Echo-Based Cardiac Function Assessment

Kerolos Nabil Mounir Fahmy
Scientific Computing Department, Faculty of Computer and Information Science,Ain Shams Unversity
Cairo, Egypt
kirolos20191700460@cis.asu.edu.eg

Kerolos Helal Zaki Soltan
Scientific Computing Department, Faculty of Computer and Information Science,Ain Shams Unversity
Cairo, Egypt
kirolos20191700461@cis.asu.edu.eg

Kerolos Waheed Kamel Abd El-Masih
Scientific Computing Department, Faculty of Computer and Information Science,Ain Shams Unversity
Cairo, Egypt
20201700616@cis.asu.edu.eg

Ganna Muhamed Ahmed Abdo
Scientific Computing Department, Faculty of Computer and Information Science,Ain Shams Unversity
Cairo, Egypt
janah20191700198@cis.asu.edu.eg

Abanoub Gamal Fekry Fawzy
Scientific Computing Department, Faculty of Computer and Information Science,Ain Shams Unversity
Cairo, Egypt
abanob20191700001@cis.asu.edu.eg

Yassa Kamille Abd El-Latif Fetian
Scientific Computing Department, Faculty of Computer and Information Science,Ain Shams Unversity
Cairo, Egypt
l20201700985@cis.asu.edu.eg

***Abstract*— The Left Ventricle Ejection Fraction (LVEF) is an important number that tells how much blood is pumped out of the left ventricle each time the heart beats. As a key indicator for predicting adverse outcomes in patients with conditions like congestive heart failure, post-heart attack, or post-revascularization procedures, accurate assessment of LVEF is essential for managing cardiovascular issues. The identification and treatment of early diseases can be aided by a quick determination of LVEF. Using deep learning models, this study proposes an efficient method for determining the ejection fraction from echocardiography videos to identify heart issues. First, the system divides up the video of the echocardiogram into frames to determine whether the heart is contracting or relaxing. After that, it uses a binary mask to find the left atrium and uses the parameters of the left atrium to figure out the size of the heart. The system then uses the heart size during systole and diastole to calculate the ejection fraction. A threshold method for the ejection fraction value is used to determine whether a person is healthy or sick. With a Dice coefficient similarity of 91.8 percent and a mean square error of 1.69, the proposed system produces effective results. This system meets the needs of doctors by providing an accurate heart failure diagnosis. It is presented as an application that is usable and easy to use. The application has the potential to improve the diagnostic process by assisting physicians in providing patients with timely and precise treatment. This advancement stands to fundamentally improve the nature of cardiovascular consideration.**

Keywords—Left Ventricle, Cardiomyopathy, Ejection Fraction, Deep Learning, Echocardiography.

I. Introduction

The heart, central to the circulatory system, maintains blood flow throughout the body, delivering oxygen and nutrients to tissues and organs while removing carbon dioxide and metabolic waste. Its function is vital for life, operating through a cycle of contraction (systole) and relaxation (diastole). During systole, the heart pumps blood out, and during diastole, it fills with blood for the next contraction [1]. The left ventricle, responsible for pumping oxygenated blood into systemic circulation, is crucial for overall health. Left ventricular failure can impair other organ systems due to insufficient blood flow, affecting oxygen and nutrient delivery.

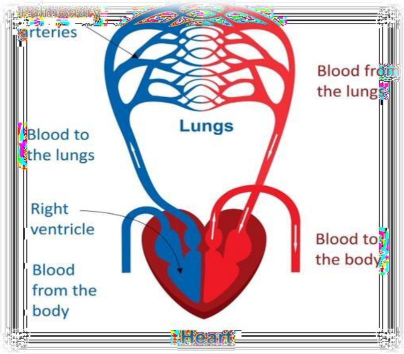


Fig.1.1 Cardiac Cycle

Echocardiography is a pivotal diagnostic tool for cardiomyopathy. This non-invasive imaging test utilizes high-frequency sound waves to produce detailed images of the heart, enabling healthcare providers to assess its structure and function. An echocardiogram can reveal crucial information about the size and shape of the heart chambers, the thickness and movement of the heart walls, and the condition of the heart valves [2].It assesses heart structure and function, providing essential measurements like Left Ventricle Ejection Fraction (LVEF), which quantifies the percentage of blood ejected with each contraction. A normal LVEF ranges from 55% to 70%. Lower values indicate heart failure or cardiomyopathy, while higher values suggest hyperdynamic states or hypertrophic cardiomyopathy.

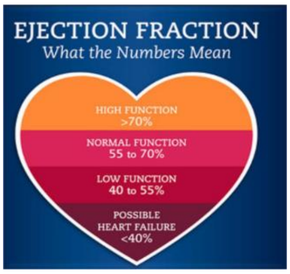
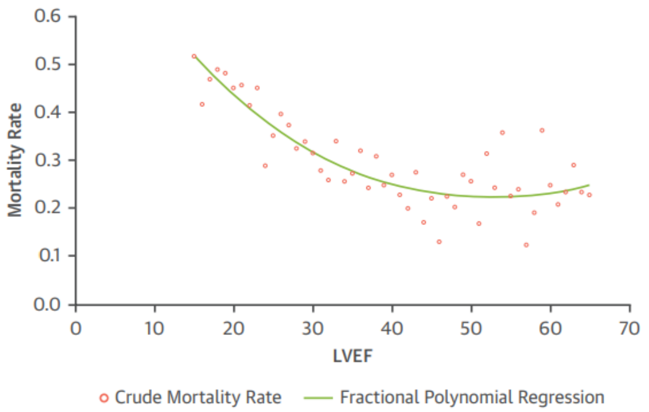


Fig.1.2. Ejection Fraction indicators

Newly graduated and less experienced cardiologists face challenges in diagnosing cardiomyopathies accurately due to the complexity of interpreting echocardiographic data. Misinterpretation can delay appropriate treatment. Advanced diagnostic tools like artificial intelligence (AI) and machine learning algorithms can assist in analyzing echocardiographic data, and providing diagnostic suggestions, enhancing accuracy and speed[3].According to the World Health Organization (WHO), over 500 million individuals worldwide are impacted by cardiovascular diseases, resulting in approximately 20.5 million fatalities in 2021. Reductions in LVEF portend worse cardiovascular outcomes, with an inverse relationship between LVEF and mortality rate plateauing at an EF of 40% to 45%. Ejection fraction assessment variability due to heart rate irregularities and manual tracing challenges leads to high inter-observer differences, ranging from 7.6% to 13.9% [4].

