Semantic Segmentation of Parasternal Short Axis View Echocardiography Images Using Unext Deep Learning Architecture

Inayathullah Ghori[1]*, Renu John[2]

Abstract

In this paper, we analyze the application of U-NeXT deep learning architecture to segment echocardiographic structures. Automation of segmentation of cardiac structures is a difficult task due to problems such as variable tissue contrast, tissue artifacts, ultrasound speckle noise (acoustic interference) and varying position, shape and movement of the cardiac structures obtained in pathological conditions. Also, the echocardiographic data is dynamic and requires optimization and appropriate selection before segmentation. We have used Echo Dataset 2.1 having 408 images containing 2601 structures. The result on UNeXT trained architecture showed IoU 0.9200. The output segmentation masks were qualitatively assessed and presented. The best contours identification accuracy by the Dice coefficient of the Right ventricular endocardium and left ventricular epicardium were 0.9692 and 0.9843 respectively. The results validates that UNeXT architecture achieves state-of-the-art performance with faster inference times.

Keywords: Cardiac Ultrasound, Cardiac Segmentation, UNeXT, Deep learning, Echocardiography, Computer vision

DOI Number: 10.4704/nq.2022.20.14. NQ880184 **Neuroquantology 2022; 20(14):1389-1392**

Introduction

Echocardiography or Cardiac Ultrasound is a low-cost, non-ionizing, affordable diagnostic method which allows us to assess the functioning of the heart and diagnose a myriad of cardiac diseases. 2D Transthoracic Echocardiography obtains images of various cardiac regions during the entire cardiac cycle through different windows (Echo Views). Certain positions on the chest wall and corresponding images have been standardized such as Parasternal Long Axis View (PLAX), Parasternal Short Axis View (PSAX), Apical Four Chamber View (A4C) and Apical Two Chamber View (A2C). From the images obtained segmentation of the structures and further analysis yield crucial information such as Left Ventricular chamber size and volumes (End-Systolic and End-Diastolic), Ejection fraction, LV Mass. However, segmentation of the structures and obtaining the geometric dimensions is not only time-consuming but also requires expertise. It is operator dependent and hence subject to variability.

Automation of segmentation of cardiac structures is a difficult task due to problems such as variable tissue contrast, tissue artifacts, ultrasound speckle noise (acoustic interference) and varying position, shape and movement of the cardiac structures obtained in pathological conditions. Unlike other ultrasounds or medical imaging techniques (MRI or CT images) the echocardiographic data is dynamic and requires optimization and appropriate selection before segmentation. [1]

Among the standard echocardiographic views, Parasternal Short Axis View (PSAX) is the most important view for quantitative measurements

***Corresponding Author:** Inayathullah Ghori
Address:1*Department of Biomedical Engineering, Indian Institute of Technology Hyderabad, Telangana, India.
Orcid ID: 0000-0003-2081-5824, Email: bm14resch11001@iith.ac.in , Email: inayath 2020@gmail.com
2Address: Department of Biomedical Engineering, Indian Institute of Technology Hyderabad, Telangana, India.
Orcid ID: 0000-0003-3254-2472, Email: renujohn@iith.ac.in

Relevant conflicts of interest/financial disclosures: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

and assessment of certain parameters such as Left ventricular mass, volumes and wall motion analysis. Left ventricular Dimensions and Volumes help to identify conditions such as dilated cardiomyopathy and hypertrophic cardiomyopathy. Regional Wall motion analysis is crucial (and the earliest detectable change) for diagnosing myocardial infarction and ischemic heart disease. Segmentation of cardiac apart from calculating cardiac function and problems can also be used to identify abnormal structures such as clots which can cause brain stroke and other systemic ischemic conditions. Analysis of PSAX view, apart from giving qualitative information, is also being increasingly adapted in other techniques such as Tissue Doppler and Speckle tracking. With respect to the final goal of clinical evaluation and cardiac functional analysis, segmentation of cardiac structures in PSAX is a quintessential step to be solved. In this paper we discuss the results of using deep learning methods for automatic segmentation of cardiac regions using U-Next Architecture [4]. U-NeXT Architecture was developed keeping in mind the need of light-weight networks with high inference speed for point of care ultrasound devices.

Experiment

Dataset

We introduce Echo 2.1, a novel annotated cardiac dataset which contains videos containing all four standard views – Apical Four Chamber View, Apical Two Chamber View, Parasternal Long Axis View, Parasternal Short Axis View – of which 408 fully annotated images from Parasternal Short Axis View were used containing 7 classes. The structures were left ventricular endocardium (inner chamber wall), left ventricular epicardium (outer wall), right ventricular endocardium and right ventricular epicardium, anterior papillary muscle, posterior papillary muscle, anterior mitral valve leaflet and posterior mitral valve leaflet. As the ultrasound view transverses different structures at different angle, certain structures such as mitral valve leaflet and papillary muscles were variably found and hence these annotations were grouped separately. In few images, due to dynamic movement there was drastic shift in position of certain structures frame-to-frame however for the purpose of our segmentation which is image-based we included all the images in the study. The annotation was done using CVAT (Computer Vision Annotation Tool) which is primarily used for annotation and label refinement and it can be integrated for scalable data processing and parallelized training pipelines.

