DOI: http://dx.doi.org/10.26483/ijarcs.v15i1.7040

Volume 15, No. 1, January-February 2024

ISSN No. 0976-5697

International Journal of Advanced Research in Computer Science

RESEARCH PAPER

Available Online at www.ijarcs.info

PREDICTING CARDIAC ISSUES FROM ECHOCARDIOGRAMS: A LITERATURE REVIEW USING DEEP LEARNING AND MACHINE LEARNING TECHNIQUES

B.Arockia Valanrani
Research Scholar, Department of Computer Science,
Vellalar College for Women,
Thindal, Tamilnadu, India.

Dr.S.Devi Suganya
Assistant Professor, Department of Computer Science,
Vellalar College for Women,
Thindal, Tamilnadu, India.

*Abstract***:** Cardiovascular disease (CVD) has a substantial impact on overall health, well-being, and life-expectancy. Echocardiography is a widely used imaging technique in cardiovascular medicine, utilizing various medical imaging technology to visualize heart chambers and valve's motion activity. In order to diagnose and treat complicated cardiovascular problems, it takes high-resolution images of the heart and its surroundings. However, it has limitations such as long procedure times, multiple measurement values, complex analyses, individualized assessments, operator subjectivity, and wide observation ranges. This makes it challenging for sonographers to accurately detect and diagnose heart diseases. In recent days, Deep Learning (DL) is increasingly used in clinical computer-assisted systems for disease detection, feature segmentation, functional evaluation, and diagnosis. It is an alternate technique for accurate detection and treatment of cardiovascular disorders; it improves the diagnostic capacities of echocardiography by identifying pathological conditions, extracting anatomically significant data, measuring cardio-motion, and calculating echo image quality. This paper presents a detailed review of various DL frameworks developed to analyse different cardiac views using echocardiography for improving the prediction and diagnosis of CVD. At the outset, a variety of echocardiography systems linked with DL-based segmentation and classification are reviewed briefly. Afterwards, a comparison research is carried out to gain insight into the shortcomings of those algorithms and provide a fresh approach to improve the accuracy of cardiac view categorization in echocardiography systems.

Keywords: Cardiovascular Disease, Echocardiography, Deep Learning, Computer-Assisted. Cardio-Motion.

I. INTRODUCTION

When it comes to global mortality, cardiovascular disease (CVD) takes the cake. Reducing mortality and treating patients at an early stage both depend on early detection of CVD [1]. Echocardiography (echo), cardiac magnetic resonance imaging (CMRI), computed tomography (CT), and multiple gated acquisition scan (MUGA) are some of the imaging modalities used to diagnose cardiovascular disease [2]. Echo is the most popular since it is accessible, cheap, portable, and does not involve any kind of intrusive procedure. Echocardiography is a kind of ultrasonography that uses M-Mode imaging to record the heart's electrical activity while it beats, giving doctors a better image of the blood flow to the organ. [3]. Echocardiography captures the heart's natural motion by means of successive frames that stand in for a three-dimensional model. The following are the dimensions of the frame: width, height, and time [4]. The sample echocardiogram is shown in Fig.1.

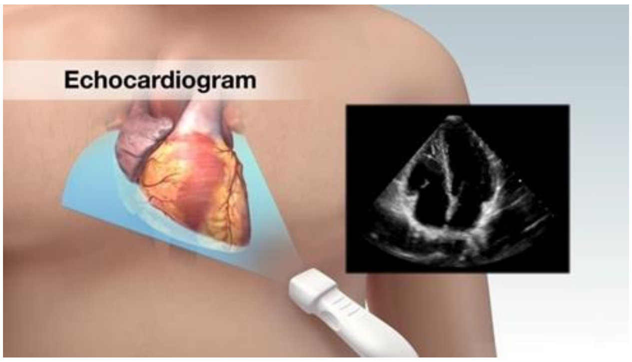

Figure 1. Echocardiography

One option for the recorded echo data is a static image taken at a certain moment of the heart, while another is a video sequence that spans the whole cardiac cycle. The contraction of the ventricles (systole) and their relaxation (diastole) are the two phases that make up a single cardiac cycle [5]. Multiple echo modes, including A-mode, B-mode, M-mode, and Doppler, are available, each with its own set of advantages and disadvantages depending on the application [6].

- **A-mode:** The most basic kind of ultrasonography is known as A-mode. The echoes are shown on the screen as a function of depth while a single transducer scans a line across the body. Another use of A-mode ultrasound is therapeutic ultrasound, which allows for the precise targeting of harmful wave energy to a particular tumor or calculus.
- **B-mode:** The B-mode ultrasound uses a two-dimensional image that is created by scanning a plane across the body using a linear array of transducers all at once.
- **M-mode:** "M" is for motion. As the organ boundaries that generate reflections shift in relation to the probe, m-mode allows physicians to see and quantify range of motion via a quick succession of B-mode scans whose images follow each other sequentially on screen.
- **Doppler mode:** Using the Doppler effect, sonographers may see the flow of blood and determine whether anatomical structures are approaching or evading the probe, as well as the relative speed of the two [7]. A blood flow jet crossing a heart valve is one example of a sample volume that may have its velocity and direction shown graphically by determining the frequency shift of that volume. Cardiovascular research and the diagnosis

of portal hypertension, which causes blood to flow backwards via the liver's vasculature, benefit greatly from this. Spectral Doppler, color Doppler, and power Doppler all provide visual representations of Doppler data. The data is often played back over stereo speakers, which create a unique, artificial, pulsating noise [8]. The overview of Doppler mode is shown in Fig. 3.