Fig.1. 3Relation between crude mortality rate and fractional Regression



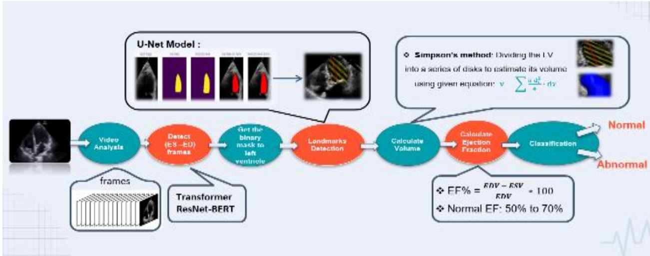
Given the significant impact of cardiovascular diseases, proactive health management is crucial. Regular heart assessments and monitoring reduce risks and improve outcomes. Early detection identifies risk factors and early signs of cardiovascular diseases, allowing for timely interventions. Personalized care through heart assessment apps offers health insights and advice on lifestyle modifications, medication adherence, and preventive measures. Digital health technologies enhance accessibility, allowing individuals to monitor heart health from home, especially in areas with limited healthcare access. Heart assessment apps empower individuals to manage their health proactively, reducing the risk of severe cardiovascular conditions.

II. DATSET

The EchoNet-Dynamic dataset is a comprehensive curated collection of 10,030 apical-4-chamber echocardiography videos that include clinical measurements and expert tracings. Videos that have been processed to remove any unnecessary information and have been standardized to 112x112 pixels have been included in the dataset, which was compiled from routine clinical care at Stanford University Hospital between the years 2016 and 2018. In order to provide insight into cardiac function, key clinical metrics, such as left ventricular ejection fraction, end-systolic volume, and end-diastolic volume, are calculated and annotated on each video. Additionally, the dataset provides valuable ground truth data for machine learning models by including expert tracings of the left ventricle's endocardial border during crucial cardiac phases. EchoNet-Dynamic fills the void left by a lack of publicly available, annotated medical video data, making it possible to create cutting-edge 3D convolutional neural networks and other machine learning architectures and use them as benchmarks. This dataset is ready to work with coordinated effort and development in the clinical local area, at last meaning to improve symptomatic precision and patient results in cardiovascular consideration.

FIG.2.1 PROPOSED SYSTEM ARCHITECTURE

III. METHODOLOGY



This chapter details the methodology for developing a system architecture for the automated assessment of echocardiogram function. The process is divided into distinct phases: video analysis, frame classification, left ventricle segmentation, landmark detection, volume calculation, ejection fraction calculation, and final classification of the echocardiogram video.

Phase 1: Video Analysis: The initial phase involves the preprocessing of echocardiogram videos. Each video is divided into individual frames to facilitate detailed analysis. This segmentation provides a basis for the subsequent classification of frames into specific cardiac cycle phases.

Phase 2: Frame Classification: Using a hyper transformer model that integrates a ResNet autoencoder and a BERT encoder, we classify the frames into three categories: End-Systole (ES), End-Diastole (ED), and transition frames. The ResNet autoencoder is responsible for extracting high-level spatial features from the frames, while the BERT encoder captures temporal dependencies and contextual information across the sequence of frames. This hybrid model enhances the accuracy of frame classification by leveraging both spatial and temporal features.

Phase 3: Left Ventricle Segmentation: After classifying the frames, we perform segmentation to isolate the left ventricle. A U-net model is employed to generate binary masks of the left ventricle for each frame. The U-net architecture is well-suited for medical image segmentation tasks due to its encoder-decoder structure, which allows precise localization and identification of the region of interest.

Phase 4: Landmark Detection: Next, we detect specific landmarks on the left ventricle using the scanline algorithm. This algorithm scans the segmented binary mask to identify anatomical landmarks critical for volume estimation. These landmarks are essential for accurately modeling the left ventricle's shape and size throughout the cardiac cycle.

Phase 5: Volume Calculation: The volume of the left ventricle is calculated using Simpson's method. This method involves approximating the volume of the ventricle by summing the volumes of multiple stacked discs, whose dimensions are derived from the detected landmarks. Simpson's method is widely used in clinical practice due to its reliability in estimating ventricular volume.

Phase 6: Ejection Fraction Calculation: The ejection fraction (EF), a key indicator of cardiac function, is calculated using the formula: EF=(EDV-ESV)/ (EDV) × 100%
where EDV is the end-diastolic volume and ESV is the end-systolic volume. This calculation provides a quantitative measure of the heart's pumping efficiency.
Phase 7: Video Classification: Finally, we classify the echocardiogram video as either normal or abnormal based on the calculated ejection fraction and other relevant parameters. This classification is crucial for diagnostic purposes, assisting clinicians in identifying patients with potential cardiac dysfunctions.
By integrating advanced machine learning models and established medical algorithms, our methodology provides a robust framework for the automated assessment of echocardiogram function, enhancing the accuracy and efficiency of cardiac diagnostics.

