Accurate Left Ventricle Segmentation and Ejection fraction estimation with Deep Learning-based Echocardiography

Madankumar Balasubramani[a], Chih-Wei Sung[b], Mu-Yang Shieh[b], Edward Pei-Chuan Huang[b], Maysam F. Abbod[c], Jiann-Shing Shieh[a,*]

[a] *Department of Mechanical Engineering, Yuan Ze University, Chung-Li District, Taoyuan City 320, Taiwan (R.O.C).*

[b] *Department of Emergency Medicine, National Taiwan University Hospital, Hsin-Chu branch, Zhubei City, Hsin-chu City 302, Taiwan (R.O.C).*

[c] *Department of Electronics and Electrical Engineering, Brunel University London, Uxbridge UB8 3PH, UK*

* *Correspondence: jsshieh@saturn.yzu.edu.tw*

Abstract

Heart failure with reduced ejection fraction is a rapidly growing public health issue with an estimated prevalence of over 37 million individuals globally. We proposed an efficient deep learning framework of automated echocardiography for predicting heart failure with reduced ejection fraction, which is most important for earlier diagnosis of heart function. The proposed deep learning framework sensibly overcame the echocardiography challenges which are domain expertise much needed for segmenting the left ventricular end-diastolic volume and end-systolic volume, high complexity in the identification of endocardial border and tiny labelled data. Due to advancements of artificial intelligence, we were able to address the challenges through a technique called data augmentation. three different deep learning models have been applied, the first model is UNet model, while the second model is Deeplab, and the third model is UNet model with backbone network for automated segmentation of left ventricle. The deep learning model has successfully segmented the left ventricle region on apical 4-chamber (A4C) views. The first model of UNet has mean IOU of 74.31%, and the second model of Deeplab has mean IOU of 79.53%. The proposed model has achieved promising mean IOU of 89.1 % in segmentation and predicted ejection fraction with a correlation coefficient r^2 of 0.71. Results show that the contactless echocardiographic approach can quantitatively estimate left ventricular chamber size and ejection fraction in humans as well as estimate the dimensions of the left ventricle in systole and diastole.

Keywords: echocardiography, left ventricle, ejection fraction, deep learning, backbone

1. Introduction

A personalized and prospective approach to care is urgently needed to address heart failure, which is the leading cause of hospitalizations and deaths worldwide. Echocardiograms (Echo) are inviolable tests that use ultrasound to produce images of the heart [1-3]. The images allow doctors to assess the size, shape, and motion of the heart [4]. Medical images can be investigated more efficiently and accurately by deep learning algorithms, which are subsets of machine learning algorithms [5-13].

Echocardiography serves as a critical imaging technique for assessing cardiac structure and function in the diagnosis, management, and monitoring of cardiovascular diseases [14]. Accurate segmentation of cardiac structures from echocardiograms is crucial for quantitative analysis, visualization, and clinical decision-making. Furthermore, the estimation of ejection fraction (EF) from echocardiograms plays a pivotal role in evaluating cardiac performance and aiding in the diagnosis and prognosis of various cardiac conditions.

The EF is calculated by measuring the amount of blood that is pumped out of the left ventricle per beat. A normal EF varies between 55-70%, while an EF of less than 40% indicates reduced left ventricle outflow and is an important marker of heart failure. Cardiac diseases can be diagnosed and managed through echocardiography. Clinical decision-making requires precise and reliable echocardiographic assessment [15].

Perhaps the most significant indicator of cardiac function is the assessment of left ventricular EF, which is the ratio variation in the left ventricular end-systolic and end-diastolic volumes. This metric indicates individuals who are eligible for life-prolonging treatments [15, 16]. An overview of the need for automated echocardiography diagnosis is presented, along with a discussion of the application of artificial intelligence (AI) to echocardiography and its prospects.

The application of AI into everyday clinical practice and its potential to become an invaluable tool for healthcare professionals who deal with cardiovascular diseases is presented [17]. Myocardial reliable diagnosis is dependent on two-dimensional echocardiography image quality. Images produced through scattering centres do not have crisp edges like natural images. Therefore, to produce clear images, algorithms are used to detect and suppress the scattering centres. This helps reduce the noise and improves the image quality, aiding in the accurate diagnosis of myocardial conditions.

Echocardiogram segmentation entails the precise delineation of specific cardiac structures such as the left ventricle (LV), right ventricle (RV), atria, and heart valves [18, 19]. Accurate segmentation enables the quantification of essential parameters including chamber volumes, wall thickness, and valve dimensions, facilitating a comprehensive assessment of cardiac function. Moreover, segmentation techniques contribute to the identification and tracking of pathological abnormalities such as ventricular hypertrophy, valvular disorders, and myocardial infarctions.

U-Net is a popularly used architecture for medical image segmentation because it could localize and produce the underlying pattern [20]. It is a convolutional neural network made up of a bottleneck layer connecting the encoder and the decoder. High-level features are extracted from the input image by the encoder, and the decoder uses these features to reconstruct the segmented image [21]. In several medical imaging tasks, including the segmentation of brain tumours [22] and vessels in the retina [23, 24], U-Net has been demonstrated to perform better than other segmentation methods. Its success can be attributed to two things which are its capacity to extract local and global information from the input image, as well as its effective use of training data via data augmentation [25] methods. U-Net has also been modified for use in various other fields, including satellite imagery [26, 27].

Moreover, deep learning has the potential to assist beginners in rapidly acquiring proficiency in quality diagnostic imaging, thereby enhancing both inter- and intra-observer consistency [28, 29]. Neural network models are trained by feeding them a sizable collection of labelled echocardiography data, where segmentation masks are established for different cardiac components as the Left Ventricle, Right Ventricle, and Mitral valve [30]. Following training, the model may be used to autonomously isolate different heart components in new echocardiography pictures [31]. These models when trained with large dataset and fine-tuned can also be used for real-time segmentation during an echocardiogram procedure which can be useful for the physicians [32].

It is important to recognize that deep learning has gained significant traction within medical imaging with particular focus on its applications towards echocardiography. Subsequently continual advancement in the area has introduced new techniques, models, and insights [33]. One notable model advancement is the success of Convolutional Neural Networks (CNNs) in autonomously segmenting echocardiogram images [34]. Further upscaling of these CNN models can be employed through fine tuning specific dataset/ applications for better performance and flexibility. Alternatively, data augmentation techniques [35-37] can also expand a models' training set and enhance adaptability.

A comparison is then made between the model's segmentation results and the ground truth labels, using metric such as the mean Intersection over Union (IoU). These measures give a quantitative assessment of the model's performance and provide a way to compare the results with other models or different methods. It is important to notice that the results of UNET on Echo segmentation vary based on the quality of images, and the dataset used for training and testing, it also needs to be reviewed by an expert cardiologist to have final verdict. Echocardiogram image segmentation involves the process of identifying and separating different structures of the heart in an echocardiogram image. The segmentation results can then be used to obtain quantitative measures of the heart's function, such as the left ventricular ejection fraction (LVEF), which is an indicator of heart health.

