Left Heart Segmentation In Echocardiographic Images Using DNN model

Submitted in partial fulfilment of the requirements for the award of the degree of

Bachelor of Technology

By

Anshika Jain

IEC2020099

Under the supervision of:

Dr. Shanti Chandra

Department of Electronics and Communication

IIIT-Allahabad

Department of Electronics and Communication

INDIAN INSTITUTE OF INFORMATION TECHNOLOGY ALLAHABAD

2024

Candidate Declaration

I **Anshika Jain**, with enrollment number **IEC2020099**, do hereby declare that the written submission entitle "**Left Heart Segmentation In Echocardiographic Images Using DNN model**", represents my ideas in my own words and where others' ideas or words have been included, I have adequately cited and referenced the original sources. I also declare that I have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in my submission. I understand that any violation of the above will be cause for disciplinary action by the Institute and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when needed.

Anshika Jain

IEC2020099

Date: ……..

Certificate from Supervisor

This is to certify that the statement made by the candidate is correct to the best of my knowledge and belief. The project titled "**Left Heart Segmentation In Echocardiographic Images Using DNN model**" is a record of candidates work carried out by her under my guidance and supervision. I do hereby recommend that it should be accepted in the fulfilment of the requirements of the Bachelor's thesis at IIIT Allahabad.

Dr. Shanti Chandra
Assistant Professor
Department of Electronics and Communication Engineering
Indian Institute of Information Technology Allahabad,
Deoghat, Jhalwa, Prayagraj-211012, Uttar Pradesh, India.

Certificate of Approval

The foregoing thesis is hereby approved as a creditable study in Information Technology and its allied areas. It is carried out and presented in a satisfactory manner to warrant its acceptance as a prerequisite to the degree for which it has been submitted. It is understood that by this approval the undersigned do not necessarily endorse or approve any statement made, opinion expressed or conclusion drawn therein but the thesis only for the purpose for which it is submitted.

Committee on final examination for the evaluation of thesis:

1. Dr. Shanti Chandra (Coordinator) …….
2. Dr. Ramesh Kumar Bhukya (Chairman) …….
3. Dr. Surya Prakash (Member) …….

Date: …… ……...
Place: …….. Dean(A&R)

ABSTRACT

Echocardiography stands as the cornerstone in evaluating the function of the left ventricle which is a crucial component responsible for pumping oxygen-rich blood throughout the body. This non-invasive imaging technique employs high-frequency sound waves to generate real-time images of the heart's chambers, valves, and blood flow. In clinical practice, echocardiography plays a pivotal role in diagnosing various cardiac conditions, including heart valve disorders, cardiomyopathies, and congenital heart defects. Accurate assessment of the left ventricle's structure and function is paramount for guiding treatment decisions and predicting patient outcomes.

In this research, we have analysed and compared the performance of 3 segmentation models: U-Net, NN-Unet and ACNN on LV endocardium, LV epicardium and left atrium. The evaluation of the different models were conducted on a CAMUS dataset comprising echocardiographic images from 500 patients with varying image quality and pathology, acquired at the St Etienne University Hospital, France.

The results demonstrate the superiority of NN-Unet for measuring the left ventricle's volume during end-diastole and end-systole, over other existing architectures, with an 2.2%, 4.9% increase in the dice similarity coefficient and mean correlation respectively compared to ACNN and a 1.2%, 4.1% improvement in the dice similarity coefficient and mean correlation respectively over the basic U-net architecture. These findings underscore the potential clinical impact of NN-Unet in facilitating more accurate segmentation of the left ventricle, thereby enhancing diagnostic accuracy and improving patient care outcomes in cardiology practice.

ACKNOWLEDGEMENTS

It is my honour and privilege to get an opportunity to study at such a great institute. "**Indian Institute Information Technology – Allahabad**", where each day was a chance to learn and grow personally and academically and professionally. The wealth of knowledge and experience it has given us is deep-rooted foundations in our various fields of study.

I would like to express our sincere gratitude to all those who provided me with the resources and guidance to complete my thesis. I am profoundly grateful to **Dr. Shanti Chandra** for his expert guidance and continuous encouragement throughout to see that this project rights its target since its commencement to its completion. His instrumental contribution in the proposed work has made it conceivable.

I would like to extend my thanks to respected Director **Prof Mukul Sharad Sutaone** and all panel members who provided me with the opportunity to work on this project. Their trust and belief in my abilities have been instrumental in driving me towards excellence.

I would also like to thank **Ms. Shivani Singh**, for her overwhelming support and for helping me out whenever and wherever I approached her. My friends, classmates who helped me to widen intellectual scope and who supported me in ways too great to forget. Special thanks to my family for their love, encouragement, interesting discussions, and continuous support throughout the course.

Finally at the end I would like to thank all the faculty members who inspired me to finish our work.

Contents

1. **Chapter 1: Introduction**……...8
   1.1 Motivation……8

2. **Chapter 2: Literature Review**……...10

3. **Chapter 3: Methodology**……12
   3.1 Dataset Description……13
   3.2 Dataset Augmentation……15
   3.3 Image Processing……16
   3.4 Theoretical Basis……16
   3.5 Model Evaluation……...17
   3.6 Model Selection……..20

4. **Chapter 4: Results**……..25

5. **Chapter 5: Comparative Analysis**……34

6. **Chapter 6: Conclusion**……...36

7. **Chapter 7: Future Scope**……37

References 38

Chapter 1

INTRODUCTION

Two-dimensional echocardiography (2D echo) stands out as the primary non-invasive tool for assessing heart disease, thanks to its high temporal resolution (typically ranging from 50 to 250 frames per second) and rapid acquisition times. This imaging technique provides grayscale images that allow clinicians to identify anatomical features and evaluate cardiac function. In particular, segmenting the left ventricular (LV) walls in 2D echo is crucial for quantifying cardiac function parameters like ejection fraction (EF) and LV volumes.

Currently, experts primarily rely on semi-automatic or manual delineation techniques to identify LV boundaries in 2D echo images. However, these approaches are time-consuming and subjective, leading to potential intra-observer and inter-observer variability.

Before the advent of deep learning techniques, automatic LV segmentation methods fell into categories such as pixel classification, image-based methods, deformable methods, active appearance and shape models (AAM/ASM), and atlas models. While these methods yielded high-quality segmentation models on magnetic resonance imaging (MRI) and computerised tomography (CT) images, they struggled with 2D echo images due to challenges like speckle noise, brightness inhomogeneities, presence of trabeculae and papillary muscles, and variations in shape and motion.

Moreover, previous methods faced difficulty in creating models that could effectively handle the diverse shapes and dynamics of the LV. In light of these challenges, this report delves into the segmentation analysis of echocardiograms using three different DNN architectures: U-Net, NN-UNet, and ACNN. We evaluate these architectures based on various metrics, including mean dice coefficient, jaccard index, MAD, MAE, correlation coefficient, Hausdorff distance, accuracy, precision, and recall.

By exploring the performance of different DNN architectures, this report aims to shed light on their potential for enhancing echocardiographic analysis. These insights could lead to advancements in automated cardiac image analysis, ultimately benefiting clinical diagnosis and patient care in the field of cardiology.

1.1 Motivation

Echocardiography, commonly referred to as an echocardiogram, serves as an invaluable diagnostic tool in cardiology due to its non-invasive nature, real-time visualisation capabilities, and versatility. By providing dynamic images of the heart's structure and function, echocardiography allows doctors to assess cardiac chambers, valves, and blood flow patterns accurately. This comprehensive assessment enables the early detection of cardiac abnormalities, even before symptoms manifest, facilitating timely intervention and treatment. Additionally, echocardiography aids in guiding treatment decisions by providing crucial

information about chamber size, wall motion, ejection fraction, and valvular function. Its ability to detect and monitor a wide range of cardiac conditions, from congenital heart defects to heart failure and valvular diseases, makes it an indispensable tool in modern cardiology practice.

Researches can enable precise delineation of cardiac structures, specially the left ventricle, from imaging data. These algorithms provide quantitative analysis, yielding parameters like volume and wall thickness, enhancing diagnostic accuracy and treatment planning. Machine learning-based segmentation offers efficiency and consistency over manual methods, streamlining workflow and reducing clinician burden. Moreover, these models can adapt and learn from large datasets, improving accuracy over time. Segmented echocardiograms also serve as valuable training data for developing robust models.

Overall, machine learning-driven segmentation of echocardiograms holds promise for advancing cardiac imaging analysis, offering rapid, standardised, and accurate assessment for improved patient care.

Chapter 2

LITERATURE REVIEW

In recent literature, several studies have explored the application of deep learning (DL) techniques to segment echocardiography images. [1] conducted a comparative analysis of various DL methods for segmenting the left ventricular endocardium and myocardium, highlighting the effectiveness of encoder-decoder-based architectures over traditional non-DL methods. [25] introduced MFP-U-Net, a modified version of the U-Net architecture, to enhance the segmentation of the left ventricle, demonstrating improved performance through additional convolution layers.

[26] aimed to segment the left ventricular endocardium and myocardium, focusing on algorithm design and testing on both porcine and human images. While their method showed promise, limitations arose from its reliance on open-chest pig images, which offer higher quality than human echocardiograms. [11] proposed a hybrid architecture combining modified U-Net and FCN encoder, emphasising improved feature extraction and error learning capabilities.

[4] introduced a bilateral segmentation network coupled with motion-enhanced segmentation and orientation-congruency of optical flow. [16] presented Cardiac-SegNet, integrating U-Net for feature extraction, a fully convolutional single-state object detector for region of interest segmentation, and a mask head network for final segmentation.

[7] proposed a dual pipeline using ResDU-Net for end-systolic and end-diastolic frame segmentation. In contrast, [17] explored self-supervised algorithms for left ventricle segmentation, aiming to address the scarcity of labelled data.

[15] utilised YOLOv3 for object detection and ventricle segmentation. Despite the advancements in these methods, they primarily focused on ventricular segmentation, neglecting other anatomical structures.

[11] proposed CU-net with the SSIM loss function as a method to achieve automated semantic segmentation achieving dice coefficient of 0.856, Hausdorff distance of 3.33, and pixel accuracy of 0.929, revealing its effectiveness whereas MAEF-Net is proposed for left ventricle segmentation identifying the end-diastolic frames and end-systolic frames achieving a dice coefficient of 93.10% and cardiac phase detection MAE was 2.36.