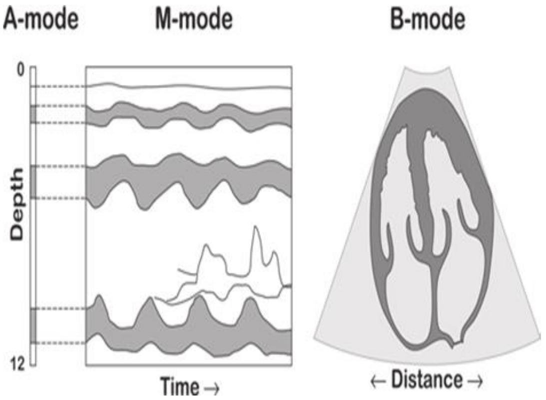

Figure 2. Sample Illustration of A, B and M-mode

Figure 3. Doppler Mode

- **Two Dimensional (2D) and Three Dimensional (3D) Echocardiogram**: 2D echocardiographic imaging offers tomographic views of cardiac structures, guiding M-mode and Doppler echocardiograms. The scan line is swept across an arc, instead of a fixed line of sight. After complex manipulation, a 2D tomographic image is generated for display, providing a comprehensive view of cardiac structures [9]. 3D echocardiographic images, obtained through a transducer, are more accurate than 2D echocardiography for localizing valvular abnormalities, Left Ventricular (LV) volume calculation, Right Ventricle (RV) assessment, guiding surgical interventions, and complex congenital heart disease. They also provide comprehensive assessment of vena contracta, improve valvular regurgitation quantification, and provide a more rapid evaluation of mitral valve area in mitral stenosis [10].

Various heart problems may be better seen and diagnosed when these modes are used together. Echocardiography imaging provides many perspectives of the heart by manipulating the transducer at various angles to record the heart's movements [11]. It is possible to identify and evaluate many anatomical structures after the doctor physically identifies the image. As shown below, there are three distinct perspectives that may be adjusted based on the position of the transducer: Location A, Location B, and Location C [12].

- There are four perspectives at Location A: apical 2 chambers (A2C), apical 3 chambers (A3C), apical 4 chambers (A4C), and apical 5 chambers (A5C) respectively [13].
- Parasternal long axis (PLA) is the only view in Location B [14].
- A total of three perspectives make up Location C: parasternal short axis of aorta (PSAA), parasternal short axis of papillary (PSAP) and parasternal short axis of mitral (PSAM) respectively [15]

In essence, the echocardiograph provides all of the aforementioned viewpoints with discriminative information based on spatial and temporal perspective. Despite its usefulness in assessing heart anatomy and function, the non-invasive imaging technique known as echocardiography isn't without its drawbacks, including a lengthy procedure time, a large range of possible measurement results, complicated analysis, and a high degree of operator subjectivity. These limitations demand medical specialist training in echocardiography, as they increase complexity and user subjectivity, and necessitate specialized training for accurate diagnosis.

Important quantitative metrics such as myocardial mass, wall thickness, left ventricular (LV), right ventricular (RV), and ejection fraction (EF) may be extracted from echo images by a critical process known as cardiac image segmentation [16]. On the most important anatomical structures, such as the ventricles, atria, and arteries, imaging techniques including CT, ultrasonography, and CMRI are utilized. Cardiac Image segmentation is a process that removes disturbances and edge enhancements, computes the region of interest (ROI) in echo images and estimates LV shape. When doctors do an LV shape estimate, it's quite similar to drawing an outline of the heart's components [17]. By segmenting the image data, this method guarantees that the final products are accurate representations of the underlying anatomy. The segmentation process in medical imaging yields an estimate of the chamber or wall form [18]. The intricacy of the heart's internal architecture, together with severe disruptions in echo images and poor contrast in fine structures, makes it difficult for the segmentation process to estimate the form of the heart valves. Evaluating the time-dependent development of LV reconstruction is extremely laborious because of the large estimate error caused by the substantial variations in LV outline created by various physicians.

Researchers have recently made strides in employing new technologies, such as Artificial Intelligence (AI), to improve the clinical assessment of echocardiography using echocardiographs [19]. An appealing alternative for early detection and treatment of cardiovascular diseases, especially for lower complexities, is Machine Learning (ML) and Deep Learning (DL) models, which are AI tools that automatically segment and categorize a variety of cardio-pathological disorders. These include disorders involving the left ventricle, mitral regurgitation, and wall motion [20]. Generally, ML approaches are used to predict malignant disease development and treatment, improving diagnostic quality in soft-tissue examination tasks [21]. They provide efficient image interpretation, segmentation, and precise decision-making, improving cardiovascular care and reducing disparities [22]. In both intra- and inter-reader contexts, ML models exhibit a great deal of variability and are prone to errors. Fetuses and babies have smaller hearts with less defined borders, making manual assessment of cardiac parameters more difficult.

When compared to the ML frameworks, DL models provides an efficient performance on heart diseases segmentation and classification. As an affordable, less subjective, and time-saving substitute for human intervention, DL systems are becoming more attractive. By using such a system, the effort may be significantly reduced, variability between and among readers can be controlled, and the lack of cardiology experts in low-resource areas can be addressed [23]. The Convolutional Neural Network (CNN), the Recurrent Neural Network (RNN), the Long-Short Time Memory (LSTM), the Deep Belief Network (DBN), and many more are examples of DL models. By using important and distinctive characteristics found in pertinent medical data, these models aid doctors in making an unbiased and trustworthy diagnosis of the condition [24]. DL's primary goal is to improve the diagnostic capacities of echocardiography via the detection of pathological conditions, the quantification of cardio-motion, and the enhancement of echo image quality. It measures cardiac motion characteristics including myocardial velocity (MV), EF, and longitudinal strain (LS), and it aids doctors in classifying cardio-views and detecting cardio-pathological illnesses like mitral regurgitation, left ventricle problems, and wall motion disorders [25]. In recent years, there have been several DL frameworks developed to segment, predict and classify various categories of CVD using echocardiograph (echo data\image).

In order to improve heart disease prediction tasks, this work aims to provide a comprehensive overview of several DL models and how they are used to the segmentation and classification of cardiac images using echocardiograph. Additionally, in order to propose future scope, a comparison research is given to address the pros and cons of those models. The rest of the sections are prepared as follows: In Section II, we'll look at a variety of models that have been developed to extract cardiac images from echocardiography data and then categorize them. The evaluation of various models in comparison is presented in Section III. The whole research is summarized in Section IV, and future scope.

II. LITERATURE SURVEY

This section divides into two modules based and DL based segmentation and classification models using the cardiac views echocardiograph data.

A. *Survey on Cardiac Image Segmentation for Cardiac View Representation*

In order to fully segment the four chamber views of pediatric echocardiography images, Hu et al. [26] created a DL model that used Bilateral Segmentation Network (BiSeNet). The BiSeNet model included the Feature Fusion Module (FFM), a geographical route, and a context path. Two routes make up this model: one for capturing low-level spatial information and another for capturing high-level context semantic elements. Then, the FFM was introduced to integrate the features learned from these two paths for automatic pediatric echocardiography segmentation.