In recent years, there have been several variations of the U-Net architecture that incorporate backbone [38] networks. The backbone, an essential component within the UNet architecture, stands as a cornerstone for achieving accurate and robust medical image segmentation. Revered for its ability to capture both local intricacies and global contextual information, the U-Net owes much of its prowess to this intricate network. Acting as a formidable feature extraction apparatus, the backbone serves to extract high-level features from input images, allowing the U-Net to acquire a deep understanding of the data and perform precise segmentation with finesse. At its core, the backbone of the U-Net embodies a pre-trained convolutional neural network (CNN) model, seamlessly integrated into the U-Net structure.

The selection of the backbone network warrants utmost consideration, as it determines the U-Net's capacity to discern meaningful and discriminative features. Among the venerated choices for the backbone lie networks such as VGG [39], ResNet [40], DenseNet [41], and EfficientNet [42], esteemed for their exemplary performance in various computer vision domains. These backbone networks are pre-trained convolutional neural networks that are used to extract features from the input image before feeding it to the U-Net model. The use of backbone networks has been shown to improve the performance of the U-Net architecture, especially when the size of the training dataset is limited. Here are some popular backbone networks that have been used with the U-Net architecture:

ResNet: ResNet is a deep residual neural network that has been used as a backbone network for the U-Net architecture. It consists of several layers of residual blocks that allow the network to learn more complex features, and it has been pre-trained on the ImageNet dataset. Densenet: DenseNet's dense connectivity and feature reuse mechanism make it a powerful deep learning architecture. It tackles the challenges of information flow and promotes effective feature utilization, leading to improved accuracy and better utilization of network parameters.

Deeplab: DeepLab framework addresses the challenges of dense pixel-level labelling by leveraging the power of deep convolutional networks. It utilizes a modified version of the popular VGG-16 or ResNet network as the backbone for feature extraction. These networks are pre-trained on large-scale image classification datasets, enabling them to capture rich and transferable features. Therefore, the purpose of this study is as follows. By using a pre-trained backbone network as the input to the U-Net architecture, the network can leverage the powerful feature extraction capabilities of the backbone network, while still maintaining the ability to perform precise localization through the symmetric expanding path of the U-Net architecture.

2. Methods

In this study, we propose deep learning-based method for LV segmentation and ejection fraction estimation. Our method is based on a U-Net architecture, which is a convolutional neural network that has been shown to be effective for medical image segmentation. The left ventricle segmentation workflow starts with dataset acquisition of echocardiogram videos obtained from the patients suspected of having heart failure and normal patients. Furthermore, the data pre-processing is followed by extracting the frames and de-identification of confidential information in the frames. The workflow for echocardiography segmentation and Ejection Fraction (EF) estimation can be summarized as follows:

2.1 Dataset

The dataset is produced and developed inhouse for this study. The inhouse dataset consists of echo videos in the form of apical four-chamber view. The data acquisitions were done from the National Taiwan University Hospital, Hsinchu branch, Taiwan. The study population includes both male and female with normal and reduced ejection fraction clinically indicated standard echocardiography.

To access the echocardiography data for research purposes, the study obtained ethical approval from the Research Ethics Committee at National Taiwan University, Hsin-chu Branch. The approval was granted under reference number 110-069-E, ensuring compliance with ethical guidelines, and safeguarding the rights and privacy of the individuals involved in the data collection. The echocardiography examinations were conducted exclusively using state-of-the-art Philips ultrasound machines, ensuring standardized imaging protocols throughout the study. The acquired images possessed a consistent frame resolution of 800 × 600 pixels, enabling a clear and detailed visualization of cardiac structures.

In consideration of the inherent physiological diversity among patients, both the frame rate (FR) and heart rate (HR) exhibited variations across the dataset. The frame rate, representing the frequency at which consecutive frames were captured, ranged between 25 and 66 Hz, accommodating diverse cardiac motion dynamics encountered during acquisition. Similarly, the heart rate, denoting the number of beats per minute, spanned a range of 60 to 150 beats per minute, reflecting the unique physiological profiles of the subjects.

By accounting for these fluctuations in frame rate and heart rate, the study acknowledged and accommodated the inherent variability in cardiac activity across the diverse patient population. This comprehensive approach ensured that the acquired data encapsulated a broad spectrum of cardiac dynamics, enhancing the applicability and robustness of the subsequent analysis and findings. The dataset was partitioned such that 80% of the data was allocated for training purposes, 10% for validation to fine-tune the model's performance, and the remaining 10% for rigorous testing to evaluate its generalization capability as shown in Table 1.

Table 1 Data distribution of Echocardiograms

Dataset	Number of patients
Training	800
Validation	100
Testing	100

In Figure 1, the distribution of ejection fraction (EF) is depicted for a cohort of 1000 patients. In accordance with clinical guidelines, when the Left ventricular ejection fraction (LVEF) falls below 40%, it is classified as reduced EF, indicating compromised cardiac function. Left Ventricular Ejection Fraction falling within the range of 40-49% is considered borderline, indicating a potential deviation from normal cardiac function. On the other hand, if the LVEF surpasses 50%, it is classified as normal, signifying a healthy level of ventricular contraction.

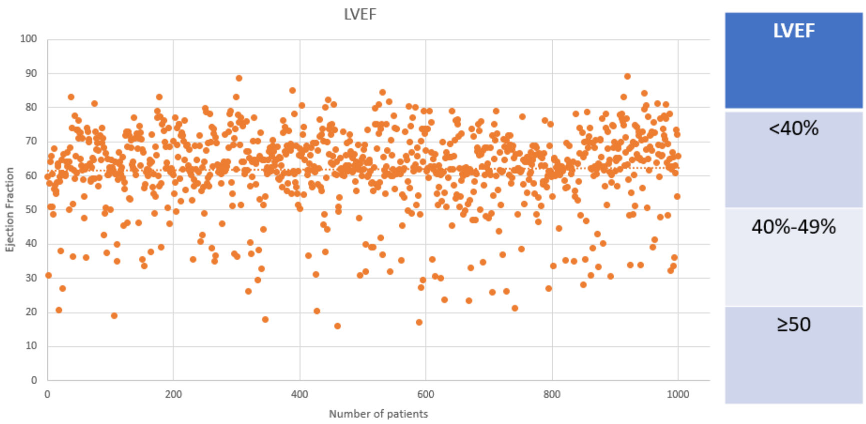


Figure 1 Ejection fraction Distribution for 1000 patients

It is noteworthy to mention that the distribution of EF values suggests a class imbalance within the data as shown in Figure 1. This implies that the number of samples representing different EF ranges may vary significantly, potentially leading to challenges in model training and evaluation. Addressing this class imbalance is crucial to ensure fair representation and accurate performance assessment of the models developed for EF estimation.

2.2 Data annotation

All Apical 4 chamber (A4C) view frames were extracted from the Digital Imaging and Communications in Medicine files and converted into images, which were cropped and resized to 128 × 128 pixels. Each pixel was then normalized between 0 and 1 for neural network training and prediction. In the first step, we have extracted the frames from clips in 50 frames per second and then selected the key frames such as end systolic (ES) and end diastolic (ED) as shown in the Figure 2, with the help of emergency room experts from National Taiwan University Hospital, Hsinchu branch. The data split into different batches and provided to the experts for key frame selection and data annotation of the left ventricle.