[25] introduced the MFP-Unet network, which not only segments the left ventricle but also predicts its long axis and area, crucial for assessing left ventricular ejection fraction (LVEF). [16] developed the pyramid local attention network (PLA-Net), which enhances feature representation by capturing information from dense and sparse contexts. Guo et al. proposed segmentation networks incorporating attention mechanisms for left ventricular segmentation and key point localization.

However, most existing ventricular segmentation methods are based on static images, overlooking the detection of end-diastolic frames (EDFs) and end-systolic frames (ESFs) throughout the cardiac cycle, essential for clinical applications. [14] addressed this by using the YOLOv3 detection network to extract regions of interest and proposing a hybrid classification framework for foetal echocardiography. [17] designed an RNN + LSTM network to identify EDFs and ESFs in apical four-chamber echocardiographic videos. Ouyang et al. introduced EchoNet-Dynamic, a large dynamic echocardiography database, and achieved promising results in left ventricular segmentation and LVEF estimation using DeepLab v3 combined with R2 + 1D. However, their method's network parameters are extensive.

[8] proposed a residual auto-encoder network based on the Transformer structure for direct LVEF prediction, though direct prediction may not align with clinical workflow, where LVEF estimation typically follows manual delineation of the left ventricular contour.

Despite the progress, two significant challenges persist. First, clinical echocardiograms often suffer from low tissue contrast, motion artefacts, and missing margins, making accurate segmentation challenging. Second, there is a lack of large-scale, dynamic echocardiographic datasets, necessitating robust evaluation across diverse databases. Addressing these issues will further enhance the efficacy of deep learning algorithms in echocardiographic analysis.

Chapter 3

METHODOLOGY

When we segment echocardiographic images, we go through a series of steps. First, we clean up the images to reduce any noise and make the details clearer. Then, we use different techniques, like setting thresholds or looking at specific regions, to pinpoint and outline important structures like the heart muscle, chambers, and valves. This process helps us get a clear picture for analysis and diagnosis based on the echocardiogram results. Detailed methodology is shown in Fig. 1.

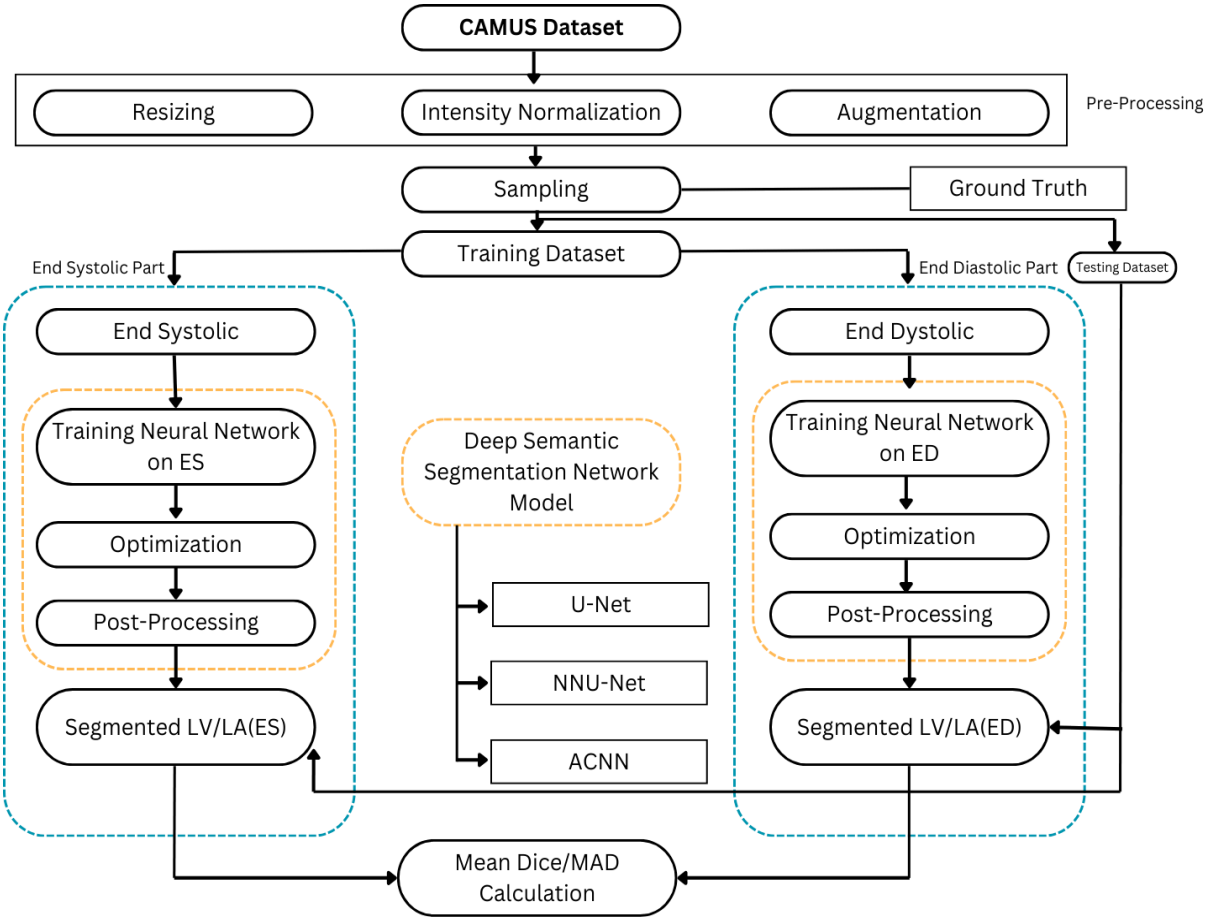


Fig. 1 illustrates the overall block diagram representing the complete methodological flow of echocardiographic images segmentation process

3.1 Dataset Preparation:

We have used the CAMUS[1] dataset for the segmentation.

3.1.1 Dataset Description:

CAMUS Dataset: The dataset comprises echocardiographic scans from 500 patients conducted at the University Hospital of St Etienne, France, adhering to ethical guidelines. Utilising GE Vivid E95 ultrasound scanners with a GE M5S probe, the images captured two-dimensional apical four-chamber and two-chamber views for each patient. The dataset was divided into 10 folds for cross-validation, with each fold containing 50 patients.

Considering the dataset consists of two and four-chamber acquisitions, this implies that each fold comprises a total of 100 echocardiographic images (50 patients x 2 acquisitions per patient).

The dataset, available to the community, includes:

1. A training set of 450 patients, accompanied by manual references derived from expert clinical analysis.
    1.1. Out of 450 patients, 400 are used as a train set.
    1.2. 50 patients were used as a validation set.
2. A testing set comprising 50 additional patients. Raw input images are provided in the raw/mhd file format.

One of the notable features of the CAMUS dataset is its large size, containing thousands of high-resolution images captured from multiple viewpoints and cardiac cycles. This extensive coverage enables researchers to explore cardiac dynamics comprehensively and facilitates the development and evaluation of advanced image analysis algorithms, such as segmentation and motion tracking techniques.

The CAMUS dataset includes annotated ground truth data, delineated by expert clinicians, which serve as reference standards for algorithm validation and performance assessment. These annotations provide precise delineation of cardiac structures, including the left ventricle endocardium, left ventricle epicardium and left atria enabling accurate evaluation of segmentation algorithms' efficacy. Annotation of End Diastole and End systole has been shown in Fig. 2. The dataset include 206 good quality images, 200 medium quality images and 44 poor quality images. The dataset consists of a total 150 female and 300 male patients varying from age 20 to 85 years. Detailed description is given in Fig. 3.

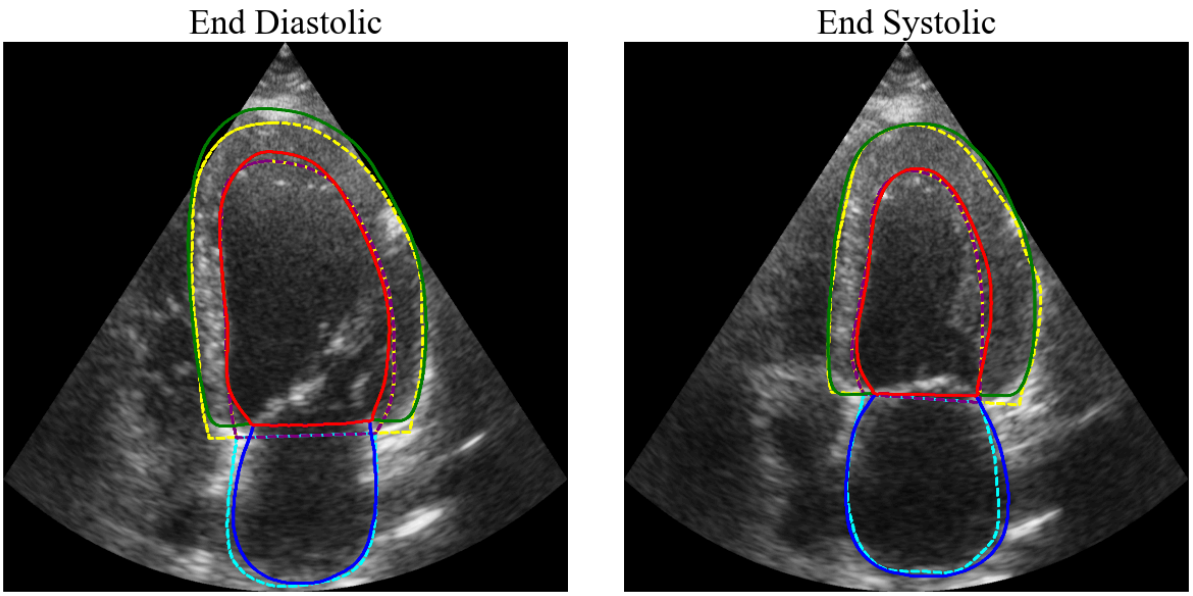
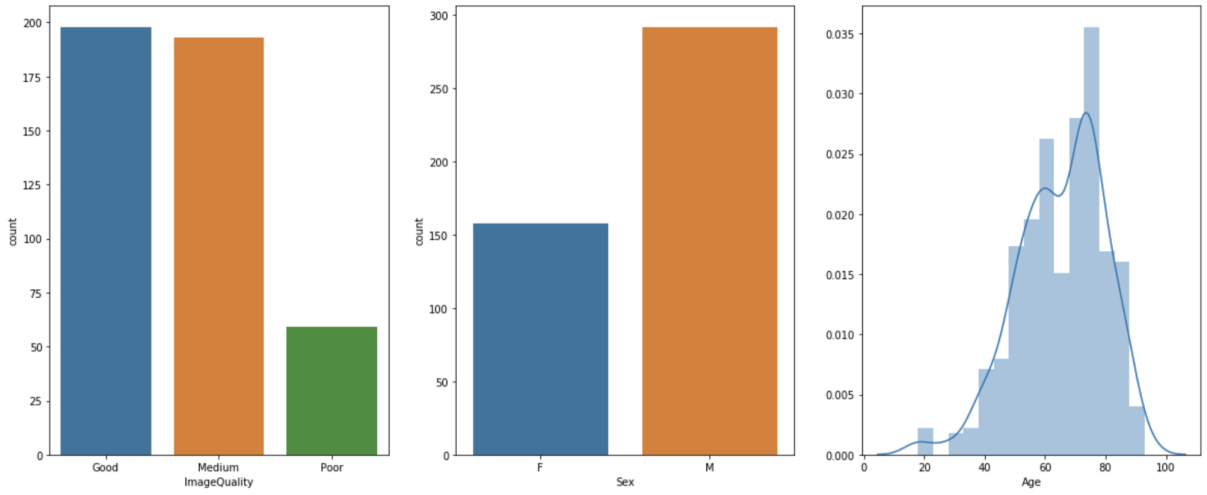


