IJISET - International Journal of Innovative Science, Engineering & Technology, Vol. 7 Issue 10, October 2020
ISSN (Online) 2348 – 7968 | Impact Factor (2020) – 6.72
www.ijiset.com

Deep Learning Based Model Architectures for Cardiac MRI Segmentation: A Survey

Surayya Ado Bala [1], Shri Kant [2]

[1]Department of Computer Science and Engineering, Sharda University,
Greater Noida, India, surayyaadob@gmail.com

[2]Research and Technology Development Centre, Sharda University,
Greater Noida, India, shrikant.ojha@gmail.com

Abstract

Cardiac MRI segmentation is one of the most critical tasks in cardiac MRI analysis. The segmented parts are used for calculating the cardiac parameters for the diagnostic of cardiovascular diseases. With the recent advancement in deep learning, many deep learning-based model architectures have been proposed by many researchers with different methodologies and optimizations of those models continuing to be an active research area. Deep Fully Convolutional Neural Networks (FCN) using 2D and 3D architecture have achieved the state of the art segmentation performance in cardiac MRI, but each architecture has some challenges. 2D convolutions cannot fully leverage the spatial information along the third dimension, while 3D convolutions suffer from high computational cost and memory consumption. This paper aims to discuss the various models used for cardiac MRI segmentation, their challenges, and how to address them. We conclude by evaluating how to measure the success of the segmentation.

Keywords: *Cardiac MRI, Deep learning, Fully Convolutional Neural Network, MRI Segmentation*

1. Introduction

Segmentation of essential structures from cardiac MRI is a prerequisite for cardiac MRI analysis. Accurate delineation of the structures is required to assess cardiac parameters such as stroke volumes, ejection fraction, left and right ventricular volumes, left ventricle mass, and the myocardium thickness[1]. Initially, expert radiologist delineates the cardiac MRI manually, which is very tedious, time-consuming, and subjected to intra and inter-observer variability[2]. Hence efficient, accurate, automated model for segmentation is needed to help facilitate the diagnosis of heart diseases.

Various difficulties are identified in [3-5], including (1) Variability from different institutions and patients. (2) The a very significant disparity in terms of pixels that belongs to the object class and the background. (3) Instability in the shape and intensity of structures from different patients and pathologies (4) Intrinsic noise causes motion artifacts and heart dynamics.

Even though these difficulties have been addressed by some Non-deep learning-based models, which include deformable models [6], [3], image-based model [7], graph-based model (graph cut) [8], atlas-based model [9], active shape and appearance model [10]. Hence, improvement is needed for it to be usable in a clinical setting. These models relied totally on handcrafted engineering features that required domain knowledge and expert. The handcrafted features have little representational capability to deal with the significant differences of anatomical structures in their appearance and shapes. Hence, deep learning-based models have been researched to seek more profoundly and useful features to solve these problems.

Deep learning variates artificial neural networks with many layers that extract a hierarchy of features from raw input images—this network work based on a mimic of human brain neurons [11]. Deep learning-based models have hyper-parameters such as activation function, learning rate, and the number of neurons, which are beneficial elements to scrutinize during model designing.

Recent researches in a medical image of cardiac MRI has shown the performance and the capabilities of the convolutional neural network and its variant in segmenting cardiac MRI. CNN achieved this success because it learns the hierarchical representation of both the 2D and 3D data without relying on handcrafted data.

2. Deep Learning Architectures for Cardiac MRI Segmentation

A deep learning-based model can be classified into two groups broadly: 2D Fully Convolution Network (2D FCN, such as U-Net architecture [12] and 3D Fully Convolutional Network (3D FCN) [13], where 2D convolutions are substituted with 3D convolution.

2.1 2D FCN: A Fully Convolutional Neural Network for Cardiac Segmentation in Short Axis MRI

FCN is a 2D deep learning-based model architecture, as illustrated in Fig.1 below. The architecture of FCN is made up of 15 stacked convolution layers and three layers of overlapping, two-pixel stride max pooling. Each convolution is followed by Rectified Linear Unit (RELU) and Mean-Variance Normalization (MVN). The architecture has almost 11 million parameters with a high tendency to overfit because of the high dimension [14].