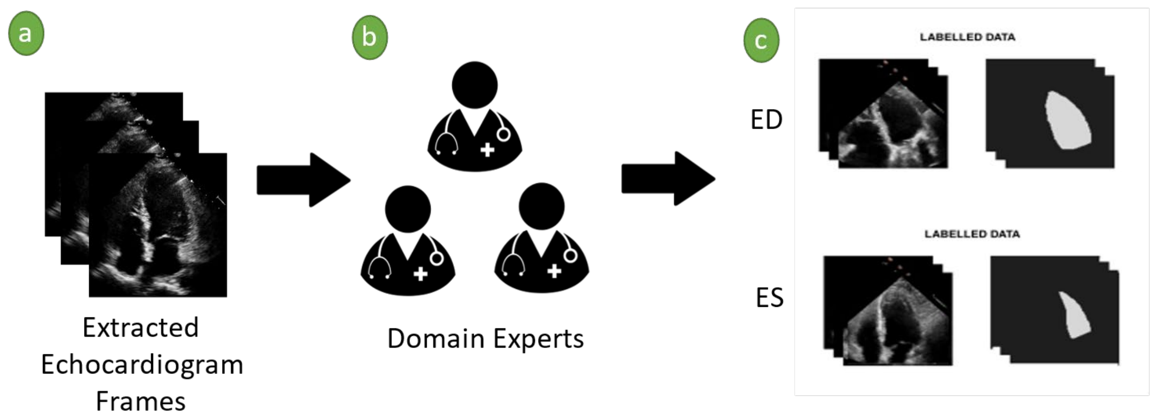


Figure 2 Data Annotation. (a) The frames are extracted from the echocardiogram video. (b) Domain experts select the key frames (ED and ES). (c) The frames are labelled to produce ground truth

2.3 Data quality assessment

One of the biggest challenges we face when trying to scale up deep learning is the need for large and high-quality annotations that describe the content of echocardiograph images. Semantic annotations involve labelling images with detailed descriptions of what's in them, like identifying objects, regions, or specific characteristics.

The need for massive and clean collections of semantic annotations is a bottleneck for a few reasons: training Data, models need a lot of labelled training data to learn patterns and make accurate predictions. Having access to large, annotated datasets directly affects how well these algorithms perform and how well they can handle different scenarios. Ensuring the quality and accuracy of semantic annotations is crucial for training reliable models. To achieve clean collections of annotations, we must be careful to minimize errors, inconsistencies, or biases in the labelling process.

This involves implementing thorough quality control measures and having humans supervise the process, which adds complexity and more time. It's important to annotate images with diverse and comprehensive semantic information to build robust and inclusive machine learning models. This means we need annotations from various domains, perspectives, and cultural contexts. Obtaining a massive and diverse collection of annotations presents additional challenges in terms of data collection, collaboration, and management.

Generating high-quality semantic annotations often requires domain expertise and skilled annotators who have a deep understanding of the task at hand. It can be expensive to acquire and retain such expertise, which can become a bottleneck when scaling up the learning process. To address these challenges, we need to develop more efficient annotation methodologies, improve automation techniques, and advance our data collection practices. Additionally, exploring alternative approaches like weakly supervised learning or active learning can help reduce our dependence on fully annotated datasets, which can help alleviate the bottleneck to some extent.

Figure 3 Illustrates the assessment of image quality using REDcap [43,44] electronic data capture tools hosted at National Taiwan University Hospital, Hsinchu branch, Taiwan, used for performing the digital survey with a team of five novice volunteers. The volunteers who are participating in the survey are provided with a short training course of assessing the quality of frames, with constraints such as clear apical four chamber, sharp LV borders, good contrast, and the overall quality for the frames. Based on the quality of frames, the volunteers will provide the rating for the frames on a scale of 1 to 10. A rating of 1-4 is considered poor quality, and a rating of 5-7 is considered good quality. A rating of 8-10 indicates good image quality.

The quality of the 2D cross-sectional echo images is variable due to dependence on the experience level of the echo operators. Furthermore, the ultrasound quality may worsen with age, disease, and obesity, which coincidentally are correlated with heart disease risk. Direct cine-based EF estimation in echo is a challenging problem. Ultrasound images are inherently noisy and yield blurry chamber boundaries, making LV size changes difficult to quantify.

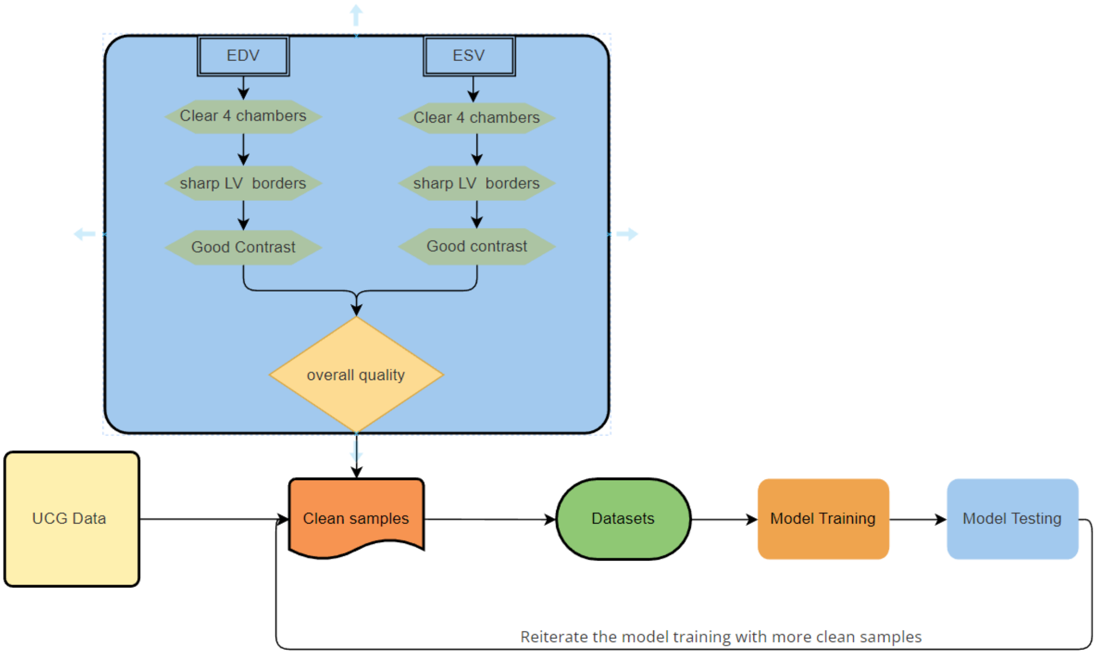


Figure 3 Image Quality Assessment survey for a team of novice using Redcap.

Cardiac anatomy and function are complex and variable, especially in the apical views. The complexity is compounded by variations in the appearance of the imaged heart in the presence of pathology. Echo images only capture 2D cross-sectional views of the heart, which may be foreshortened. LV foreshortening may inevitably lead to incorrect measurements, even if the image analysis task is accurate.

2.4 Model

Our deep neural network architecture was designed in python[45] using the TensorFlow [46] and Keras [47] frameworks. Our model consists of convolutional layers with activation functions. Our model consisted of a succession of modest 3 × 3 convolutional filters connected with max-pooling layers spread across 2 × 2 windows. To counteract overfitting [48], dropout was used in training for both the convolutional and fully connected layers.

