Automated Left Atrial Segmentation from MRI Sequences

Anushka Raj[1] and Anuraag Ghosh[2]
[1]*Student, Department of Electronics and Communication Engineering, R V College of Engineering*
[2]*Student, Department of Electronics and Communication Engineering, R V College of Engineering*
[1]*anushkaraj.ec18@rvce.edu.in*, [2]*anuraagghosh.ec18@rvce.edu.in*

Abstract: *This paper presents a novel algorithm for automated left atrial segmentation, from cine MRI sequences, using a modified deep CNN model. Segmentation of the LA is necessary for a variety of medical and similar applications. The analysis of the LA often involves manual tracing of the boundary of the chamber, which is subject to human errors, and is a complex and time-consuming process. An automated and precise LA segmentation model is thus quite desirable in society. The objective is to build a neural network based on the pre-trained CNN Inception V4 architecture and to predict a compressed vector by applying a multi-layer autoencoder, which is then to be back-projected into the segmentation contour of the LA to perform delineation using open contours. Quantitative evaluations are performed to compare the proposed method with the current state-of-the-art U-net method. MR images are made to undergo Late Gadolinium Enhancement before it is used as input to the CNN model, improving the quality of the image. The CNN architecture was enhanced by adding convolution and ReLU layers which help in object-background separation. The model was trained using images from 1000 MRI patients.*

Keywords: Convolutional Neural Network (CNN); Late Gadolinium Enhancement (LGE); Left Atrial Segmentation; Magnetic Resonance Imaging (MRI); Long axis cine sequences

1. Introduction

Most existing methods of automated and semi-automatic cardiac chamber delineation focus on segmentation of the left and right ventricles. Automatic segmentation of the left atrium is much more difficult compared to the segmentation of the ventricles for a number of reasons. The size of the LA is relatively small compared to the ventricles. In addition, the boundaries are not clearly defined when the blood pool of the LA enters the pulmonary veins. Last but not the least, the variability of the structure of the LA is significant among different subjects.

1.1 MRI

Magnetic Resonance Imaging (MRI) is a crucial technology in clinical diagnostics. It allows people to view and analyze a variety of tissue characteristics, blood flow and distribution, and different physiological and metabolic functions. Much of this power comes from the ability to adjust the imaging process to be particularly sensitive to each of the characteristics being evaluated. It can be interpreted as a multipurpose imaging and analysis procedure that can be configured and optimized to provide answers to a wide range of clinical questions for virtually all parts and systems of the body.

1.2 LA segmentation

LA segmentation is needed to assess atrial size and function, which are important imaging biomarkers for a wide range of cardiovascular conditions, such as atrial fibrillation, stroke, and diastolic dysfunction. LA segmentations are currently done manually, which is time consuming and observer dependent.

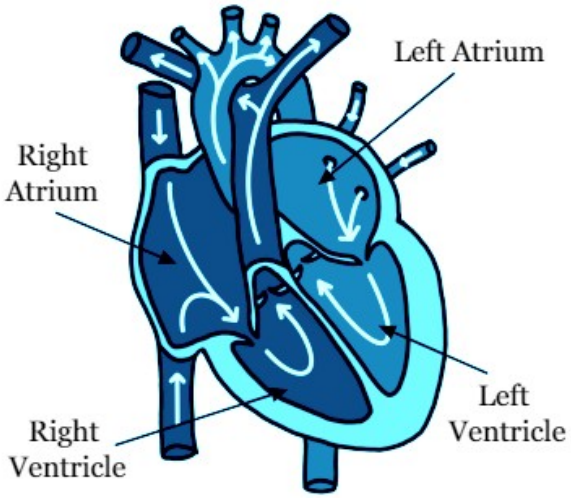

Figure 1. Chambers of the Heart

Figure 1 shows the chambers of the heart and also the pattern in which blood flow happens within the heart. It can be observed that the ventricles are larger than the LA. The LA is at the top right corner of the figure as has been labelled.

