Cardiac MRI Segmentation Using Deep Learning

Niharika Das (✉ niharikavnee@gmail.com)

Maulana Azad National Institute of Technology <https://orcid.org/0000-0002-9173-715X>

Sujoy Das

Maulana Azad National Institute of Technology

Research

Keywords: Medical Imaging, cardiac MRI Segmentation, Deep Learning, Neural Networks.

Posted Date: February 1st, 2022

DOI: <https://doi.org/10.21203/rs.3.rs-1271768/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Cardiovascular diseases (CVDs) remain the principal cause of all global death and disabilities worldwide. Cardiac MR Images play an important role in diagnosing and treating cardiac ailments in patients. Automatic segmentation of Cardiac Magnetic Resonance Imaging (Cardiac MRI) is an essential application in clinical practice. In this paper, Cardiac MRI segmentation is performed using a convolutional neural network. ACDC Challenge 2017 dataset is used the training and testing purpose. It consists of data of 100 subjects, including the End Systole and End Diastole phase. The model's performance is measured using the Dice coefficient, achieving an accuracy of 0.90. The results for basal as well as with apical slices are pretty encouraging.

Introduction

Cardiovascular diseases (CVDs) remain the leading cause of all global causalities and disabilities overall. Several people die annually because of cardiovascular diseases other than any other cause of death [1]. Different types of technology have been developed for the non-invasive examination of CVDs, such as computed tomography (CT), cardiovascular magnetic resonance (CMR), echocardiography, etc., [2]. Today recognized as a reference, CMR is a non-invasive benchmark to measure different parameters such as cardiac chamber volume, ejection fraction, mass and wall thickness, wall motion abnormality of several CVDs because of superior quality of image, non-availability of ionizing radiation and tremendous tissue contrast [3]. At present, chamber segmentation is carried out via manual outline by experts in ideal clinical practice. Manual segmentation is monotonous, tend to intra and inter-observer variability, and is time-consuming. Hence, this system needs to be automated to expedite and smooth the progress and process of segmentation [4]. Recently several techniques have been developed for semi-automatic and fully automatic segmentation, based on image processing techniques for the segmentation of either one or both, left ventricle (LV) and right ventricle (RV). The LV has been usually derived the clinical interest for the characterization of the disease progression. Nowadays, RV has achieved more attention from the research community because of numerous essential findings for managing circumstances such as dysplasia cardiomyopathies, coronary heart disease and pulmonary hypertension [4–5]. However, RV segmentation is considered as much more challenging compared to LV. As a result of leaving the problem of RV segmentation wide open, RV segmentation has long been evaluated as secondary in contrast to the LV [6–7].

In recent years, deep learning techniques are becoming more popular with automatic features detection capabilities. Deep learning could be used to extract features from a database based on the need of the particular application. Traditional feature extraction methods involve the use of prior knowledge, which helps in the better extraction of features for an individual application [8]. Deep learning can discover significant features that were not possible to extract earlier. Convolutional networks are best suitable for classification applications in which there is a single class label for a given image. In many applications, such as biomedical image processing, there is a need to assign a class label to each pixel. Also, one of the most challenging tasks for biomedical image processing is collecting massive amounts of training

data [9]. Automatic Cardiac MRI segmentation of left and right ventricles is performed using a convolutional neural network. The data of 100 subjects, including End Systole and End Diastole phase. A brief discussion on previous work done by the different research communities on RV segmentation in cardiac MRI is discussed. The methodology is described in detail with the outcome and analysis.