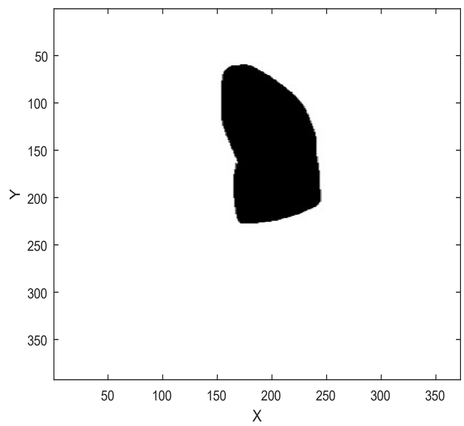
UNet utilizes an encoder-decoder structure with skip connections for image segmentation, while ResNet34 employs residual blocks and skip connections to enable training of very deep networks. UNet focuses on capturing hierarchical features and recovering spatial information, while ResNet34 excels at handling deep networks and learning residual mappings. The combination of these two architectures in the UNet+ResNet34 model allows for efficient feature extraction, depth, and skip connections, leading to improved performance in left ventricle segmentation task. In our Model, the ResNet34 encoder is used to extract features from the input image, and the U-Net decoder is used to reconstruct the output image.

2.5 Ejection fraction estimation

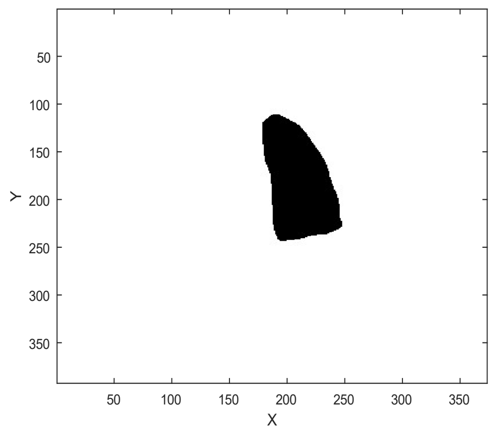
In semantic segmentation, the model will segment the area of left ventricle to estimate the ejection fraction using the equation as mentioned in Equation 1. The estimate of ejection fraction is based on the segmentation of echocardiograms' key frames. Based upon the ejection fraction, the heart function is classified. If the EF is between 50% to 70%, heart function is normal. Whereas the EF is between 41% to 49%, heart function is borderline. Reduced EF leads to heart failure if it is below 40%.

Left ventricular ejection fraction (LVEF) is one of the most reported measures of left ventricular (LV) systolic function. It is the ratio of blood ejected during systole (stroke volume) to blood in the ventricle at the end of diastole (end-diastolic volume). If the LV end-diastolic volume (EDV) and end-systolic volume (ESV) are segmented or annotated as shown in the Figures 4(a) and 4(b), LVEF can be determined using Equation 1:

$$LVEF = \frac{(EDV - ESV)}{EDV} \times 100 \qquad (1)$$



(a)



(b)

Figure 4 (a) End-diastolic (ED) and (b) End-systolic (ES) frames.

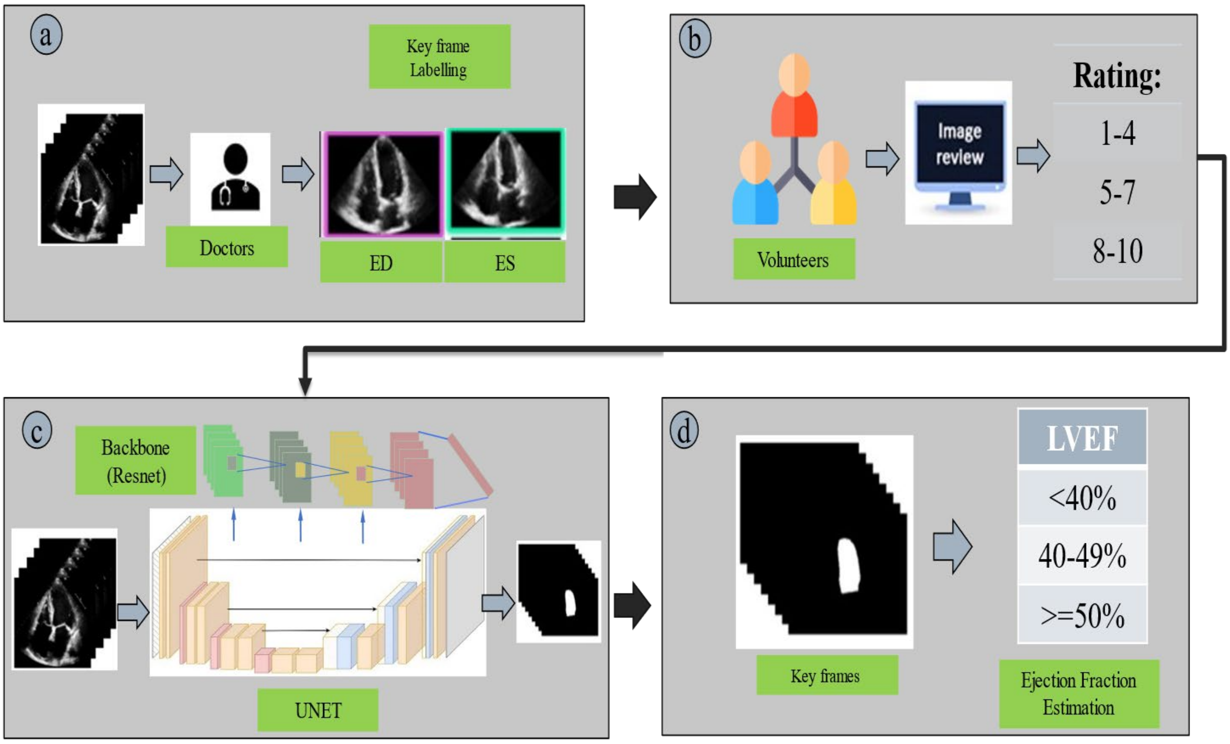


Figure 5 Echocardiography segmentation and EF estimation workflow. (a) Echocardiography videos are converted to frames and the key frames, End-Diastolic (ED) and End-Systolic (ES). (b) The frames are rated by the volunteers according to the quality of the images. (c) After the pre-processing step, data was segmented and predicted. (d) Key frames are the used for Ejection Fraction (EF) estimation.

Finally, the whole procedure is summarised in Figure 5. Echocardiography videos are converted to frames, and the key frames at End-Diastolic (ED) and End-Systolic (ES) phases are identified by emergency room experts (i.e., medical doctors). These videos are processed to extract individual frames, and the frames at the ED and ES phases represent the maximum and minimum volumes of the heart, respectively as shown in Figure 5(a).

Volunteers or experts meticulously rate the frames on a scale of 1 to 10, assessing the quality of the images with utmost precision as shown in Figure 5(b). If the rating is in the range between 1 to 4, it is considered noisy data and frames with poor quality can be excluded from further analysis or flagged for additional scrutiny. If the rating falls within the range of 5 to 7, it is deemed acceptable. When the rating falls within the range of 8 to 10, it serves as an indication that the data is impeccably clean and boasts exceptional quality. This quality assessment step helps filter out frames that may be affected by artifacts, motion blur, or other factors that could potentially affect the accuracy of segmentation and EF estimation.