2. Materials

The tools for the model were chosen with the idea that it makes implementation of the design easier. The bulk of the needed tools are software tools and with regard to hardware tools, the only requirement is a PC with appropriate specifications. The tools utilized contained various functions and features which aided in various aspects, such as image processing, building of the model, testing the accuracy of the model etc.

2.1 Programming Language

The programming language Python has been used for this project. It is often used to create websites and software applications, for automation of tasks, and to perform analysis of data. Python is a general-purpose language, meaning it can be used to build a variety of different programs and is not specialized for any particular problem. This versatility, along with its user-level friendliness, makes it very appropriate for this work.

2.2 Libraries Used

Python Libraries are a collection of related modules. They contain content bundles that can be used repeatedly in different applications. They make Python Programming easy and convenient for the editor. Python Libraries play a very important role in the fields of Mechanical Learning, Data Science, Data Viewing, etc. For the proposed model, the important libraries which were imported are opencv, scikit, tensorflow, numpy, and matplotlib.

3. Methodology

This model uses Inception-V4 pre-trained on ImageNet Large Scale Visual Recognition Competition (ILSVRC) datasets and updates network parameters to be used with MRI datasets. The first convolutional layer and the last fully connected layers are modified by varying their initial weights. Inception-V4 is trained on 3-channel

RGB color images different from the MRI dataset used in this study, which consists of gray scale values. For this reason, CNN's 3 channels were reduced to 1 channel.

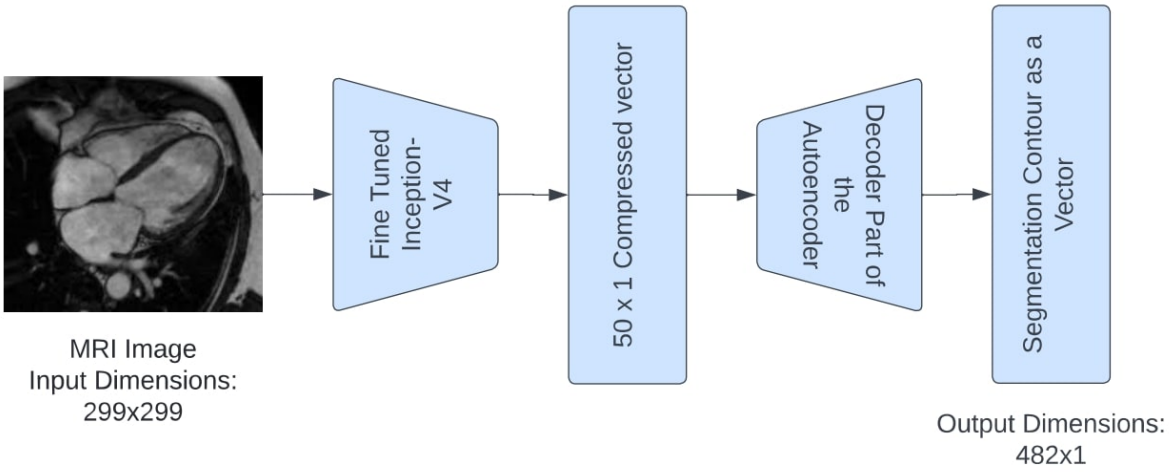

Figure 2. Methodology

The ImageNet version of the algorithm comes with 1000 categories and a softmax layer as the last layer. However, it does not include an approach for estimating object boundaries, and therefore the task of estimating the segmentation contour (i.e., vector 482_1) must be developed for the proposed problem. Therefore, the last layer of the pre-trained network (softmax layer) is cut and replaced with a new softmax layer of size 482. The flow of the methodology is shown in Figure 2. In the proposed neural network approach, the mean square error (MSE) is used as the loss. Stochastic gradient descent function (Adam version) was used as the optimizer.

3.1 Dataset

The proposed method and the U-net approach were trained and tested on a dataset of over 1000 MRI patients obtained at the University of Alberta hospital. The study was approved by the University of Alberta Health Research Ethics Committee. Evaluations were performed on 20,000 images obtained from 200 patients compared to manual segmentation of LA from 2-, 3-, and 4-chamber long-axis MR image arrays. Manual segmentation of the basic truth was initially performed by a medical student using semi-automatic software, and the contours were corrected by an experienced radiologist.