To improve left ventricular segmentation in two-dimensional echocardiography, Leclerc et al. [27] suggested a Localization U-Net (LU-Net). Through end-to-end learning problems, this model seeks to detect the LV before segmenting the endocardial and epicardial boundaries. In this model, two U-Nets work together to train three tasks at once: first, U-Net segmentation in the region proposal network (RPN); second, U-Net segmentation on an ultrasound image cut from the localization network; and last, U-Net segmentation on the LV bounding box. Using this method, gradients may be transferred from the output to the input of the network. Additionally, an attention model was employed to enhance segmentation and clinical index estimation in 2DE.

In order to quantify cardiac indices from echocardiography images, segment the heart based on regions, and retrieve echocardiograms in real time, Zamzmi et al. [28] created Trilateral Attention Network (TaNet). This model learns the context link among regions of interest (ROIs) and uses a Spatial Transformer Network (STN) module to localize them. Then, the low-level, textural, and high-level characteristics that were rich were encoded using the lightweight routes. In order to segment the cardiac area, the feature embeddings from each of these paths were combined. Finally, the TaNet was applied to jointly to segment, quantify and localize the cardiac indices.

In order to get strain measurements from echocardiography footage, Deng et al. [29] created a DL model for myocardial segmentation and motion estimation. In order to segment and extract the target frame's centerline, the model employs a 3D cardiac segmentation network (3D-CSN). At each instant, the motion network generates a velocity vector based on its estimation of the movement field of pixels in the ROI. In order to estimate strain and segment cardiac tissue, these velocities compute global and local line arc lengths and propagate the centerline location.

Lal [30] developed an automatic two-chamber segmentation model called TC-Signets for echocardiography. This model integrates U-Net with modified skip connection (MSC), Atrous Spatial Pyramid Pooling (ASPP) modules, squeeze and excitation modules (SEM). The U-Net model includes an encoder, bridge, and decoder for echocardiogram segmentation tasks. The ASPP extract multiple scale features by expanding convolutional layers. MSC was incorporated between the encoder and decoder levels to reduce the path size and convolutional blocks along with the skip connections. The SEM was used to apply the weighting functions to channels based on their importance to capture the relevant features from echo data.

To segment echocardiogram sequences, Sirjani et al. [31] created EchoRCNN, a semi-automated neural network that combines a CNN cardiac image segmentation structure based on mask regions with a reference-guided mask propagation video object segmentation network. The network learns to distinguish between ventricles from ultrasound cardiac data. From the output of the network, the estimated EF for the LV was derived. The output area for robust image segmentation was finally identified for fractional area changes for the right ventricle.

A DL segmentation model for myocardial T1 mapping was built by Bhatt et al. [32] utilizing MRI relaxation-based synthetic contrast augmentation. To automate the segmentation of native and postcontrast T1 maps as well as the measurement of global and segmental myocardial T1 and ECV, this technique developed a fully automated segmental relaxometry method using synthetic contrast augmentation with neural networks (FASTR-SCANN). During the training and testing of CNN for the cardiac T1 segmentation task, the synthetic inversion recovery images were obtained from the fitted T1 relaxation signal model. These images were generated with strong blood-myocardium contrast.

The TF-Unet algorithm was proposed by Fu et al. [33] for the automated segmentation of cardiac MRI images. This approach uses U-Net and Transformer to automate the process of medical image segmentation, making it powerful. Full use of convolutional advantages in detail grasping was then achieved by using CNN for feature extraction and spatial encoding of inputs. Lastly, in order to make the most of Transformer for effective cardiac image segmentation, it was utilized to model features at various sizes and add distant dependencies to high-level features.

In order to analyze myocardial strain, Barbaroux et al. [34] used spatio-temporal CNN to automate the segmentation of long- and short-axis DENSE cardiovascular MRI. This technique involves training 2D+time nnU-Net models to segment the left ventricle myocardium using DENSE magnitude data in both short- and long-axis images. A dataset including short-axis and long-axis slices from both healthy individuals and patients with different diseases was used to train the networks. When it comes to cardiac myocardial segmentation, our spatio-temporal model delivers consistent performance throughout the cine stream.

An open-source, fully-automated hybrid cardiac substructure segmentation system (OFHCSS) was developed by Finnegan et al. [35]. An automated segmentation of the whole heart, its chambers, major vessels, valves, coronary arteries, and conduction nodes was achieved by the use of a multi-stage approach that combined DL with multi-atlas mapping and geometric modeling. In order to guarantee the performance dependability for appropriately delineating cardiac substructures, the atlases used to create automated segmentations were incompatible with the assessment images.

In order to separate the left heart's anatomy from echocardiography images, Mortada et al. [36] created a DL-based method. In order to automatically segment an echocardiographic image into LVendo, LVepi, and LA, it was created as a mix of two convolutional neural networks, namely a U-Net and the YOLOv7 algorithm. To aid in cardiological clinical practice, doctors obtained and annotated apical two- and four-chamber images at end-systole and end-diastole for each patient. This allowed for automated segmentation of the left heart's anatomical features.

B. Survey on DL Interpretation of Echocardiography for Cardiac View Classifications

Shahin & Almotairi [37] developed a DL model for classifying and identifying physiological locations in echocardiography views. In the beginning, the neutrosophic sets retrieved spatial CNN characteristics from every frame of the echo-motion activity, and new temporal features were identified. After that, in order to get the deep features, this model triggers a deep ResNet model that has already been trained. Using the features concatenation approach, spatial and neutrosophic temporal CNN features were merged in the end. At last, the LSTM network was trained using the fused CNN features in order to locate and categorize echo-cardio images.

Kusunose et al. [38] suggested a CNN model for accurate view classification of echocardiography images. This approach used the building and training of three CNN models to identify the echocardiographic images. As its input, the first model made use of the time-averaged image. The second model used selected images which were trained independently with and the averaged probability as input. Finally, the third model was employed alongside with the view classification model in the prediction phase to classify the echocardiography image views.