Fig. 2 Standard images taken from dataset. They show the endocardium and epicardium of the left ventricle and left atrium wall, marked in green, red, and blue, respectively. On the left side is the annotated ED, and on the right is the annotated ES.



(a) Image Quality and count (I) (b) Sex and count (II) (c) Age and count (III)

Fig. 3 illustrates the distribution of images on the bases of image quality (a), sex (b) and age (c) w.r.t total number of echocardiogram images of CAMUS dataset.

3.1.2 Data Cleaning:

Low-quality echocardiographic images were excluded from the training dataset which constitutes 19% of total images, to prevent potential negative effects on the performance of deep learning models. By removing these images, the training process could focus on high-quality data, enhancing the model's ability to learn relevant features and patterns

effectively. Eliminating low-quality images helped maintain the robustness of the deep learning models against noise and variability in image quality. This approach ensured that the models could generalise well to unseen data and cope effectively with the variability present in echocardiographic images.

Cross-Validation for Error Detection:

1. By dividing the dataset into 10 folds for cross-validation, the process enabled the identification of potential inconsistencies or outliers in the data distribution.
2. Through this iterative validation approach, any data cleaning issues could be detected and resolved, ensuring the dataset's overall integrity for subsequent analyses.

3.1.3 Equal distribution of data:

1. Cross-Validation Approach:
    a. The researchers adopted a cross-validation technique by dividing the dataset into 10 folds. This method ensured through training and evaluation coverage across different data subsets.
    b. Utilising this strategy enhanced the reliability of the model assessment by validating its performance on diverse data partitions.
2. Allocation of Training and Validation Sets:
    a. The dataset was segregated into distinct training and validation sets. This segregation facilitated model training and performance evaluation.
    b. Training on one subset while validating on another aided in preventing overfitting and provided insights into the model's ability to generalise to unseen data.
3. Progressive Model Assessment:
    a. Models underwent training on varying patient sample sizes, with evaluations conducted on fold 5 to monitor segmentation performance.
    b. This iterative evaluation process allowed for a comprehensive analysis of model performance with increasing training dataset sizes, offering valuable insights into the model's learning progression and convergence patterns.

3.2 Data Augmentation:

1. Data augmentation techniques were crucial for enhancing the robustness and generalisation of the deep learning models trained on the echocardiographic dataset.
2. The study employed various augmentation strategies such as rotation, scaling, and flipping to artificially increase the diversity of the training data [1].
3. These augmentations help the models learn invariant representations of the cardiac structures, making them more adaptable to variations in orientation and size within the images.

4. By augmenting the dataset, the models can better capture the underlying patterns in the echocardiographic images, leading to improved segmentation performance and clinical index estimation [3].
5. Various techniques includes -
    a. Rotation: Images were rotated to different angles to simulate variations in orientation, helping the models learn invariant representations of cardiac structures.
    b. Scaling: Resizing images to different scales allowed the models to capture variations in size and perspective, enhancing their ability to segment structures accurately.
    c. Flipping: Horizontal and vertical flipping of images provided additional training samples, aiding the models in learning robust features regardless of the image orientation.

3.3 Image Processing:

1. The study implemented preprocessing techniques to enhance the echocardiographic images before feeding them into the deep learning models.
2. Techniques like normalisation were used to standardise pixel values across images i.e. their original scale (ranging from 0 to 255 for 8-bit images) to a grey scale, promoting consistency and aiding model training.
3. Resizing to 256 * 256 pixels and cropping were applied to ensure uniform image dimensions, facilitating model input standardisation.
4. Noise reduction methods were employed to improve image quality by reducing artefacts and enhancing the visibility of cardiac structures.
5. Tailored preprocessing steps addressed challenges unique to echocardiographic imaging, such as contrast issues and brightness variations, to improve segmentation accuracy.
6. Overall, the preprocessing steps aimed to optimise the images for accurate segmentation of cardiac structures and clinical index estimation, enabling the deep learning models to effectively learn from the data.

3.4 Theoretical basis:

While evaluating and building our segmentation model we will require various mathematical and general terms which are explained below.

3.4.1 Activation Function:

1. ReLU Activation Function: Rectified Linear Unit activation function is a popular choice in neural network which can be represented as:

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

   Here, x is the input to the function and f(x) is the output. The ReLU function outputs the input value *x* if it is positive, and zero otherwise. This means that the function passes through any positive input unchanged, while setting all negative inputs to zero. The key advantage of ReLU lies in its ability to introduce non-linearity into the network, allowing it to learn complex patterns and relationships in the data. Additionally, ReLU's computational efficiency and avoidance of the vanishing gradient problem.

2. SoftMax Activation Function: It is a type of sigmoid function applied to the output of a neural network to produce a probability distribution for each pixel across different classes or categories. For every pixel in the image, the neural network computes raw scores (logits) indicating the likelihood of belonging to each class. These logits are then processed through the softmax function independently for each pixel position. Essentially, the softmax function normalises the logits into probabilities, ensuring that they sum up to one. This transformation enables pixel-wise classification, where each pixel is assigned a probability of belonging to each segmentation class, facilitating the segmentation process by identifying the most probable class for each pixel. Softmax function is given by:

   $$p(y = j/x) = \exp(z\_j) / \Sigma \exp(z\_k)$$

   It maps the linear combination of input features to a probability distribution over the different classes. For j = 1,2,...,K classes, where K is the number of classes. Here, z_j is the linear combination of input features for class j, and $\Sigma \exp(z\_k)$ sums over all K classes.

3.5 Model Evaluation:

3.5.1 Dice Loss: It measures the similarity or overlap between two sets. It quantifies how well the predicted region (e.g., object boundaries) aligns with the ground truth region (the actual location of objects in the image). The Mean dice coefficient ranges from 0 to 1. Higher the dice coefficient higher the similarity between predicted and ground truth mask.

$$\text{Dice coefficient} = 2 * |X \cap Y| / (|X| + |Y|)$$

Where,

|X| represents the number of pixels in predicted mask

|Y| represents the number of pixels in Ground Truth mask

|X ∩ Y| represents the number of overlapping pixels in both masks

3.5.2 Mean Hausdorff: It measures the average distance between the boundaries of the predicted segmentation mask and the ground truth mask. And then averages these distances over all points in the mask in both masks. The value ranges from 0 to infinity. Lower values indicating better segmentation accuracy. Mathematically, the Mean Hausdorff Distance between two sets X and Y is defined as:

$$d_{AHD(x,\,y)} = \left(\frac{1}{X} \sum_{x \in X} \min d\,(x,\,y) + \frac{1}{Y} \sum_{y \in Y} \min\, d\,(x,\,y) \right) / 2$$

Where,

x represents the set of points or contour pixels in the predicted segmentation mask.

y represents the set of points or contour pixels in the ground truth segmentation mask.

|x| and |y| denote the cardinalities of sets X and Y, respectively.

d(x, y) calculates the distance between points x, y in Euclidean space.

3.5.3 MAD: It is the average distance between the observation points and the mean. The value ranges from 0 to infinity. Lower values are associated with closely related data points.

$$MAD = \sum \frac{|x_i - \overline{x}|}{n}$$

Where,

MAD = average value of the data set
n = number of data values
xi = data values in the set
$\bar{x}$ = mean value

3.5.4 Correlation Coefficient: It is used to measure the extent of the relationship between two or more variables. The correlation coefficient ranges from -1 to + 1. -1 indicates the predicted and ground truth masks are uncorrelated, whereas +1 indicates the predicted and ground truth masks are completely correlated.

$$r_{xg} = \frac{n\sum_{i=1}^{n} x_i g_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} g_i}{\sqrt{n\sum_{i=1}^{n} x^2 - \left(\sum_{i=1}^{n} x_i\right)^2}.\sqrt{n\sum_{i=1}^{n} g_{i^2} - \left(\sum_{i=1}^{n} g_i\right)^2}}$$

Where,

n = Data quantity or number of data available

Σx = Total of the predicted data point

Σg = Total of the ground truth data point

Σxg = Sum of the Product of predicted & ground truth data points

Σx2 = Sum of the Squares of the predicted data point

Σg2 = Sum of the Squares of the ground truth data points

3.5.5 Precision: It measures the accuracy of the model's positive predictions by evaluating the proportion of correctly segmented pixels relative to all pixels predicted as positive. Higher precision values indicate lesser false positives and better accuracy.

Precision = (pred_mask * ground_mask) / pred_mask

Here, pred_mask and ground_mask are the predicted mask and ground truth mask respectively.