Video Analysis

In the video analysis phase, the process begins by converting the video into individual frames using Python code, typically with libraries such as OpenCV. This allows for detailed examination and manipulation of each moment captured in the footage. This frame-by-frame analysis is crucial for achieving precision and clarity in the next phase which is Detecting the ES and ED Frames.

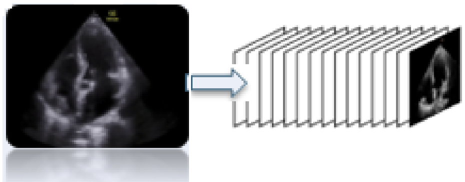


FIG.3.1 CONVERTING 3D VIDEO TO FRAMESET

End-Systolic and End-Diastolic Frame Detection

In order to facilitate batch processing and accurately identify the ES (end-systolic) and ED (end-diastolic) frames, we require fixed-length videos for training. Our dataset contains just single ES and ED outline names per video, leaving many genuine casings unlabeled. We made successions of 128 casings, in view of the run of the mill distances among ES and ED outlines in the preparation information. Groupings longer than 128 edges are down-inspected by an element of 2 [6].
To resolve the issue of unlabeled ES and ED outlines, which would adversely influence transformer execution, we utilized two methodologies are Guidedd Irregular Examining and Reflecting to line up with standard practices for transformer models, guaranteeing all seen ES and ED outlines are marked accurately while keeping up with spatio-transient soundness without presenting void casings.

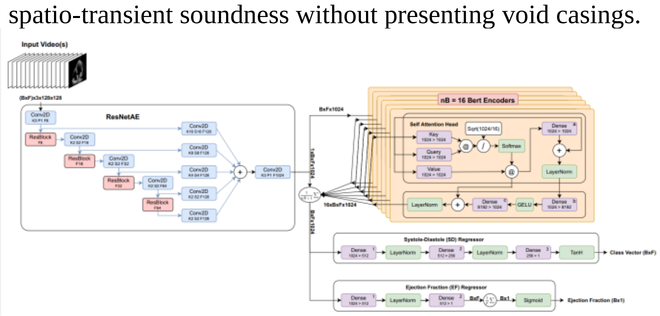


FIG.3.2 RESNETAE + BERT ENCODERS ARCHITECTURE

This Hyper model Consist of 2 Main Parts:
1. ResNetAE is an encoder based on ResNet it consists of 4 Conv2D layers each of them followed by ResBlock and each result of them going to Conv2D layer then summed and going to Conv2D layer.
2. It encodes the feature effectively and reduces the size of the resulted encoded feature to reduce the computation of the followed transformer (BERT Encoders) Part.

Find the Binary Mask of Left Ventricle (LV)
U-Net Network

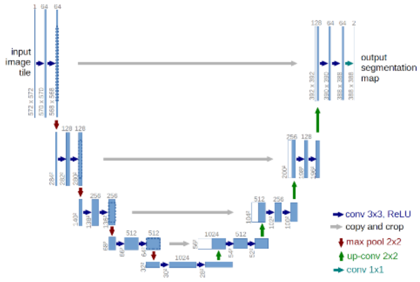


FIG.3.3 U-NET NETWORK ARCHITECTURE

The contracting network is another name for the encoder network. This network attempts to answer our first question, "what" is in the image, by learning a feature map of the input image. With the exception of the fact that in a U-Net, we do not have any fully connected layers at the end, as the output we require now is a mask of the same size as our input image [7], this task is comparable to any classification task we perform with convolutional neural networks.

The expansive network is another name for the decoder network. Our plan is to increase the sample size of our feature maps to match our input image. Using skip connections, this network uses the feature map from the bottleneck layer to create a segmentation mask. The decoder network attempts to respond to our second inquiry—where is the object in the image? There are four decoder blocks in it.

In the model architecture, a grey arrow denotes skip connections. Skip connections enable us to generate our segmentation map by making use of the contextual feature data gathered in the encoder blocks. Through skip connections, our high-resolution encoder block features will be used to help us project our feature map (the output of the bottleneck layer).

Left Ventricle Landmark Detection

Recognizing the Vital Vertical Line First Track down the Upper Point: Distinguish the most extreme y-coordinate in the twofold veil. The vertical line's highest point will be at this point. Second Look for the Key Points: The leftmost bottom point can be obtained by locating the binary mask's minimum (x, y) coordinate. To determine the rightmost bottom point, look for the point with the maximum x-coordinate and the minimum y-coordinate (x, y). From these two points, determine the midpoint. Third Check to see that every point is within the binary mask: Verify that the upper point, the point closest to the left, the point closest to the right, and the midpoint are all components of the binary mask.

Choosing the Most Important Horizontal Lines First Divide the vertical line into twenty equal parts: Part the upward line acquired in the initial step into 20 equivalent fragments. Second Include Parallel Lines: Draw a line that is perpendicular to the vertical line at each division point. Make certain that these lines reach the binary mask's edges.

Adding Points of Interest First Recognize Edges of the Double Veil: At every opposite line, decide the convergence focuses with the twofold veil edges. Second Mark Important Points: As landmarks, mark the intersection points from the previous step. This will produce 21 lines or 42 points, which include the ends of the vertical line and the points where the perpendicular lines meet the mask edges.

Using Simpson's method, you can precisely determine the volume of the left ventricle from a binary mask by following these steps. For an accurate volume calculation to obtain the EF from Simpson's Method by obtaining the ED and ES volumes, this method requires the precise identification of key points and the careful division of the mask into segments.