Østvik et al. [40] developed the EchoPWC-Net model for motion estimation and improved myocardial function imaging in echocardiography. The model employs a CNN to distinguish between diastole and systole in B-mode images and a cardiac view classification network (CVC-Net) to follow legitimate acoustic windows. The U-Net architecture segments the myocardium, and centerline extraction extracts the myocardia's contour and endo and epi cardial borders. Finally, EchoPWC-Net was used for motion estimation and enhanced myocardial function imaging in echocardiography.

Khan et al. [40] presented a Cycle Generative Adversarial Network (GAN) model for developing the high quality electrocardiograph images simultaneously reducing the speckle noise and blocking artifacts. The model uses a generator network to minimize variation between fake and target image distributions, and an additional generator to convert clean images to real ones. Additionally, a block artifact model is integrated to eliminate block-artifact and speckle noise resulting in high-quality echocardiograph images.

For the purpose of echocardiography video categorization, Feng et al. [41] proposed a spatio-temporal network with two streams of attention. The network consists of spatial context and motion streams, which input echo frames and optical flow. Visual features are extracted from each stream and fed into a CNN through a pre-trained auto-encoder and LSTM classifier. An attention module is employed to enhance task-related time steps for better network performance. The learned features are concatenated for normal and abnormal echo video classification.

An echocardiography model for self-supervised cardiac view synchronization was created by Dezaki et al. [42]. The model makes use of two types of supervisory signals: interdependencies between different cines and intra-view self-supervision for a single cine. The repetition counting method (RepNet) was used to approximate a single period for synchronization. The encoder network improved the embeddings for temporal synchronization, and a 2D network was used to insert temporally stacked frames into a 3D convolutional network. Finally, the ResNet model was adopted to classify and synchronize 2DE videos.

Using a combination of DL and natural language processing (NLP), Hagberg et al. [43] created a model that can identify the size and function of the RV from echocardiographic images. A view classifier was created to choose the 4-chamber or RV-focused view, and this model used 2DE video loops to determine the size and function of the RV. Annotation by humans and NPL algorithms worked together to accurately categorize RV functions and sizes.

Evain et al. [44] developed the EchoPWC-Net model for motion estimation using echocardiography. A Convolutional Neural Sub-Network (CNSN) estimator and normalized cross-correlations from 2DE images are used by the model to forecast dense displacement maps. By using prior flows and dilation convolutions, the estimator improves flow accuracy and iteratively obtains a displacement field that contains a fourth of the input image. In order to improve the network's parameters, which are used to detect motion in 2DE images, a multi-scale loss function and the calculated distance between intermediate estimated flows are used.

In order to automatically recognize regional wall motion abnormalities (RWMAs) in echocardiography data, Lin et al. [45] built a DL model for myocardial infarction patients. Initially, the view selection of myocardium was processed using Xception model. Next, LSTM-Unet separates the Xception output frames. In the end, the segment and the original video were used as inputs in the 3D-CNN model, which can identify and categorize abnormalities in regional wall motion and quantify heart function in myocardial infarction.

Darmawahyuni et al. [46] developed a generalization model of deep learning for ECG signal categorization in intra and inter-patients' scenarios. This technique reduced artifacts and noise in ECG data by pre-processing them using discrete wavelet transforms. The last step was to segment the pre-processed data according to rhythm and beats. This allowed us to detect not only the next heartbeat, but also the waveforms included inside each beat. In the end, the ECG signal classification process made use of 1D-CNNs for feature extraction and classifier learning the properties of each rhythm and beat event.

To identify mitral regurgitation (MR) in children using echocardiography, Edwards et al. [47] built a CNN model. In order to identify the perspective and the existence of MR, the echocardiograms were tagged using clip and frame. Using the annotated data, two CNNs were trained to categorize parasternal long-axis color Doppler images according to view and the presence of any MR, including physiological MR. Frames were used to create the view classification model for MR detection.

Tokodi et al. [48] suggested a DL model to analyze the right ventricular EF (RVEJ) using 2DE. This model utilizes 2DE videos and 3DE RVEF values as input. The RVEF was predicted using the pre-processed image sequences and their associated binary masks by means of the spatiotemporal neural networks. Each frame's characteristics were extracted using the ShuffleNet. To combine the data and assess changes over time in the image sequence, a temporal convolutional layer was used. The last step was to use occlusion sensitivity visualization (OSV) to hide parts of the input image and see how much of an effect the specified area had on RVEJ prediction.

III. COMPARATIVE ANALYSIS

Tables 1 and 2 below offer a comparison of the pros and cons of using various DL methods for segmentation and classification in cardiac views derived echocardiography, as discussed briefly in the previous section.

Table I. Comparison of Different Segmentation Models for Cardiac Views Interpretation

Ref No.	Technique	Merits	Demerits	Performance Evaluation
[26]	BiSeNet, FFM	This model effectively removes the presence of shadowing and inter-variability in the images.	It takes significant amount of time to train the model	Accuracy = 91.3% Precision = 90.5% Recall = 92.8% Dice Coefficient (DC) = 0.908; Jaccard Index (JI) = 0.832
[27]	LU-Net, RPN, Attention model	This model was highly optable to apply in the real time medical treatment applications	The parameters were sensitive and the execution time was high	Correlation Coefficient (CC) = 0.96; Mean Absolute Error (MAE) = 8.1; Limit Of Agreement (LOA) = 1.7
[28]	TaNet, STN, Lightweight Pathways	Retrieves high-quality echos, segmenting complex chambers, high inference speed	This model was slightly automated and trained with limited dataset	F1-Score = 0.96; Intersection Over Union (IOU) = 0.97; CC rate = 0.99
[29]	3D-CSN, Motion Network,	Highly automated, enhanced reproducibility, high processing speed	This model required large synthetic data for training and slightly opts to overfitting issues	Average DC = 0.82; End-point error = 0.05 per frame Spearman correlation = 0.90
[30]	U-Net, MSC, ASPP, SEM,	This model eliminates the training issues like vanishing gradients and degradation problems.	The parameter were not optimized properly which yield high generalizability error	Accuracy = 95% F1-Score = 94% DC = 0.92; JI = 0.83
[31]	Mask region-based CNN, Video object Segmentation Network.	This method aids radiologists in efficiently tracking heart function over time.	Error propagation was high throughout the echocardiography series	Precision = 97.03 ± 0.01; Recall = 93.14 ± 0.04; Dice Score = 94.97 ± 0.02
[32]	FASTR-SCANN, Postcontrast T1 map segmentation model	High network performance and lower error rate	Manual expert labelling of training data needs to be devised properly	Accuracy = 92.8%;
[33]	CNN, U-Net and Transformer	Data interpretation and convergence rate was efficient	Learning rate was insufficient which decreases the training speed of the model	Dice score = 91.72%; Accuracy = 87.45%
[34]	2D+time nnU-Net, CNN	Lesser computational complexity and burden	However, this model provides poor contrast between blood and myocardium with potential spiral streaking artifacts.	DICE = 0.83 ± 0.05;
[35]	DL, multi-atlas mapping and geometric modelling	Even for the smaller structures, robust and anatomically consistent segmentations were provided.	The acquisition parameters were consecutively low so specific substructures were not contoured properly	Accuracy = 79%; Dice Score = 0.89; F1-Score = 78%
[36]	YOLOv7, U-Net,	This method was easily-reproducible and machine-independent	High manual leads to maximized inter-cardiologist variability and subjectivity	Dice = 87.57%; Accuracy = 88%