3.2 Data Preprocessing

MR images are available in different sizes or dimensions. Therefore, the images need to be resized to fit the network input shape. The input for the proposed mesh is a 2D MR image of size 299_299 and the output is the segmentation boundary, i.e., the stacked coordinate points (x; y) of the boundary. 241 even points (x; y) have been used and stacked to form a vector of dimension 482_1. In the proposed approach, CNN directly estimates the hidden representation of the segmentation using an explicit boundary. Therefore, scaling raw images to any size does not affect the estimation of stroke positions. For U-net, each input image and segmentation mask are resized to 256_256.

3.3 Convolutional Neural Network

A Convolutional Neural Network, also known as a CNN or ConvNet, is a class of neural networks that specializes in processing data with a grid-like topology, such as an image. A digitized image is the representation of an image in binary values. It contains a series of pixels arranged in a grid-like fashion that contains pixel values to indicate the brightness and color of each pixel.

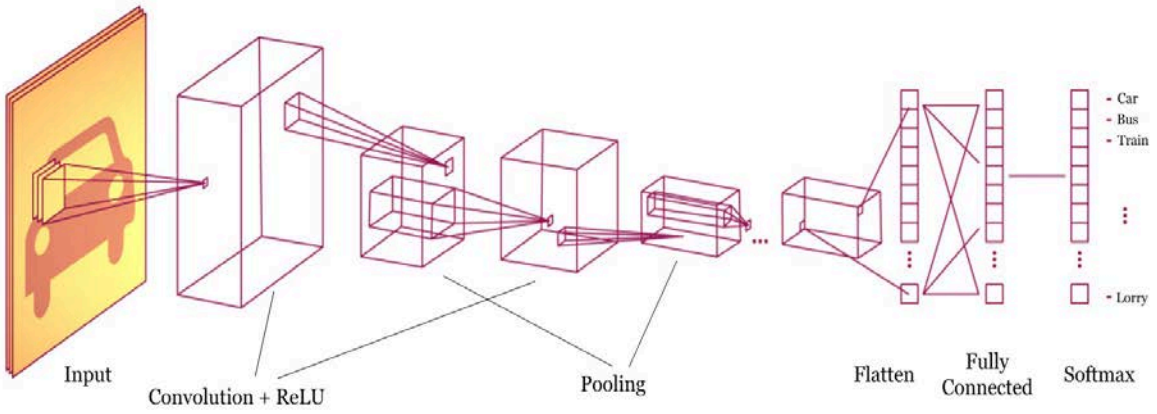

Figure 3. Depiction of a Convolutional Neural Network

Figure 3 shows the brief structure of the convolution layer that has been used for the proposed model. The human brain processes an enormous amount of information the moment it sees an image. Each neuron works in its own receptive field and is connected to other neurons in order to cover the entire visual field. Just as each neuron responds to stimuli only in the narrow region of the visual field called the receptive field in the biological vision system, each neuron in a CNN also processes data only in its receptive field. The layers are arranged in such a way as to detect simpler patterns (lines, curves, etc.) first and more complex ones (faces, objects, etc.) later on. Using a CNN, it is possible to enable the view to computers.

3.4 Dimensionality Reduction Using Autoencoders

An autoencoder is a type of artificial neural network used to learn efficient coding in unsupervised learning. The validation of the encoding is done by trying to recreate the input from the encoding procedure itself. The autoencoder encodes for a data set, typically for dimensionality reduction, training the network to ignore noise. There are variants that aim to force the learned representations to take on useful properties.

An autoencoder consists of two major parts, firstly an encoder which maps the input into the code and a decoder which maps the code back to a recreation of the input. The simplest way to perform the copy task perfectly is to duplicate the signal. Autoencoders are forced to reconstruct the input in a rough manner, preserving only the most significant aspects of the data.