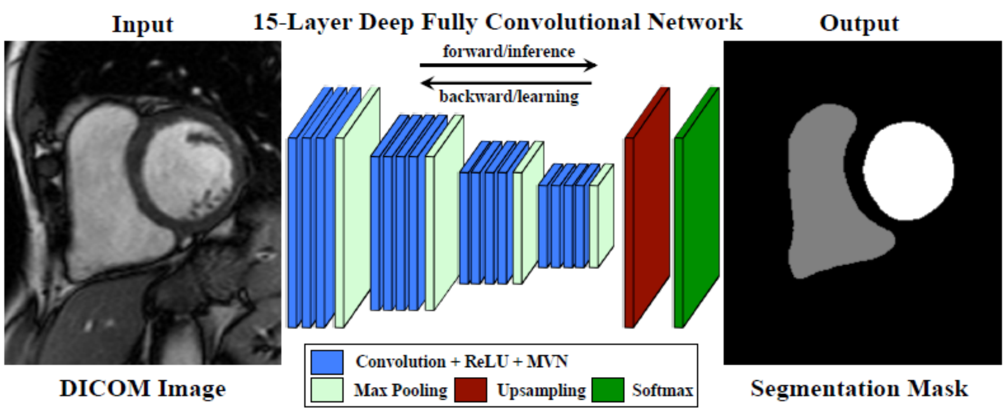

Fig. 1 2D FCN Architecture [14]

2.2 U-Net: Convolutional Network for Biomedical Image Segmentation

It is a variant of FCN, a 2D deep learning-based model. Fig. 2 shows that the architecture consists of upsampling (contracting path) at the left side, the bottleneck, and downsampling (expensive path) on the right side. The contracting contains many contracting blocks which take an input and applies two 3 x 3 convolution layer followed by a Rectified Linear Unit Layer (RELU), a 2x2 max pooling. After each contracting block, the number of feature maps doubles itself to enable the architecture to learn the complex structures. The bottleneck lies in the middle of the contracting path and the expansion path. The expansion path uses 3x3 CNN layers followed by a 2 x 2 up convolution layer. The architecture was trained end-to-end with a limited dataset that relies on substantial data augmentation [12].

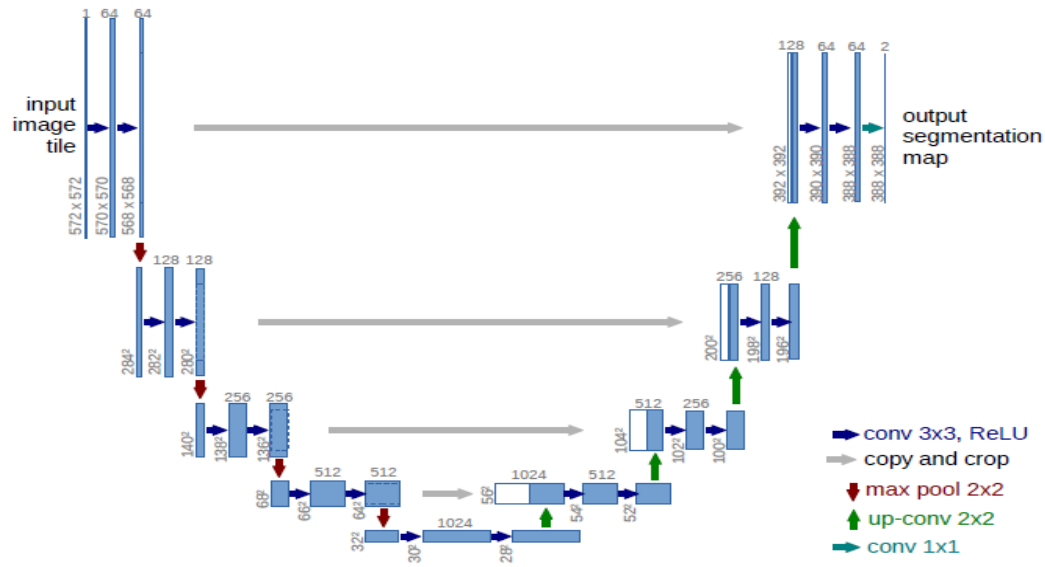

Fig.2 2D Unet [12]

2.3 3D U-Net: 3D Deeply-Supervised U-Net Based Whole Heart Segmentation

This is a deep learning-based model using deep supervised 3D U-Net for fully automating whole heart segmentation. The architecture is shown in Fig. 3 is based on coarsely detecting the whole heart and segmenting it into the region of interest (ROI) using a 3D U-Net trained using 3D image data. The 3D U-Net is accomplished by replacing the 2D layers with 3D layers [15].

