Deep Learning Based Volumetric Segmentation of Heart Ventricles for Assessment of Cardiac Disease Using MRI

Anika Pallapothu
anikap@novaltech.org
Novaltech
Hayward, CA, USA

ABSTRACT

Diagnosis of cardiovascular diseases through cardiac MRI imaging plays a crucial role. Manual evaluation is time consuming and prone to errors. With the help of deep learning, a lot of traction has been developed for cardiac imaging diagnosis. In this study, we present a fully automated pipeline for the segmentation of left ventricle, right ventricle, myocardium, and classification of cardiovascular diseases into five classes using the cardiac MRI scans from the ACDC dataset. We adopted Segnet architecture for segmentation and made a comparative analysis using 2D and 3D approach. Best results were obtained using 2D approach with dice scores of 0.877(RV), 0.877(MYO), 0.937(LV) on the test set. We later on use the segmentation outputs to extract quantitative features to develop a robust classifier that gave us an overall accuracy of 85% on the test set and 0.81,0.89 scores of precisions and recall. Our proposed approach is computationally efficient and can be used for making critical decisions during diagnosis.

KEYWORDS

Cardiovascular, MRI, Heart Ventricles, Segmentation, Artificial Intelligence, Deep Learning

ACM Reference Format:
Anika Pallapothu. 2024. Deep Learning Based Volumetric Segmentation of Heart Ventricles for Assessment of Cardiac Disease Using MRI. In . ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

Cardiovascular diseases are one of the major factors that contribute towards the death globally as per WHO[1]. Cardiovascular magnetic resonance (CMR) is often used for diagnosis and management of cardiovascular diseases as it helps in giving a detailed and quantitative analysis of the parameters associated with heart's anatomy. Clinical changes can be quantitatively analyzed with the help of CMR imaging using which doctors can monitor and strategize further diagnosis. Manual delineation of quantitative features from these images is often time-consuming, error-prone and introduces inter-observer variability which can affect the diagnosis of patients and becomes infeasible in real life scenario where the footfall of patients is high. With the advent of deep learning, a lot of progress has been made in the automated analysis of medical imaging for various tasks which in turn can help the doctors in treatment planning. Automated delineation of cardiac features from CMRI plays a crucial role in analyzing the normal and abnormal parameters at large scale. Convolution neural networks have shown to be of great use especially for medical image segmentation-based tasks that primarily uses encoder-decoder based architectures to localize the features of interest. Current methods tend to use convolution-based approach to segment heart anatomy features and then extract quantitative features to come up with an automated diagnosis. In this study, we will be using cardiac MRI data provided as part of automated cardiac diagnosis challenge (ACDC) 2017 to segment for left ventricle (LV), right ventricle (RV) and myocardium and also to detect the presence of five types of classes namely normal, dilated cardiomyopathy, hypertrophic cardiomyopathy, prior myocardial infarction and abnormal right ventricle using deep learning and machine learning models. Our aim will be to come with a robust deep learning framework using Segnet architecture for the segmentation task and then build on the segmentation maps obtained to build a predictive model using machine learning algorithms for the prediction of five types of classes as mentioned. We will finally validate our results against the ground truths provided from the challenge and reports our results using evaluation metrics.

2 LITERATURE REVIEW

A lot of advancement has been made in the field of medical image segmentation and classification tasks recently especially after the development of Unet architectures[2]. In 2018 [3] combined Unet and M-net[4] architecture to come up with an automated cardiac segmentation model. Data augmentation in form of rotation was used after which the final dice score improved significantly for right ventricle. [5] developed a fully automated framework for segmentation of heart anatomical structures like left ventricles, right ventricles and myocardium using multiple 2D and 3D convolutional tasks using the cardiac MRI images provided as part of ACDC 2017 challenge. Out of various combination of networks, it was found that 2D networks outperformed other 3D networks due to presence of large slice thickness in terms of dice evaluation metric. [6]developed a fully automated processing pipeline for segmentation and classification using cardiac cine MRI data. An ensemble model of 2D and 3D Unet architecture was used with dice loss used as optimizer to come up with robust segmentation outputs. Geometrical features were extracted from the segmentation outputs and further prediction models were developed using an ensemble of classifiers in which an overall accuracy of 92% was achieved. [7]proposed a multi-task cardiac segmentation and diagnosis training from CMR images that