Left Ventricle Volume Calculation

we only have A4C measurements available, the monoplane Simpson's method is used, which only makes use of A4C measurements. The contour points, also known as volume tracings, are derived from the segmented mask that represents the left ventricle at the end-diastolic and end-systolic frames, respectively, in order to extract features from the segmented images. The mitral valve and apex points are identified during the localization process, which makes it possible to identify the left ventricle's major axis [8].

The distance between these two points is the major axis. The disks are then spaced at the positions obtained by dividing the major axis into 20 equal parts and placing them orthogonally to it. A disk shape centered at each disk position is used to define a region of interest (ROI), and the pixels within this ROI are extracted. The maximum distance between any two points within the ROI is calculated to determine each disk's diameter. The segmented images and their associated feature extractions at the end-systole and end-diastole, respectively, are depicted in Fig. 3.4.

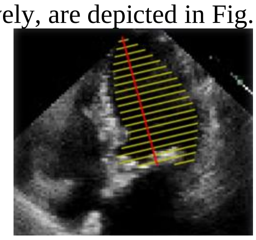


FIG.3.4 LEFT VENTRICLE LANDMARKS

$$LV\ Volume = \sum \frac{\pi d_i^2}{4} .dx$$

Left Ventricle Ejection Fraction (LVEF) Calculation

One important metric for assessing cardiac function is the left ventricular (LV) ejection fraction (EF). Eq. 1 shows the percentage change in volume between the end diastolic volume (EDV) and the end systolic volume (ESV). EF is a measure of the heart's efficiency because it shows how much left ventricle volume is pumped out in each heartbeat. It has been demonstrated that a low, or even slightly reduced, EF is a strong indicator of heart disease [9].

$$LVEF\% = 100 * \frac{EDV - ESV}{EDV}$$

Threshold Classification

As recommended by the American Society of Echocardiography and the European Association of Cardiovascular Imaging [10], we obtained LVEF using straightforward thresholding Normal ranges for two-dimensional echocardiography.

LVEF (%) among the male population:

- 52% to 72% normal range
- 41% to 51 mildly abnormal
- 30% to 40% moderately abnormal
- Less than 30% severely abnormal

LVEF (%) among the female population:

- 54% to 74% normal range
- 41% to 53 mildly abnormal
- 30% to 40% moderately abnormal
- Less than 30% severely abnormal

IV. EXPERIMENTAL RESULTS

The Echo-Net dataset consists of 10,030 apical-4-chamber echocardiography videos from Stanford University Hospital, recorded between 2016 and 2018. Each video, standardized to 112x112 pixels, shows the heart from different angles and has been processed to remove extraneous information. For each video, the left ventricle is traced at end-systole and end-diastole, with coordinates provided for these tracings. These coordinates include the length and direction of the long axis and short axis distances from the heart apex to the mitral apparatus, linked to the video file name and frame number. The dataset also includes clinical measurements verified by sonographers and echocardiographers, with a focus on Left Ventricular Ejection Fraction (LVEF), a key metric for cardiac function calculated as (EDV−ESV)/EDV(EDV - ESV) / EDV(EDV−ESV)/EDV. This dataset supports the development of algorithms for cardiac function analysis and diagnostics. [11]

Table 1 Echo-Net dataset label variables

Variable	Description
FileName	Hashed file name used to link videos, labels, and annotations
EF	Ejection fraction calculated from ESV and EDV
ESV	End systolic volume calculated by method of discs
EDV	End diastolic volume calculated by method of discs
Height	Video Height
Width	Video Width
FPS	Frames Per Second
NumFrames	Number of Frames in whole video
Split	Classification of train/validation/test sets used for benchmarking

he primary measurement criteria for segmentation in the experiments are the Dice-Sørensen coefficient (DSC) and Mean Square Error (MSE). The DSC measures the similarity between two samples and is calculated as **DSC** = $\frac{2|\boldsymbol{A} \cap \boldsymbol{B}|}{|\boldsymbol{A}|+|\boldsymbol{B}|}$, where |A| and |B| are the sizes of the two sets. For segmentation, the average DSC values were 92.7 for diastolic and 90.9 for systolic, resulting in a total average of 91.8.

MSE, which measures the average squared difference between estimated and actual values, is calculated as **MSE =** $\frac{1}{n}\sum_{i=1}^{n}(y_i-\hat{y}_i)^2$ The average MSE values were 1.88 for diastolic and 1.51 for systolic, resulting in a total average of 1.69. Classification accuracy is evaluated by the average accuracy metric, calculated as **Average Accuracy** = $\frac{1}{n}\sum_{i=1}^{n}\frac{\textbf{Number of Correct Predictions in i−th Epoch}}{\textbf{Totlal Number of Predictions in i−th Epoch}}$, which provides an overall performance measure across different epochs.

The Left Ventricular Ejection Fraction (LVEF) is a crucial metric for assessing cardiac function, defined as the percentage of blood pumped out of the left ventricle with each heartbeat, calculated as LVEF = $100\times\frac{EDV-ESV}{EDV}$. LVEF helps diagnose heart efficiency, with low values indicating potential heart disease. For classification, we used simple thresholding based on normal ranges from the American Society of Echocardiography and the European Association of Cardiovascular Imaging. For males, an LVEF of 52-72% is normal, 41-51% is mildly abnormal, 30-40% is moderately abnormal, and below 30% is severely abnormal. For females, normal ranges are 54-74%, mildly abnormal is 41-53%, moderately abnormal is 30-40%, and severely abnormal is below 30%. This system enables doctors to efficiently assess and diagnose heart conditions based on gender-specific LVEF thresholds.