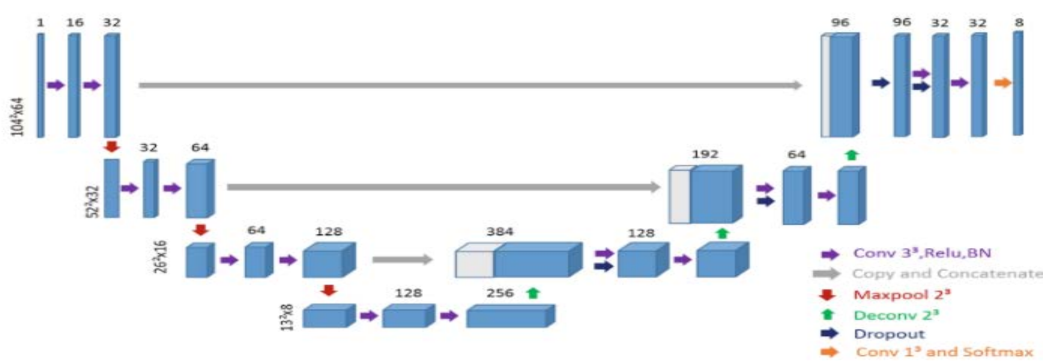

Fig. 3 3D UNET ARCHITECTURE [15]

IJISET - International Journal of Innovative Science, Engineering & Technology, Vol. 7 Issue 10, October 2020
ISSN (Online) 2348 – 7968 | Impact Factor (2020) – 6.72
www.ijiset.com

2.4 3D FCN: Deeply-Supervised 3D FCN

This model is based on 3D FCN. The 2D FCN extended to 3D volumetric data to ensure an efficient volume to volume inference, as shown in Fig. 4. The model employs deep supervision to speed up the network's convergence and exploits 3D – HOL (dilated convolutional layer) layers.

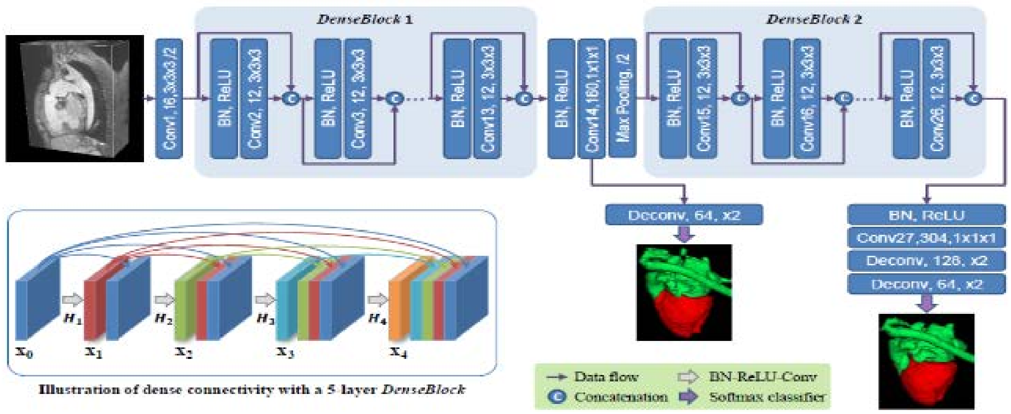

Fig. 4 3D FCN Architecture [13]

2.5 3D Fractalnet: Dense Volumetric Segmentation for Cardiovascular MRI Volumes

3D Fractal Net [16] is a 3D deeply supervised deep learning based model architecture based on fully convolutional architecture [17] which appraise 3D image information. The architecture segments the whole heart and great vessel in MRI volumes using multi-paths with different receptive fields. The architecture shown in Fig. 5 below is organized in a self-similar fractal scheme to apprehend the myocardium and vessels' hierarchical features. This architecture is designed with a deep supervision scheme to alleviate the vanishing gradient problems during the training process, which also serves as a regularization technique to improve the dataset's training performance [18].
3D Fractal Net and other 3D ConvNets such as 3D U-Net [19], the model is designed to generate many features channels generated in each layer with many parameters to be tuned during training. Although these networks introduce different skip connections to ease the training, an efficient model with limited MR images for heart segmentation is still challenging.