The size of the hidden layer is smaller than the input layer in autoencoders, and in this case it is 50. By reducing the hidden layer size, the network is forced to learn important properties of the dataset. The main reason autoencoders are used is to reduce dimensionality and reconstruct the original contour from hidden representation. The nonlinearity of autoencoders captures spatial information about the shape and sharp edges of the segmentation contour. Before calculating the latent representation, the nonlinear activation function is applied, the rectified linear unit (ReLU), to introduce the nonlinears in the linear output.

In the proposed application, first a multilayer autoencoder was used and trained with the segmentation contour, namely the vector 482_1 (241 x coordinates and 241 y coordinates) and the bottleneck dimension is 50_1. The CNN is then trained with the target as the bottleneck vector (i.e., 50_1 latent representation) of the autoencoder. After the CNN's prediction, the decoder part of the autoencoder is added and the CNN's prediction (i.e., vector 50_1) is passed through the decoder to reconstruct the original 482_1 vector that produces the segmentation contour.

3.5 ReLU Layer

The rectified linear unit (ReLU) has become very popular in recent years. Calculate the function $f(\kappa) = \max(0, \kappa)$. In other words, activation is simply a zero threshold. Compared to sigmoid and tanh, ReLU is more reliable and accelerates convergence by six times.

Unfortunately, one drawback is that ReLU can be fragile during training. A large gradient running through it can update it in such a way that the neuron is never updated further. However, work can be done with this by setting an appropriate learning rate.

4. Results and Discussions

The targeted accuracy in this project was set at 95%. The model was made with the focus that it should plot a contour more accurate than existing CNN models. The performance comparisons will be discussed in further detail. The predictability of the model is perceived by the accuracy it achieves on the validation set as well as the visualization of the predicted vector. The imported libraries provide useful tools and functions to find out the accuracy of the model.

4.1 Results

A loss function helps to optimize a machine learning algorithm. Training and validation dataset helps to calculate loss, and its interpretation depends on the model performance on the two clusters. It is the sum of the errors made for each sample in the training or validation sets. The loss value indicates how badly or well a model behaves after each optimization iteration.

An accuracy metric is used to measure the performance of the algorithm in an interpretable way. The accuracy of a model is usually determined after the model parameters and is calculated as a percentage. It is a measure of how accurate the model's prediction is compared to actual data.

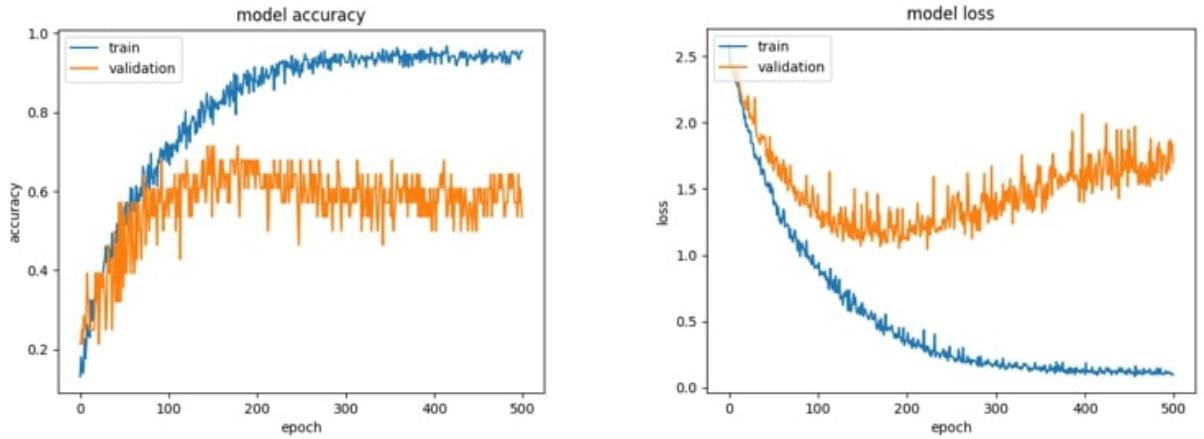

Figure 4. Accuracy and Loss graphs for a Regular CNN Model

Figure. 4 represents the accuracy and loss graphs for a regular 3 Layer CNN model. It can be inferred from these graphs that even after 500 epochs, the model isn't being able to reach a desirable validation accuracy. The batch size used was 16. The validation accuracy is at 54% while the validation loss is at the 1.7 mark. These are undesirable and unsatisfactory values which cannot give accurate results. Even though the training accuracy and loss values are appropriate, standing at 95% and 0.0979 respectively, the model fails the test phase.