The process of annotation was performed manually where the expert echocardiographers marked key points and polygons onto the desired contours. Although time-consuming, the process required all the annotations to be reviewed by another cardiologist.

The Echo 2.1 dataset contains total of 8 classes (including ultrasound view area) and contains total of 2601 annotated structures. Sample parasternal short axis images with its annotated masks are shown in figure 1.

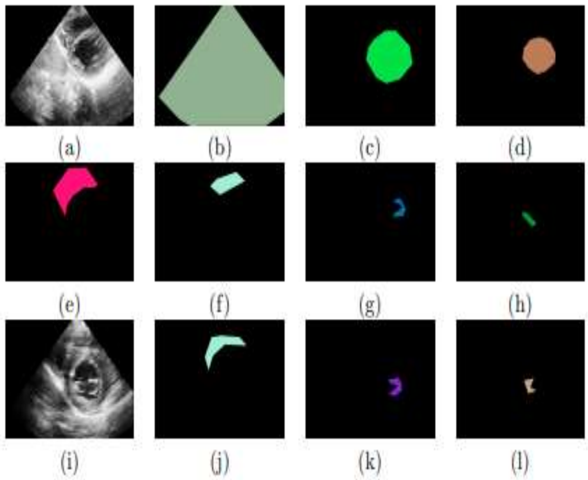

Figure 1: Different structures annotated - (a) Parasternal Short Axis View Image frame (PSAX), (b) Ultrasound View area, (c) Left ventricular epicardial wall, and (d) Left ventricular endocardial wall, (e) Right Ventricular epicardial wall, (f) Right Ventricular endocardial wall (g) Anterior Papillary Muscle, (h) Posterior Papillary Muscle. In certain parasternal images, views that were taken by transversing mitral valve structures labeled were anterior mitral leaflet and posterior mitral leaflet (i) Another parasternal short axis image with mitral valve (j) Right ventricular endocardial wall (k) Anterior Mitral valve Leaflet (l) Posterior Mitral valve leaflet. Best viewed in colour.

Method

Many approaches to segmentation of echocardiographic images have been proposed – however keeping in mind the need for good performance and point-of- care use we have used UNeXt which is a convolutional and MLP-based network. Similar to UNet [2], it is a U-shaped encoder-decoder network architecture, which consists of five-layer encoder blocks and decoder blocks that are connected via a bridge along with skip connections. The encoder network (contracting path) halves the spatial dimensions and doubles the number of filters (feature channels) at each encoder block [5][6]. Likewise, the decoder network doubles the spatial dimensions and half the number of feature channels.

However, unlike UNet, the convolutional stage is followed by a MLP stage. UNeXt introduces several new methods – a) MLP blocks to learn meaningful information for segmentation b) a novel Tokenized MLP block which projects the features into an abstract token and then uses MLP c) In order to extract local information corresponding to different axial shifts, a shifting operation in MLPs is introduced. The UNeXt has shown to reduce the number of parameters by 72x, decrease the computational complexity by 68x and improve the inference speed by 10x. Figure 2 shows an illustration of UNet with MLP architecture.

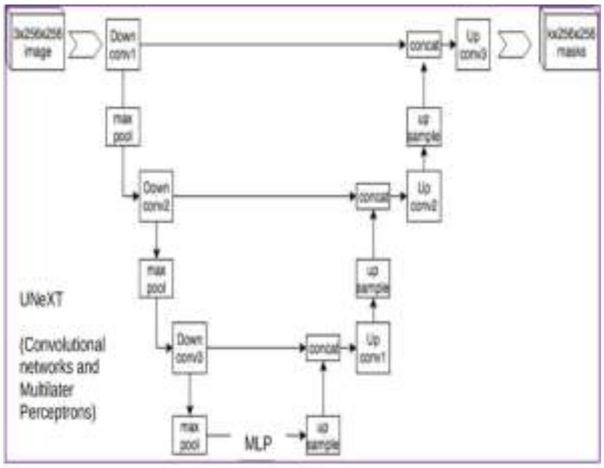

Figure 2: Illustration of UNeXT Architecture. Adapted and modified from Mehrdad Yazdan [3].

Many other differences exist between UNext and UNet, for example number of channels across each block is comparatively less than UNet, the use of bilinear interpolation instead of transpose convolution for up-sampling, use of shifted MLP similar to Swin transformer to induce locality into the block - thus adding more meaningful feature information with less computation and parameters.