Conference'17, July 2017, Washington, DC, USA

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM
https://doi.org/10.1145/nnnnnnn.nnnnnnn

had a better convergence rate. Densenets[8] and Unet models are used for classification and segmentation training in which the use of handcrafted features was completely avoided, which is generally used in clinical diagnosis. Classification error was reduced from 32% to 22% by incorporating the segmentation training block. [9] used Fourier analysis and circular Hough transform to get the region of interest in cardiac MRI images and deployed a FCN based architecture based on Densenets[8] along with long skip and short-cut connections in the up-sampling path to avoid feature map explosion. Multiscale processing of input was done at initial layers of the network in parallel paths and later ensembled as in inception networks. Weighted cross-entropy and dice loss was used as optimizers. The proposed architecture achieved an overall accuracy of 100% for cardiac disease diagnosis on the ACDC 2017 challenge.

3 METHODOLOGY

3.1 Dataset Description

The data in this study has been taken from ACDC 2017 challenge that consists of cine-MRI scans acquired at University Hospital of Dijon (France). The training set consists of almost 100 patient scans with slice thickness varying from 5mm to 10mm, with each scan having a corresponding end-systolic(ES) and end-diastolic(ED) phase volumes. Ground truths masks are provided for myocardium, left ventricle and right ventricle for ES and ED phases of each patient. The data consists of equal distribution of normal, dilated cardiomyopathy, hypertrophic cardiomyopathy, prior myocardial infarction and abnormal right ventricle cases which have been pre-annotated.

In this study, we will be using 70 cases for training, 10 cases for validation and 20 cases for testing on which the final evaluation metrics in terms of dice and accuracy will be reported. Five-fold cross validation strategy will be adopted for model hyperparameter tuning and finally the model results across all the models will be ensembled for final reporting.

3.2 Segmentation

3.2.1 Data Preprocessing. In the segmentation part of the ACDC challenge, we normalize the whole image to zero mean and unit variance. For 2D segmentation, we do not resample the image in the axial plane, but, select slices from the whole volume and resize it to a fixed size. Each scan comprised of at least 18 slices from both ED and ES phases, so a total of 700 slices was used as part of our training and a total of 190 slices was used as part of testing. Random augmentations like cropping, rotating and flipping were performed to avoid overfitting and also to overcome lack of training data. For 3D segmentation, we resample the whole volume to 1.25x1.25x10mm. We will randomly crop the images for a patch size of 128*128*128 to be input to the model. From each scan comprising of both ED and ES phases, two random crops were selected based on a positive to negative ratio of 0.4 where probability of less than 0.4 will focus on patches containing regions where organ of interest will be found and probability greater than 0.4 will focus on patches containing regions organ of interest will not be found to make sure the model sees multiple instances of each scan.