Related Work

In recent years, cardiac ventricle segmentation methods have been developed based on deep learning techniques. Many researchers proposed cardiac segmentation methods based on deep learning; for example, Avendi et al. [10] proposed a method to segment the left ventricle from magnetic resonance images, combining a deep learning approach and deformable models. The results had an outstanding closeness to the ground truth. Another method was proposed for the automated segmentation of the left ventricle from magnetic resonance images by Ngo et al. [11]. They used a level set algorithm and a deep learning approach. Ronneberger et al. [9] suggested UNET for biomedical image segmentation. In UNET, classification on every pixel was performed to localize and distinguish borders. The U-Net architecture used the Fully Convolutional Network as the base architecture and was modified to provide better segmentation in medical imaging. In 2016, Rupprecht et al. [12] suggested a method combining a deep learning architecture in the combination of patch-based representation with an active contour framework for interactive boundary extraction. The segmentation method could be trained with comparatively small graphics cards and could provide acceptable results with limited computational resources. Poudel et al. [13] suggested a recurrent fully-convolutional network (RFCN) in the same year. In a recurrent fully-convolutional network, a stack of 2D slices was used to train the image representations. It was a combination of cardiac ventricles detection and segmentation into one framework. It was trained end-to-end resulting in reduced computational time, which was an essential parameter for real-time applications. Later, Patravali et al. [14] proposed 2D and 3D segmentation models. The models were trained end-to-end from the start. Their segmentation models accomplished an excellent result in terms of distance metrics and have acceptable accuracy in terms of cross-entropy loss and dice loss. Same year Baumgartner et al. [15] presented a fully automated framework for segmentation of the left ventricle (LV), the myocardium (Myo) and right (RV) ventricle. They claimed that it was advantageous if the images were processed in a slice-by-slice form using 2-dimensional networks. It was beneficial due to a relatively large slice thickness. In 2018, Bai [16] proposed an automated analysis method that accomplished outperforming results in segmentation of the LV and RV on short-axis CMR images and the left atrium (LA) and right atrium (RA) on long-axis CMR images. Zotti et al. [17] proposed a convolutional neural network (CNN) model based on the UNET architecture. It used shape prior knowledge for the segmentation task. It also implied a loss function. Both high- and low-level features were used to train the network to determine the shape prior and used to the localization of the position of the cardiac ventricles with precision. Results achieved from experiments indicate that the model segmented multi-slices CMRI in significantly less time with high accuracy.

Methodology

The segmentation of LV and RV is performed using the deep learning approach. Cardiac MRI datasets from ACDC challenge [18] is used for training and testing purpose. Pre-processing of data is done to map the problem into a single-channel problem. In this network architecture, the contracting path involves two convolutional blocks, two batch normalization layers. It also has a K number of filters. A rectified linear unit (ReLU) and a 2x2 max pooling operation follow the batch normalization layer. The Dropout layers are also used. Max pooling is added after individual convolution layers. Max pooling operation helps in reducing the resolution of the image leading to evaluate larger portions of the image in one go. It helps the network to reduce the number of parameters and ultimately decreases computation time. Batch normalization is used to standardize the inputs to a layer in the network automatically. Batch normalization also helps in providing some regularization and reducing generalization error. The normalization layer is used after the activation function of the previous layer to standardize inputs. Dropout is used to temporarily remove randomly selected neurons during training. When they are dropped out, their values are not used for activation on the forward pass. Also, an update on any weight is not applied to the neuron on the backward pass. When the selected neurons are dropped out, other neurons in the network make a prediction on behalf of missing neurons. The purpose is to reduce the effect of the specific weights of neurons so that the network can provide better generalization efficiently. There are minimum chances to overfit the training data. The ReLU activation function is used here. The essential advantage of using the ReLU activation function is that it consists of little computation.

This activation function also removes the problem of co-adaption by taking advantage of sparsity. It uses this simple formula:

$$f(x) = \max(0, x)$$

ReLU function is monotonic. When the function receives any negative input, it returns 0. when it receives any positive value x; it returns that value back. The range of the output of this activation function has a range from 0 to infinity. The ReLU also helps in faster convergence of the model.

Dataset

The cardiac cine-MRI training dataset of the ACDC challenge [18] are used for the experiments. The dataset comprises of data of 100 patients in short axis cine-MRI captured on 1.5T and 3T system and has resolutions ranging from 0.70X0.70 mm to 1.92X1.92 mm in-plane and 5mm to 10mm through-plane. Also, segmentation masks are available for the left ventricle (LV), the right ventricle (RV) and the myocardium (Myo) for the end-diastolic (ED) and end-systolic (ES) phases of each patient. The dataset is divided into five groups, each containing 20 patient's data. The group includes data of normal patients (NOR), patients with systolic heart failure with infarction (MINF), patients with dilated cardiomyopathy (DCM), patients with hypertrophic cardiomyopathy (HCM) and patients with abnormal right ventricles (ARV). The dataset is divided randomly for the purpose of training, validation and testing. Data of 70 patients are used for the training, data of 10 subjects are used for validation and data of 20 subjects is used for the testing. In summary, 1328 combinations of input and label images are used for the training

purpose, 177 combinations of input and label images are used for the validation purpose, and 366 combinations of input and label images are used for the testing.