3.5.6 Recall: It assesses the model's ability to capture all positive instances, measuring the proportion of correctly segmented pixels relative to all the pixels that should have been identified as positive. A higher recall value specifies lesser missed detections and better performance in capturing relevant pixels of the target class.

Recall = (pred_mask * ground_mask) / ground_mask

Here, pred_mask and ground_mask are the predicted mask and ground truth mask respectively.

3.5.7 Jaccard Index (I0U): It measures the similarity between finite sample sets A,B as the Intersection over Union (IoU). Its value ranges from 0 to 1 where 0 indicates no overlap and

1 indicates complete overlap between predicted and ground truth mask.

$$\text{IoU} = \frac{\text{predMask * groundtruthMask}}{\text{predMask} + \text{groundtruthMask - predMask * groundtruthMask}}$$

3.5.6 **MAE** - It measures the average of the errors' magnitude between the predicted and actual values. Its value ranges from 0 to infinity where lower values are better.

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \overline{y_i}|$$

Where,
n: number of observation
yi: the actual value of the ith observation
yi': the predicted value of the ith observation

3.6 Model Selection:

In this study, our aim is to evaluate the effectiveness of Convolutional Neural Networks (CNNs) in segmenting 2D echocardiographic images. To accomplish this, we decided to concentrate on encoder-decoder networks (EDNs), which have proven to be fundamental in various CNN architectures that have shown success in medical imaging tasks [23]. EDNs utilise a two-stage convolutional network design, making them particularly suitable for segmentation tasks. The encoder, the initial segment, comprises a sequence of convolutional layers and downsampling steps. These processes are aimed at capturing image features while reducing spatial dimensions, facilitating the extraction of abstract features. Following the encoder, the decoder utilises the extracted features and employs a series of convolutional layers and upsampling steps. This gradual transformation of feature maps by the decoder ultimately produces the final segmentation output.

The selection criteria include model performance, computational efficiency and scalability. Each model is implemented using popular DL framework TensorFlow, and experimentally tuned hyperparameters.

Model Architecture:
For cardiac image segmentation, following fully connected layers are used:

A. ReLU activation functions to introduce non-linearity.
B. Batch normalisation layers to stabilise and accelerate training.
C. Softmax activation function in the final layer for semantic segmentation.
D. Dropout layers are removed to enhance the model capacity to learn intricate features.
E. Global average pooling layers to reduce spatial dimensions and extract features.

Training Parameters:
Training consists of 30 epochs with variable batch size, indicating that the entire dataset is divided into batches of samples each for training.

ADAM optimizer is used with the following parameters:
Learning rate: 1e4
Decay: 1e6
Momentum: 0.9

A. U-Net:
The U-Net architecture is a widely-used convolutional neural network (CNN) architecture primarily employed for image segmentation tasks, particularly in biomedical imaging. Its distinctive structure consists of an encoder-decoder framework, characterised by a contracting path followed by an expansive path.

At the core of the U-Net architecture lies its encoder, which is responsible for extracting features from input images. A total of 3 downsampling layers in encoder were used which comprises a series of 3*3 convolutional layers while doubling the number of features for downsampling interspersed with rectified linear unit (ReLU) activation functions, which introduce non-linearity to the network. These convolutional layers are followed by max pooling layers, which downsample the feature maps, thereby reducing spatial dimensions and extracting high-level features.

After the encoder, the network transitions to the decoder phase, where the upsampled feature maps are gradually reconstructed to produce the final segmentation output. In the decoder, total of 3 upsampling layers were used with 3*3 convolutional layers with ReLU activation functions are employed to learn intricate spatial patterns, while upsampling operations are utilised to restore spatial dimensions. Number of features are reduced in half per sampling. Additionally, skip connections are incorporated between corresponding encoder and decoder layers, facilitating the retention of fine-grained spatial information and aiding in the recovery of detailed features. The final 3*3 convolutional layers of 384 dimension each are used and softmax classification technique is used. Detailed Architecture of U-net is shown in Fig. 4.

Batch size [3] of 8 has been implemented. The loss function [5] used in the training of this architecture is multi-class dice loss. Adam [12] optimizer is used in loss minimization. Total number of trainable parameters for this architecture is 18M. Standard evaluation matrices in the form of dice coefficient, mean hausdorff, MAD, accuracy, recall and precision are used in both training and testing data.

Overall, the U-Net architecture leverages an encoder-decoder structure, featuring convolutional layers, max pooling layers, ReLU activation functions, and normalisation layers. This design enables the network to effectively capture and integrate spatial information from input images while preserving fine details crucial for accurate segmentation outputs.

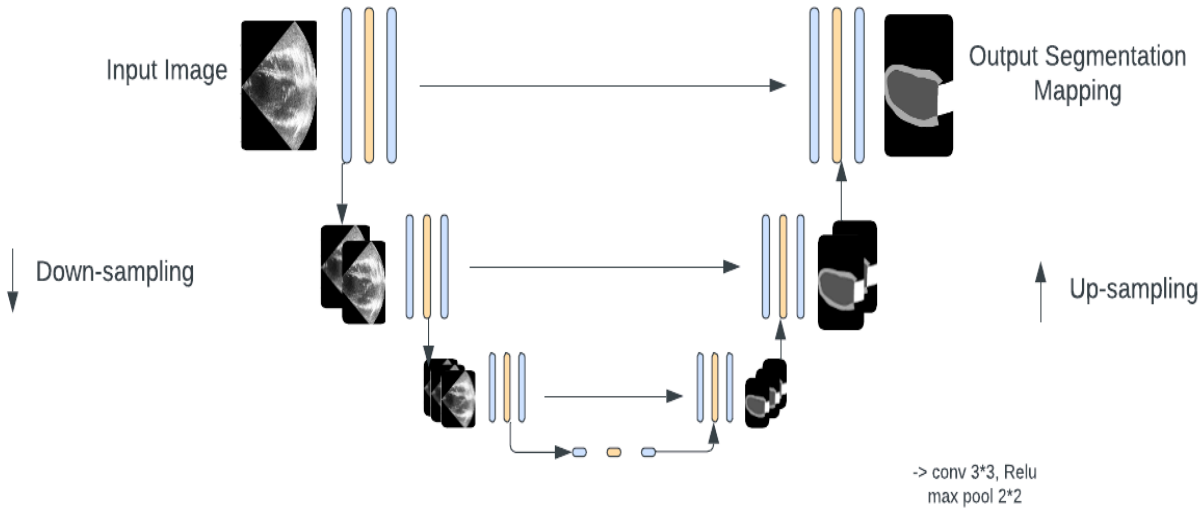


Fig. 4 The U-Net encoder-decoder deep learning architecture representation

B. ACNN [3]:

Attention Convolutional Neural Network, introduces a novel approach aimed at achieving comprehensive feature processing at full resolution. It strategically employs an optimal dilation setting to expand the receptive field, all while keeping parameter count minimal. This innovation allows ACNN to outperform alternative techniques, yielding superior segmentation Intersection over Union (IoU) scores with significantly fewer trainable parameters and overall model sizes. This underscores the advantages associated with employing full-resolution feature maps for thorough feature processing tasks.

The ACNN (Auxiliary Convolutional Neural Network) model implemented in the study integrated an auxiliary loss to enhance segmentation accuracy by aligning the segmentation output with a compact representation of the underlying anatomy[5]

The ACNN model utilised the U-Net 1 architecture as the segmentation module, incorporating specific design choices to optimise performance on the dataset. Receptive field (RF) values are chosen as 1537, giving us a total of 192 Atrous II-blocks[3].

Key parameters in the ACNN model included a code of 32 coefficients for the auto-encoder network, enabling an average reconstruction accuracy of 97, and a hyperparameter balancing segmentation and shape regularisation losses for improved accuracy**[**12]

With a total of 5.5 million parameters, the ACNN model aimed to refine the segmentation process by leveraging anatomical constraints derived from the auto-encoder network, enhancing the model's ability to accurately delineate cardiac structures in 2D echocardiographic images.

Through these tailored design elements, the ACNN model sought to improve segmentation precision and shape regularisation, contributing to more accurate and reliable analysis of cardiac structures in clinical applications. Detailed architecture is shown in Fig. 5.

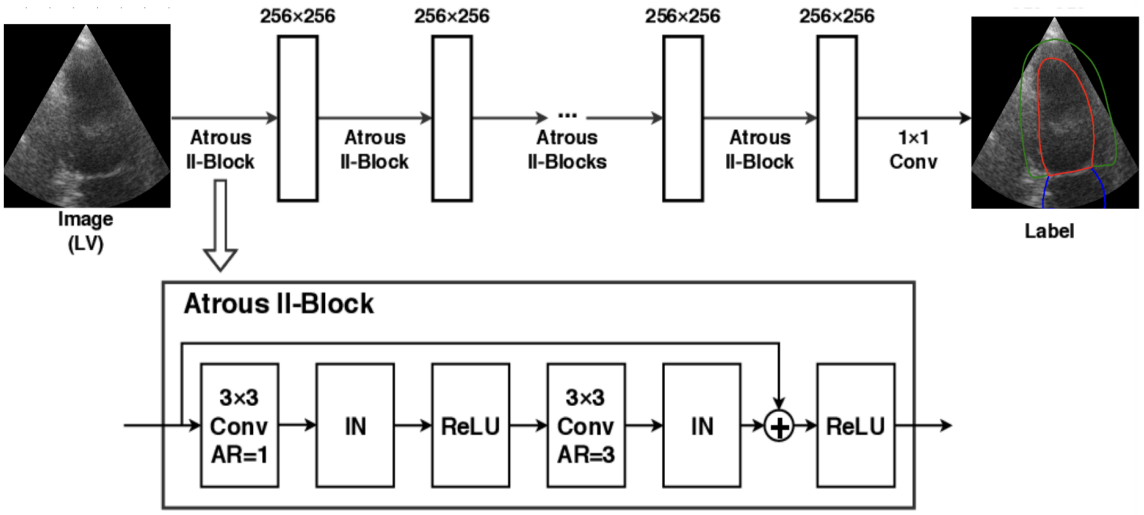


Fig. 5 The architecture of ACNN. It involves a series of residual II-blocks, with the number of these blocks determined by a formula (RF - 1) / 8, based on the targeted receptive field (RF). The atrous rate (AR) is used for atrous convolutions, with a kernel size of 3 for 3x3 Conv and a kernel size of 1 for 1x1 Conv1.