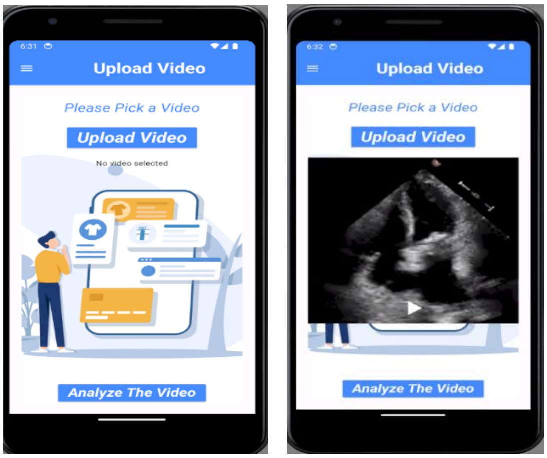


FIG. 3.5 UPLOAD WINDOW

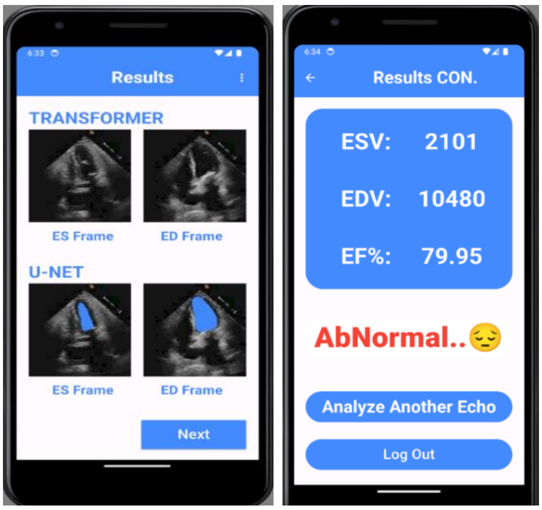


FIG. 3.6 RESULT WINDOW

V. PREPROCESSING

The video analysis process begins by converting the video into individual frames. This step, typically executed using Python and libraries such as OpenCV, enables detailed examination and manipulation of each captured moment. Converting the video into discrete frames achieves the precision and clarity needed for subsequent phases, such as detecting End-Systolic (ES) and End-Diastolic (ED) frames.

In deep learning, preprocessing is crucial for preparing raw data for effective training. This involves transforming the data into a suitable format for learning, including tasks such as numerical normalization, handling missing data, encoding categorical variables, and augmenting data to enrich training samples. Preprocessing enhances data quality by reducing noise and ensures models can efficiently discern underlying patterns. Techniques like scaling, normalization, and data augmentation further bolster model accuracy and generalization, aligning input data with real-world scenarios.

For video analysis, where fixed-length sequences are required for batch processing and accurate identification of ES and ED frames, the challenge of insufficiently labeled frames per video arises. To address this, two innovative strategies are employed: Guided Random Sampling and Mirroring. Guided Random Sampling involves sampling labeled frames and intermediate frames within specified distances. To maintain the integrity of transformer attention, sequences are padded to 128 frames with masked black frames, ensuring the model can effectively focus on relevant frames while maintaining a consistent sequence length. Mirroring extends the transition frames between labeled frames to ensure temporal coherence in the sequences. Additionally, random cropping is used to maintain sequence length adherence. This approach enhances transformer performance by accurately labeling ES and ED frames seen during training, while maintaining coherence and avoiding empty frames.

By employing these strategies, the accuracy and reliability of the deep learning model in identifying key frames in video sequences improve. The combination of Guided Random Sampling and Mirroring ensures that the model can handle the variability and complexity inherent in video data, leading to more precise and robust cardiac function assessment.
Features
The proposed approach utilizes a novel transformer model designed to interpret echo videos of variable lengths with high accuracy and reasoning capability. The methodology comprises three main components: (a) an initial encoder for dimensionality reduction using a ResNetAE to distill echo frames into compact embeddings, (b) a Bidirectional Encoder Representations from Transformers (BERT)-based module for spatio-temporal reasoning, integrating self-attention and attention mechanisms within the transformer architecture, and (c) two regression modules to predict crucial cardiac parameters, specifically end-systolic (ES) and end-diastolic

(ED) frames, as well as Left Ventricular Ejection Fraction (LVEF).

To manage computational complexity effectively, the echo frames undergo dimensionality reduction through the ResNetAE encoder before being input into the transformer model. The transformer, along with the initial encoder, is trained end-to-end to optimize model weights and enhance prediction accuracy. Post feature extraction by BERT encoders, outputs are averaged with those from the ResNetAE embeddings and passed to the regression modules. These regressors predict ES and ED frame indices and estimate LVEF. LVEF estimation involves dimensionality reduction for each frame, averaging predictions, and incorporating regularization techniques to balance training objectives.

The overall loss function combines Mean Squared Error (MSE) and Mean Absolute Error (MAE) metrics, augmented by a regularization term to prioritize accurate LVEF predictions. This comprehensive methodology ensures robust performance in accurately estimating critical cardiac parameters from echo videos, leveraging the strengths of both convolutional and transformer-based architectures.

VI. Features

The proposed approach utilizes a novel transformer model designed to interpret echo videos of variable lengths with high accuracy and reasoning capability. The methodology comprises three main components: (a) an initial encoder for dimensionality reduction using a ResNetAE to distill echo frames into compact embeddings, (b) a Bidirectional Encoder Representations from Transformers (BERT)-based module for spatio-temporal reasoning, integrating self-attention and attention mechanisms within the transformer architecture, and (c) two regression modules to predict crucial cardiac parameters, specifically end-systolic (ES) and end-diastolic (ED) frames, as well as Left Ventricular Ejection Fraction (LVEF).