Pre-processing steps are applied to the selected frames to enhance their quality and facilitate accurate segmentation as shown in Figure 5(c). This involves image denoising, contrast adjustment. The pre-processing stage aims to remove noise and artifacts that could hinder the subsequent segmentation and EF estimation processes.

The pre-processed frames are then segmented to identify and delineate specific structures of interest, the LV as shown in Figure 5(d). The segmentation step aims to precisely outline the boundaries of the structures, providing accurate spatial information for subsequent analysis. Once the segmentation is complete, the key frames (ED and ES) are utilized for EF estimation. Ejection Fraction is a crucial metric used to assess cardiac function and is calculated by comparing the volume of blood pumped out of the LV during systole (ES phase) with the volume at maximum filling during diastole (ED phase). The segmented regions corresponding to the LV are used to measure the volumes, and the EF is calculated as the ratio of stroke volume to end-diastolic volume, often expressed as a percentage as shown in Figure 5 (d).

By following this workflow, echocardiography videos are processed into frames, key frames are identified, image quality is assessed, pre-processing is performed, segmentation is executed, and finally, EF estimation is conducted using the key frames. This workflow enables the assessment of cardiac function based on echocardiographic data, providing valuable insights for diagnosis, monitoring, and treatment planning in cardiovascular medicine.

3. Results

In this research, we show that left ventricle segmentation has been effectively clarified with a CNN-based U-Net architecture. IoU is a measure of how well two objects overlap [49]. The IoU formula requires an understanding of two key terms: True Positive (TP) and False Positive (FP). True Positive is when the model correctly predicts a pixel as being part of an object when it is part of the object. It is a False Positive when the model predicts that a pixel is part of an object when in fact it is part of the background. It is the ratio of the area of overlap between the two objects to the area of their union according to the Equation 2. It is used in object detection and image segmentation tasks to measure the accuracy of the model's prediction.

$$IoU = TP\ /\ (TP + FP + FN) \tag{2}$$

Conventional methods cannot obtain such an IoU using clustering [50], threshold segmentation [51], or machine learning algorithms [52]. The reason for this is that regardless of which conventional method is used, there will always be a manual selection process to determine colours, textures, or shapes. Despite this, different LV borders in the dataset exhibit clearly distinct characteristics. However, by using an image segmentation approach such as U-Net, this manual selection process can be avoided, as U-Net is able to automatically identify and segment target objects from an image.

This makes U-Net far more suitable for determining LV borders, as it can achieve a much higher IoU score than conventional methods. Figure 6 illustrates (a) grey image is input image to the model, (b) red region annotated by the experts (c) yellow region shows model could perform well on the left ventricle segmentation task. Left ventricle boundary definition is crucial for treatment, but it is not easily accomplished during segmentation of medical images, which is known as segmentation. Table 2 shows the comparison of model performance, UNet has performed the automated LV segmentation of echocardiogram 2D frames with mean IOU of 74.31%.

The proposed model, UNet with Resnet as backbone [38] network has outperformed the Deeplab [53] and UNet base model with highest mean IoU of 89.12% on the LV segmentation task for the 2D echo images. In comparison with the proposed model, MaskRCNN [54], Echonet [55], and 3D-UNet [56] has higher mean IOU of 92.21%, 92%, and 93% respectively. According to Echonet [55], They have used massive dataset of video recordings represent unique individuals because the dataset contains videos recorded from 10,036 random echocardiography procedures performed between the year of 2006 and 2018.

According to Table 2, to ensure data-centric models perform effectively, we require vast amounts of high-quality data. This means having datasets that are extensive and diverse, enabling the models to learn from a wide range of examples. However, it's not just about quantity; the data also needs to be meticulously cleaned, removing any noise, errors, or inconsistencies. This ensures that the models can rely on accurate and trustworthy information during their training.

Expert annotation is another vital aspect of data-centric AI. Domain experts with deep knowledge in the relevant field provide precise labels or tags to the data. Their expertise ensures that the models learn from ground truth labels, enabling them to make accurate predictions and perform complex tasks.

Maintaining data quality is essential throughout the annotation process. Rigorous quality control measures are implemented to ensure the correctness and consistency of the annotations. Addressing biases and ensuring agreement among experts are critical steps in this process. Working together with domain experts plays a crucial role in developing comprehensive datasets. Their insights and understanding of the specific application area are essential for guiding the annotation process and capturing the domain's nuances.

Table 2 Model Performance comparison

Base Model	**Backbone**	**mean IoU (%)**
UNet	-	74.31
Deeplab [53]	VGG16	79.53
MaskRCNN [54]	-	92.21
Echonet [55]	-	92.0
3D U-Net [56]	-	93.0
Proposed Model	Resnet50	89.12

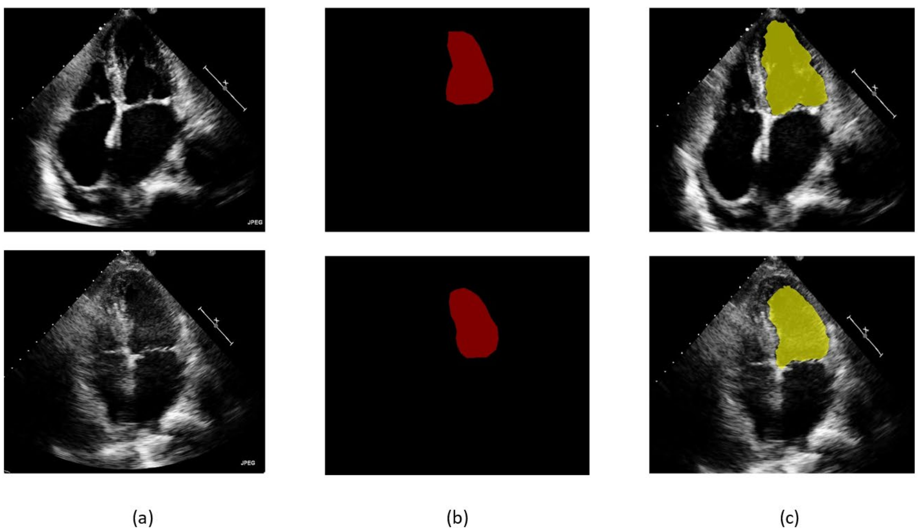
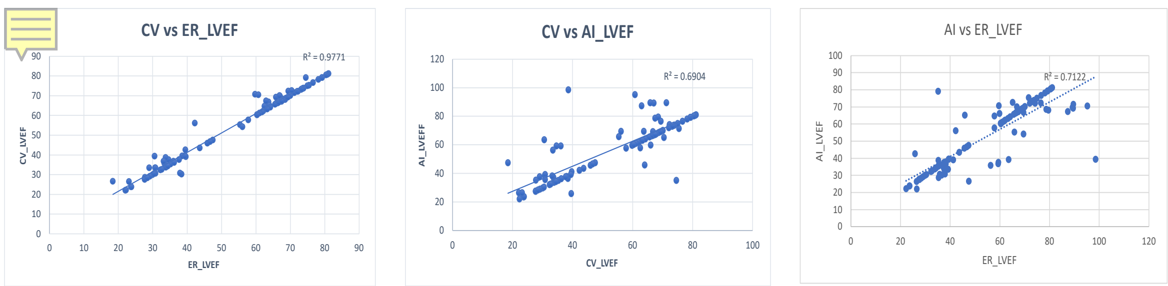