C. NN-Unet:

The NN-Unet architecture is an extension of the U-Net architecture, designed to enhance its performance in medical image segmentation tasks. Similar to U-Net, NN-Unet also follows an encoder-decoder framework but incorporates additional features to improve segmentation accuracy.

In the encoder part of NN-Unet, a total of 5 times down-sampling is done until the feature map reaches either 4 voxels or the feature map space becomes anisotropic, followed by a rectified linear unit(ReLU). This down-sampling process occurs separately for each axis, with high-resolution axes down-sampled until their resolution is approximately half that of the lower resolution axis. Each axis is handled individually during the down-sampling process, stopping when the feature map constraints are met. By default, convolutions are performed with a kernel size of 3×3 for 2D U-Net. This increased depth allows for better feature extraction and representation, which is particularly beneficial for complex segmentation tasks in medical imaging. Additionally, batch normalisation layers are included after each convolutional layer to stabilise and accelerate the training process by normalising the activations.

Rather than traditional max pooling layers used in U-Net, NN-Unet employs more sophisticated downsampling techniques. In this case we have used 2*2 strided convolutions [7]. These alternatives help preserve spatial information better while reducing the spatial dimensions of feature maps, thereby enhancing the network's ability to capture fine details.

In the decoder part of NN-Unet, upsampling operations are performed to gradually reconstruct the spatial dimensions of feature maps. Skip connections, similar to those in U-Net, are utilised to concatenate feature maps from the encoder to the corresponding decoder layers. This facilitates the integration of low-level and high-level features, aiding in the precise localization of segmentation boundaries. Model is combined through ensemble averaging, where softmax probabilities are averaged. Standard evaluation matrices in the form of dice coefficient, mean hausdorff, MAD, accuracy, recall and precision are used in both training and testing data. Detailed architecture is shown in Fig. 6.

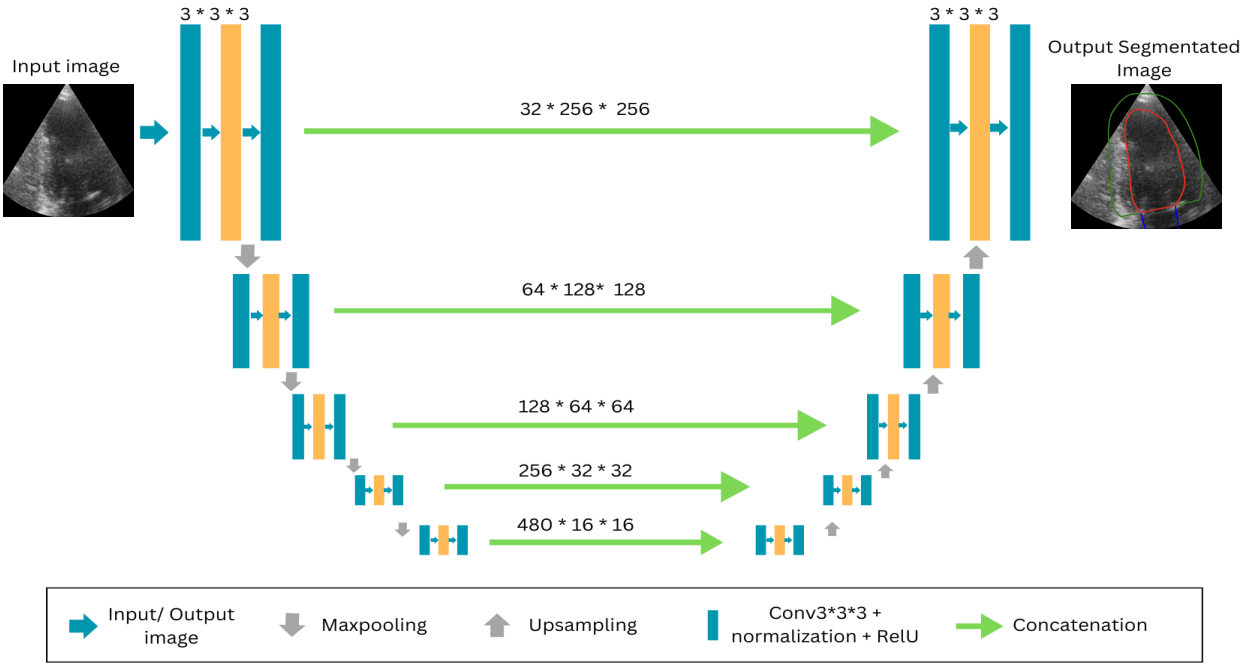


Fig. 6 illustration of NN-Unet network architecture.

TABLE I

Main characteristics of all 3 DNN segmentation model architecture including U-Net, ACNN and NN-Unet

Architectures	Lowest Resolution	Upsampling Scheme	Downsampling Scheme	Normalisation Scheme	Batch Size	Learning Rate	Loss Function	Trainable Parameters	Optimizer
U-Net	24 * 24	3 * 3 repeats	3 * 3 repeats	None	32	1e-4	Multi-class Dice Loss	18M	ADAM
ACNN	32 * 32	Deconvolution	3 * 3 Atrous convolution	Batch normalisation	32	1e-4	Multi-class Dice Loss/ Auxiliary Loss	5.5M	ADAM
NNU-Net	16 * 16	Deconvolution	convolution	Batch normalisation	32	1e-4	Multi-class Dice Loss	2.4M	ADAM

Chapter 4

Experimental Results

We excluded 19% of low-quality images from the analysis to ensure the accuracy of the metrics presented in this section. Additionally, to maintain consistency across acquisition settings, we opted to train a single model for each machine learning approach using two-chamber (2Ch) views, regardless of the specific time point in the cardiac sequence. The trained models are evaluated on the basis of accuracy, precision, recall along with mean dice coefficient, MAE, MAD, mean correlation, jaccard index and hausdorff distance.

1. **U-Net:**

U-Net achieved the accuracy of 95.7% along with a mean dice score of 94.7% for the left ventricle (endocardium). Similar results are also achieved for the left ventricle (epicardium) and left atrium. Mean hausdorff value of 4.1 which is significantly less. It also achieved the correlation for LVED of 0.959. Mean Dice Plot's results have been shown in Fig. 7. Fig. 8 shows the training and testing accuracy, loss curve and dice coefficient graph for ED and ES respectively.

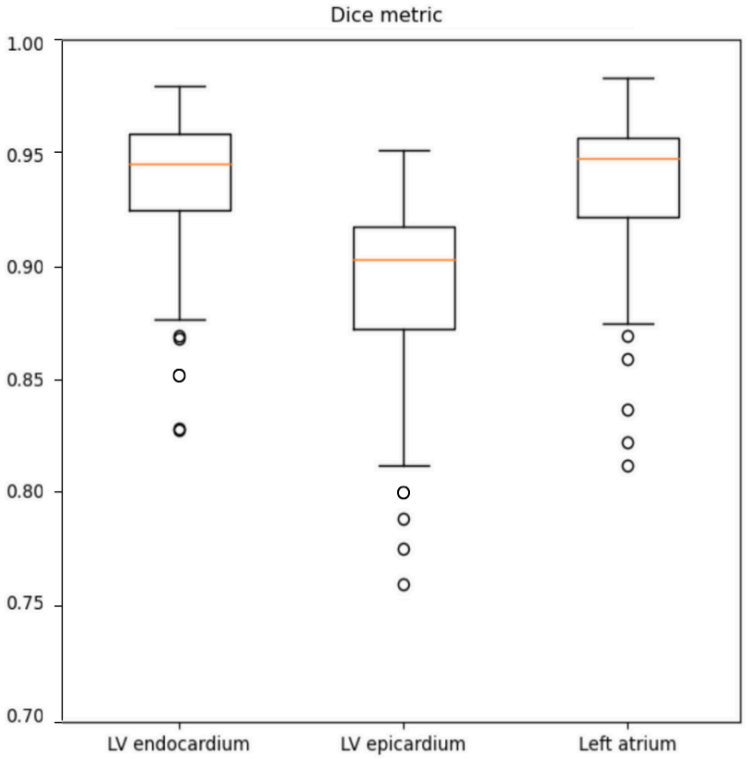
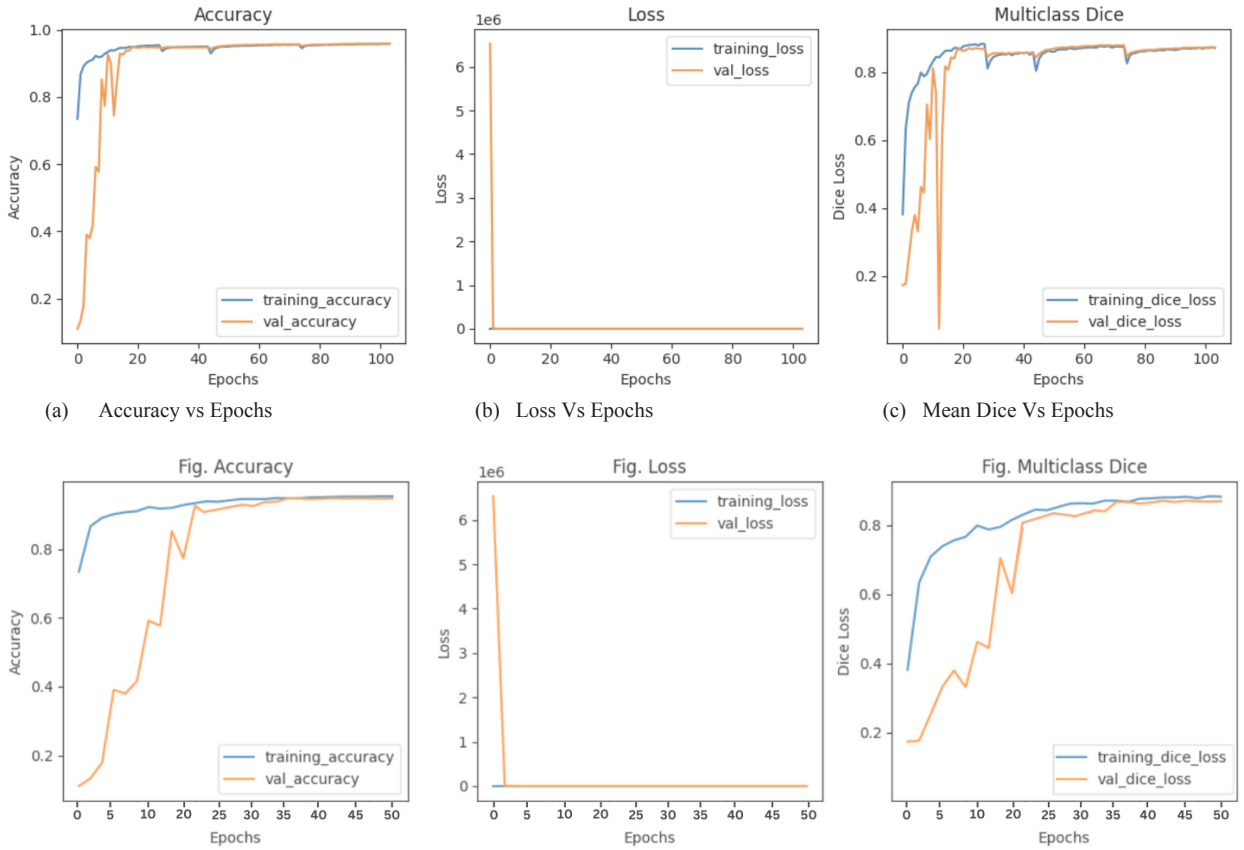