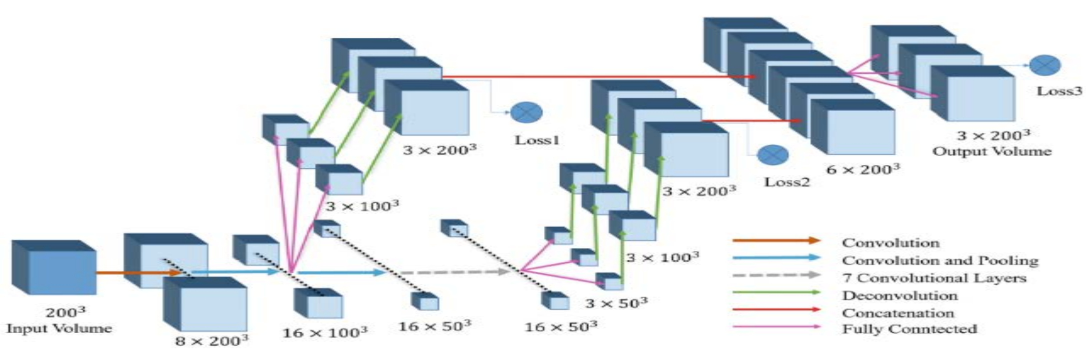

Fig.5 3D FRACTALNET ARCHITECTURE [16]

2.6 Densevoxnet: Automatic 3D Cardiovascular MRI Segmentation with Densely-Connected Volumetric Convnets

DenseVoxNet is a 3D deep learning based model that adopts the 3D fully convolutional architecture for robust volume-to-volume prediction. As shown in Fig. 6 below, the architecture was proposed to eliminate the limited data problem of 3D ConvNets to segment the cardiac and vascular structures in the cardiac MRI images. It has up and down sampling components enclosed in densely connected blocks to achieve end-to-end learning. The blocks in the architecture comprise of 12 transformations with a dense connection. Each transformation is sequentially composed of 3 x 3 x 3 convolutional layers, a RELU, and a growth rate k [20]. The DenseVoxNet proposed by [21] uses dense connectivity of learning perspective due to its numerous advantages. These include; direct connections within all the layers in the architecture, fewer feature maps in each layer, which are essential during training ConvNets with minimal data. The network architecture substantially minimizes over-fitting problems compared to other architecture.

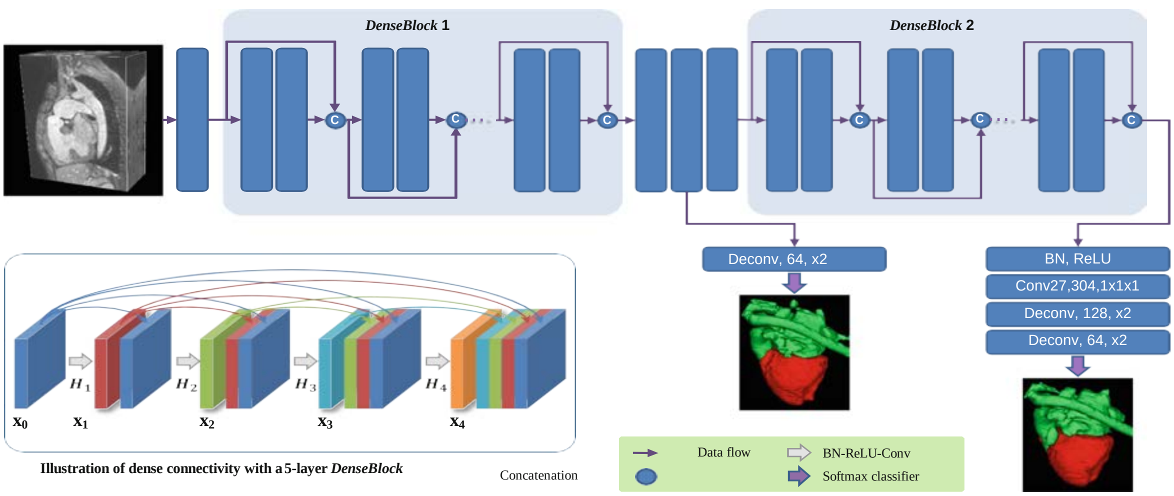

Fig.6 DENSEVOXNET ARCHITECTURE [20]

