Left Atrium Segmentation Using Deep Learning Model

Rishav Aryan*[1], Vaibhav Kejriwal[1], Vaishnavi Patel[1], Ansh Aggarwal[1], Vibhum Khanna[1], Shweta B Thomas[1] and Sangeetha S[1]

Department of Electronics and communication engineering

Vellore Institute of Technology, Vellore, Tamil Nadu

India - 632014

rishavsanju@gmail.com*

2022 IEEE 19th India Council International Conference (INDICON) | 978-1-6654-7350-7/22/$31.00 ©2022 IEEE | DOI: 10.1109/INDICON56171.2022.10039741

***Abstract*— Atrium segmentation is a very important task for cardiologists, especially in the left atrium (LA). Since the chances of narrowing are common in that area where blood stops flowing and which causes a heart attack. To tackle this problem, Artificial Intelligence can play an important role in segmenting the blockage area. There are many models for segmenting the left atrium such as Convolutional Neural Network (CNN), bidirectional Long Short Term Memory (LSTM) etc. Hybridize method can be an efficient method for segmenting the left atrium The most common method used is to hybridize CNN and LSTM. By analyzing the above models, it was found that the dice score was not working well even after making the model hybridized. The dice score is considered in the range from 0.85 to 0.92 for segmenting the left atrium. In our work, a U-Net model has been developed for segmenting the left atrium and the model has been developed with a change in parameters. It was observed that this model achieved a dice score of 0.94, which was found comparatively more accurate than any other hybridized model.**

***Keywords*—** *Atrium segmentation, artificial intelligence, CNN, dice score, U-Net*

I. INTRODUCTION

Artificial Intelligence (AI) has a lot of applications in medical fields. There are various methods for analysing the patient's condition using AI. Atrial fibrillation is the most widely recognized constant arrhythmia which has an extraordinary gamble of causing cardiovascular breakdown, vascular embolism and surprisingly abrupt passing [1]. Initial analysis and treatment of arrhythmia can assist with working on the capability of the heart and the anticipation of patient's condition and successfully diminish the mortality of patients. To get the physical construction of the left atrium chamber perfectly and aid interventional medical procedures and post operative perception, the division of the chamber is a key stage.

Pictures from late gadolinium improvement attractive reverberation imaging (MRI) are broadly used to assess the construction and capability of the heart in light of its benefits, for example, no radiation harm, the capacity to perform heart and blood vessel imaging without contrast specialists and can be utilized to identify and measure the scar tissue in the atrial wall. Despite the fact that it is feasible to reproduce the chamber and investigate its design by physically fragmenting the pictures, it frequently requires proficient space information and parts of work costs [2]. Accordingly, the examination of the programmed division strategy for the left chamber has significant logical importance and application esteem. There is significant inconstancy across subjects physical designs of left atrium (LA), that is to say, profoundly individualized contrasts in atrial shape and size [3]. The picture commotion furthermore, the powerful development of the heart will likewise cause extraordinary trouble in picture division. Simultaneously, the volume of the chamber is more modest than that of the ventricle, and the atrial myocardium is more slender. The left chamber division is substantially more testing. The division techniques for clinical pictures can be isolated into picture-driven strategies in view of no earlier information or on the other hand frail earlier information, model-driven techniques based on solid earlier information, and profound learning strategies [4]. For picture-driven strategies, there is thresholding, locale developing [5], grouping, and so forth. For model-driven techniques, there are chart book-directed strategies [6,7], factual shape model-based techniques like ASM [8] and AAM, and so on.