Table II. Comparison of Different DL Associated Echocardiography Model for Cardiac Views Classification

Ref No.	Technique	Merits	Demerits	Performance Evaluation

[37]	CNN, ResNet, LSTM	The suggested technique has significantly reduced the time cost required for training and feature extraction.	The hyper-parameters were not regularized or configured properly results with high classification errors	Accuracy = 96.3%; Precision = 96.4%; Sensitivity = 95.7%; Specificity = 99.4%
[38]	CNN, View classification model	Cost-effective and reduced training time	This model required large clinical data to train the model	F1-Score = 94.1%; Error rate = 1.9%;
[39]	EchoPWC-Net, U-Net, CNN, CVC-Net, Centreline Extraction	This model reduces the image noises and robust to overfitting issues	It restricted the values within the measured strain making it harder to spot localized problems.	Runtime = 18.9 ± 0.7 frames per second (FPS); Standard Deviation (SD) = 0.16
[40]	CycleGAN, Block artifact model	Rapid reconstruction time, low powered high visual quality and enhanced resolution effect on echocardiograph images	The convergence rate was slower and in some cases the training was unstable	Average Reconstruction Time = 6.59 milliseconds; Peak Signal-To-Noise Ratio (PSNR) = 26.16;
[41]	CNN-LSTM, Auto-encoder, Attention Module	This model works well on larger dataset with lesser complexities	Increase in the data, increases the significance of high misclassification error	Accuracy = 91.18%; Sensitivity = 94.11%; Specificity = 88.24
[42]	RepNet, 2D network, 3D convolutional network, ResNet	Accurately, it can generalize to views that weren't in the training set.	The technique assumes no specific echo view making it insensitive to different cardiac views.	SD = 4.6; MAE = 3.9
[43]	NLP, DL model	Rapid and cost-effective expansion of training dataset utilized for medical diagnostic domains with limited labeled image data	Biased selection towards higher-quality as the view classifier couldn't find relevant views for image classification.	Accuracy on RV size = 83%; Accuracy on RV function = 82%
[44]	EchoPWC-Net, CNSN	Better generalization ability and eliminates overfitting issues	Hugh computational time and lower learning rate	Mitral Annular Plane Systolic Excursion (MAPSE) = 0.85; Global Longitudinal Strain (GLS) = 0.71
[45]	Xception, LSTM-UNet, 3D-CNN	High data interpretability and memory capacity	The well-defined boundaries close to the imaging area's edge yielded a relatively lower IOU.	Accuracy = 94%; DC = 89%; Area Under Curve (AUC) = 91%
[46]	1D-CNN	This model was much robust to uncertainty and noisy images	The partition of the different wave categories from the ECG signals was not conducted before the categorization.	Accuracy = 92.17% Sensitivity = 96.97% Precision = 92.23%; F1-score = 94.39%
[47]	CNN model	This model result with lower computational cost and time	Higher generalization error due to smaller training data	Accuracy = 89% Precision = 93%; Recall = 86%; F1-Score = 88%
[48]	ShuffleNet, OSV, temporal convolutional layer	Dynamic clinical tasks, cost-effective and less memory space	Retrospective design using 3DE data might cause selection bias in RV analysis due to limited subjects.	Accuracy = 77.4%; AUC = 0.92

IV. Result and Discussion

This section examines the literature on efficient procedures and compares their performance to find out how they stack up against one another. For the cardiac image segmentation, BiSeNet [26], U-Net [30], FASTR-SCANN [32], U-Net-Transformer [33], OFHCSS [35] and U-Net-YOLOv7 [36] are considered for the evaluation. Then, CNN-ResNet [37], CNN-LSTM [41], Xception [45], 1D-CNN [46], CNN [47] and ShuffleNet [48] are taken for the cardiac view classification. For the evaluation, both segmentation and classification compared in terms of accuracy.

Accuracy: It is the proportion of the number of exact identifications of positive and negative cases over the total cases tested.

$$Accuracy = \frac{True\ Positive\ (TP) + True\ Negative\ (TN)}{TP+TN+False\ Positive\ (FP)+False\ Negative\ (FN)} \quad (1)$$

Eqn. (1) states that TP is the number of positive instances that were properly recognized as normal, and TN is the number of negative cases that were correctly identified as abnormal. Another difference between FP and FN is that FP represents the amount of false positives while FN represents the number of false negatives.

Fig. 4 shows the comparison of various DL based segmentation models in terms of accuracy. It is observed that the U-Net [30] 4.05%, 2.37%, 8.63%, 20.25% and 7.95% higher than BiSeNet, FASTR-SCANN, U-Net-Transformer, OFHCSS and U-Net-YOLOv7. This results indicates that the U-Net [30] has better segmentation accuracy than other models as it employs integrates U-Net with MSC, ASPP and SEM. The model includes an encoder, bridge and decoder for echocardiogram segmentation extracting multiple scale features through convolutional layers, and incorporating MSC to reduce path size and convolutional blocks. SEM applies weighting functions to channels to capture relevant features from echo data for automatic segmentation task.