Preprocessing

The image file format of the dataset is NIFTI [19]. The dataset is converted to PNG format. The dataset comprises images of the various resolution, so all the images are brought to a common resolution of 256X256. Images with higher resolutions are not possible to process due to the GPU memory limitations on the machine the experiment is performed. The problem is mapped to a single class problem. The exact pixel values to the left ventricle (LV) and the right ventricle (RV) are provided for the label images. No changes were made in myocardium pixel values in the labelled images.

Network Architecture

The network architecture is inspired by [9, 20]. Like a typical UNET architecture, this architecture contains a contracting path and an expansive path. It features an input size of 256X256. The contracting path consists of two convolutional blocks, two batch normalization layers. These convolutional blocks and batch normalization layers are followed by an activation function which is rectified linear unit (ReLU). After the activation layer, there is a 2x2 max pooling operation. The padding is set to "same". The above process is repeated five times. The Dropout layers are also used. The value for the dropout is set to 0.3. At each downsampling step, the number of feature channels is doubled. The transposed convolution is used with an additional concatenation layer and the two convolutional and batch normalization layers in the expensive path. The transposed convolution halves the number of feature channels during the process. Here also, rectified linear unit (ReLU) activation is used in every layer except for the last except the output layers where the sigmoid function is used as an activation function. The padding is again set to "same". This process is repeated five times.

Training

For the training purpose, a combination of input images and their segmentation maps are used. There is a total of 1328 image and label pairs are used for the training purpose. Also, for validation purpose total of 172 image and label pairs are used. HP OMEN series personal computer with Intel® Core™ i9-9880H CPU @ 2.30 GHz, 2304 Mhz, 8 Cores, with 32 GB RAM and NVIDIA GeForce RTX 2080 is used for the training purpose. The training is performed for 15 epochs. It took 7 hours to complete the training. The total number of parameters is 34,535,810, out of which there are 34,524,034 trainable and 11,776 are non-trainable parameters.

Optimization function

In this experimental setup, the following optimization functions are used:

Categorical Cross Entropy Loss

One of the most widely used loss functions is categorical cross-entropy. In multi-class classification, an object can only belong to one of the many classes, and it's the task of the model to decide which class

the object will belong. The categorical cross-entropy is constructed in a way to evaluate the difference between two probability distributions.

The categorical cross-entropy loss function is calculated as follows:

$$\text{Loss} = - \sum_{i=1}^{\text{Output size}} A_i \cdot \log \hat{A}_i$$

Where $\hat{A}_i$ is the i-th scalar value in the model output, A_i is the parallel target value; the output size is the number of scalar values in the model output.

Dice Coefficient and Dice Loss

The Dice's Coefficient [21] and Dice Loss are metrics used to evaluate the similarity between the resultant segmented image and the ground truth of the same image. These are defined by the area overlap between two images. Suppose A defines the area enclosed by the segmentation algorithm and B define the area enclosed by the ground truth. The dice coefficient can be defined as follows:

$$D(A,B) = 2 |A \text{ and } B| / (|A|+|B|)$$

These numbers may also be written as $2 TP / ((FP + TP) + (TP + FN))$ in terms of true positive, true negative, false positive, and false negative. The value of D varies from 0 to 1. A higher value of D indicates higher accuracy, i.e. 0 represent the total mismatch whereas 1 represents a perfect match [22]. Similarly, 1-DC can be used as Dice loss to maximize the overlap between two images. i.e.

$$\text{Dice Loss} = 1 - (2 |A \text{ and } B| / (|A|+|B|))$$

It can be said that Dice loss evaluate the loss information on a local and global scale. The Dice loss is an important parameter to measure the accuracy.