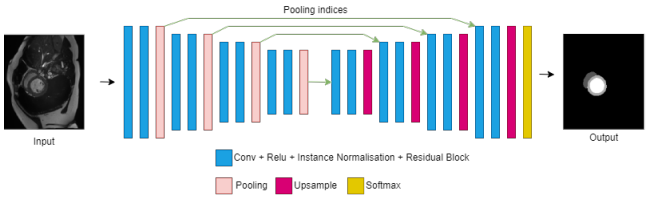

Figure 1: Standard Segnet architecture

3.2.2 Network Architecture. We approach the segmentation problem by using 2D and 3D based approaches. We will make use of Segnet as proposed in[10] , that works quite similar to traditional Unet [2] which is a standard go to algorithm that has been used mainly for biomedical image segmentation problems. Unet's capability lies in its architecture that has an encoder and a decoder path that captures the context and then successfully manages to localize it and also the skip connections that helps to efficiently integrate low level features with high level semantic information. This helps in getting better pixel level classification. Segnet works quite similar to Unet but with small changes in the decoder path that tends to use the pooling indices that have been obtained from the encoder path to eliminate the need for learning to upsample. This will help us in reducing parameter count and make the model memory efficient than Unet. Also, as our main focus is to get the overall classification accurate, we primarily use the segments as an additional input later on to the classification model We will also make slight changes in the encoding path as proposed in[11] in 2019. Here, the encoder will use ResNet blocks, where each bock consists of two convolutions, normalization and ReLu followed with a skip connection. Instance Normalization[12] will be used instead of traditional batch normalization. We will make use of four down sampling blocks and four up sampling blocks and initial filters being set to 16 followed by a final softmax layer to get final predictions (see Figure 1). We use the same architecture for 3D and 2D approaches with the difference being only in the convolution dimensions.

We used a weighted multiclass dice loss[13] which tries to optimize the overlap of prediction and ground truth and also overcomes the class imbalance problem.

$$L_{dsc} = -\frac{2}{|K|}\sum_{k\in K}\frac{\sum_{i\in I} u_i^k v_i^k}{\sum_{i\in I} u_i^k + \sum_{i\in I} v_i^k} \quad (1)$$

Both the approaches were trained for 300 epochs with five-fold cross validation using Adam optimizer. Dice score efficient was used as the evaluation metric which tends to measure the overlap of two volumes. Final results were obtained after ensembling the outputs from each fold and resampling to original voxel resolution.

$$\text{Dice Score Coeffecient} = \frac{2TP}{2TP + FP + FN} \quad (2)$$

3.3 Cardiac disease classification

3.3.1 Feature Engineering. In the cardiac disease classification part of ACDC challenge, our goal was to accurately classify the cardiac MRI images into five classes namely normal, dilated cardiomyopathy, hypertrophic cardiomyopathy, prior myocardial infarction and

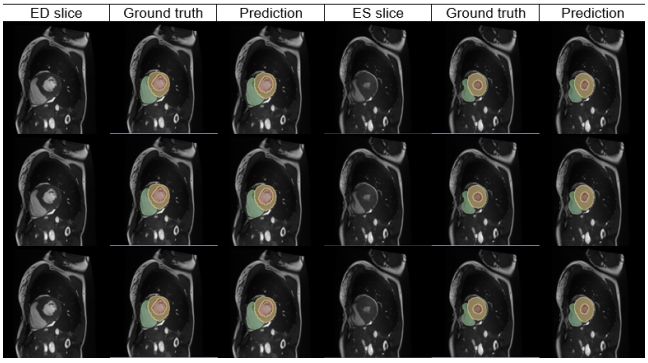

Figure 2: Segmentation results on test set using 2D Segnet model at both ED and ES phases

abnormal right ventricle. To achieve this purpose, we extract cardiac features from the ground truth segmentations as proposed in[9] . Primary features like myocardial wall thickness, volume of right ventricle, left ventricle, myocardium at end diastole and systole phases were extracted using the segmentation masks. Derive features like ejection fraction, volumetric ratios and variation profile of myocardial wall thickness were obtained from the primary features.

3.3.2 Classification. In addition to the primary and derived features extracted, we used the cardiac MRI images at end diastole and systole phases provided along with the segmentation outputs as input to the classification model. Classification model used was an ensemble of 3d resnet50 [14]and layer multilayer perceptron (MLP) with 2 hidden layers, each containing 100 units. The output from each branch was finally passed through a linear layer to get the final output. A five-fold cross-validation approach was used along with Adam optimizer. Each fold was trained for 20 epochs. Modal value across all the folds was taken to obtain the final ensemble prediction.