Profound learning organizations, for example, CNN are additionally frequently utilized for picture division undertakings and have accomplished great outcomes. Among them, the examination of the left chamber division strategy has likewise gained great headway. In [9], the creator summed up the exploration status of the profound learning strategy in the field of the left chamber division from LGE-MRI. In [11], the authors have made the CNN model for atrium segmentation using MRI images. In this paper authors have done two experiments, one using both stacks of 2-D axial slices and the other using 3-D data. The dice coefficients of the 2D slice are 0.89, while the dice coefficients of the 3D slice are 0.92. The Hausdorff distances for the 2D slice were 8.98 mm and the 3D slice was 8.34 mm. According to the model of the authors the proposed method shows the accurate segmentation of the left atrium. It also shows that the 3D model performs better than 2D model. Here in the paper, the authors didn't tell why the 3D Model gives the best result than that of the 2D model

A new type of model is made in paper [12].The AB-ResUNet+ architecture. The Architecture of the author primarily focus on high accuracy when provided with small datasets. The AB-ResUNet+ architecture uses, residual learning, squeeze and excitation operations,. The proposed architecture of the author improves the segmentation of complex cardiovascular structure. The Author's model is evaluated based on 11 datasets of different cardiovascular structures. The dice coefficient of AB-ResUNet+ was good. In paper [13], the Author's model proposed the "Automatic Cardiac Diagnosis Challenge" dataset (ACDC). It has large datasets. The author here did the assessment of Cardiac MRI (CMR)datasets. The author's shown that through machine learning algorithms patients' data can easily be classified and also gave high accuracy. During segmentation Author's results also reveal that better the Convolutional Neural Network model, the accuracy increases and we get accurate correlation scores on clinical metrics and low bias and standard deviation on the LVEDV and LVEF. But the problem with the author;'s model is that it is falling at the base and apex. Especially in Hausdorff distance. Another way of making a better model is making hybridized model in [14] the Authors has made the hybrid model from U-Net and Bidirectional LSTM. The author has made the automatic segmentation of left atrium in two-step learning process, i.e. U-Net and Bidirectional LSTM. The author concluded that the segmentation process depends upon the model , according to the authors, the combination of methods has better scope of giving accurate segmentation.

The above-described model provides a 0.85 to 0.92 dice score for left atrium segmentation which is not sufficient for the cardiologist for its proper analysis and its treatment, during the cardiovascular breakdown, it was thought necessary to develop a new model, which can provide more accuracy in segmenting the left atrium, which will be beneficial to the cardiologist during the cardiovascular breakdown, for initial analysis and its treatment. deep learning methods to work on the left atrium segmentation. Atrium classified each voxel of MRI images into Non left Atrium or Left Atrium. If there is any change in Atrial Volume, it means there is a certain disorder in the cardinal part such as atrial fibrillation in patients or the mitral valve (which is basically the narrowing of the mitral valve orifice, which causes the blocking of blood flow). So for the analysis, the most common process of analysis is MRI images. The segmentation process helps in creating anatomical models, which helps in the treatment process of patients.

In Section II, deep learning model and data requirements for atrium segmentation are explained. In Sections III carries the results and discussion. Finally conclusion is made in Section IV.

II. DEEP LEARNING MODEL FOR LEFT ATRIUM SEGMENTATION

A. DATA SOURCES

We took 1932 training datasets for training the model and 339 image datasets to test the model. These photographs have been obtained from the decathlon [15], which includes MRI images of left atrial segmentation. Following that, we performed data pre-processing using the Pytorch package. We chose a 352*352 picture size. Figure 1 depicts the R.A.S. depiction of the cardiac MRI scans. In Figure 1, the x-axis represents the right, the y-axis represents the anterior, and the z-axis indicates the superior. So the y-axis is oriented from posterior to anterior or back to front, but the x-axis is oriented from inferior to superior or bottom to top. Therefore, on Y-Axis, The top rows of the image imply the back of the patient, while on X-axis, the left columns imply the stomach and the rightmost neck.