To manage computational complexity effectively, the echo frames undergo dimensionality reduction through the ResNetAE encoder before being input into the transformer model. The transformer, along with the initial encoder, is trained end-to-end to optimize model weights and enhance prediction accuracy. Post feature extraction by BERT encoders, outputs are averaged with those from the ResNetAE embeddings and passed to the regression modules. These regressors predict ES and ED frame indices and estimate LVEF. LVEF estimation involves dimensionality reduction for each frame, averaging predictions, and incorporating regularization techniques to balance training objectives.

The overall loss function combines Mean Squared Error (MSE) and Mean Absolute Error (MAE) metrics, augmented by a regularization term to prioritize accurate LVEF predictions. This comprehensive methodology ensures robust performance in accurately estimating critical cardiac parameters from echo videos, leveraging the strengths of both convolutional and transformer-based architectures.

VII. Detection and Classification

In this study, the researchers employed a U-Net architecture for left ventricle (LV) area segmentation in cardiac images. The unique design of the U-Net, characterized by an encoder-decoder structure with skip connections, facilitated accurate segmentation by addressing both "what" and "where" questions in image analysis. The encoder network, featuring multiple convolutional layers and max pooling, learned high-level feature representations of input frames, which were crucial for identifying LV boundaries. The bottleneck layer then condensed these features, serving as a bridge to the decoder network. This expansive network utilized transpose convolutions and skip connections to up sample feature maps and refine segmentation masks, leveraging contextual information from earlier encoding stages. The final 1x1 convolution with sigmoid activation produced pixel-wise classification, yielding precise LV area masks aligned with the input dimensions. The researchers' approach demonstrates the efficacy of U-Net in medical image segmentation tasks, showcasing its ability to integrate feature extraction and spatial localization effectively.

The study also proposes a method for accurately quantifying the left ventricular volume from a binary mask using Simpson's method, which is essential for calculating the ejection fraction (EF). Initially, the vertical line of the left ventricle is identified by locating its uppermost point and the leftmost and rightmost bottom points, followed by determining their midpoint. These points are verified to ensure they lie within the binary mask. Subsequently, the vertical line is divided into 20 equal segments, and perpendicular lines are drawn at each division point to intersect the edges of the binary mask. These intersection points serve as landmark points, totaling 21 lines or 42 points, which include the endpoints of the vertical line and the intersections of the perpendicular lines with the mask edges. This approach ensures precise delineation of the left ventricle's volume, crucial for accurate EF calculation using Simpson's method.

Due to the availability of only A4C measurements, the monoplane Simpson's method was employed in this study. This method utilizes contours derived from segmented images of the left ventricle at end-diastolic and end-systolic frames. The segmentation process involves extracting contour points (volume tracings) to delineate the left ventricle. Mitral valve and apex points are localized to establish the major axis of the ventricle, defined as the distance between these points. Disks perpendicular to the major axis are positioned at intervals derived by dividing the axis into 20 equal parts. Each disk defines a region of interest (ROI), and pixel data within each ROI are extracted. The diameter of each disk is determined by computing the maximum distance between any two points within its ROI.

Left ventricular ejection fraction (EF) is a critical metric assessing cardiac function, defined as the percentage change in volume between end-diastolic volume (EDV) and end-systolic volume (ESV), expressed as $EF = [(EDV - ESV) / EDV] * 100\%$. It quantifies the efficiency of the heart's pumping action, indicating the proportion of blood ejected with each heartbeat. Even a slight decrease in EF is strongly correlated with cardiovascular disease, making EF a pivotal clinical indicator. According to the American Society of Echocardiography and the European Association of Cardiovascular Imaging, normal ranges for left ventricular ejection fraction (LVEF) obtained through two-dimensional echocardiography differ between males and females. For males, an LVEF of 52% to 72% is considered normal, 41% to 51% is mildly abnormal, 30% to 40% is moderately abnormal, and less than 30% is severely abnormal. For females, an LVEF of 54% to 74% is normal, 41% to 53% is mildly abnormal, 30% to 40% is moderately abnormal, and less than 30% is severely abnormal.

VIII. Results and Discussions

Accurately assessing left ventricular systolic function is vital for diagnosing and treating cardiovascular diseases, with LVEF and GLS being key metrics. Echocardiography is the preferred imaging method due to its non-invasiveness and real-time processing capabilities. However, it relies heavily on the sonographer's expertise, resulting in variability. AI can mitigate these issues by automating the extraction of cardiac features and estimating ventricular volume and motion from echocardiographic images. This review discusses recent advancements in AI for assessing cardiac function and diagnosing heart diseases, highlighting the promises and challenges of this technology. Machine learning and deep learning, subsets of AI, enhance this process by learning from data and improving over time, with deep learning particularly excelling in complex medical image analysis

Based on the image segmentation, researchers have devoted themselves to developing fully automatic models to assess LVEF.

Ouyang et al. [5] introduced EchoNet Dynamic, a video-based AI model for evaluating cardiac function. Using apical four-chamber views from 10,030 patients, they applied a weak supervision approach with annotations for end-systole and end-diastole frames. This method reduced labeling costs and utilized an atrous convolution model to provide frame-level semantic segmentation throughout the cardiac cycle. EchoNet-Dynamic accurately segmented the left ventricle, achieving a Dice similarity coefficient over 0.9, and created left ventricular volume curves to assess LVEF. External validation with 2,895 patients showed that automatic measurements matched expert assessments and were highly repeatable. The study also addressed clinical challenges, such as measuring cardiac function in patients with arrhythmias and variations in video quality and imaging conditions, demonstrating EchoNet-Dynamic's robustness under different scenarios.