3. Challenges and Way Forward

Choosing a suitable deep learning architecture model and its corresponding hyper-parameters for different modalities for getting a good result is a great challenge. By comparing the 3D and 2D deep learning model based architectures, there are some challenges that each of them faces. Firstly, 3D remarkably reduces the number of training data. Secondly, 3D relies on 3D convolution, which may compromise the information in intermediate representations. Thirdly, 3D suffers from memory consumption that limits the depth of the network as well as the filter's field of view [22]. At the same time, 2D convolutions cannot fully leverage the spatial information along the third dimension.
Despite its benefits, deep learning-based models face certain complications during training. First, voluminous training data is required to train models to avoid the over-fitting problem. Second, the required data is challenging to obtain in the medical field due to an expert annotated disease-specific dataset cost. These challenges made some researchers use either small or unannotated datasets. These unbalanced datasets are inaccurate predictions, as a result, are biased towards the majority samples, thus leading to overfitting. Various strategies are employed by researchers to minimize over-fitting as well

as small training dataset problems. A commonly used method is data augmentation [23] to increase the size of the dataset artificially.

Presently, many researchers have paid more attention to deep learning-based models and have obtained relatively good results. Yet, there is still a need to address the problems mentioned above.

To address these problems, we need to propose a hybrid deep learning-based model that will explore the intra slice and interslice features, which will keep the computational assets of 2D and solve the problems above in 3D models.

Deep learning models can be optimized by tuning hyper-parameters such as learning rate, network architectures, activation functions, and more. These parameters are essential for controlling learning behavior and must be determined before the training process.

4. Metrics for Segmentation Accuracy

Different evaluation metrics are used to measure the performance of deep learning models used in medical image segmentation. In cardiac MRI segmentation, Dice Metric and the Hausdorff Distance are generally used. While the Dice metric calculates the spatial overlap between two discretely labeled objects, the Hausdorff distance maximizes two labeled contours.

5. Conclusion

Researchers in medical imaging continues to show interest in applying deep learning models to delineate cardiac MRI. Even though many successes have been achieved, especially in accuracy and speed, the models did not reach the perfection stage. Both 2D and 3D architecture exhibit some limitations. While 3D degrades intermediate layer information and high memory consumption, the 2D architecture does not leverage third dimension spatial information. Selecting hyper-parameters values poses additional challenges in choosing a model to use. On a larger scale, the benefits of using deep learning models for cardiac MRI segmentation proved promising. Furthermore, with continuous research and perfection of existing models, we expected a dramatic improvement in accuracy and performance in the near future.