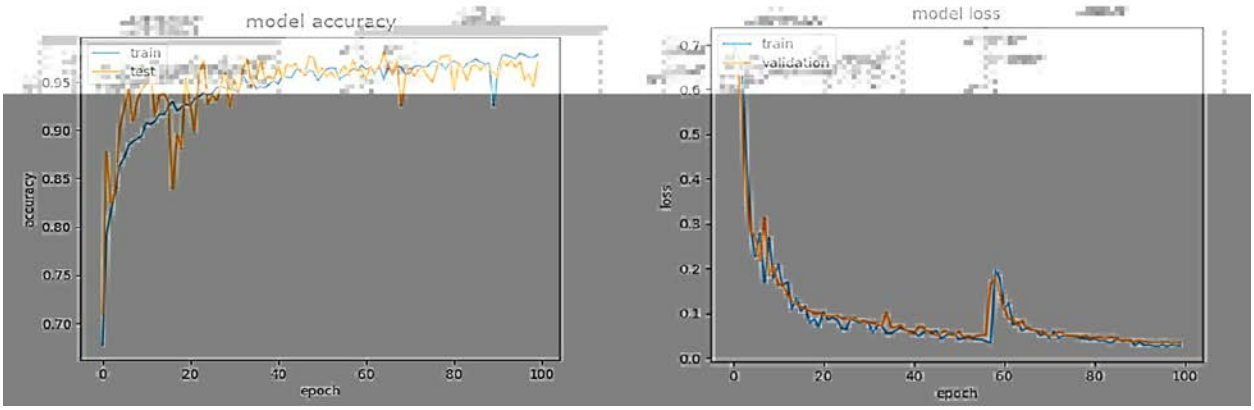

Figure 5. Accuracy and Loss graphs after Training the Model

Above are the accuracy and loss graphs, depicting the accuracy and loss respectively. The training accuracy starts logarithmically growing over the duration of the training and stabilizes at about 97% after 100 epochs. The test accuracy shows a similar trend but stabilizes at 96%. An exponential decrease in the loss graphs for both train and validation were seen. Loss for the train dataset stabilizes at about 0.0305 and that of the validation dataset stabilizes at about 0.0372. Both sets of values are desirable and satisfy the initial set targets, and they are significant improvements to the ones obtained from a regular CNN model.

4.2 ResNet18

The proposed model was trained for 100 epochs, and the training accuracy stabilized at 97% while the validation accuracy stabilized at about 96%. The Resnet18 is a CNN which is trained on more than a million images from the ImageNet database, and its accuracy is about 91%.

```
Resnet18 confusion matrics= [[17  0  0  0  0  0  0  0  0  0  0  0  0]
 [ 0  8  0  0  0  0  0  0  0  0  0  0  0]
 [ 0  0 44  2  0  0  0  0  0  0  0  0  0]
 [ 0  0  4  6  0  0  0  0  0  1  0  0  0]
 [ 0  0  0  0 20  3  0  0  0  0  0  0  0]
 [ 0  0  0  0  2 11  0  0  0  0  0  0  0]
 [ 0  0  0  0  0  0  9  0  0  0  0  0  0]
 [ 0  0  0  0  0  0  0 20  0  0  0  0  0]
 [ 0  0  0  0  0  0  0  0  8  0  1  0  0]
 [ 0  0  1  0  0  0  1  0  0 32  0  0  0]
 [ 0  0  1  0  0  0  0  0  0  0 46  0  0]
 [ 0  0  0  0  0  0  0  0  0  0  7  7  0]
 [ 1  1  0  0  0  0  0  0  0  0  0  0 18]]

Resnet18 accuracy= 90.7749077490775
>>>
```