Figure 6 Models Segmentations samples. (a) Original image (b) Ground truth (c) Segmentation Result

According to our approach for the help of emergency room experts from National Taiwan University Hospital, Hsinchu branch, the data split into different batches and provided to the experts for key frame selection and data annotation of the left ventricle. Hence, we want to see what the differences between cardiovascular experts and emergency room experts for labelling area then calculating LV of EF. Figure 7 shows the comparison of correlation coefficients, (a) comparison of cardiovascular experts and emergency room experts; (b) Comparison of cardiovascular experts and model prediction; (c) comparison of model prediction and emergency room experts. The correlation coefficient between the cardiovascular expert and emergency room experts is 0.9771, which exhibits the strong correlation coefficient and expertise of domain knowledge. Therefore, we use emergency room experts to label then calculate LV of EF is reasonable and acceptable. However, the correlation coefficient between the cardiovascular expert and model prediction is 0.6904. Also, the correlation coefficient between the model prediction and emergency room experts is 0.7122. This shows the proposed model is needed to further improve the accurate estimates of LV of EF based on the given input and output data in the near future. Although our proposal model has been shown a little bit of less than accuracy in comparison with previous research Echonet [55] model for using large database of 10,036 patients which are over 10 times of our patient data.



(a) (b) (c)

Figure 7 Comparison of correlation coefficients. (a) Comparison of Cardiovascular experts and Emergency room experts (b) Comparison of Cardiovascular experts and model prediction (c) Comparison of model prediction and Emergency room experts.

It is clinically significant and timely to conduct this research. It has become increasingly apparent that biomarkers such as EDV, ESV and EF are not optimal for stratifying patients with heart failure. A single biomarker, EF, is currently recommended as the main biomarker for stratifying heart failure patients [57]. Therefore, a better biomarker is needed to provide a more accurate assessment of the disease. This research aims to examine the feasibility of using a combination of biomarkers to stratify heart failure patients. The results of this study will provide valuable insight into the efficacy of using multiple biomarkers for stratifying heart failure patients.

4. Discussion

Our results suggest that deep learning is a promising technology for LV segmentation and estimation of ejection fraction. Our method is more accurate and efficient than manual segmentation, and it has the potential to improve the diagnosis and management of heart disease. Deep learning has emerged as a powerful tool for medical image analysis, and it has been shown to be particularly effective for segmentation of the LV from the echocardiograms.

Technically, deep learning models can learn to identify complex patterns of the LV from the images, even in the presence of noise and other challenges. The use of deep learning for LV segmentation has several potential benefits. First, it can automate the segmentation process, saving time and improve the accuracy of measurements. Second, deep learning models can be trained on large image datasets, which can improve the accuracy of segmentation even in challenging cases. Third, deep learning models can be used to segment images from different imaging modalities, which can improve the precision of diagnosis and treatment planning.

Clinically, in the crowded emergency department, distinguishing potential patients with moderate to severe heart failure and further to avoid impending cardiac collapse are difficult but crucial. Traditional physical examinations, such as heart sound auscultation and heart border percussion, may not identify true emergency because of limited time in diagnosis and subjective judgment of the physicians. Automatic real-time detection of the left ventricle ejection fraction provided objective assessment of the heart condition. Our proposed model, consistent with previous studies, assessed the ejection fraction using the deep learning algorithm with excellent prediction performance [58, 59]. However, some population and setting variations caused different model prediction powers. In Liu's study in China, two open-source data sets and one clinical data set were collected. The sample size of the mixed database was large, but the heterogeneity thereby existed [58]. In Asch's study, which compared the agreement of the reference ejection fraction (by experienced cardiologists) and the prediction of the interclass correlation ranged from 0.86 to 0.95 with biases less than 2% [28]. In the study conducted in Norway, they indicated that the ejection fraction was considerably affected by apical foreshortening using three-dimensional ultrasound, and the mean absolute difference in their result was measured at 7.2%, which are comparable with related studies [59]. The above results, whether 2D or 3D echocardiography, implied that automatic deep learning-based prediction may work in some circumstances without specialists.

Some limitations were noted in this study. First, deep learning models require large datasets of labelled images to train, which can be expensive and time consuming to collect. Second, deep learning models can be sensitive to changes in image quality, which can lead to errors in segmentation. Third, deep learning models can be computationally expensive to train and deploy, which can limit their use in resource-limited settings. Deep learning is a promising tool for segmenting the LV from echocardiograms. It has the potential to automate the process of segmentation, improve the accuracy of measurements, and make it possible to segment images from different imaging modalities.

In the coming future, in addition to the prediction of the left ventricle ejection fraction, the prediction of the right heart function is drawing attention. The predictive significance of RV has been well recognized [60]. It is not only the main determinant of clinical symptoms and survival in patients with hypertension, but also has an independent predictive value for left heart disease. Approximately half of patients with heart failure and reduced left ventricular output (LVEF) suffer from RV dysfunction, which is associated with a double increase in the risk of side effects. The implementation of right ventricular ejection fracture by the UNet model could be further anticipated.

5. Conclusion

In summary, the present research demonstrated that the proposed model, which was adapted for echocardiogram usually used in the clinical setting, enables automated evaluation of LV of EF and, more specifically, detection of heart failure with accuracy comparable to through imagery analysis by experienced imaging cardiac specialists. Based on the results, the method shows clean data is essential for accurate and reliable AI model training, while noisy data can introduce biases and inaccuracies. This technique is likely to further improve for increasing more patient data so it can help more health care providers to reliably conduct and analyse cardiac ultrasound assessments.

References

1. McDonagh, T.A., M. Metra, M. Adamo, R.S. Gardner, A. Baumbach, M. Böhm, H. Burri, J. Butler, J. Čelutkienė, and O. Chioncel, *2021 ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure: Developed by the Task Force for the diagnosis and treatment of acute and chronic heart failure of the European Society of Cardiology (ESC) With the special contribution of the Heart Failure Association (HFA) of the ESC.* European heart journal, 2021. 42(36): p. 3599-3726.
2. Nishimura, R.A. and A.J. Tajik, *Evaluation of diastolic filling of left ventricle in health and disease: Doppler echocardiography is the clinician's Rosetta Stone.* Journal of the American College of Cardiology, 1997. 30(1): p. 8-18.
3. Nagueh, S.F., C.P. Appleton, T.C. Gillebert, P.N. Marino, J.K. Oh, O.A. Smiseth, A.D. Waggoner, F.A. Flachskampf, P.A. Pellikka, and A. Evangelisa, *Recommendations for the evaluation of left ventricular diastolic function by echocardiography.* European journal of echocardiography, 2009. 10(2): p. 165-193.
4. Rudski, L.G., W.W. Lai, J. Afilalo, L. Hua, M.D. Handschumacher, K. Chandrasekaran, S.D. Solomon, E.K. Louie, and N.B. Schiller, *Guidelines for the echocardiographic assessment of the right heart in adults: a report from the American Society of Echocardiography: endorsed by the European Association of Echocardiography, a registered branch of the European Society of Cardiology, and the Canadian Society of Echocardiography.* Journal of the American society of echocardiography, 2010. 23(7): p. 685-713.
5. Topol, E.J., *High-performance medicine: the convergence of human and artificial intelligence.* Nature medicine, 2019. 25(1): p. 44-56.
6. Esteva, A., B. Kuprel, R.A. Novoa, J. Ko, S.M. Swetter, H.M. Blau, and S. Thrun, *Dermatologist-level classification of skin cancer with deep neural networks.* nature, 2017. 542(7639): p. 115-118.
7. Ting, D.S.W., C.Y.-L. Cheung, G. Lim, G.S.W. Tan, N.D. Quang, A. Gan, H. Hamzah, R. Garcia-Franco, I.Y. San Yeo, and S.Y. Lee, *Development and validation of a deep*