Then, for each subject, we perform Z-Normalization, which means we computed the mean and standard deviation. Following that, we standardised the normalised subject into the 0 and 1 ranges ([0,1]) by conducting minimum and maximum scaling, which may be expressed as [13].

$$R_a = \frac{R - R_{\min}}{R_{\max} - R_{\min}} \tag{1}$$

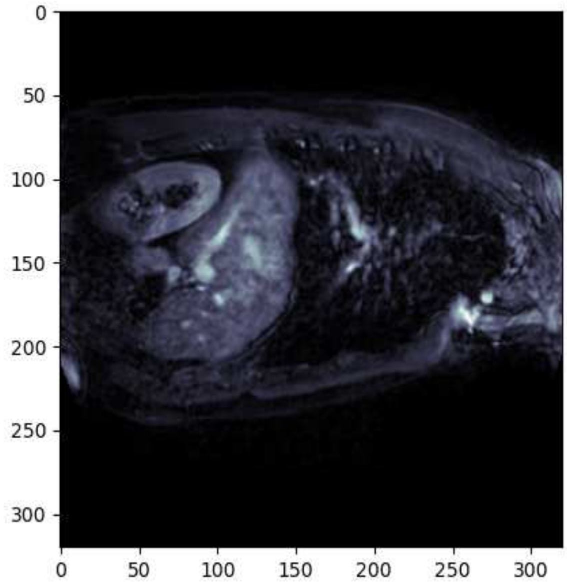

Figure 1. R.A.S representation of cardiac MRI SCANS

Ra is the normalised value, which means that we first deduct the minimum value, Rmin, and divide it by the difference between Rmax and Rmin to get the normalised value.

We built a list of 2D slices from Custom data and then extracted and loaded slices with the associated label mask. Following pre-processing, the image dataset should be constructed in such a way that the model can predict any orientation of the image dataset. As a result, Data Augmentation is required for this purpose, so that when testing on any random data, it can predict well in any direction.

B .DATA AUGMENTATION

Data Augmentation is utilized to Artificially expand the size of the preparation dataset. In clinical imaging, this is regularly finished with changes that are applied to the pictures, making distorted forms of the preparation information. Increase techniques of data augmentation usually utilize changes like pivots, reflections, and flexible disfigurements, which produce, the preparation of pictures that intently look like one specific preparation model.

While the natural inspiration driving growth procedures is appealing, a recently proposed method, 'mistake,' works by preparing on direct blends of existent preparation data: the preparation markings are likewise direct blends of the ground-truth names. Despite the fact that the photos created in this manner are clearly not the same as preparing pictures (they appear to be two pictures super-represented), this improved technique has been demonstrated to further develop execution on a variety of AI tasks. Sample augmentation of MRI scans are shown in Figure 2.

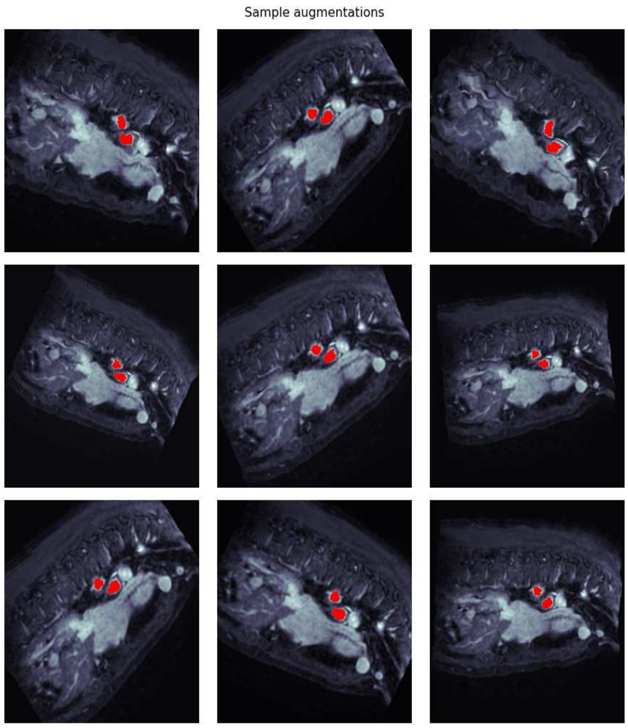

Figure 2. Sample Augmentations of MRI SCANS

We can clearly see the augmentation section in Figure 2. We scaled the value for Augmentation to (0.85, 0.15), rotated it from -45 degrees to +45 degrees, and applied the elastic transformation. Elastic transformation enhances an image by moving the pixels about the displacement field locally. After completing the augmentation part, the image data is uploaded to the U-Net model.