Figure 6. ResNet18 Accuracy

Thus, the proposed model outperforms the existing Resnet18. Figure 5 shows the accuracy obtained on the dataset by the Resnet18 model. As can be seen, this model obtained an accuracy of 90.77% which is more than 600 bps lower than the proposed model. Thus, a very prominent difference exists between the two methods and it can clearly be concluded that a significant improvement has been made in the proposed model.

4.3 Dice Coefficient and Hausdroff Distance

The dice function is none other than the F1 score. This loss function directly tries to optimize the F1 score. Likewise, the direct IOU score can also be used to perform optimization. The Hausdroff distance is a technique used to measure the similarity between the boundaries of basic truth and the predicted ones. It is calculated by finding the maximum distance from any point in one boundary to the closest point in the other. Direct reduction of boundary loss function is a recent trend and has been shown to provide better results especially in use cases such as medical image segmentation where exact boundary identification plays a key role.

Table 1. Dice Coefficients and Hausdroff Distance

	Dice Coefficients in %		Hausdroff Distance in mm	
No. of Chambers	U-net	Proposed Model	U-net	Proposed Model
2	93.3 ± 3.1	94.2 ± 3.2	7.7 ± 6.3	4.1 ± 3.1
3	92.5 ± 3.9	92.4 ± 4.0	16.2 ± 22.8	4.8 ± 3.9
4	88.7 ± 11.1	92.7 ± 3.8	11.9 ± 11.0	3.8 ± 3.2

Table 1 shows the evaluation of the results of the automated segmentation by the proposed method and U-net in terms of dice coefficient (DC) and Hausdorff distance (HD) compared to expert manual segmentation on 20000 images. The mean values ± standard deviation is reported for each metric. The higher the dice coefficient or the shorter the Hausdorff distance, the better the results.

4.4 Predictability

The predictability offered by automated ML models enables people to derive meaningful insights from data and improves decision-making abilities that were previously limited by conditional fixed rules. It is not probability that needs to be improved by these ML models, but the ability to predict that probability that needs to be improved to make actionable decisions.

Figure 6 shows the predicted segments obtained in 2-, 3- and 4-chamber images. The model predicts a compressed vector which is visualized. The green contour is the prediction of the proposed model while the red one is that of the existing U-net method. The yellow contour shows the prediction of a 3-layer CNN model. It can be inferred from the above figure that the proposed model predicts a more appropriate contour and segments the left atrium more accurately. This higher accuracy is achieved by LGE and further layers added in the CNN.

The proposed neural net based on pre-trained CNN Inception V4 architecture was built and trained over the dataset successfully. An accuracy of 97% was achieved on the training dataset and an accuracy of 96% was achieved on the validation dataset. Appropriate contours were drawn by the model highlighting the left atrium of the heart accurately.