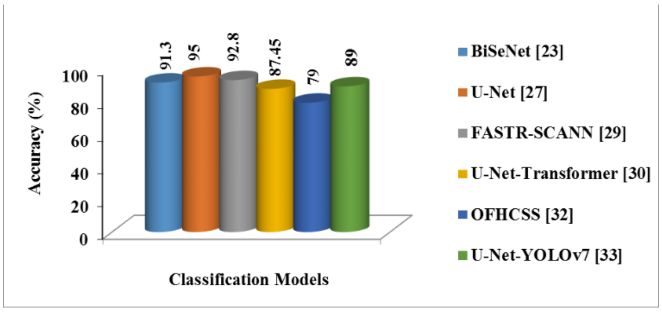

Figure 4. DL based Segmentation models for cardiac view Representation

Fig. 5 shows the comparison of DL based classification models in terms of accuracy. It is observed that the CNN-ResNet [37] model 5.62%, 2.45%, 4.48%, 8.20% and 24.42% higher than CNN-LSTM, Xception, 1D-CNN, CNN and ShuffleNet. This results indicates that the CNN-ResNet [37] has better classification accuracy than other models as it employs CNN and LSTM model to classify cardioviews and identify three cardio locations. The deep CNN features were extracted using pre-trained networks, and the concatenation method was used to integrate the features. The next step was to use the ResNet classifier to locate and categorize each echo-clip into cardio-views. This model saves time and costs for training and features extraction, providing a rapid cardio-views classification model.

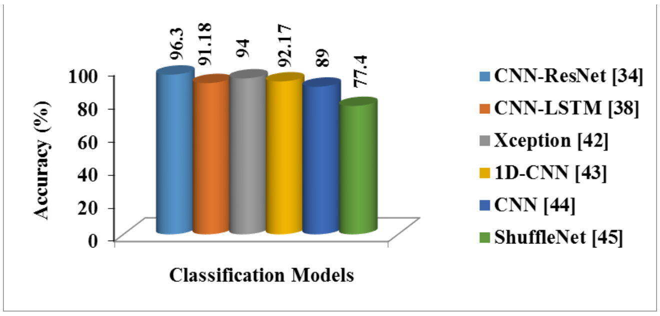

Figure 5. DL based Classification models for cardiac view identification

V. CONCLUSION

Echocardiography is a common screening modality for healthy, asymptomatic patients and diagnostic tool for complex CVD. Several approaches have been offered in the literature by various researchers employing DL techniques to address the problem of erroneous diagnosis. In this paper, a comprehensive review on different DL methods for segmenting and classifying the cardiac views using echocardiograph according to their strengths, weaknesses and prediction efficiencies are provided. The discussed challenges and performances are key access for the researchers to develop fully functional models that could help in improving in CVD prediction and diagnosis and provides ultimately personalized treatments for heart disease patients. Therefore, in the future, researchers will concentrate on building more complex computational models for use in rapid mobile apps that can diagnose cardiac conditions and abnormalities utilizing electrographic data in real-time at the point of treatment.

VI. REFERENCES

[1] T. A. Gaziano, “Cardiovascular disease in the developing world and its cost-effective management,” Circulation, vol. 112, no. 23, 2005, pp. 3547-3553 doi:10.1161/CIRCULATIONAHA.105.59179.

[2] V. Panis and E. Donal, “Imaging techniques for cardiac function,” Applied Sciences, vol. 11, no. 22, 2021, pp. 10549, doi:10.3390/app112210549.

[3] M. T. Jensen, P. Sogaard, I. Gustafsson, J. Bech, T. F. Hansen, T. Almdal and P. Rossing, “Echocardiography improves prediction of major adverse cardiovascular events in a population with type 1 diabetes and without known heart disease: the Thousand & 1 Study,” Diabetologia, vol. 62, 2019, pp. 2354-2364, doi:10.1007/s00125-019-05009-2.

[4] B. Pinamonti, E. Abate, A. De Luca, G. Finocchiaro and R. Korcova, “Role of Cardiac Imaging: Echocardiography,” Dilated Cardiomyopathy: From Genetics to Clinical Management, 2019, pp. 83-111, doi:10.1007/978-3-030-13864-6_7 .

[5] R. Alizadehsani, J. Habibi, Z. A. Sani, H. Mashayekhi, R. Boghrati, A. Ghandeharioun and F. Alizadeh-Sani, “Diagnosing coronary artery disease via data mining algorithms by considering laboratory and echocardiography features,” Research in cardiovascular medicine, vol. 2, no. 3, 2013, pp. 133, doi:10.5812%2Fcardiovascmed.10888 .

[6] A. A. Mohamed, A. A. Arifi and A. Omran, “The basics of echocardiography,” Journal of the Saudi Heart Association, vol. 22, no. 2, 2010, pp. 71-76, doi:10.1016/j.jsha.2010.02.011.

[7] E. Naredo and I. Monteagudo, “Doppler techniques,” Clin Exp Rheumatol, 32(1 Suppl 80), 2014, pp. S12-9.

[8] N. S. Anavekar and J. K. Oh, “Doppler echocardiography: a contemporary review,” Journal of cardiology, vol. 54, no. 3, 2009, pp. 347-358, doi:10.1016/j.jjcc.2009.10.001.

[9] M. A. Fagiry, I. A. Hassan and M. Z. Mahmoud, “Two-dimensional echocardiography in the diagnosis of ischemic heart disease,” Journal of Radiation Research and Applied Sciences, vol. 12, no. 1, 2019, pp. 177-185, doi:10.1080/16878507.2019.1594129.

[10] K. L. Mowers, J. B. Fullerton, D. Hicks, G. K. Singh, M. C. Johnson and S. Anwar, “3D echocardiography provides highly accurate 3D printed models in congenital heart disease,” Pediatric Cardiology, vol. 42, 2021, pp. 131-141, doi:10.1007/s00246-020-02462-4.

[11] R. Kumar, F. Wang, D. Beymer and T. Syeda-Mahmood, “Echocardiogram view classification using edge filtered scale-invariant motion features,” In 2009 IEEE Conference on Computer Vision and Pattern Recognition, Jun. 2009, pp. 723-730, IEEE, doi:10.1109/CVPR.2009.5206838.