C. DEEP LEARNING MODEL

The medical image dataset tends to only have a small amount of labelled data. The boundary of the target structure is often blurred, and the gradient is complex. The U-Net network [10] was first presented at MICCAI 2015. For the learning of the small dataset, a great segmentation result can also be obtained. Simultaneously, it can supplement the contextual semantic information of the segmentation target at the high level of the network and provide more detailed information such as gradient for segmentation at the low level of the network through skip connection and concatenation.

Therefore, it is very suitable for the segmentation of medical images. In this work, our basic structure adopted the U-Net network. The basic structure of the U-Net model is represented in Figure 3. The U-Net network adopts a U-shaped structure. The first half of the U-Net network is an encoder process with four coding blocks. Each coding block contains two 3*3 convolutions with stride 1. After each convolutional layer, the rectified linear unit (ReLU) is used as the activation function. Each coding block uses a max-pooling layer for down-sampling. The coding process performs four down-sampling. In the decoder part, the resolution of the features is restored to the resolution of the block of the original image by block through up-sampling.

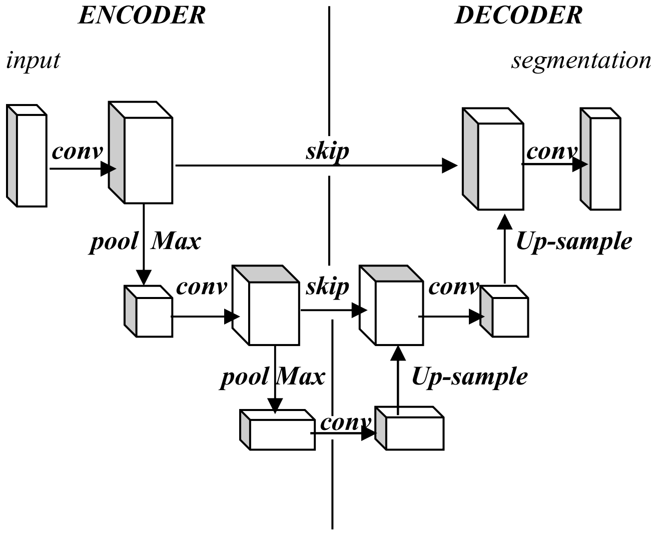

Figure 3. Representation of the U-Net model

Correspondingly, there are four decoding blocks. Each of them contains two convolutions with stride 1 and similarly uses ReLU as the activation function. The skip connection structure of the network connects the shallow feature map output by each encoding block in the encoder path and the deep feature map input by the corresponding decoding block in the decoder path. The batch normalization layers were added to the network. The data can be normalized so that the training process can be speed up, and it increased the generalization ability of the model to a certain extent. Data augmentation and dropout regularization were used to prevent overfitting. All layers except the last layer used the ReLU activation function. The last layer of the network used the sigmoid activation function. Adam optimizer with a learning rate of 0.001 was selected for the learning. It can be seen from Figure 3, that it consists of an encoder and decoder architecture with a skip connection. The encoder reduces the feature maps by using down convolutions and max-pooling, while the decoder reconstructs the segmentation mask by Up-sampling

Skip Connections allow the information to flow from the encoder to decode which reduces the problem of gradient descent. So through skip connection, we can directly forward our information. This allows for a high-quality mask and simplifies the training process. The encoder of the U-Net contain four layers. The first layer of the encoder is assigned 1 input channel and 64 output channels, after the second layer contains 64 input channels and 128 channels, the third layer contains 128 channels and 256 output channels, and the fourth layer contains 256 input channels and 512 output channels. Now in decoder it also has 4 layers, so counting serially, layer 5 has 512 plus skip connections, which contains 256 input channels so the total input channels in the 5th layer are 512 + 256 and we wanted to have 256 output channels. Similarly, the 6th layer contains 256 + 128 input channels and 128 output channels,7th layer contains 128+64 input channels and 64 output channels and last layer i.e. 8th layers contains 64 input channels and 1output channels. So basically due to skipping connection, extra channels are added. After making the proper U-Net model, the image datasets were trained properly in the model.