Results And Discussion

Different statistics are calculated for the segmentation results produced by this network architecture. Also, the analysis of the visual difference between the ground truth and segmentation results are done. Figures 2 and 3 show the dice coefficient and loss graphs of 15 epochs for training and validation phases. Figure 2 shows a comparison of the Dice coefficient and numbers of training epochs. From the graph, it can be analyzed that there is a significant increase in dice with the increasing number of training steps. Also, Figure 3 shows a comparison of the loss versus the number of training epochs. From the graph, it can be interpreted that there is a gradual decrease in loss with the increasing number of training steps. The visual examples of prediction are shown in Figure 4, Figure 5 and Figure 6. Figure 4 shows images of basal slices, Figure 5 shows images of middle slices, and Figure 6 shows images of slices from the apex. It can be interpreted from the visual inspection that the prediction matches the ground

truth for the left and right ventricle in almost all cases. The model has achieved a significant dice score of 0.90 and a variance of 0.088. It took a total of 3 minutes to segment 366 images. Images from the apex are difficult to segment because there is intensity inhomogeneity present in the area. Also, the ventricle boundaries are fuzzy, and it isn't easy to distinguish the boundaries from the surrounding tissues. But, in this experiment, the model performance in apical slices and basal slices is impressive. Also, it can be visually inspected in figure 5 that the model has performed outstanding segmentation even for the smallest of the contour.

Conclusion

Deep learning architecture is used for left ventricle (LV) and right ventricle (RV) segmentation of the human heart in the short axis of MRI scans. For the same purpose, a full convolution network model is designed and implemented with learning parameters. A total of 1328 image and label pairs for training, 172 image and label pairs are used for validation and 366 image and label pairs for testing are used for the algorithm. It took a total of 7 hours to train the network. After training, for testing, the algorithm took 3 minutes to segment the images. The algorithm reached an accuracy of 0.90 in terms of the dice coefficient.

In the future, cardiac heart segmentation can be mapped as a multiple class problem also to include the myocardium. Also, integration of U-NET with other architecture such as Attention Gates or Squeeze-and-Excitation, etc., can further improve the overall system's performance.

Declarations

Acknowledgement

Department of Science and Technology supported this research, New Delhi, KIRAN division under Women Scientist Scheme (WOS) File No. SR/WOS-B/19/2016 dated 30/05/2017. This support is gratefully acknowledged.

Authors Contributions

Ms. Niharika Das: Visualization and Formal analysis, Writing-Original draft,

Professor (Dr.) Sujoy Das Conceptualization, Review & Editing

Data Availability Statement

I on behalf of all authors declare that authors are ready to share the Data and Materials supporting the findings with the reviewer.

Declaration of Conflict of Interest

I hereby declare that all the authors have mutually agreed for the submission of this article. We like to draw your attention on the present work which highlights accurate computer-aided segmentation methods for complexity of RV/LV segmentation that are tackled using various segmentation methods. Review on Cardiac MRI segmentation operation were future research direction and approaches can be drawn through analysing reports. It is to declare that the authors do not have any conflict of interest.

