Proceedings of the 11th World Congress on Electrical Engineering and Computer Systems and Sciences (EECSS'25)
Paris, France – August, 2025
Paper No. ICBES 205
DOI: 10.11159/icbes25.205

Revealing End-Systolic Right Ventricle Segmentation Strengths of EfficientNetB3 in DeepLabv3+

Muniba Ashfaq[1], Zoran Bosnić[1]
[1]University of Ljubljana, Faculty of Computer and Information Science, Večna pot 113, 1000 Ljubljana, Slovenia
muniba.ashfaq@fri.uni-lj.si; zoran.bosnic@fri.uni-lj.si

***Abstract* -** Accurate segmentation of cardiac anatomical structures in cardiac magnetic resonance imaging (MRI) is vital for early diagnosis and treatment planning in cardiovascular diseases. In particular, the right ventricle (RV) during the end systolic (ES) phase is critical, as RV size is a strong indicator of cardiovascular health. Unlike the LV, the RV has a more complex geometry and thinner walls, making it difficult to delineate even manually. We propose to evaluate the performance of DeepLabv3+ using different backbone networks including EfficientNetB3, ResNet50, ResNet101, DesNet121, Xception, InceptionV3, VGG16, and VGG19 for multi-class segmentation of left ventricle (LV), right ventricle (RV), and myocardium (MYO). EfficientNetB3 as a backbone architecture in DeepLabv3+ outperformed with average score of Dice (0.913) and Jaccard (0.84). Moreover, it demonstrated the best performance in segmenting the RV during the challenging end-systolic phase structure often misclassified as MYO. This highlights the clinical potential of EfficientNetB3-integrated DeepLabv3+ for end-systolic challenging RV delineation.

***Keywords*:** Cardiac image segmentation, ACDC, DeepLabv3+.

1. Introduction

Cardiovascular diseases are one of the main causes of mortality in Europe. According to the health in the European Union - facts and figures [1], 1.71 million deaths are due to diseases of the circulatory system. A decision support system for the early diagnosis of fatal diseases can improve healthcare efficiency and reduce the economic burden of hospitalization. The improved non-invasive approach involves the advancements in medical diagnostic imaging and techniques to assess the qualitative and quantitative measurements of heart anatomical structures. The major anatomical structures of heart include LV, RV, and MYO [2]. These anatomical measurements are further used to diagnose a particular cardiac disease [3]. Multiple imaging modalities are used to annotate the anatomic structures, especially cardiac MRI. However, manual annotation of cardiac MRI is time-consuming, subjective, and prone to inter-observer variability. Deep learning–based semantic segmentation techniques, particularly DeepLabv3+ [4], have shown promise in automating this task.

2. Methodology

We used DeepLabv3+ architecture, which is designed to handle the dual challenge of capturing multi-scale contextual features and recovering spatial resolution for accurate segmentation. DeepLabv3+ combines atrous spatial pyramid pooling (ASPP) with an encoder-decoder structure. The ASPP module extracts rich contextual features using atrous convolutions at multiple scales, while the decoder module refines object boundaries through progressive upsampling and fusion, as shown in Fig. 1.

We evaluate and compare the performance of DeepLabv3+ integrated with different ImageNet pretrained backbone Deep Convolutional Neural Networks (DCNN) including EfficientNetB3 [5], ResNet50, ResNet101 [6], DesNet121 [7], Xception [8], InceptionV3 [9], VGG16, and VGG19 [10]. Our objective is to identify the most effective feature extractor for accurate delineation of key cardiac structures. The encoder begins by processing the input cardiac MRI images which, due to their native grayscale format, are replicated across three channels to meet the input dimensionality requirements of the pre-trained backbones. Prior to training, images undergo preprocessing including intensity normalization, and resizing to 256x256 spatial resolution.