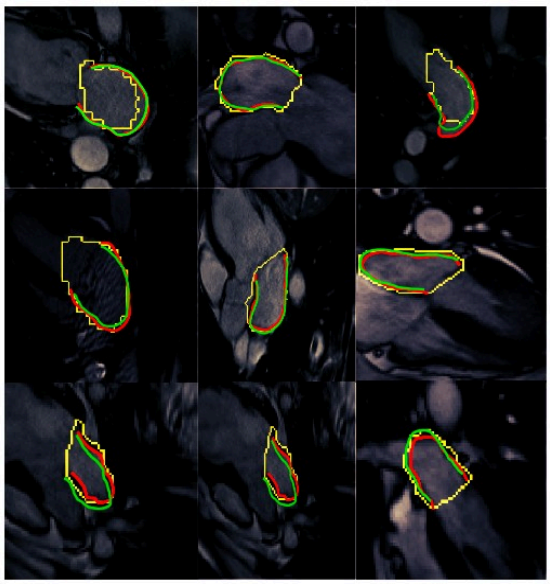

Figure 7. Predicted segments

5. Conclusions and Future Scope

This paper proposes a new method for automatic segmentation of the left atrium (LA) using deep CNN powered by autoencoder. LA segmentation is very difficult even for specialists as there is no strong demarcation line to segment the structure on MRI. By applying the autoencoder, a robust structured regression approach was applied to the proposed model. The proposed approach includes autoencoders to achieve delineation using open contours. Clear contour-based segmentation is essential in clinical calculation. Measures such as strain and strain rate that exclude the mitral valve in the calculation, whereas current machine learning-based methods such as U-net rely on semantic segmentation, which may include the mitral valve and is not suitable for generating such clinical measurements.

5.1 Conclusions

- The proposed neural net model based on pre-trained CNN Inception V4 architecture was successfully built.
- An accurate compressed vector was predicted by applying a multi-layer autoencoder, which was then back-projected into the segmentation contour of the LA to perform the delineation using open contours.
- Quantitative assessments were performed to compare the performances of the proposed method and the current state-of-the-art U-net method.

5.2 Future Scope

Future developments include the design and development of a new custom loss function and separation of the structures of PVs from the common segmentation of LA and PVs, evaluation of the performance measure of the LA chamber alone without PVs. In addition, having 3-D LA models and patient-specific fibrosis distribution, the approach can be used as a first step for 3-D fibrosis measurement. The use of an edge detection technique and a redesigned problem optimization in a three-dimensional space, i.e., a stack of all cine frames. Other limitations, such

as differences in image quality and resolution, or the emergence of scanner manufacturer-specific image artifacts, should be taken into account for future clinical application.

Acknowledgements

We are indebted to our guide, Dr. Ramavenkateshwaran.N., Associate Professor, RV College of Engineering. for the wholehearted support, suggestions and invaluable advice throughout our project work and also helped in the preparation of this thesis. We also express our gratitude to our panel members Dr. K. A. Nethravathi, Associate Professor and Dr. Ramavenkateshwaran.N., Assistant Professor, Department of Electronics and Communication Engineering for their valuable comments and suggestions during the phase evaluations. Our sincere thanks to the project coordinators Prof. Subrahmanya K N, Dr Nithin M and Dr Veena Devi for their timely instructions and support in coordinating the project. Our gratitude to Prof. Narashimaraja P for the organized latex template which made report writing easy and interesting. Our sincere thanks to Dr. K S Geetha, Professor and Head, Department of Electronics and Communication Engineering, RVCE for the support and encouragement. We express sincere gratitude to our beloved Principal, Dr. K. N. Subramanya for the appreciation towards this project work. We thank all the teaching staff and technical staff of Electronics and Communication Engineering department, RVCE for their help. Lastly, we take this opportunity to thank our family members and friends who provided all the backup support throughout the project work.