learning system for diabetic retinopathy and related eye diseases using retinal images from multiethnic populations with diabetes. Jama, 2017. 318(22): p. 2211-2223.
8. Bejnordi, B.E., M. Veta, P.J. Van Diest, B. Van Ginneken, N. Karssemeijer, G. Litjens, J.A. Van Der Laak, M. Hermsen, Q.F. Manson, and M. Balkenhol, *Diagnostic assessment of deep learning algorithms for detection of lymph node metastases in women with breast cancer.* Jama, 2017. 318(22): p. 2199-2210.
9. Zhu, H., C. Cheng, H. Yin, X. Li, P. Zuo, J. Ding, F. Lin, J. Wang, B. Zhou, and Y. Li, *Automatic multilabel electrocardiogram diagnosis of heart rhythm or conduction abnormalities with deep learning: a cohort study.* The Lancet Digital Health, 2020. 2(7): p. e348-e357.
10. De Fauw, J., J.R. Ledsam, B. Romera-Paredes, S. Nikolov, N. Tomasev, S. Blackwell, H. Askham, X. Glorot, B. O'Donoghue, and D. Visentin, *Clinically applicable deep learning for diagnosis and referral in retinal disease.* Nature medicine, 2018. 24(9): p. 1342-1350.
11. Raghunath, S., A.E. Ulloa Cerna, L. Jing, D.P. VanMaanen, J. Stough, D.N. Hartzel, J.B. Leader, H.L. Kirchner, M.C. Stumpe, and A. Hafez, *Prediction of mortality from 12-lead electrocardiogram voltage data using a deep neural network.* Nature medicine, 2020. 26(6): p. 886-891.
12. Courtiol, P., C. Maussion, M. Moarii, E. Pronier, S. Pilcer, M. Sefta, P. Manceron, S. Toldo, M. Zaslavskiy, and N. Le Stang, *Deep learning-based classification of mesothelioma improves prediction of patient outcome.* Nature medicine, 2019. 25(10): p. 1519-1525.
13. Yim, J., R. Chopra, T. Spitz, J. Winkens, A. Obika, C. Kelly, H. Askham, M. Lukic, J. Huemer, and K. Fasler, *Predicting conversion to wet age-related macular degeneration using deep learning.* Nature Medicine, 2020. 26(6): p. 892-899.
14. LONG, K.J., J. Bonagura, and P. Darke, *Standardised imaging technique for guided M-mode and Doppler echocardiography in the horse.* Equine veterinary journal, 1992. 24(3): p. 226-235.
15. Loehr, L.R., W.D. Rosamond, P.P. Chang, A.R. Folsom, and L.E. Chambless, *Heart failure incidence and survival (from the Atherosclerosis Risk in Communities study).* The American journal of cardiology, 2008. 101(7): p. 1016-1022.
16. Kusunose, K., A. Haga, T. Abe, and M. Sata, *Utilization of artificial intelligence in echocardiography.* Circulation Journal, 2019. 83(8): p. 1623-1629.
17. Kusunose, K., *Steps to use artificial intelligence in echocardiography.* Journal of echocardiography, 2021. 19(1): p. 21-27.
18. Badano, L.P., T.J. Kolias, D. Muraru, T.P. Abraham, G. Aurigemma, T. Edvardsen, J. D'Hooge, E. Donal, A.G. Fraser, and T. Marwick, *Standardization of left atrial, right*

ventricular, and right atrial deformation imaging using two-dimensional speckle tracking echocardiography: a consensus document of the EACVI/ASE/Industry Task Force to standardize deformation imaging. European Heart Journal-Cardiovascular Imaging, 2018. 19(6): p. 591-600.

19. Lorenzo-Valdés, M., G.I. Sanchez-Ortiz, A.G. Elkington, R.H. Mohiaddin, and D. Rueckert, *Segmentation of 4D cardiac MR images using a probabilistic atlas and the EM algorithm.* Medical image analysis, 2004. 8(3): p. 255-265.
20. Ronneberger, O., P. Fischer, and T. Brox. *U-net: Convolutional networks for biomedical image segmentation*. in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18*. 2015. Springer.
21. Seo, H., C. Huang, M. Bassenne, R. Xiao, and L. Xing, *Modified U-Net (mU-Net) with incorporation of object-dependent high level features for improved liver and liver-tumor segmentation in CT images.* IEEE transactions on medical imaging, 2019. 39(5): p. 1316-1325.
22. Noori, M., A. Bahri, and K. Mohammadi. *Attention-guided version of 2D UNet for automatic brain tumor segmentation*. in *2019 9th international conference on computer and knowledge engineering (ICCKE)*. 2019. IEEE.
23. Li, D., D.A. Dharmawan, B.P. Ng, and S. Rahardja. *Residual u-net for retinal vessel segmentation*. in *2019 IEEE International Conference on Image Processing (ICIP)*. 2019. IEEE.
24. Li, D. and S. Rahardja, *BSEResU-Net: An attention-based before-activation residual U-Net for retinal vessel segmentation.* Computer Methods and Programs in Biomedicine, 2021. 205: p. 106070.
25. Van Dyk, D.A. and X.-L. Meng, *The art of data augmentation.* Journal of Computational and Graphical Statistics, 2001. 10(1): p. 1-50.
26. Alsabhan, W. and T. Alotaiby, *Automatic building extraction on satellite images using Unet and ResNet50.* Computational Intelligence and Neuroscience, 2022. 2022.
27. Sha, Y., Y. Zhang, X. Ji, and L. Hu, *Transformer-unet: Raw image processing with unet.* arXiv preprint arXiv:2109.08417, 2021.
28. Asch, F.M., V. Mor-Avi, D. Rubenson, S. Goldstein, M. Saric, I. Mikati, S. Surette, A. Chaudhry, N. Poilvert, and H. Hong, *Deep learning–based automated echocardiographic quantification of left ventricular ejection fraction: a point-of-care solution.* Circulation: Cardiovascular Imaging, 2021. 14(6): p. e012293.
29. Schneider, M., P. Bartko, W. Geller, V. Dannenberg, A. König, C. Binder, G. Goliasch, C. Hengstenberg, and T. Binder, *A machine learning algorithm supports ultrasound-naïve novices in the acquisition of diagnostic echocardiography loops and provides*

accurate estimation of LVEF. The International Journal of Cardiovascular Imaging, 2021. 37: p. 577-586.