Grant Duffy et al. [6] proposed a method that enhances segmentation by predicting the z-depth of the left ventricle (LV) at each pixel, setting any pixels outside the LV to zero. Figure 3 illustrates an example LV depth map, which is generated for every frame of an echocardiogram video. The volume for each frame is calculated by summing the pixel depths, and these volumes are used to identify end-systole and end-diastole frames, from which the ejection fraction (EF) is calculated. The depth maps can be visualized as a 3D surface over the LV, akin to performing a regression on depth for each pixel rather than a classification as in segmentation models. Convolutional neural networks (CNNs) have been commonly used to predict depth information, a prevalent approach in computer vision and autonomous AI. Summing the pixel depths to calculate the volume aligns with human methods, making it easily interpretable.

Xin Liu et al. [7] introduced a new feature extraction and fusion strategy to improve the accuracy of automatic LVEF assessment using multiview 2-D echocardiographic sequences. The DPS-Net model demonstrated high diagnostic performance in determining heart failure across various heart disease phenotypes and achieved excellent results in left ventricle segmentation, indicating its potential for broader application in interpreting 2D echocardiographic images.

Ashley P. Akerman et al. [8] developed a novel AI model for detecting HFpEF using a single routinely acquired TTE video clip. This model accurately identified HFpEF, produced fewer nondiagnostic outputs compared to current clinical scores, and identified patients with poorer survival rates. The classifier's application in HFpEF screening, especially when the diagnosis is uncertain, could automate the detection process for this complex clinical syndrome, ensuring more patients receive a correct and timely diagnosis.

Anesth Analg et al. [9] reported that 0.41% of adult patients surviving major non-cardiac surgery were newly diagnosed with HFrEF within 730 days postoperatively. Using machine learning, the study modeled patterns in preoperative data and intraoperative response profiles to detect early-stage HFrEF. The technique of segmenting intraoperative records by overlapping anesthetic and surgical interventions could be valuable for analyzing granular intraoperative data. The findings suggest potential for developing perioperative systems for early HFrEF diagnosis and management. However, further studies are needed to externally validate detection algorithms, assess their feasibility within EHRs at the point of care, and understand the clinical effectiveness of such decision support tools.

Francesc Formiga et al. [10] demonstrated that HFpEF is a prevalent condition linked with high morbidity and mortality. Diagnosing HFpEF is challenging due to the numerous comorbidities that can mimic it. Additionally, HFpEF is diagnosed based on a cluster of parameters, including elevated natriuretic peptide levels and specific echocardiographic changes, rather than a single criterion. The authors present an easily applicable algorithm for real-world patients to help confirm or rule out HFpEF diagnoses.

Sunita Pokhrel Bhattarai et al. [11] conducted a comprehensive analysis of ECG features to predict reduced LVEF, emphasizing the potential of automated 12-lead ECG analysis for improving heart failure diagnosis. Using the LASSO model, they identified key ECG features significantly contributing to the estimation of LVEF <30%, achieving high sensitivity and moderate specificity. The study acknowledges limitations due to its retrospective design and stresses the need for validation in broader, prospective clinical trials. Nonetheless, the results highlight the promise of ECG-based diagnostic tools and AI in cardiology, potentially enabling more accurate, real-time assessments of cardiac function and risk stratification, which could improve patient outcomes.

Bing Feng et al. [12] introduced a method to calculate LVEF without edge detection. In computer simulations, their method performed better for small LVs compared to the QGS method, which tended to overestimate LVEF. In patient studies, both methods yielded similar results for large hearts, but the new method proved more accurate for small hearts, with QGS overestimating LVEF by over 9%. This method enhances the accuracy of functional parameter calculations for patients with small hearts. While the study focused on LVEF, extending this method to other parameters, such as heart wall thickness, would be feasible.

According to the current studies in the literature presented in the previous section, the following observations can be made:
Dataset Size: The size of the dataset plays a crucial role in classification performance. The EchoNet-Dynamic dataset, which contains a large number of 3D videos, is more complex and has yielded higher performance.
Classification Model: Deep learning models are versatile and suitable for domain-specific problems, making them the most widely used approach.

VII. CONCLUSION AND FUTURE WORK

This project addresses a critical issue facing heart failure patients. The objective was to develop software that can determine the ejection fraction of a heartbeat from echocardiography videos, which was successfully achieved using deep learning models (DL). Throughout the development process, we utilized a variety of programming languages, development tools, and third-party services to ensure that our application is both user-friendly and reliable. Additionally, we developed comprehensive technical and user documentation to support the development and deployment process. Our goal is to create a system that meets the needs of doctors everywhere, aiding them in providing accurate diagnoses of heart failure to patients. We believe that our app has the potential to make a positive impact on the diagnosis process. We are excited to see it in action, helping more people monitor their condition and ultimately saving lives. In the future, Efforts should be made to enhance the performance of the proposed system without increasing the processing time. This could involve optimizing algorithms, utilizing more efficient hardware, or applying advanced data processing techniques. The system should be integrated with echocardiogram devices to provide real-time results to doctors, minimizing the need for additional operations or manual input from healthcare professionals. It is essential to make this system available on devices in hospitals worldwide, ensuring broad accessibility and standardization of care across different regions. Developing a mobile application to track patient progress is crucial. This app should offer features for monitoring vital signs, scheduling follow-up appointments, and providing educational resources to patients. Once all these improvements are implemented, the application can be launched to assist public and private hospitals, patients, and doctors in making accurate and timely diagnoses. This comprehensive approach will improve healthcare delivery and patient outcomes.

REFERENCES

[1] Lang, R. M. et al. Recommendations for cardiac chamber quantification by echocardiography in adults: an update from the American Society of Echocardiography and the European Association of Cardiovascular Imaging. Eur. Heart J. Cardiovasc.Imaging 16, 233–271 (2015).