References

1. D. P. Ripley, T. A. Musa, L. E. Dobson, S. Plein, and J. P. Greenwood, Cardiovascular magnetic resonance imaging: What the general cardiologist should know. *Heart* 102(19) (2016) 1589–1603.
2. R. B. Stacey, & W. G. Hundley, The Role of Cardiovascular Magnetic Resonance (CMR) and Computed Tomography (CCT) in facilitating heart failure management. *Current treatment options in cardiovascular medicine*, 15(4) (2013) 373-386.
3. E. V. Gelfand, S. Hughes, T. H. Hauser, S. B. Yeon, L. Goepfert, K. V. Kissinger, W. J. Manning, Severity of mitral and aortic regurgitation as assessed by cardiovascular magnetic resonance: optimizing correlation with Doppler echocardiography. *Journal of Cardiovascular Magnetic Resonance*, 8(3) (2006)503-507.
4. N. Savioli, M. S. Vieira, P. Lamata, G. Montana, A generative adversarial model for right ventricle segmentation. *arXiv preprint arXiv:1810.03969* (2018).
5. J. Chen, L. Yang, Y. Zhang, M. Alber, D. Z. Chen, Combining fully convolutional and recurrent neural networks for 3d biomedical image segmentation. *Advances in neural information processing systems* (2016) 3036-3044.
6. O. M. Maier, D. Jiménez, A. Santos, M. J. Ledesma-Carbayo, Segmentation of RV in 4D cardiac MR volumes using region-merging graph cuts. *IEEE Computing in Cardiology* (2012) 697-700.
7. J. Caudron, J. Fares, V. Lefebvre, P. H. Vivier, C. Petitjean, J. N. Dacher, Cardiac MRI Assessment of Right Ventricular Function in Acquired Heart Disease. *Factors of Variability, Acad. Radiol.*, 19(8) (2012) 991–1002.
8. Hira, Z. M., Gillies, D. F. A Review of Feature Selection and Feature Extraction Methods Applied on Microarray Data. In: *Advances in Bioinformatics*, (2015).
9. O. Ronneberger, P. Fischer, and T. Brox. U-net: Convolutional networks for biomedical image segmentation. In *MICCAI*, pages 234–241. Springer, 2015.
10. Avendi, M., Kheradvar A., and Jafarkhani, H. A combined deep-learning and deformable-model approach to fully automatic segmentation of the left ventricle in cardiac MRI. In: *Med. Image Anal.* **30**, 108-119 (2016).
11. Ngo, T. A., Lu, Z., Carneiro, G. Combining deep learning and level set for the automated segmentation of the left ventricle of the heart from cardiac cine magnetic resonance. In: *Med. Image Anal.* **35**, 159-171(2017).

12. C. Rupprecht, E. Huaroc, M. Baust, and N. Navab. Deep active contours. arXiv preprint arXiv:1607.05074, 2016.
13. Rudra P K Poudel, Pablo Lamata, and Giovanni Montana. Recurrent fully convolutional neural networks for multi-slice mri cardiac segmentation. In International Workshop on Reconstruction and Analysis of Moving Body Organs, pages 83–94. Springer, 2016.
14. Patravali J, Jain S, Chilamkurthy S. 2D-3D fully convolutional neural networks for cardiac MR segmentation. 2017. <https://arxiv.org/abs/1707.09813> (accessed Aug 20, 2018)
15. Christian F Baumgartner, Lisa M Koch, Marc Pollefeys, and Ender 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, pages 111–119. Springer, 2017.
16. Bai W, Sinclair M, Tarroni G, Oktay O, Rajchl M, Vaillant G, Lee AM, Aung N, Lukaschuk E, Sanghvi MM, Zemrak F. Automated cardiovascular magnetic resonance image analysis with fully convolutional networks. *J Cardiovasc Magn Reson* 2018;20:6.
17. 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.
18. O. Bernard, A. Lalande, C. Zotti, F. Cervenansky, et al. "Deep Learning Techniques for Automatic MRI Cardiac Multi-structures Segmentation and Diagnosis: Is the Problem Solved?" in *IEEE Transactions on Medical Imaging*, vol. 37, no. 11, pp. 2514-2525, Nov. 2018.
19. Cox, R. W., J. Ashburner, H. Breman, K. Fissell, C. Haselgrove, C. J. Holmes, J. L. Lancaster, D. E. Rex, S. M. Smith, J. B. Woodward, and S. C. Strother. "A (sort of) new image data format standard: NiFTI-1." 10th Annual Meeting of Organisation of Human Brain Mapping, Budapest, Hungary, June 2004.
20. S. M. Ibrahim, et al., Heart Segmentation From MRI Scans Using Convolutional Neural Network, in 13th International Conference on Mathematics, Actuarial Science, Computer Science and Statistics (MACS), 2019.
21. Dice LR. Measures of the amount of ecologic association between species. *Ecology*. 1945;26(3):297–302.
22. Petitjean, C., Zuluaga M.A., Bai W., Dacher J.N., Grosgeorge D. et al. 2015. Right ventricle segmentation from cardiac MRI: A collation study. *Medical Image Analysis* 19(1) 187–202.

Figures

Figure 1

2D model of CNN Architecture

Figure 2

Dice Coef. For train and validation data sets

Figure 3

Loss For train and validation data sets

Figure 4

Images of Original, ground truth and prediction from the basal slices

Figure 5

Images of Original, ground truth and prediction from the middle slices

Figure 6

Images of Original, ground truth and prediction from the apical slices