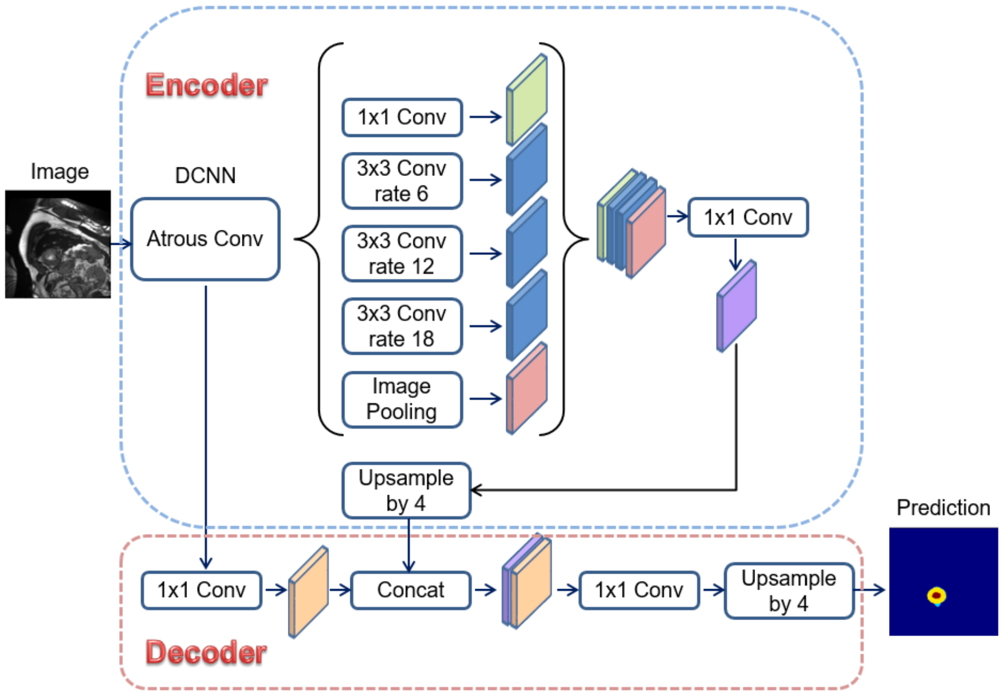

Fig. 1: DeepLabv3+ architecture with varying DCNN as backbone architecture

3. Results

The proposed method is evaluated on the Automated Cardiac Diagnosis Challenge (ACDC) [11] dataset for segmentation of RV, MYO, and LV. It consists of cine-MRI scans from 150 patients. Each patient have two annotated volumes (ED and ES) with 8 to 10 slices per volume. The dataset contains five subgroups: normal, myocardial infarction, dilated cardiomyopathy, hypertrophic cardiomyopathy, and abnormal right ventricle. It is divided into 100 patients for training and 50 for testing, resulting in 1902 training slices. We split the training set into 80% (1521 slices) for training and 20% (381 slices) for validation, while the test set consists of 1076 slices. All images are resized to 256×256 and normalized. We employ DeepLabv3+ with EfficientNetB3 backbone, trained using the Adam optimizer, sparse categorical cross-entropy loss, a batch size of 8, and 50 epochs. The loss and accuracy curves are shown in the Fig. 3 for 50 epochs. The model demonstrates an improved segmentation accuracy as shown in Fig. 2, particularly for the small and challenging RV structure in the ES phase, while other models misclassify RV as MYO. Table 1 and 2 show the average multi-class Dice and Jaccard score respectively of EfficientNetB3 based DeepLabv3+. The overall performance of EfficientNetB3 based DeepLabv3+ in multi-class segmentation is the best as compared to the other models.