REFERENCES

[1] *Mohammadali Habibi, Joao AC Lima, Irfan M Khurram, Stefan L Zimmerman, Vadim Zipunnikov, Kotaro Fukumoto, David Spragg, Hiroshi Ashikaga, John Rickard, Joseph E Marine, et al., "Association of left atrial function and left atrial enhancement in patients with atrial fibrillation: cardiac magnetic resonance study," Circulation: Cardiovascular Imaging, vol. 8, no.2, pp. e002769, 2015.*
[2] *Liu S, Guan Z, Zheng X, Meng P, Wang Y, Li Y, ZhangY, Yang J, Jia D, and Ma C, "Impaired left atrial systolic function and inter-atrial dyssynchrony may contribute to symptoms of heart failure with preserved left ventricularejection fraction: A comprehensive assessment by echocardiography," International Journal of Cardiology, vol. 257, pp. 177–181, 2018.*
[3] *Erberto Carluccio, Paolo Biagioli, Anna Mengoni, Maria Francesca Cerasa, Rosanna Lauciello, Cinzia Zuchi, Giuliana Bardelli, Gianfranco Alunni, Stefano Coiro, Edoardo G Gronda, et al., "Left atrial reservoir function and outcome in heart failure with reduced ejection fraction: the importance of atrial strain by speckle tracking echocardiography," Circulation: Cardiovascular Imaging, vol. 11, no. 11, pp. e007696, 2018.*
[4] *Peng Peng, Lekadir K, Gooya A, Shao L, Petersen SE and Frangi AF, "A review of heart chamber segmentation for structural and functional analysis using cardiac magnetic resonance imaging," Magnetic Resonance Materials in Physics, Biology and Medicine, vol.29, no. 2, pp. 155–195, 2016.*
[5] *Liangjia Zhu, Yi Gao, Anthony Yezzi, Rob MacLeod,Joshua Cates, and Allen Tannenbaum, "Automatic segmentation of the left atrium from MRI images using salient feature and contour evolution," in Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE, 2012, pp. 3211–3214.*
[6] *Aliasghar Mortazi, Rashed Karim, Kawal Rhode,Jeremy Burt, and Ulas Bagci, "Cardiacnet: Segmentation of left atrium and proximal pulmonary veins from MRI using multi-view CNN," in MICCAI, 2017, pp.377–385.*
[7] *Chen Chen, Wenjia Bai, and Daniel Rueckert, "Multitask learning for left atrial segmentation on GE-MRI," in Statistical Atlases and Computational Models of the Heart. Atrial Segmentation and LV Quantification Challenges, 2019, pp. 292–301.*
[8] *Xiaoran Zhang, David Glynn Martin, Michelle Noga, and Kumaradevan Punithakumar, "Fully automated left atrial segmentation from MR image sequences using deep convolutional neural network and unscented Kalman filter," in IEEE International Conference on Bioinformatics and Biomedicine (BIBM). IEEE, 2018, pp. 2316–2323.*
[9] *Olaf Ronneberger, Philipp Fischer, and Thomas Brox, "U-net: Convolutional networks for biomedical image segmentation," in MICCAI, 2015, pp. 234–241.*
[10] *Shrimanti Ghosh, Pierre Boulanger, Scott T Acton, Silvia S Blemker, and Nilanjan Ray, "Automated 3D muscle segmentation from MRI data using convolutional neural network," in IEEE International Conference on Image Processing (ICIP), 2017, pp. 4437–4441.*

[11] *Christian Szegedy, Sergey Ioffe, Vincent Vanhoucke, and Alexander A Alemi, "Inception-V4, Inception-ResNet and the impact of residual connections on learning,"in Thirty-First AAAI Conference on Artificial Intelligence, 2017.*
[12] *Diederik P Kingma and Jimmy Ba, "Adam: A method for stochastic optimization," arXiv preprintarXiv:1412.6980, 2014.*
[13] *WeiWang, Yan Huang, YizhouWang, and LiangWang,"Generalized autoencoder: A neural network framework for dimensionality reduction," in The IEEE Conference on Computer Vision and Pattern Recognition(CVPR) Workshops, June 2014.*
[14] *Xiaoran Zhang, David Glynn Martin, Michelle Noga,and Kumaradevan Punithakumar, "Fully automated left atrium segmentation from anatomical cine long-axis MRI sequences using deep convolutional neural network with unscented Kalman filter," Medical Image Analysis, under review, 2019.*
[15] *Chollet Francois et al., "Keras," published by GitHub,2015.*
[16] *Ismail Ben Ayed, Kumaradevan Punithakumar, Shuo Li,Ali Islam, and Jaron Chong, "Left ventricle segmentation via graph cut distribution matching," in MICCAI,2009, pp. 901–909.*
[17] *U-net: Convolutional networks for biomedical image segmentation - in MICCAI, 2015.*
[18] *Generalized autoencoder: A neural network framework for dimensionality reduction - in The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops, June 2014.*
[19] *Left ventricle segmentation via graph cut distribution matching - in MICCAI, 2009.*
[20] *Lifetime risk for development of atrial fibrillation: the Framingham Heart Study - Circulation.;110(9):1042–6, 2004.*