Result

The dataset images were resized to a resolution of 512x512. Initially 4 classes were selected for building the network. We perform a 80-20 random split thrice across the dataset and report the mean and variance. The UNext framework was implemented using PyTorch, using a combination of binary cross entropy and dice loss to train the network. Adam optimizer with a learning rate of 0.0001 and momentum of 0.9 were used. A cosine annealing learning rate scheduler with a minimum learning rate upto 0.00001 was set. The batch size was set to 8.

The UNext was trained for 300 epochs. The results obtained showed IoU: 0.9200. The best contours identification accuracy by the Dice coefficient of the Right ventricular endocardium and Left ventricular epicardium were 0.9692 and 0.9843 respectively.

Figure 3 shows the input images and generated masks. Our results validate UNeXt on echocardiography dataset achieving the state-of-the-art performance along with fast inference times and less number of parameters.

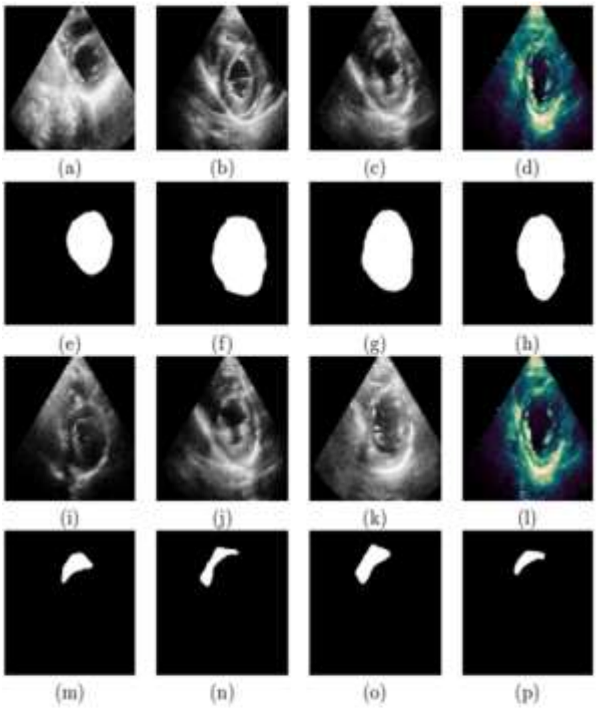

Figure 3: Different masks generated: Parasternal Short Axis View frame (PSAX) resized images(a,b,c,d). Left Ventricular Outer wall segmentation masks (e,f,g,h), PSAX frames (i,j,k,l) Right Ventricular endocardium segmentation masks (m,n,o,p). Best viewed in colour.

Conclusion and Future Work

We have presented the results of using UNeXt architecture for segmentation of cardiac structures. Future work entails segmentation of multiple classes with overlapping masks, segmentation using other views, using poor quality images, performing segmentation using other deep learning networks such as DeepLabv3, other encoders such as ResNet, Inception, DenseNet and comparison with them, further work would also be to perform clinical validation and assess computational effectiveness in point-of-care scenarios.

Acknowledgment

We like to thank our industry partner, Cardiologists and the Echocardiography team for their effort to develop the dataset, for providing ethical clearance and for allowing us to access it.

Conflict of Interest

None

Funding Source

None

References

I. Ghori, D. Roy, R. John and K. M. Chalavadi. Echocardiogram Analysis Using Motion Profile Modeling. IEEE Transactions on Medical Imaging 2020; 39(5): 1767-74. doi:10.1109/TMI.2019. 2957290

N. Siddique, S. Paheding, C. P. Elkin and V. Devabhaktuni. U-Net and Its Variants for Medical Image Segmentation: A Review of Theory and Applications. IEEE Access 2021; 82031-57. doi: 10.1109/ACCESS.2021.3086020.

Mehrdad Yazdani, Example architecture of U-Net for producing k 256-by- 256 image masks for a 256-by-256 RGB image, Wikimedia-commons. Available from https://commons.wikimedia.org/wiki/File:Example_architecture_of_U-Net_for_producing_k_256-by-256_image_masks_for_a_256-by 256_RGB_image.png [Accessed on 12/06/2021]

Valanarasu, Jeya Maria Jose and Patel, Vishal M. UNeXt: MLP- based Rapid Medical Image Segmentation Network. ArXiv preprint arXiv:2203.04967, 2022; https://doi.org/10.48550/arXiv.2203.04967

Zongwei Zhou, Md Mahfuzur Rahman Siddiquee, Nima Tajbakhsh, Jianming Liang UNet++: A Nested U-Net Architecture for Medical Image Segmentation. arXiv:1807.10165v1 2018; https://doi.org/10. /48550arXiv.1807.10165

Enze Xie, Wenhai Wang, Zhiding Yu, Anima Anand kumar, Jose M. Alvarez, Ping Luo. Seg Former: Simple and Efficient Design for Semantic Segmentation with Transformers. 2021; https://doi.org/10.48550/ arXiv.2105.15203