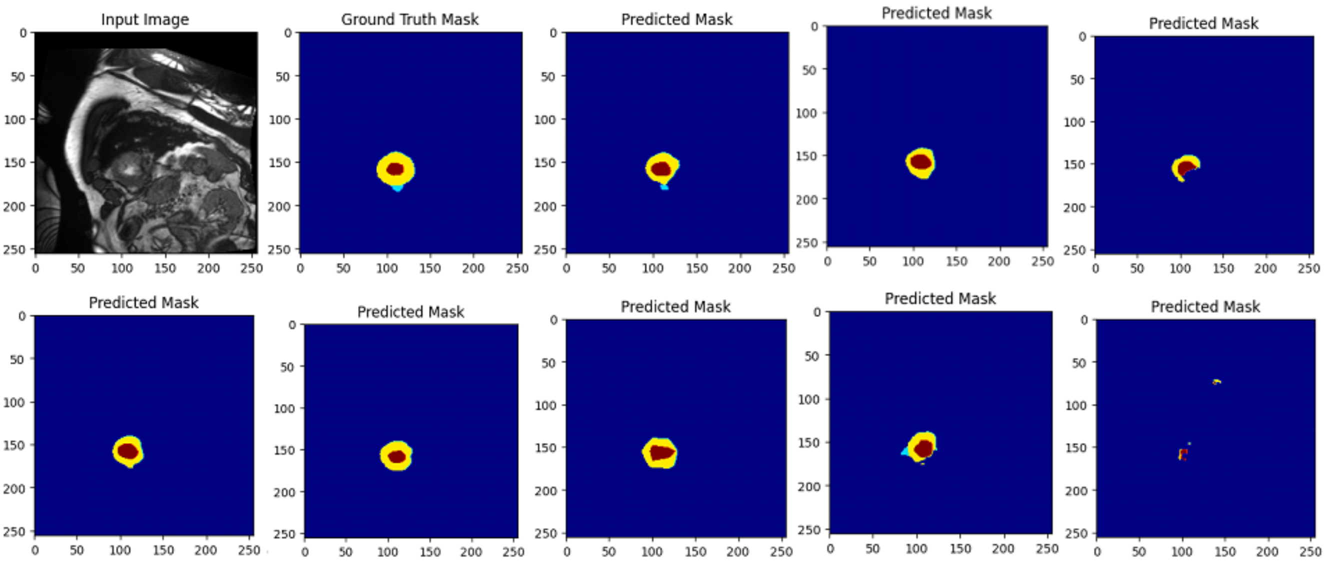

Fig. 2: Top row from left to right input image, ground truth mask, predicted mask from EfficientNetB3, ResNet50, ResNet101. Bottom row from left to right predicted mask from DesNet121, Xception, InceptionV3, VGG16 and VGG19

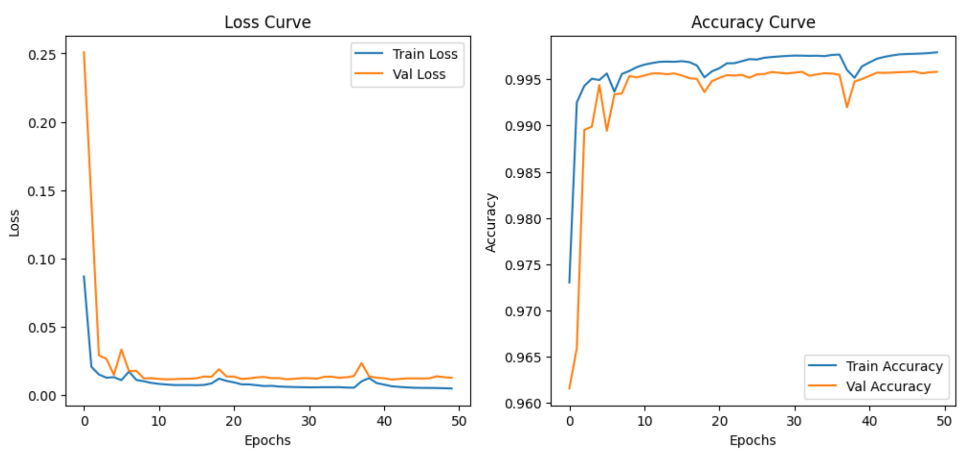

Fig. 3: Loss and Accuracy curve of EfficientNetB3 based DeepLabv3+ with epochs = 50