30. Sirjani, N., S. Moradi, M.G. Oghli, A. Hosseinsabet, A. Alizadehasl, M. Yadollahi, I. Shiri, and A. Shabanzadeh, *Automatic cardiac evaluations using a deep video object segmentation network.* Insights into Imaging, 2022. **13**(1): p. 1-14.
31. Vafaeezadeh, M., H. Behnam, A. Hosseinsabet, and P. Gifani, *A deep learning approach for the automatic recognition of prosthetic mitral valve in echocardiographic images.* Computers in Biology and Medicine, 2021. **133**: p. 104388.
32. Sharma, A., K. Sharma, and A. Kumar, *Real-time emotional health detection using fine-tuned transfer networks with multimodal fusion.* Neural Computing and Applications, 2022: p. 1-14.
33. Zhou, S.K., H. Greenspan, C. Davatzikos, J.S. Duncan, B. Van Ginneken, A. Madabhushi, J.L. Prince, D. Rueckert, and R.M. Summers, *A review of deep learning in medical imaging: Imaging traits, technology trends, case studies with progress highlights, and future promises.* Proceedings of the IEEE, 2021. **109**(5): p. 820-838.
34. Smistad, E. and A. Østvik. *2D left ventricle segmentation using deep learning*. in *2017 IEEE international ultrasonics symposium (IUS)*. 2017. IEEE.
35. Taylor, L. and G. Nitschke. *Improving deep learning with generic data augmentation*. in *2018 IEEE symposium series on computational intelligence (SSCI)*. 2018. IEEE.
36. Zhang, Y., Z. Mu, L. Yuan, and C. Yu, *Ear verification under uncontrolled conditions with convolutional neural networks.* IET Biometrics, 2018. **7**(3): p. 185-198.
37. Savarese, G., D. Stolfo, G. Sinagra, and L.H. Lund, *Heart failure with mid-range or mildly reduced ejection fraction.* Nature Reviews Cardiology, 2022. **19**(2): p. 100-116.
38. Ma, X. and Z. Yang, *A new multi-scale backbone network for object detection based on asymmetric convolutions.* Science Progress, 2021. **104**(2): p. 00368504211011343.
39. Simonyan, K. and A. Zisserman, *Very Deep Convolutional Networks for Large-Scale Image Recognition.* CoRR, 2014. **abs/1409.1556**.
40. He, K., X. Zhang, S. Ren, and J. Sun, *Deep Residual Learning for Image Recognition.* 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015: p. 770-778.
41. Huang, G., Z. Liu, and K.Q. Weinberger, *Densely Connected Convolutional Networks.* 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016: p. 2261-2269.
42. Tan, M. and Q.V. Le, *EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks.* ArXiv, 2019. **abs/1905.11946**.
43. Harris, P.A., Taylor, R., Thielke, R., Payne, J., Gonzalez, N. and Conde, J.G., 2009. Research electronic data capture (REDCap)—a metadata-driven methodology and

workflow process for providing translational research informatics support. Journal of biomedical informatics, 42(2), pp.377-381.
44. Harris, P.A., Taylor, R., Minor, B.L., Elliott, V., Fernandez, M., O'Neal, L., McLeod, L., Delacqua, G., Delacqua, F., Kirby, J. and Duda, S.N., 2019. The REDCap consortium: building an international community of software platform partners. Journal of biomedical informatics, 95, p.103208.
45. Rossum, G.v. *The Python Language Reference: Release 3.6.4*. 2018.
46. Abadi, M., A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G.S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I.J. Goodfellow, A. Harp, G. Irving, M. Isard, Y. Jia, R. Józefowicz, L. Kaiser, M. Kudlur, J. Levenberg, D. Mané, R. Monga, S. Moore, D.G. Murray, C. Olah, M. Schuster, J. Shlens, B. Steiner, I. Sutskever, K. Talwar, P.A. Tucker, V. Vanhoucke, V. Vasudevan, F.B. Viégas, O. Vinyals, P. Warden, M. Wattenberg, M. Wicke, Y. Yu, and X. Zheng, *TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems.* ArXiv, 2016. **abs/1603.04467**.
47. Chollet, F. *Keras: The Python Deep Learning library*. 2018.
48. Webb, G.I., *Overfitting*, in *Encyclopedia of Machine Learning*, C. Sammut and G.I. Webb, Editors. 2010, Springer US: Boston, MA. p. 744-744.
49. Rezatofighi, S.H., N. Tsoi, J. Gwak, A. Sadeghian, I.D. Reid, and S. Savarese, *Generalized Intersection Over Union: A Metric and a Loss for Bounding Box Regression.* 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2019: p. 658-666.
50. Jin, X. and J. Han, *K-Means Clustering*, in *Encyclopedia of Machine Learning*, C. Sammut and G.I. Webb, Editors. 2010, Springer US: Boston, MA. p. 563-564.
51. Al-amri, S.S., N.V. Kalyankar, and D. KhamitkarS., *Image Segmentation by Using Threshold Techniques.* ArXiv, 2010. **abs/1005.4020**.
52. Samuel, A.L., *Some Studies in Machine Learning Using the Game of Checkers.* IBM J. Res. Dev., 1967. **44**: p. 206-227.
53. Chen, L.-C., G. Papandreou, I. Kokkinos, K.P. Murphy, and A.L. Yuille, *DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs.* IEEE Transactions on Pattern Analysis and Machine Intelligence, 2016. **40**: p. 834-848.
54. De Alexandria, A.R., P.C. Cortez, J.A. Bessa, J.H. da Silva Félix, J.S. De Abreu, and V.H.C. De Albuquerque, *pSnakes: A new radial active contour model and its application in the segmentation of the left ventricle from echocardiographic images.* Computer methods and programs in biomedicine, 2014. **116**(3): p. 260-273.

55. Ouyang, D., B. He, A. Ghorbani, N. Yuan, J. Ebinger, C.P. Langlotz, P.A. Heidenreich, R.A. Harrington, D.H. Liang, and E.A. Ashley, *Video-based AI for beat-to-beat assessment of cardiac function.* Nature, 2020. **580**(7802): p. 252-256.
56. Chen, Y., X. Zhang, C.M. Haggerty, and J.V. Stough. *Assessing the generalizability of temporally coherent echocardiography video segmentation*. in *Medical Imaging 2021: Image Processing*. 2021. SPIE.
57. Gu, H., B.S. Sidhu, L. Fang, J. Webb, T. Jackson, S. Claridge, E. Einarsen, R. Razavi, N. Papageorgiou, A. Chow, S. Bhattacharyya, P. Chowienczyk, and C.A. Rinaldi, *First-Phase Ejection Fraction Predicts Response to Cardiac Resynchronization Therapy and Adverse Outcomes.* JACC: Cardiovascular Imaging, 2021. **14**(12): p. 2275-2285.
58. Liu, X. et al. Deep learning-based automated left ventricular ejection fraction assessment using 2-D echocardiography. Am J Physiol Heart Circ Physiol 321, H390-H399, doi:10.1152/ajpheart.00416.2020 (2021).
59 Smistad, E. et al. Real-Time Automatic Ejection Fraction and Foreshortening Detection Using Deep Learning. IEEE Trans Ultrason Ferroelectr Freq Control 67, 2595-2604, doi:10.1109/TUFFC.2020.2981037 (2020).
60 Tokodi, M. et al. Deep Learning-Based Prediction of Right Ventricular Ejection Fraction Using 2D Echocardiograms. JACC Cardiovasc Imaging 16, 1005-1018, doi:10.1016/j.jcmg.2023.02.017 (2023).