Fig. 7 Dice score for LV endocardium, LV epicardium and left atrium for U-Net architecture



(a) Accuracy vs Epochs (b) Loss Vs Epochs (c) Mean Dice Vs Epochs

(e) Accuracy vs Epochs (f) Loss Vs Epochs (g) Mean Dice Vs Epochs

Fig. 8 (a), (b), (c) shows the accuracy, loss and mean dice plots Vs epochs for LVED. Whereas (e), (f), (g) shows the accuracy, loss and mean dice plots Vs epochs for LVES for U-Net architecture.

.

2. ACNN:

U-Net achieved the accuracy of 93.6% along with a mean dice score of 93.3% for the left ventricle (endocardium). Similar results are also achieved for the left ventricle (epicardium) and left atrium. Mean hausdorff value of 5.4 which is significantly less. It also achieved the correlation for LVED of 0.930. Mean Dice Plot's results have been shown in Fig. 9. Fig. 10 shows the training and testing accuracy, loss curve and dice coefficient graph for ED and ES respectively.

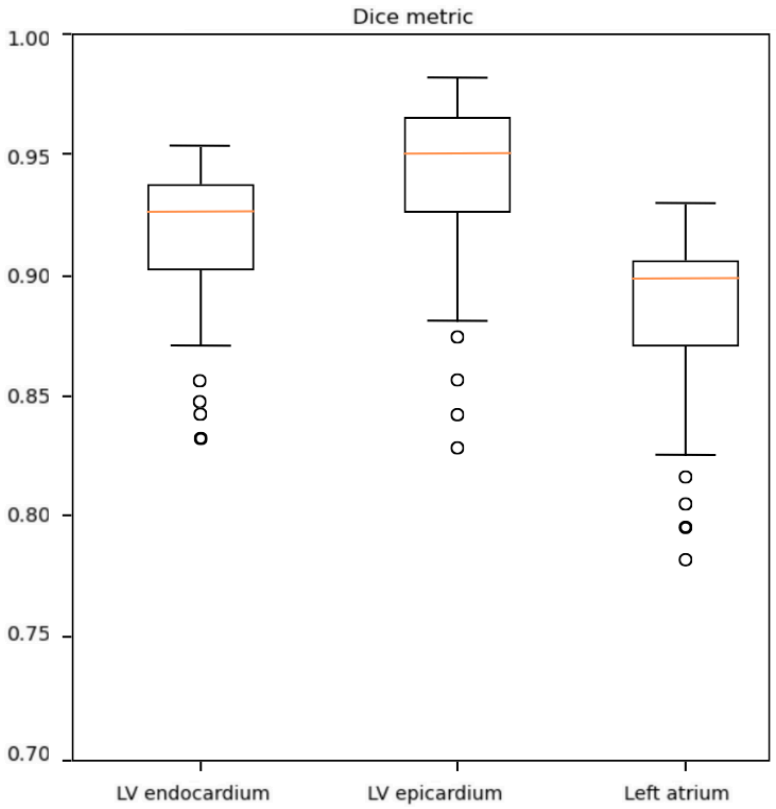
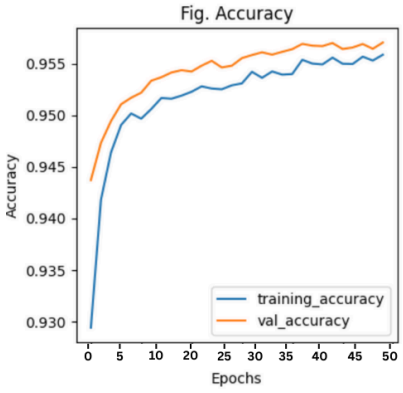
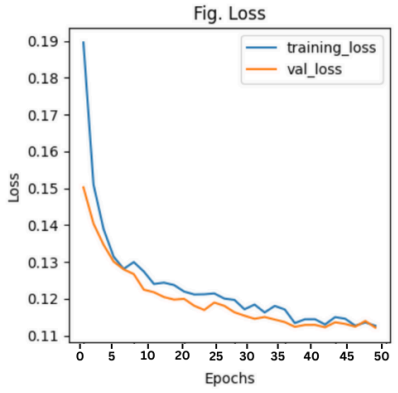


Fig. 9 Dice score for LV endocardium, LV epicardium and left atrium for ACNN architecture

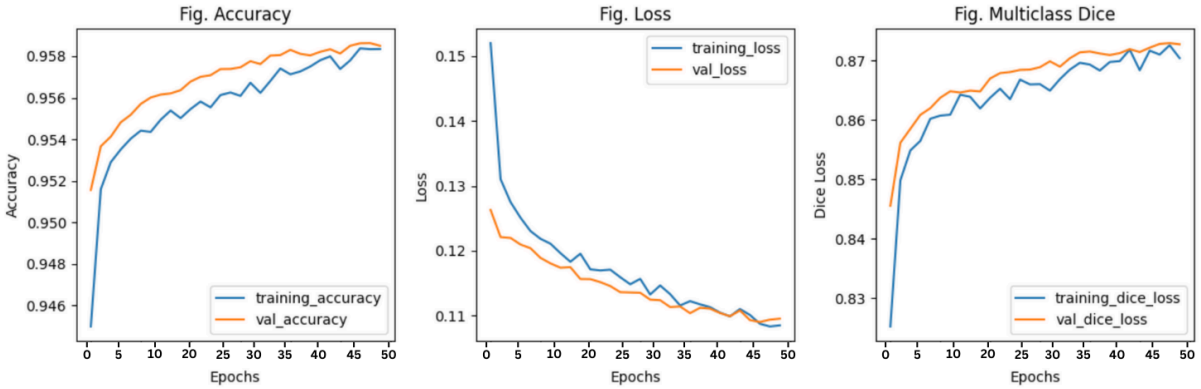


(a) Accuracy vs Epochs



(b) Loss Vs Epochs

(c) Mean Dice Vs Epochs



(e) Accuracy vs Epochs (f) Loss Vs Epochs (g) Mean Dice Vs Epochs

Fig. 10 (a), (b), (c) shows the accuracy, loss and mean dice plots Vs epochs for LVED. Whereas (e), (f), (g) shows the accuracy, loss and mean dice plots Vs epochs for LVES for ACNN architecture.

3. NN-Unet:

NN-Unet performed the best as compared to U-net and ACNN.It achieved the accuracy of 96.7% along with a mean dice score of 95.4% for the left ventricle (endocardium). Similar results are also achieved for the left ventricle (epicardium) and left atrium. Mean hausdorff value of 4.5 which is significantly less. It also achieved the correlation for LVED of 0.983. Mean Dice Plot's results have been shown in Fig. 11. Fig. 12 shows the training and testing accuracy, loss curve and dice coefficient graph for ED and ES respectively.

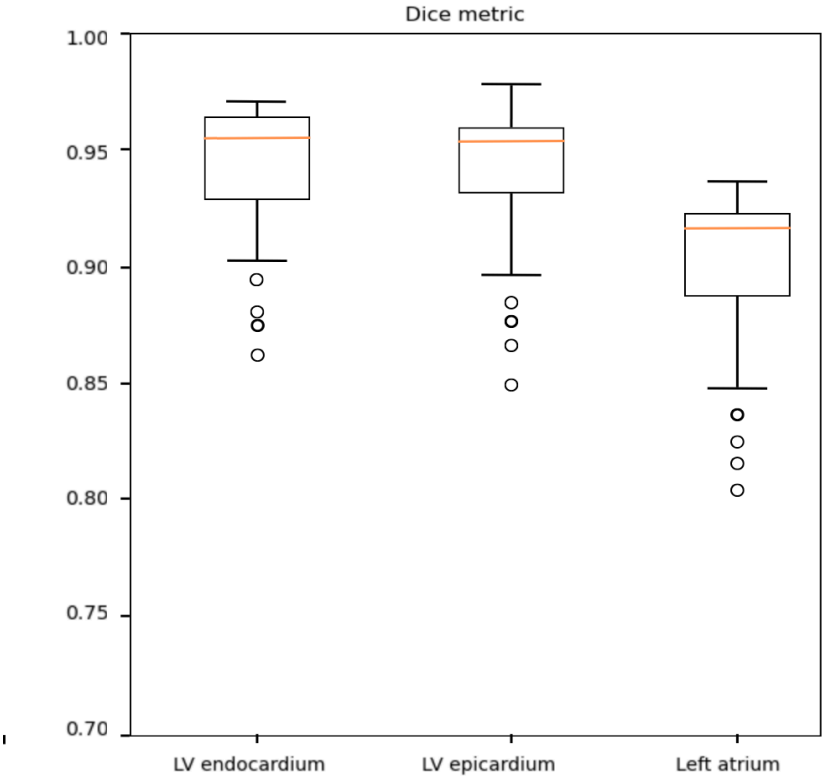
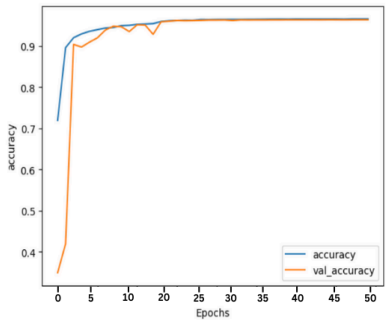
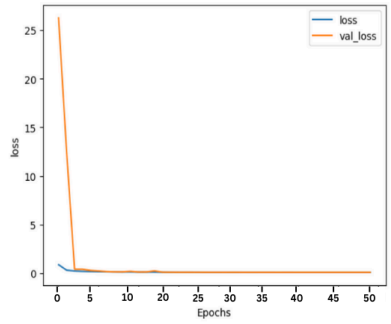