References

[1] P.Peng., et al., A review of heart chamber segmentation for structural and functional analysis using cardiac magnetic resonance imaging. 2016. 29(2): p. 155-195.
[2] C.A.Miller et al., Quantification of left ventricular indices from SSFP cine imaging: Impact of real-world variability in analysis methodology and utility of geometric modeling. 2013. 37(5): p. 1213-1222.
[3] S.Queirós et al., Fast automatic myocardial segmentation in 4D cine CMR datasets. 2014. 18(7): p. 1115-1131.
[4] C. Petitjean and J. Dacher, A review of segmentation methods in short axis cardiac MR images. 2011. 15(2): p. 169-184.
[5] V.Tavakoli, V.Amini, and I. Understanding, A survey of shaped-based registration and segmentation techniques for cardiac images. 2013. 117(9): p. 966-989.
[6] T.Chen, et al., Semiautomated segmentation of myocardial contours for fast strain analysis in cine displacement-encoded MRI. 2008. 27(8): p. 1084-1094.
[7] H.Liu, et al., Automatic left ventricle segmentation in cardiac MRI using topological stable-state thresholding and region restricted dynamic programming. 2012. 19(6): p. 723-731.
[8] I.B. Ayed et al., Max-flow segmentation of the left ventricle by recovering subject-specific distributions via a bound of the Bhattacharyya measure. 2012. 16(1): p. 87-100.
[9] W.Bai., et al., Multi-atlas segmentation with augmented features for cardiac MR images. 2015. 19(1): p. 98-109.
10] S.C.Mitchell et al., 3-D active appearance models: segmentation of cardiac MR and ultrasound images. 2002. 21(9): p. 1167-1178.
[11] Y.Jing et al., Deep learning for drug design: an artificial intelligence paradigm for drug discovery in the big data era. 2018. 20(3): p. 58.
[12] O.Ronneberger, P. Fischer, and T. Brox. U-net: Convolutional networks for biomedical image segmentation. International Conference on Medical image computing and computer-assisted intervention. 2015. Springer.
[13] J.Li et al., Automatic whole-heart segmentation in congenital heart disease using deeply-supervised 3D FCN, in Reconstruction, Segmentation, and Analysis of Medical Images. 2016, Springer. p. 111-118.
[14] P.Tran, A fully convolutional neural network for cardiac segmentation in short-axis MRI. 2016.

[15] Q.Tong et al. 3D Deeply-Supervised U-Net Based Whole Heart Segmentation. in International Workshop on Statistical Atlases and Computational Models of the Heart. 2017. Springer.
[16] L Yu et al., 3D FractalNet: dense volumetric segmentation for cardiovascular MRI volumes, in Reconstruction, segmentation, and analysis of medical images. 2016, Springer. p. 103-110.
[17] J.Long, E. Shelhamer, and T. Darrell. Fully convolutional networks for semantic segmentation. In Proceedings of the IEEE conference on computer vision and pattern recognition. 2015.
[18] C.Y.Lee et al. Deeply-supervised nets. In Artificial Intelligence and Statistics. 2015.
[19]O. Çiçek et al. 3D U-Net: learning dense volumetric segmentation from sparse annotation. In International conference on medical image computing and computer-assisted intervention. 2016. Springer.
[20] L.Yu, et al. Automatic 3D cardiovascular MR segmentation with densely-connected volumetric convnets. International Conference on Medical Image Computing and Computer-Assisted Intervention. 2017. Springer.
[21]G. Huang et al. Densely connected convolutional networks. In Proceedings of the IEEE conference on computer vision and pattern recognition. 2017.
[22] K.Simonyan and A.J. Zisserman, Very deep convolutional networks for large-scale image recognition. 2014.
[23] H.C.Shin, et al. Medical image synthesis for data augmentation and anonymization using generative adversarial networks. In International Workshop on Simulation and Synthesis in Medical Imaging. 2018. Springer.

First Author received a bachelor's degree in computer science from Bayero University Kano, Nigeria, in 2010, a master's degree in software engineering from Sharda University, India in 2015. She is currently pursuing a Ph.D. degree with the Computer Science and Engineering Department, Sharda University, India. Her current research interests include medical image analysis, deep learning, machine learning, and data mining. She is a member of the IEEE and the IEEE Women in Engineering.

Second Author received Ph.D. (Mathematics) from the Institute of Technology, Banaras Hindu University, India. He worked for more than 35 years in Defense Research and Development Organization (DRDO), M/O Defense in various capacities viz: Scientists, Coordinator, and Director of a DRDO lab. During this period, he has guided a team of scientists working on Pattern Recognition, Cluster Analysis, and Soft Computing application in the field of cryptology, mainly cryptanalysis. Currently, he is working as a Professor at Research and Technology Development Centre (RTDC), Dept. of Computer Science and Engineering of Sharda University, India, and is involved actively in teaching and research, mainly in the area of cybersecurity and application of machine learning in various other areas. His areas of interest are Special Functions, Cryptology, Pattern Recognition, Cluster Analysis, and Data Mining. He has published more than eighty research papers in international and national journals and conferences, published fifteen confidential technical reports, and obtained eight biography copyright for different algorithms developed during his career with DRDO. He has guided many PG projects and Ph.D. theses. He has received commendation certificates and scientists of the Lab award for exhibiting excellence in pattern recognition application to cryptology.