III. RESULTS AND DISCUSSION

So to train the model, we have used an Adam optimizer. Adaptive Moment Estimation is a calculation for improvement procedure for gradient descent. The strategy is truly proficient while working with the enormous issue including a ton of information or boundaries. It requires little memory and has a high learning rate. According to research, Adam has demonstrated superior exploratory execution in DNN over a wide range of enhancers such as AdaGrad, SGD, RMSP, and so on [16].

This kind of analyzer is valuable for huge datasets. As we probably aware, this enhancer is a mix of Momentum and RMSP improvement calculations. This strategy is essentially direct, simple to utilize, and requires less memory. It is a strong enhancer and appropriates for non-raised improvement issues in the field of Machine Learning and Deep Learning. Learning rate for U-Net model is given as [13]:

$$A_r = 1e^{-a} \quad (2)$$

Here a is the learning coefficient whose value is lies from 0 to 1. To calculate the loss for the segmentation mask, the dice loss is given as [13]

$$k\,(\hat{v}:v) = 1 - \frac{2\left(\hat{v} \cap v\right)}{\left|\hat{v}\right| + \left|v\right|} \quad (3)$$

Dice Loss is broadly utilized in clinical picture segmentation undertakings to address the information awkwardness issue. Nonetheless, it just addresses the irregularity issue among closer view and foundation yet disregards one more unevenness among simple and hard models that likewise seriously influences the preparation interaction of a learning model where, k is the dice loss value, $\hat{v}$ is the prediction and v is the label. In our model, we have achieved the high dice score value, which is 0.94. The dice Score is given as [13],

$$D = \frac{2\left(\hat{v} \cap v\right)}{\left|\hat{v}\right| + \left|v\right|} \quad (4)$$

So when we compared value of dice score obtained by the presented work with the already published results in [1] and [2], it is found that the dice score value is more here. It shows that the U-Net model work more accurately as compared to CNN and hybridized model. MRI scan of one of the segmented data out of four test data is shown in Figure 4. It can be observed from Figure 4 that, the segmentation of the left atrium has been segmented properly. The red color in Figure 4 shows the segmented area of left atrium.

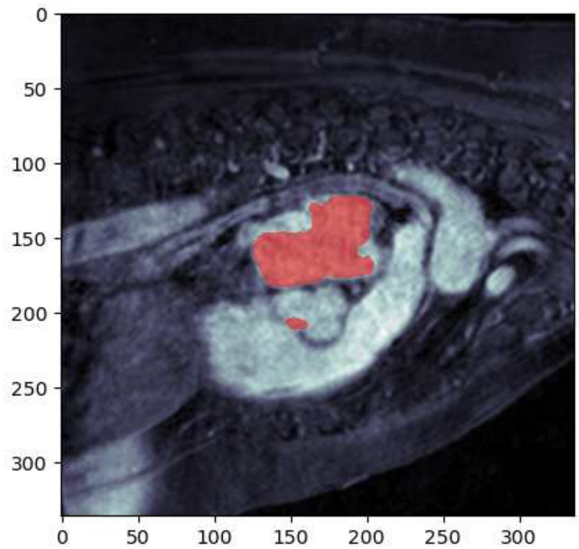

Figure 4. Segmented left atrium

TABLE I

Number of Epochs tested and their corresponding dice score

Serial No.	No. of Epochs	Dice Score
1	25	0.89
2	50	0.90
3	75	0.94
4	100	0.91

Table I shows the total number of epoch tested and their corresponding dice score. It can be seen from Table I that, fundamental outcomes on the preparation dataset without normalization showed a dice score of 0.94.Aafter 4 epochs, with the help of standardization, the dice score comes to 0.9±0.02 (under cross-validation), with an average of 0.91 and a limit of 0.94. The advancement of the dice score as per the number of epochs shows that the network converges rapidly and really, 4 epochs are enough to reach a dice score of 0.94.