[12] B. Remenyi, K. Davis, A. Draper, N. Bayley, E. Paratz, B. Reeves and J. R. Francis, “Single parasternal-long-axis-view-sweep screening echocardiographic protocol to detect rheumatic heart disease: a prospective study of diagnostic accuracy,” Heart, Lung and Circulation, vol. 29, no. 6, 2020, pp. 859-866, doi:10.1016/j.hlc.2019.02.196.

[13] R. Senior, H. Becher, M. Monaghan, L. Agati, J. Zamorano, J. L. Vanoverschelde and P. Lancellotti, “Clinical practice of contrast echocardiography: recommendation by the European Association of Cardiovascular Imaging (EACVI) 2017,” European Heart Journal-Cardiovascular Imaging, vol. 18, no. 11, 2017, pp. 1205-1205af, doi:10.1093/ehjci/jex182.

[14] S. J. Karia and R. H. Kassamali, “Transthoracic echocardiogram in a parasternal long axis window,” BMJ, 2014, 348, doi:10.1093/ehjci/jex182.

[15] L. Lindström, U. M. Wilkenshoff, H. Larsson and B. Wranne, “Echocardiographic assessment of arrhythmogenic right ventricular cardiomyopathy,” Heart, pp. 86, no. 1, 2001, pp. 31-38, doi:10.1136/hrt.86.1.31.

[16] A. Skalski and P. Turcza, “Heart segmentation in echo images,” Metrology and Measurement Systems, vol. 18, no. 2, 2011, pp. 305-314, doi:10.2478/v10178-011-0012-y.

[17] S. Mazaheri, P. S. B. Sulaiman, R. Wirza, F. Khalid, S. Kadiman, M. Z. Dimon and R. M. Tayebi, “Echocardiography image segmentation: A survey,” In 2013 International Conference on Advanced Computer Science Applications and Technologies, Dec. 2013, pp. 327-332, IEEE, doi:10.1109/ACSAT.2013.71.

[18] M. Habijan, D. Babin, I. Galić, H. Leventić, K. Romić, L. Velicki and A. Pižurica, “Overview of the whole heart and heart chamber segmentation methods,” Cardiovascular Engineering and Technology, vol. 11, 2020, pp. 725-747, doi:10.1007/s13239-020-00494-8.

[19] A. Karužas, K. Šablauskas, L. Skrodenis, D. Verikas, E. Rumbinaitė, D. Žaliaduonytė-Pekšienė and J. Plisienė, “Artificial intelligence in echocardiography-Steps to automatic cardiac measurements in routine practice,” In European heart journal: ESC Congress 2019 365: Paris, France, 31 August-4 September 2019/European Society of Cardiology. Oxford: Oxford University Press, 2019, vol. 40, suppl. 1, October, doi:10.1093/eurheartj/ehz748.0230.

[20] K. C. Siontis, P. A. Noseworthy, Z. I. Attia and P. A. Friedman, “Artificial intelligence-enhanced electrocardiography in cardiovascular disease management,” Nature Reviews Cardiology, vol. 18, pp. 7, 2021, pp. 465-478, doi:10.1038/s41569-020-00503-2.

[21] N. Kumar and D. Kumar, “Machine learning based heart disease diagnosis using non-invasive methods: A review,” In Journal of Physics: Conference Series, vol. 1950, no. 1, Aug. 2021, pp. 012081, IOP Publishing, doi:10.1088/1742-6596/1950/1/012081.

[22] S. Narula, K. Shameer, A. M. Salem Omar, J. T. Dudley and P. P. Sengupta, “Machine-learning algorithms to automate morphological and functional assessments in 2D echocardiography,” Journal of the American College of Cardiology, vol. 68, no. 21, 2016, pp. 2287-2295, doi:10.1016/j.jacc.2016.08.062.

[23] A. Ghorbani, D. Ouyang, A. Abid, B. He, J. H. Chen, R. A. Harrington and J. Y. Zou, “Deep learning interpretation of echocardiograms,” NPJ digital medicine, vol. 3, no. 1, 2020, pp. 10, doi:10.1038/s41746-019-0216-8.

[24] H. Vaseli, Z. Liao, A. H. Abdi, H. Girgis, D. Behnami, C. Luong and T. Tsang, “Designing lightweight deep learning models for echocardiography view classification,” In Medical Imaging 2019: Image-Guided Procedures, Robotic Interventions, and Modeling, vol. 10951, March. 2019, pp. 93-99. SPIE, doi:10.1117/12.2512913 .

[25] R. Poplin, A. V. Varadarajan, K. Blumer, Y. Liu, M. V. McConnell, G. S. Corrado and D. R. Webster, “Prediction of cardiovascular risk factors from retinal fundus photographs via deep learning,” Nature biomedical engineering, vol. 2, no. 3, 2018, pp. 158-164, doi:10.1038/s41551-018-0195-0.

[26] Y. Hu, L. Guo, B. Lei, M. Mao, Z. Jin, A. Elazab and T. Wang, “Fully automatic pediatric echocardiography segmentation using deep convolutional networks based on BiSeNet,” In 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), July. 2019, pp. 6561-6564, IEEE, doi:10.1109/EMBC.2019.8856457.

[27] S. Leclerc, E. Smistad, A. Østvik, F. Cervenansky, F. Espinosa, T. Espeland and O. Bernard, “LU-Net: a multistage attention network to improve the robustness of segmentation of left ventricular structures in 2-D echocardiography,” IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control, vol. 67, no. 12, 2020, pp. 2519-2530, doi:10.1109/TUFFC.2020.3003403.

[28] G. Zamzmi, S. Rajaraman, L. Y. Hsu, V. Sachdev and S. Antani, “Real-time echocardiography image analysis and quantification of cardiac indices,” Medical image analysis, 80, 2022, 102438, doi:10.1016/j.media.2022.102438.

[29] Y. Deng, P. Cai, L. Zhang, X. Cao, Y. Chen, S. Jiang and B. Wang, “Myocardial strain analysis of echocardiography based on deep learning,” Frontiers in Cardiovascular Medicine, vol. 9, 2022, pp. 1067760, doi:10.3389/fcvm.2022.1067760.

[30] S. Lal, “TC-SegNet: robust deep learning network for fully automatic two-chamber segmentation of two-dimensional echocardiography,” Multimedia Tools and Applications, 2023, pp. 1-1, doi:10.1007/s11042-023-15524-5.