4 RESULTS

We evaluated our segmentation model with respect to the ground truth segmentation provided as part of the ACDC challenge on ED and ES phases using dice score evaluation metric on the separately reserved test set. We ensembled the final outputs obtained from five folds to get the final output. 2D model clearly outperformed the 3d model in terms of final dice scores, this can be attributed to the presence of poor resolution of the z-axis. 2d model achieved an overall dice score of 0.877, 0.877, 0.937 for right ventricle, myocardium and left ventricle respectively whereas 3d model got a dice score of 0.73, 0.725 and 0.852. Figure 2 shows the results of 2d segmentation model on both ED and ES phases. Detailed results for both 2d and 3d model in dice scores are tabulated in Table 1 and 2.

We used the 2d segmentation outputs to extract the features as described in earlier section to perform classification on the test set. Along with the extracted features, input images of both the phases and their respective segmentation outputs from the 2d model were given as input to the classification model. We ensembled the results from all the folds and took the model value across the five folds to get the final prediction.We achieved an overall accuracy of 85% on the test set and 0.81,0.89 and 0.9 scores of precision, recall and auc-score respectively. Confusion matrices are provided in figure 3. We can see from figure 3, a clear difficulty in differentiating DCM from MINF and RV patients, also, a normal patient being misclassified as RV on the test set. Misclassifying DCM for MINF can have clinical implications as both diseases have different diagnosis and also misclassifying DCM for RV will have clinical implications too, as DCM would require more focus on left ventricle functioning than right ventricle as seen in RV.

Table 1: Dice scores of 2d segmentation model

	Phase	Dice		
		RV	MYO	LV
DCM	ED	0.904	0.86	0.969
	ES	0.808	0.866	0.96
HCM	ED	0.926	0.908	0.957
	ES	0.776	0.922	0.866
MINF	ED	0.916	0.868	0.944
	ES	0.788	0.9	0.933
RV	ED	0.926	0.84	0.958
	ES	0.876	0.863	0.896
NOR	ED	0.95	0.87	0.96
	ES	0.912	0.892	0.92

Table 2: Dice scores of 3d segmentation model

	Phase	Dice		
		RV	MYO	LV
DCM	ED	0.743	0.667	0.875
	ES	0.571	0.633	0.837
HCM	ED	0.81	0.8	0.894
	ES	0.65	0.687	0.662
MINF	ED	0.837	0.742	0.915
	ES	0.72	0.786	0.901
RV	ED	0.788	0.738	0.917
	ES	0.705	0.758	0.855
NOR	ED	0.91	0.787	0.932
	ES	0.762	0.81	0.855

5 DISCUSSION

In this study, we presented a fully automated pipeline to classify cardiovascular diseases into five classes on cardiac MRI images using deep learning. We initially started of with a segmentation model that would accurately segment left ventricle, right ventricle and myocardium using end diastole and end systole phases of cardiac MRI images. We tested our approach with 2D and 3D models of Segnet architecture with 2D model achieving the best dice score of 0.877, 0.877, 0.937 for right ventricle, left ventricle and myocardium respectively on the test set. We used the segmentation outputs to extract quantitative features which were used as an additional input to the ensembled classification model to predict the cardiovascular diseases. We achieved an overall accuracy of 85% on the test set with a recall score of 0.86. As our main focus is to improve recall

Figure 3: Confusion matrix

of each disease against normal, a normal case being misclassified as RV would mean giving extra precaution to patient than missing the disease out. There were some cases of misclassification, which can lead to significant errors in patient management and diagnosis.

Further improvements to this study can be made by using more patient records apart from the one provided by the ACDC challenge along with different cases of cardiovascular diseases. In addition to images, patient's clinical history can be considered in future, to get more better results. Also, more robust segmentation architecture can be used with different data processing techniques to yield better results in terms of segmentation. As there were scenarios of misclassification, accurate classification needs to be achieved for ensuring appropriate patient disease management.