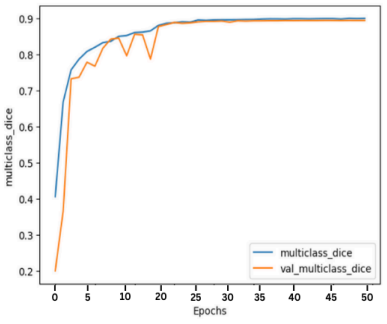
Fig. 11 Dice score for LV endocardium, LV epicardium and left atrium for NN-Unet architecture



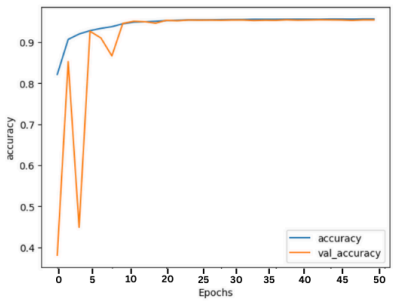
(a) Accuracy vs Epochs



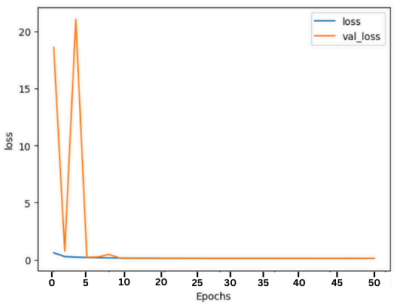
(b) Loss Vs Epochs



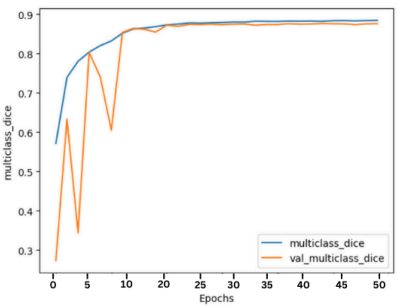
(c) Mean Dice Vs Epochs



(e) Accuracy vs Epochs



(f) Loss Vs Epochs



(g) Mean Dice Vs Epochs

Fig. 12 (a), (b), (c) shows the accuracy, loss and mean dice plots Vs epochs for LVED. Whereas (e), (f), (g) shows the accuracy, loss and mean dice plots Vs epochs for LVES for NN-Unet architecture.

Overall NN-Unet provided the best results of LVEndo, LVEpi and Left Atrium segmentation. Table II, Table III and Table IV shows the mean dice score, mean hausdorff, MAD, correlation, accuracy, precision, recall, MAE and jaccard Index for left ventricle endocardium (LVEndo), left ventricle epicardium (LVEpi) and left atrium ED and ES respectively.

Table II:

Segmentation accuracy for LV Endocardium of the 3 evaluated segmentation methods i.e. U-Net, ACNN and NN-Unet on 50 test patients having good and medium image quality. The values in bold refer to the best performance for each measure for LVEndo. All the scores obtained with NN-Unet show that it is the best performing method overall.

	Mean Dice ED	**Mean Dice ES**	**Mean Hausdorff ED**	**Mean Hausdorff ES**	**MAD ED**	**MAD ES**	**Correlation ED**	**Correlation ES**
U-Net	0.947	0.925	**4.1**	4.4	1.5	**1.3**	0.938	0.959
ACNN	0.933	0.915	5.4	5.2	1.7	1.7	0.930	0.947
NN-Unet	**0.954**	**0.949**	4.5	**4.3**	**1.5**	1.5	**0.978**	**0.983**

	Accuracy ED (%)	**Accuracy ES (%)**	**Precision ED**	**Precision ES**	**Recall ED**	**Recall ES**	**MAE ED**	**MAE ES**	**Jaccard Index ED**	**Jaccard Index ES**
U-Net	95.7	**95.8**	0.938	0.927	0.873	0.912	11.2	7.5	0.889	0.873
ACNN	93.6	94.8	0.932	0.911	0.890	0.934	9.7	6.9	0.905	**0.901**
NN-Unet	**96.7**	94.6	**0.950**	**0.943**	**0.953**	**0.947**	**5.9**	**4.0**	**0.915**	0.891

Table III:

Segmentation accuracy for LV Epicardium of the 3 evaluated segmentation methods i.e. U-Net, ACNN and NN-Unet on 50 test patients having good and medium image quality. The values in bold refer to the best performance for each measure for LVEpi. All the scores obtained with NN-Unet show that it is the best performing method overall.

	Mean Dice ED	**Mean Dice ES**	**Mean Hausdorff ED**	**Mean Hausdorff ES**	**Mean MAD ED**	**Mean MAD ES**	**Correlation ED**	**Correlation ES**
U-Net	0.958	0.936	5.2	5.7	1.7	1.9	0.915	0.902
ACNN	0.946	0.953	5.8	5.7	1.8	2.0	0.928	0.910
NN-Unet	**0.968**	**0.928**	**4.6**	**4.4**	**1.5**	**1.5**	**0.948**	**0.953**

	Accuracy ED (%)	**Accuracy ES (%)**	**Precision ED**	**Precision ES**	**Recall ED**	**Recall ES**	**MAE ED**	**MAE ES**	**Jaccard Index ED**	**Jaccard Index ES**
U-Net	93.8	95.6	0.925	0.933	0.907	0.917	9.9	6.7	0.926	0.915
ACNN	94.6	93.2	0.945	**0.949**	0.869	0.892	10.1	7.8	0.893	0.897
NN-Unet	**97.1**	**96.3**	**0.963**	0.915	**0.947**	**0.955**	**6.2**	**5.4**	**0.945**	**0.938**

Table IV:

Segmentation accuracy for left atrium of the 3 evaluated segmentation methods i.e. U-Net, ACNN and NN-Unet on 50 test patients having good and medium image quality. The values

in bold refer to the best performance for each measure for LA. All the scores obtained with NN-Unet show that it is the best performing method overall.

	Mean Dice ED	**Mean Dice ES**	**Mean Hausdorff ED**	**Mean Hausdorff ES**	**Mean MAD ED**	**Mean MAD ES**	**Correlation ED**	**Correlation ES**
U-Net	0.890	0.921	5.6	5.4	2.1	2.0	0.899	0.895
ACNN	0.881	0.870	5.9	6.0	2.3	2.3	0.926	**0.945**
NN-Unet	**0.912**	**0.929**	**5.1**	**4.3**	**2.0**	**1.8**	**0.987**	0.927

	Accuracy ED (%)	**Accuracy ES (%)**	**Precision ED**	**Precision ES**	**Recall ED**	**Recall ES**	**MAE ED**	**MAE ES**	**Jaccard Index ED**	**Jaccard Index ES**
U-Net	93.1	92.6	0.874	0.917	0.889	0.845	13.4	10.7	0.889	**0.895**
ACNN	93.8	91	**0.969**	**0.958**	0.937	0.943	10.9	9.8	0.88	0.890
NN-Unet	**94.2**	**94.1**	0.902	0.931	**0.950**	**0.949**	**7.4**	**6.0**	**0.894**	0.883

Runtime Performance

The three segmentation methods were implemented in Python using the same versions of TensorFlow and Keras libraries, and they were run on an Apple M2 chip with 8-core GPU, a 16-core Neural Engine and 100GB/s memory width. Due to the higher number of trainable parameters in each segmentation method (as shown in Table I), the training times for these networks vary. Training on 400 patients takes approximately 150 ± 5 minutes for U-Net, it takes about 85 ± 5 minutes for ACNN, whereas it takes 54 ± 1 minutes for NN-Unet. For testing, segmenting a single image takes around 1.23 ± 0.03 seconds with U-Net, 0.43 ± 0.06 seconds with ACNN and 0.21 ± 0.04 seconds with NN-Unet.

Chapter 5

COMPARATIVE ANALYSIS

We have compared our results with methods mentioned in other researches including the dataset used along with methodology implemented and different evaluation methods and its comparative analysis has shown in Table V.

Table V:
Segmentation accuracy comparison for different evaluated segmentation methods. The values in bold refer to the best performance for each measure. All the scores obtained with NN-Unet show that it is the best performing method overall.

Title	Dataset	Method Used	Dice	Haus-dorff	MAD	Correlation	Jaccard Index(IOU)	MAE
[1]	STACOM	ResDUnet	0.867	6.5	1.6	-	-	-
[2]	Self-Collected	ACNN + DBN	-	-	-	0.86	0.8	-
[3]	UK Digital Heart Project	ACNN	0.866	-	-	-	-	-
[4]	CAMUS	SOCOF	0.814	-	-	0.918	-	-
[5]	Cardiac US image of canine LV	LVnet	0.902	-	-	-	**0.823**	-
[6]	Self-Collected	BEAS	0.825	-	-	-	-	-
[7]	Self-Collected	TransV-Net	0.902	4.9	1.5	-	-	-
[8]	EchoNet-Dynamic	CSS	0.919	4.17	-	-	-	4.9
[9]	CAMUS	PWC Net	0.79	-	-	0.84	-	**2.8**

	CAMUS	U-Net	0.947	**4.1**	**1.5**	0.938	0.889	11.2
	CAMUS	ACNN	0.933	5.4	1.7	0.930	0.905	9.7
	CAMUS	NNU-Net	**0.954**	4.5	**1.5**	**0.978**	**0.915**	**5.9**

Chapter 6

CONCLUSION

In conclusion, this research demonstrates efficacy of encoder-decoder networks, particularly the NN-Unet architecture, demonstrate high accuracy in segmenting cardiac structures in 2D echocardiography, showcasing a favourable trade-off between model complexity and performance, moreover striking a balance between the number of parameters and the achieved performance. Notably, it accurately replicated expert analysis for left ventricular volumes, demonstrating strong correlations for end-diastole (ED) and end-systole (ES), along with minimal absolute mean distance of 1.5 for ED and 1.3 for ES. Moreover, NNU-Net yielded impressive mean dice scores for left atrial end-diastole and end-systole. These findings underscore the effectiveness of NNU-Net for accurate and precise segmentation in 2D echocardiography applications.