[31] N. Sirjani, S. Moradi, M. G. Oghli, A. Hosseinsabet, A. Alizadehasl, M. Yadollahi and A. Shabanzadeh, “Automatic cardiac evaluations using a deep video object segmentation network,” Insights into Imaging, vol. 13, no. 1, 2022, pp. 1-14, doi:10.1186/s13244-022-01212-9.

[32] N. Bhatt, V. Ramanan, A. Orbach, L. Biswas, M. Ng, F. Guo and N. R. Ghugre, “A deep learning segmentation pipeline for cardiac T1 mapping using MRI relaxation–based synthetic contrast augmentation,” Radiology: Artificial Intelligence, vol. 4, no. 6, 2022, pp. e210294, doi:10.1148/ryai.210294.

[33] Z. Fu, J. Zhang, R. Luo, Y. Sun, D. Deng and L. Xia, “TF-Unet: An automatic cardiac MRI image segmentation method,” Math. Biosci. Eng, vol. 19, 2022, pp. 5207-5222, doi:10.3934/mbe.2022244.

[34] H. Barbaroux, K. P. Kunze, R. Neji, M. S. Nazir, D. J. Pennell, S. Nielles-Vallespin and A. A. Young, “Automated segmentation of long and short axis DENSE cardiovascular magnetic resonance for myocardial strain analysis using spatio-temporal convolutional neural networks,” Journal of Cardiovascular Magnetic Resonance, vol. 25, no. 1, 2023, pp. 1-17, doi:10.1186/s12968-023-00927-y.

[35] R. N. Finnegan, V. Chin, P. Chlap, A. Haidar, J. Otton, J. Dowling and L. Holloway, “Open-source, fully-automated hybrid cardiac substructure segmentation: development and optimisation,” Physical and Engineering Sciences in Medicine, vol. 46, no. 1, 2023, pp. 377-393, doi: 10.1007/s13246-023-01231-w.

[36] M. J. Mortada, S. Tomassini, H. Anbar, M. Morettini, L. Burattini and A. Sbrollini, “Segmentation of Anatomical Structures of the Left Heart from Echocardiographic Images Using Deep Learning,” Diagnostics, vol. 13, no. 10, 2023, pp. 1683, doi:10.3390/diagnostics13101683.

[37] A. I. Shahin and S. Almotairi, “An accurate and fast cardio-views classification system based on fused deep features and LSTM,” IEEE Access, vol. 8, 2020, pp. 135184-135194, doi:10.1109/ACCESS.2020.3010326.

[38] K. Kusunose, A. Haga, M. Inoue, D. Fukuda, H. Yamada and M. Sata, “Clinically feasible and accurate view classification of echocardiographic images using deep learning,” Biomolecules, vol. 10, no. 5, 2020, pp. 665, doi:10.3390/biom10050665.

[39] A. Østvik, I. M. Salte, E. Smistad, T. M. Nguyen, D. Melichova, H. Brunvand and L. Lovstakken, “Myocardial function imaging in echocardiography using deep learning,” ieee transactions on medical imaging, vol. 40, no. 5, 2021, pp. 1340-1351, doi:10.1109/TMI.2021.3054566.

[40] S. Khan, J. Huh and J. C. Ye, “Unsupervised deep learning for accelerated high quality echocardiography,” In 2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI), April. 2021, pp. 1738-1741, IEEE, doi:10.1109/ISBI48211.2021.9433770.

[41] Z. Feng, J. A. Sivak and A. K. Krishnamurthy, “Two-stream attention spatio-temporal network for classification

of echocardiography videos," In 2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI), April. 2021, pp. 1461-1465, IEEE, doi:10.1109/ISBI48211.2021.9433773.

[42] F. T. Dezaki, C. Luong, T. Ginsberg, R. Rohling, K. Gin, P. Abolmaesumi and T.Tsang, "Echo-SyncNet: self-supervised cardiac view synchronization in echocardiography," IEEE Transactions on Medical Imaging, vol. 40, no. 8, 2021, pp. 2092-2104, doi:10.1109/TMI.2021.3071951.

[43] E. Hagberg, D. Hagerman, R. Johansson, N. Hosseini, J. Liu, E. Björnsson and O. Hjelmgren, "Semi-supervised learning with natural language processing for right ventricle classification in echocardiography—a scalable approach," Computers in Biology and Medicine, vol. 143, 2022, pp. 105282, doi:10.1016/j.compbiomed.2022.105282.

[44] E. Evain, Y. Sun, K. Faraz, D. Garcia, E. Saloux, B. L. Gerber and O. Bernard, "Motion estimation by deep learning in 2D echocardiography: synthetic dataset and validation," IEEE transactions on medical imaging, vol. 41, no. 8, 2022, pp. 1911-1924, doi:10.1109/TMI.2022.3151606.

[45] X. Lin, F. Yang, Y. Chen, X. Chen, W. Wang, X. Chen and K. He, "Echocardiography-based AI detection of regional wall motion abnormalities and quantification of cardiac function in myocardial infarction," Frontiers in Cardiovascular Medicine, vol. 9, 2022, pp. 903660, doi:10.3389/fcvm.2022.903660.

[46] A. Darmawahyuni, S. Nurmaini, M. N. Rachmatullah, B. Tutuko, A. I. Sapitri, F. Firdaus and A. Predyansyah, "Deep learning-based electrocardiogram rhythm and beat features for heart abnormality classification," PeerJ Computer Science, vol. 8, 2022, pp. 1-26, doi:10.7717/peerj-cs.825.

[47] L. A. Edwards, F. Feng, M. Iqbal, Y. Fu, A. Sanyahumbi, S. Hao and J. Luo, "Machine learning for pediatric echocardiographic mitral regurgitation detection," Journal of the American Society of Echocardiography, vol. 36, no. 1, 2023, pp. 96-104, doi:10.1016/j.echo.2022.09.017.

[48] M. Tokodi, B. Magyar, A. Soos, M. Takeuchi, M. Tolvaj, B. K. Lakatos and A. Kovács, "Deep learning-based prediction of right ventricular ejection fraction using 2D echocardiograms. JACC: Cardiovascular Imaging, 2023, doi:10.1016/j.jcmg.2023.02.017.