6 CONCLUSION

The present research can be elaborated by incorporating more features and patient records to design a more robust model. The patient's data can be recorded over time and create a sequential model to predict CAD with better accuracy. Along with prediction, further research can lead to the identification of the underlying cause of the disease so that better preventative measures can be taken.

ACKNOWLEDGMENTS

I would like to express my deepest gratitude to Dr.Michael Singer for his invaluable guidance, insightful feedback, and continuous encouragement throughout the research process. His expertise and support were instrumental in the completion of this work.

REFERENCES

[1] World Health Organization. Cardiovascular diseases. https://www.who.int/health-topics/cardiovascular-diseases#tab=tab_1, 2024. Accessed: 2024-08-30.

[2] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III*, pages 234–241. Springer, 2015.

[3] Yeonggul Jang, Yoonmi Hong, Seongmin Ha, Sekeun Kim, and Hyuk-Jae Chang. Automatic segmentation of lv and rv in cardiac mri. In *Statistical Atlases and Computational Models of the Heart. ACDC and MMWHS Challenges: 8th International Workshop, STACOM 2017, Held in Conjunction with MICCAI 2017, Quebec City, Canada, September 10-14, 2017, Revised Selected Papers*, pages 161–169. Springer, 2018.

[4] Xiao-Yun Zhou and Guang-Zhong Yang. Normalization in training u-net for 2-d biomedical semantic segmentation. *IEEE Robotics and Automation Letters*, 4(2):1792–1799, 2019.

[5] 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 *Statistical Atlases and Computational Models of the Heart. ACDC and MMWHS Challenges: 8th International Workshop, STACOM 2017, Held in Conjunction with MICCAI 2017, Quebec City, Canada, September 10-14, 2017, Revised Selected Papers*, pages 111–119. Springer, 2018.

[6] Fabian Isensee, Paul F Jaeger, Peter M Full, Ivo Wolf, Sandy Engelhardt, and Klaus H Maier-Hein. Automatic cardiac disease assessment on cine-mri via time-series segmentation and domain specific features. In *Statistical Atlases and Computational Models of the Heart. ACDC and MMWHS Challenges: 8th International Workshop, STACOM 2017, Held in Conjunction with MICCAI 2017, Quebec City, Canada, September 10-14, 2017, Revised Selected Papers*, pages 120–129. Springer, 2018.

[7] Gerard Snaauw, Dong Gong, Gabriel Maicas, Anton Van Den Hengel, Wiro J Niessen, Johan Verjans, and Gustavo Carneiro. End-to-end diagnosis and segmentation learning from cardiac magnetic resonance imaging. In *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)*, pages 802–805. IEEE, 2019.

[8] Gao Huang, Zhuang Liu, Laurens Van Der Maaten, and Kilian Q Weinberger. Densely connected convolutional networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 4700–4708, 2017.

[9] Mahendra Khened, Varghese Alex Kollerathu, and Ganapathy Krishnamurthi. Fully convolutional multi-scale residual densenets for cardiac segmentation and automated cardiac diagnosis using ensemble of classifiers. *Medical Image Analysis*, 51:21–45, 2019.

[10] Vijay Badrinarayanan, Alex Kendall, and Roberto Cipolla. Segnet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(12):2481–2495, 2017.

[11] Andriy Myronenko. 3d mri brain tumor segmentation using autoencoder regularization. In *BrainLes: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries: 4th International Workshop, BrainLes 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, September 16, 2018, Revised Selected Papers, Part II*, pages 311–320. Springer, 2019.

[12] Yuxin Wu and Kaiming He. Group normalization. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pages 3–19, 2018.

[13] Fausto Milletari, Nassir Navab, and Seyed-Ahmad Ahmadi. V-net: Fully convolutional neural networks for volumetric medical image segmentation. In *2016 Fourth International Conference on 3D Vision (3DV)*, pages 565–571. IEEE, 2016.

[14] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 770–778, 2016.