IV. CONCLUSION

In the presented work, for segmenting the left atrium from MRI images, a deep learning model named U-Net is developed. The model is trained with 1932 image dataset and tested with 339 image dataset from decathlon. The model, dice score was good, it gave a 0.94 dice score. After testing with the input test data, segmentation was done. We can conclude from the work presented here is that, if the layer of the U-Net model and optimizer was provided accurately, there is a high chance of getting better accuracy with proper segmentation.

V. REFERENCES.

[1] H. Calkins et al., “2012 HRS/EHRA/ECAS expert consensus statement on catheter and surgical ablation of atrial fibrillation: Recommendations for patient selection, procedural techniques, patient management and follow-up, definitions, endpoints, and research trial design,” Europace, vol. 14, no. 4, pp. 528–606, Apr. 2012.

[2] R. S. Oakes et al., "Detection and quantification of left atrial structural remodeling using delayed enhancement MRI in patients with atrial fibrillation," Circulation, vol. 119, no. 13, pp. 1758, 2009.

[3] A. M. Maceira et al., "Reference left atrial dimensions and volumes by steady state free precession cardiovascular magnetic resonance," Journal of cardiovascular magnetic resonance, vol. 12, no. 1, pp. 65, 2010.

[4] L. Wang et al., "Principles and methods for automatic and semi-automatic tissue segmentation in MRI data," Magnetic Resonance Materials in Physics, Biology and Medicine, vol. 29, no. 2, pp. 95-110, 2016.

[5] L. Zhu et al., “Automatic segmentation of the left atrium from MR images via variational region growing with a moments-based shape prior,” IEEE Trans. Image Process, vol. 22, no. 12, pp. 5111–5122, Dec. 2013.

[6] I. Isgum et al., "Multi-Atlas-Based Segmentation with Local Decision Fusion—Application to Cardiac and Aortic Segmentation in CT Scans," IEEE Transactions on Medical Imaging, vol. 28, no. 7, pp. 1000-1010, 2009.

[7] Z. Sandoval et al., "Multi-atlas-based segmentation of the left atrium and pulmonary veins," International Workshop on Statistical Atlases and Computational Models of the Heart, Springer, Berlin, Heidelberg, pp. 24-30, Sep. 2013.

[8] H. C. van Assen et al., "Cardiac LV segmentation using a 3D active shape model driven by fuzzy inference," International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, Berlin, Heidelberg, pp. 533-540, Nov. 2003.

[9] K. Jamart et al., "Mini Review: Deep Learning for Atrial Segmentation from Late Gadolinium-Enhanced MRIs," Frontiers in Cardiovascular Medicine, Jul. 2020.

[10] O. Ronneberger et al., "U-net: Convolutional networks for biomedical image segmentation," International Conference on Medical image computing and computer-assisted

[11] Borra D, Andalò A, Paci M, Fabbri C, Corsi C. A fully automated left atrium segmentation approach from late gadolinium-enhanced magnetic resonance imaging based on a convolutional neural network. Quant Imaging Med Surg. 2020 Oct;10(10):1894-1907. DOI: 10.21037/aims-20-168. PMID: 33014723; PMCID: PMC7495320.

[12] Harijanijan, M.; Galić, I.; Romić, K.; Leventić, H. AB-ResUNet+: Improving Multiple Cardiovascular Structure Segmentation from Computed Tomography Angiography Images. *Appl.Sci.* **2022**, *12*,3024.https://doi.org/10.3390/app12063024

[13] O. Bernard *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, DOI: 10.1109/TMI.2018.2837502.

[14] Z. Zhang et al., "Automatic Segmentation of the Left Atrium from LGE-MRI Based on U-Net and Bidirectional Convolutional LSTM," 2020 Computing in Cardiology, 2020, pp. 1-4, DOI: 10.22489/CinC.2020.288.

[15] http://www.medicaldecathlon.com/

[16] https://optimization.cbe.cornell.edu/index.php?title=Adam#Applications/