ha, S., Kim, S., Chang, H.-J.: Automatic segmentation of LV and RV in cardiac MRI. In: Pop, M., et al. (eds.) STACOM 2017. LNCS, vol. 10663, pp. 161–169. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-75541-0_17
10. Khened, M., Alex, V., Krishnamurthi, G.: Densely connected fully convolutional network for short-axis cardiac cine MR image segmentation and heart diagnosis using random forest. In: Pop, M., et al. (eds.) STACOM 2017. LNCS, vol. 10663, pp. 140–151. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-75541-0_15
11. Isensee, F., Jaeger, P.F., Full, P.M., Wolf, I., Engelhardt, S., Maier-Hein, K.H.: Automatic cardiac disease assessment on cine-MRI via time-series segmentation and domain specific features. In: Pop, M., et al. (eds.) STACOM 2017. LNCS, vol. 10663, pp. 120–129. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-75541-0_13
12. Yan, W., Wang, Y., van der Geest, R.J., Tao, Q.: Cine MRI analysis by deep learning of optical flow: adding the temporal dimension. *Comput. Biol. Med.* **111**, 103356 (2019). <https://doi.org/10.1016/j.combiomed.2019.103356>
13. He, Y., et al.: Automatic left ventricle segmentation from cardiac magnetic resonance images using a capsule network. *J. X-Ray Sci. Technol.* **28**(3), 541–553 (2020)
14. Simantiris, G., Tziritas, G.: Cardiac MRI segmentation with a dilated CNN incorporating domain-specific constraints. *IEEE J. Sel. Top. Signal Process.* **14**(6), 1235–1243 (2020). <https://doi.org/10.1109/JSTSP.2020.3013351>
15. Huang, G., Liu, Z., van der Maaten, L., Weinberger, K.Q.: Densely connected convolutional networks, pp. 4700–4708 (2017). Accessed 01 May 2021
16. Badrinarayanan, V., Kendall, A., Cipolla, R.: SegNet: a deep convolutional encoder-decoder architecture for image segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **39**(12), 2481–2495 (2017). <https://doi.org/10.1109/TPAMI.2016.2644615>
17. Zhao, H., Shi, J., Qi, X., Wang, X., Jia, J.: Pyramid scene parsing network, pp. 2881–2890 (2017). Accessed 01 May 2021
18. He, K., Zhang, X., Ren, S., Sun, J.: Deep Residual Learning for Image Recognition, pp. 770–778 (2016). Accessed 02 May 2021
19. Zhang, J., Du, J., Liu, H., Hou, X., Zhao, Y., Ding, M.: LU-NET: an improved U-Net for ventricular segmentation. *IEEE Access* **7**, 92539–92546 (2019). <https://doi.org/10.1109/ACCESS.2019.2925060>
20. Bernard, O., et al.: Deep learning techniques for automatic MRI cardiac multi-structures segmentation and diagnosis: is the problem solved? *IEEE Trans. Med. Imaging* **37**(11), 2514–2525 (2018). <https://doi.org/10.1109/TMI.2018.2837502>
21. Zuiderveld, K.: Contrast limited adaptive histogram equalization. *Graph. Gems* **4**, 474–485 (1994)
22. Kingma, D.P., Ba, J.: Adam: A Method for Stochastic Optimization (2017). <https://arxiv.org/abs/1412.6980>. Accessed 03 May 2021