[2] McMurray, J. J. et al. ESC guidelines for the diagnosis and treatment of acute and chronic heart failure 2012. Eur. J. Heart Fail. 14, 803–869 (2012).

[3] Loehr, L. R., Rosamond, W. D., Chang, P. P., Folsom, A. R. & Chambless, L. E. Heart failure incidence and survival (from the Atherosclerosis Risk in Communities study). Am. J. Cardiol. 101, 1016–1022 (2008).

[4] Bui, A. L., Horwich, T. B. & Fonarow, G. C. Epidemiology and risk profile of heart failure.Nat. Rev. Cardiol. 8, 30–41 (2011).

[5] Hadrien Reynaud, Athanasios Vlontzos, Benjamin Hou, Arian Beqiri,Paul Leeson, and Bernhard Kainz,” Ultrasound Video Transformers for Cardiac Ejection Fraction Estimation”

[6] Leclerc, S.; Smistad, E.; Pedrosa, J.; Ostvik, A.; Cervenansky, F.; Espinosa, F.; Espeland, T.; Berg, E.A.R.; Jodoin, P.-M.; Grenier, T.; et al. Deep Learning for Segmentation Using an Open Large-Scale Dataset in 2D Echocardiography. IEEE Trans. Med. Imaging 2019, 38, 2198–2210.

[7] Leclerc, S.; Smistad, E.; Ostvik, A.; Cervenansky, F.; Espinosa, F.; Espeland, T.; Rye Berg, E.A.; Belhamissi, M.; Israilov, S.; Grenier,T.; et al. LU-Net: A Multistage Attention Network to Improve the Robustness of Segmentation of Left Ventricular Structures in 2-D Echocardiography. IEEE Trans. Ultrason. Ferroelectr. Freq. Control. 2020, 67, 2519–2530.

[8] Ouyang, D.; He, B.; Ghorbani, A.; Yuan, N.; Ebinger, J.; Langlotz, C.P.; Heidenreich, P.A.; Harrington, R.A.; Liang, D.H.; Ashley,E.A.; et al. Video-based AI for beat-to-beat assessment of cardiac function. Nature 2020, 580, 252–256.

[9] Asch, F.M.; Poilvert, N.; Abraham, T.; Jankowski, M.; Cleve, J.; Adams, M.; Romano, N.; Hong, H.; Mor-Avi, V.; Martin, R.P.; et al.Automated Echocardiographic Quantification of Left Ventricular Ejection Fraction without Volume Measurements Using a Machine Learning Algorithm Mimicking a Human Expert. Circ. Cardiovasc. Imaging 2019, 12, e009303.

[10] Zhang, J.; Gajjala, S.; Agrawal, P.; Tison, G.H.; Hallock, L.A.; Beussink-Nelson, L.; Lassen, M.H.; Fan, E.; Aras, M.A.; Jordan,C.; et al. Fully Automated Echocardiogram Interpretation in Clinical Practice. Circulation 2018, 138, 1623–1635.

[11] https://stanfordaimi.azurewebsites.net/datasets/834e1cd1-92f7-4268-9daa-d359198b310a

[12] Ouyang, David & He, Bryan & Ghorbani, Amirata & Yuan, Neal & Ebinger, Joseph & Langlotz, Curtis & Heidenreich, Paul & Harrington, Robert & Liang, David & Ashley, Euan & Zou, James. (2020). Video-based AI for beat-to-beat assessment of cardiac function. Nature. 580. 10.1038/s41586-020-2145-8.

[13] Duffy G, Jain I, He B, Ouyang D. Interpretable deep learning prediction of 3d assessment of cardiac function. Pac Symp Biocomput.27:231-241. (2022)

[14] Liu, Xin & Fan, Yiting & Li, Shuang & Chen, Meixiang & Li, Ming & Hau, William & Zhang, Heye & Xu, Lin & Lee, Alex. Deep learning-based automated left ventricular ejection fraction assessment using 2D echocardiography. American Journal of Physiology-Heart and Circulatory Physiology. (2021)

[15] Ashley P. Akerman, Mihaela Porumb, Christopher G. Scott, Arian Beqiri, Agisilaos Chartsias. Automated Echocardiographic Detection of Heart Failure With Preserved Ejection Fraction Using Artificial Intelligence, JACC Adv. 2 (6), (2023)

[16] Anesth Analg, Mathis MR, Engoren MC, Joo H, Maile MD, Aaronson KD, Burns ML, Sjoding MW, Douville NJ, Janda AM, Hu Y, Najarian K, Kheterpal S. Early Detection of Heart Failure With Reduced Ejection Fraction Using Perioperative Data Among Noncardiac Surgical Patients: A Machine-Learning Approach. (2020)

[17] Francesc Formiga, Julio Nuñez, María José Castillo Moraga, Marta Cobo Marcos, María Isabel Egocheaga, Concha F. García-Prieto, Angel Trueba-Sáiz, Arantxa Matalí Gilarranz & José María Fernández Rodriguez. Diagnosis of heart failure with preserved ejection fraction: a systematic narrative review of the evidence. Heart Failure Reviews. 29, pp. 179–189, (2024)

[18] Bhattarai, Sunita & Dzikowicz, Dillon & Xue, Ying & Block, Robert & Tucker, Rebecca & Bhandari, Shilpa & Boulware, Victoria & Stone, Breanne & Carey, Mary. Estimating Ejection Fraction from the 12 Lead ECG among Patients with Acute Heart Failure. medRxiv : the preprint server for health sciences. (2024)

[19] Feng, Bing & Sitek, Arkadiusz & Gullberg, Grant. Calculation of the left ventricular ejection fraction without edge detection: Application to small hearts. Journal of nuclear medicine: official publication, Society of Nuclear Medicine. 43. 786-94, (2002).