Table 1: Average Dice score of different DCNN as backbone network in DeepLabv3+

EfficientNetB3	**ResNet50**	**ResNet101**	**DesNet121**	**Xception**	**Inceptionv3**	**VGG16**	**VGG19**
0.913	0.904	0.902	0.911	0.910	0.864	0.891	0.887

Table 2: Average Jaccard score of different DCNN as backbone network in DeepLabv3+

EfficientNetB3	**ResNet50**	**ResNet101**	**DesNet121**	**Xception**	**Inceptionv3**	**VGG16**	**VGG19**
0.840	0.825	0.824	0.838	0.837	0.763	0.806	0.798

4. Conclusion

We analysed the performance of DeepLabv3+ with different backbone DCNN in both ED and ES phases. Quantitative evaluation with the average Dice (0.913) and Jaccard (0.840) scores showed that EfficientNetB3 outperformed other backbone networks. The most significant advantage was observed in the segmentation of the RV during the ES phase with a small structure misclassified by other models as MYO. The results indicate that EfficientNetB3 is a promising backbone for clinically relevant segmentation tasks that require precision across cardiac phases.

Acknowledgements

This publication is supported by the European Union's Horizon Europe research and innovation programme under the Marie Sklodowska-Curie Postdoctoral Fellowship Programme, SMASH co-funded under the grant agreement No. 101081355. The operation (SMASH project) is co-funded by the Republic of Slovenia and the European Union from the European Regional Development Fund. The authors gratefully acknowledge the HPC RIVR consortium [(www.hpc-rivr.si)](https://www.hpc-rivr.si) and EuroHPC JU [(eurohpc-ju.europa.eu)](https://eurohpc-ju.europa.eu/) for funding this research by providing computing resources of the HPC system Vega at the Institute of Information Science [(www.izum.si)](https://www.izum.si/en/home/).

References

[1] eurostat: Cardiovascular diseases statistics [Online]. Available: https://ec.europa.eu/eurostat/statistics-explained, last accessed 2025/04/03

[2] C. Chen, C. Qin, H. Qiu, G. Tarroni, J. Duan, W. Bai, and D. Rueckert, "Deep learning for cardiac image segmentation: a review," *Frontiers in cardiovascular medicine* 7, 25, 2020.

[3] A. Sirajuddin, S.M. Mirmomen, S.J. Kligerman, D.W. Groves, A.P. Burke, F. Kureshi, C.S. White, and A.E Arai, "Ischemic heart disease: noninvasive imaging techniques and findings," *Radiographics* 41(4), 990–1021, 2021.

[4] L.C. Chen, Y. Zhu, G. Papandreou, F. Schroff, and H. Adam, "Encoder-decoder with atrous separable convolution for semantic image segmentation," in *Proceedings of the European conference on computer vision (ECCV)*, pp. 801–818, 2018.

[5] M. Tan, and Q. Le, "Efficientnet: Rethinking model scaling for convolutional neural networks," in *International conference on machine learning*, pp. 6105-6114, 2019.

[6] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770–778, 2016.

[7] G. Huang, Z. Liu, L. Van Der Maaten, and K.Q. Weinberger, "Densely connected convolutional networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 4700-4708, 2017.

[8] F. Chollet, "Xception: Deep learning with depthwise separable convolutions," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1251-1258, 2017.

[9] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 2818-2826, 2016.

[10] K. Simonyan, and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv*, 1409.1556, 2014.

[11] O. Bernard, A. Lalande, C. Zotti, F. Cervenansky, X. Yang, P.A. Heng, I. Cetin, K. Lekadir, O. Camara, M.A.G. Ballester, and G. Sanroma, "Deep learning techniques for automatic MRI cardiac multi-structures segmentation and diagnosis: is the problem solved?," in *IEEE transactions on medical imaging*, 37(11), pp.2514-2525, 2018.