Overall, This study makes significant strides in the realm of computer-aided diagnosis in echocardiography, paving the way for the creation of reliable new clinical research. The results highlight the crucial role of selecting suitable deep learning architectures, prioritising computational efficiency, and delving into alternative image representations to enhance model efficacy and real-world applicability.

Despite the remarkable results achieved, there is room for improvement to match expert manual annotations, especially in estimating left ventricular volumes and ejection fraction. The study highlights the potential for accurate and fully-automatic analysis of 2D echocardiographic images, paving the way for enhanced clinical diagnosis and treatment planning in cardiology.

Chapter 7

FUTURE SCOPE

Future research could aim to develop more advanced deep learning models to further automate the analysis of 2D echocardiographic images, striving for even higher accuracy in segmenting cardiac structures such as right ventricle, right atrium, myocardium, chambers, valves. This improved automation could reduce the manual effort required by experts and advanced models can provide immediate analysis during echocardiographic examination.

Further enhancing current encoder-decoder architectures to improve the accuracy of left ventricular volume and ejection fraction estimations, with the goal of matching or surpassing the accuracy of clinical inter-observer assessments. This could involve techniques such as hyperparameter tuning, attention mechanisms, or using loss functions tailored for medical image segmentation.

Expanding the CA-MUS dataset with more cases and annotations, which would support the development of more robust and generalizable models for echocardiographic image analysis. A larger and more diverse dataset would enable the creation of models that are more generalised, reducing the risk of overfitting and thus improving the performance across different patient populations and imagining conditions. Enhanced dataset could also incorporate 3D echocardiographic images providing better foundation for training and testing segmentation models.

The results of this research could pave the way for integrating deep learning algorithms into clinical practice, enhancing the efficiency and accuracy of echocardiographic assessments and potentially leading to faster and more precise diagnoses in cardiology.

REFERENCES

[1] S. Leclerc *et al*., "Deep Learning for Segmentation Using an Open Large-Scale Dataset in 2D Echocardiography," in *IEEE Transactions on Medical Imaging*, vol. 38, no. 9, pp. 2198-2210, Sept. 2019, doi: 10.1109/TMI.2019.2900516.

[2] Amer, Alyaa, et al. "ResDUnet: Residual dilated UNet for left ventricle segmentation from echocardiographic images." 2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC). IEEE, 2020.

[3] Carneiro, Gustavo, Jacinto C. Nascimento, and António Freitas. "The segmentation of the left ventricle of the heart from ultrasound data using deep learning architectures and derivative-based search methods." IEEE Transactions on Image Processing 21.3 (2011): 968-982.

[4] Oktay, Ozan, et al. "Anatomically constrained neural networks (ACNNs): application to cardiac image enhancement and segmentation." IEEE transactions on medical imaging 37.2 (2017): 384-395.

[5] Xue, Wufeng, et al. "Improved segmentation of echocardiography with orientation-congruency of optical flow and motion-enhanced segmentation." IEEE Journal of Biomedical and Health Informatics 26.12 (2022): 6105-6115.

[6] Awasthi, Navchetan, et al. "LVNet: Lightweight model for left ventricle segmentation for short axis views in echocardiographic imaging." IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control 69.6 (2022): 2115-2128.

[7] Morais, Pedro, et al. "Fast segmentation of the left atrial appendage in 3-D transesophageal echocardiographic images." IEEE transactions on ultrasonics, ferroelectrics, and frequency control 65.12 (2018): 2332-2342.

[8] Zhang, Jiapeng, et al. "Dual-branch TransV-Net for 3D echocardiography segmentation." IEEE Transactions on Industrial Informatics (2023).

[9] Lin, Zihan, et al. "CLA-U-Net: Convolutional Long-short-term-memory Attention-gated U-Net for Automatic Segmentation of the Left Ventricle in 2-D Echocardiograms." 2022 IEEE International Ultrasonics Symposium (IUS). IEEE, 2022.

[10] Dai, Weihang, et al. "Cyclical self-supervision for semi-supervised ejection fraction prediction from echocardiogram videos." IEEE Transactions on Medical Imaging (2022).

[11] Mortada MJ, Tomassini S, Anbar H, Morettini M, Burattini L, Sbrollini A. Segmentation of Anatomical Structures of the Left Heart from Echocardiographic Images Using Deep Learning. Diagnostics (Basel). 2023 May 9;13(10):1683. doi: 10.3390/diagnostics13101683. PMID: 37238168; PMCID: PMC10217142.

[12] Yang, Tingyang, et al. "Segmentation of five components in four chamber view of fetal echocardiography." 2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI). IEEE, 2020.

[13] Østvik, Andreas, et al. "Myocardial function imaging in echocardiography using deep learning." ieee transactions on medical imaging 40.5 (2021): 1340-1351.

[14] Ahn, Shawn S., et al. "Multi-frame attention network for left ventricle segmentation in 3d echocardiography." Medical Image Computing and Computer Assisted

Intervention–MICCAI 2021: 24th International Conference, Strasbourg, France, September 27–October 1, 2021, Proceedings, Part I 24. Springer International Publishing, 2021.
[15] Naghne, R., et al. "An Efficient Capsule-based Network for 2D Left Ventricle Segmentation in Echocardiography Images." 2023 45th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC). IEEE, 2023.
[16] Cao, Jiawei. (2021). Research on crack detection of bridge deck based on computer vision. IOP Conference Series: Earth and Environmental Science. 768. 012161. 10.1088/1755-1315/768/1/012161.
[17] Monkam, Patrice, et al. "A Disentanglement and Fusion Data Augmentation Approach for Echocardiography Segmentation." 2022 IEEE International Ultrasonics Symposium (IUS). IEEE, 2022.
[18] Zhou, Guang-Quan, et al. "DSANet: Dual-branch shape-aware network for echocardiography segmentation in apical views." IEEE Journal of Biomedical and Health Informatics (2023).
[19] Sustersic, Tijana, Milos Anic, and Nenad Filipovic. "Heart left ventricle segmentation in ultrasound images using deep learning." 2020 IEEE 20th Mediterranean Electrotechnical Conference (MELECON). IEEE, 2020.
[20] T. Sustersic, M. Anic and N. Filipovic, "Heart left ventricle segmentation in ultrasound images using deep learning," *2020 IEEE 20th Mediterranean Electrotechnical Conference (MELECON)*, Palermo, Italy, 2020, pp. 321-324, doi: 10.1109/MELECON48756.2020.9140527.
[21] Y. Guo, Y. Wang, S. Nie, J. Yu and P. Chen, "Automatic Segmentation of a Fetal Echocardiogram Using Modified Active Appearance Models and Sparse Representation," in *IEEE Transactions on Biomedical Engineering*, vol. 61, no. 4, pp. 1121-1133, April 2014, doi: 10.1109/TBME.2013.2295376.
[22] O. S. Al-Kadi, "Spatio-Temporal Segmentation in 3-D Echocardiographic Sequences Using Fractional Brownian Motion," in *IEEE Transactions on Biomedical Engineering*, vol. 67, no. 8, pp. 2286-2296, Aug. 2020, doi: 10.1109/TBME.2019.2958701.
[23] Y. Zhu, C. Xiong, H. Zhao and Y. Yao, "SAM-Att: A Prompt-Free SAM-Related Model With an Attention Module for Automatic Segmentation of the Left Ventricle in Echocardiography," in *IEEE Access*, vol. 12, pp. 50335-50346, 2024, doi: 10.1109/ACCESS.2024.3384383.
[24] M. C. Brindise, B. A. Meyers, S. Kutty and P. P. Vlachos, "Automated Peak Prominence-Based Iterative Dijkstra's Algorithm for Segmentation of B-Mode Echocardiograms," in *IEEE Transactions on Biomedical Engineering*, vol. 69, no. 5, pp. 1595-1607, May 2022, doi: 10.1109/TBME.2021.3123612.
[25] S. Leclerc *et al*., "LU-Net: A Multistage Attention Network to Improve the Robustness of Segmentation of Left Ventricular Structures in 2-D Echocardiography," in *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 67, no. 12, pp. 2519-2530, Dec. 2020, doi: 10.1109/TUFFC.2020.3003403.
[26] W. -Y. Hsu, "Automatic Left Ventricle Recognition, Segmentation and Tracking in Cardiac Ultrasound Image Sequences," in *IEEE Access*, vol. 7, pp. 140524-140533, 2019, doi: 10.1109/ACCESS.2019.2920957.

[27] G. Zamzmi, L. -Y. Hsu, W. Li, V. Sachdev and S. Antani, "Harnessing Machine Intelligence in Automatic Echocardiogram Analysis: Current Status, Limitations, and Future Directions," in *IEEE Reviews in Biomedical Engineering*, vol. 14, pp. 181-203, 2021, doi: 10.1109/RBME.2020.2988295.
[28] Moradi S, Oghli MG, Alizadehasl A, Shiri I, Oveisi N, Oveisi M, Maleki M, Dhooge J. MFP-Unet: A novel deep learning based approach for left ventricle segmentation in echocardiography. Phys Med. 2019 Nov;67:58-69. doi: 10.1016/j.ejmp.2019.10.001. Epub 2019 Oct 28. PMID: 31671333.
[29] Ondersma SJ, Martino S, Svikis DS, Yonkers KA. Commentary on Kim et al. (2017): Staying focused on non-treatment seekers. Addiction. 2017 May;112(5):828-829. doi: 10.1111/add.13736. PMID: 28378329; PMCID: PMC6552680.